Image analysis in coagulation process control

Bildeanalyse anvendt i prosesskontroll av koagulering

Nataliia Sivchenko
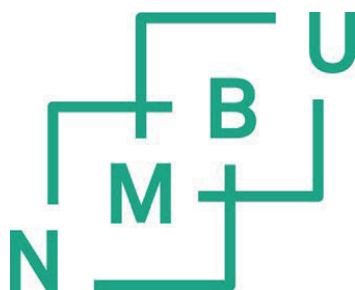
Image analysis in coagulation process control

Bildeanalyse anvendt i prosesskontroll av koagulering

Philosophiae Doctor (PhD) Thesis

Nataliia Sivchenko

Norwegian University of Life Sciences
Faculty of Science and Technology

Ås (2017)

Supervisory team

Harsha Ratnaweera, Professor (main supervisor)
Faculty of Science and Technology
Norwegian University of Life Sciences

Knut Kvaal, Professor (co-supervisor)
Faculty of Science and Technology
Norwegian University of Life Sciences

Evaluation committee

Volodymyr Tarabara, Professor (first opponent)
Department of Civil and Environmental Engineering
Michigan State University

Maths Halstensen, Associate Professor (second opponent)
Faculty of Technology, Natural Sciences and Maritime Sciences
Department of Electrical Engineering, Information Technology and Cybernetics
University College of Southeast Norway

Ingunn Burud, Associate Professor (committee coordinator)
Faculty of Science and Technology
Norwegian University of Life Sciences

Abstract

Coagulation-flocculation is a conventional process in drinking water, municipal and industrial wastewater treatment. This stage of wastewater treatment has become popular in Norway due to its superior ability to remove particles and phosphates. However, a conceptual model of the process still does not exist in the sense of an established and universally accepted model, due to the complex nature of the system. The treatment efficiency and economics of the process are mainly based on chemical (coagulants and flocculants) consumption, thereby emphasising the need for optimal coagulant dosing; in addition to reaching the required treatment efficiencies, optimal coagulant dosing reduces the operational costs related to chemicals and waste management as a result of reduced amounts of sludge. Optimisation of the process highly depends on the dosage control concept. A vast majority of treatment plants operate the process with a flow-proportional dosage concept, at best combined with pH, despite the fact that the optimal coagulant dosage highly depends on particles and phosphates. Innovative concepts that include these parameters, directly or indirectly, have reported savings of up to 30 % of coagulants.

Despite the potential savings, the limited use of such systems may be explained partly due to the lack of a conceptual model and partly due to high costs of appropriate multi-parameter dosage control strategies. Additionally, not all the desired parameters can be measured affordably online, nor can they be used in feed-back control strategies due to the long time lag between the inlet and outlet measurements. Thus, sensors that can be placed inside the flocculation chamber to evaluate the process *in situ* would significantly increase the efficiency of the existing feed-forward systems. This would help to improve the dosage prediction models, as well as decrease the number of expensive and complicated sensors in existing multi-parameter based systems. The author's PhD work concentrated on solving these issues. Such sensor concepts also contribute to the prediction of outlet qualities, which is another challenge in the water industry.

Thus, the objective of this PhD work was to address the above challenges and needs by developing a low-cost sensor prototype based on image analysis of the flocs – particles aggregating during the coagulation-flocculation process. Different image analysis techniques were evaluated including conventional particle recognition methods and texture analysis methods, which are broadly used in other fields than water treatment. Texture image analysis methods were found to be a successful solution to challenges associated with wastewater flocs.

The concept of characterising flocs images by texture image analysis techniques was first tested in the laboratory scale batch process (jar tests) with model wastewater. The non-intrusive image acquisition system was established to capture images of flocs during the slow mixing stage of the coagulation process. It was proven that the images of flocs have distinct texture features correlating to the coagulation conditions (type and amount of chemicals, time after the start of a slow mixing phase, etc.) and inlet wastewater parameters. The changes in flocs images, coagulant dosages and treatment efficiencies were studied. The correlations between textural features of the flocs images, coagulant dosages and treatment efficiencies were found using multivariate statistics.

After successful laboratory studies, the full-scale experiments were conducted in Skiphelle municipal wastewater treatment plant (Drøbak, Norway). The texture image analysis concept was proven to be applicable for real municipal wastewater flocs. The changes in inlet wastewater parameters and coagulation conditions were traceable with the images of flocs. The system was proven to predict the outlet turbidity values, which can potentially be used for troubleshooting as an early indication of coagulation failure.

The low-cost floc sensor prototype, consisting of a single-board computer and camera module, was developed and tested at the same municipal wastewater treatment plant. Customised software was written to control the camera and adjust settings. The investigations documented that the images of flocs captured by the low-cost camera module could be used for optimal coagulant dosage predictions.

Overall, the results of this PhD work confirmed the potential for the floc sensor to be a stand-alone online digital image analysis device that could increase the accuracy and affordability of the existing multi-parameter based coagulant dosage control systems.

Sammendrag

Koagulering-flokkulering, også kjent som kjemisk felling, er en konvensjonell prosess i behandling av drikkevann og kommunalt og industrielt avløpsvann. Prosessen er blitt svært populær for avløpsbehandling i Norge på grunn av høy rensegrad for partikler og fosfor. Imidlertid finnes det fortsatt ingen etablert og universelt akseptert konseptuell modell for prosessen, noe som må tilskrives prosessens iboende kompleksitet. Renseeffektiviteten og økonomien i prosessen er i hovedsak kontrollert av kjemikalieforbruket (koagulanter og flokkulanter). Dette understreker behovet for optimal dosering av kjemikalier, som i tillegg til å sørge for at rensekravene overholdes, vil redusere driftskostnadene knyttet til kjemikalier og slambehandling pga. mindre slamproduksjon. Optimering av prosessen er nært knyttet til styring og regulering av kjemikaliedoseringen. De aller fleste avløpsrenseanlegg benytter i dag mengde-proporsjonal kjemikaliedosering, i beste fall kombinert med pH overstyring, til tross for at den optimale doseringen er sterkt avhengig av vannets innhold av partikler og fosfor. Innovative tilnærminger som inkluderer disse parameterne direkte eller indirekte i styringen av doseringen, har kunnet vise til en reduksjon kjemikalieforbruket på opptil 30 %.

Til tross for besparelsene er slike modeller lite brukt i praksis, noe som kan forklares delvis med mangelen på en konseptuell modell og delvis med de høye kostnadene knyttet til å etablere doseringsstyring basert på flere parametere. Dessuten kan ikke alle de ønskede parameterne måles i sanntid til en overkommelig kostnad, og de kan heller ikke benyttes i en *feed-back* styringsstrategi på grunn av strømningstiden som gir en tidsforsinkelse mellom målinger i innløpet og utløpet til renseanlegget. Sensorer som kan plasseres i flokkuleringskammeret og benyttes til å evaluere prosessen *in situ* vil derfor kunne vesentlig øke effektiviteten til eksisterende *feed-forward* systemer. Dette ville bidra til å forbedre modeller for doserings-prediksjon, og redusere antall kostbare og kompliserte sensorer i eksisterende multi-parameter-baserte systemer. Doktorgradsarbeidet har hatt fokus på å løse disse problemene. Slike sensorkonsepter vil også kunne bidra til å predikere vannkvalitet i utløpet av renseanlegget, som er en annen utfordring i vannbransjen.

Målet med dette doktorgradsarbeidet har altså vært å løse utfordringene nevnt ovenfor ved å utvikle en lav-kostnads prototype for en sensor basert på billedanalyse av fnokker – partiklene som aggregerer i koagulering-flokkulerings-prosessen. Forskjellige teknikker for billedanalyse ble vurdert, inkludert konvensjonelle metoder for partikkelgjenkjenning og metoder for teksturanalyse, som er i utstrakt bruk i andre fagfelt utenfor vannbehandling. Billedanalyse basert på tekstur ble funnet å være en god løsning på utfordringene knyttet til fnokker i avløpsvann.

Karakterisering av fnokkbilder ved hjelp av teksturanalyse ble først testet i laboratorieskala i en «batch»-prosess (jar-tester) med syntetisk avløpsvann. Et system for å ta bilder uten å forstyrre fnokkene ble etablert for å kunne ta bilder under sakteomrøringsfasen av fellingsprosessen. Det ble demonstrert at fnokkbildene har teksturegenskaper som korrelerer med fellingsbetingelsene (type og dose av kjemikalier, tiden etter oppstart av sakteomrøring osv.) og kvalitetsparametre til det innkommende avløpsvannet. Endringene i fnokkbildene som funksjon av kjemikaliedosering og renseeffekt ble studert. Korrelasjonene mellom teksturegenskapene til fnokkbildene, kjemikaliedoseringen og renseeffekter ble funnet ved hjelp av multivariate statistiske metoder.

Etter vellykkede laboratorieforsøk ble det gjennomført fullskala forsøk i Skiphelle kommunale avløpsrenseanlegg i Drøbak, Norge. Billedanalyse basert på tekstur ble funnet å være anvendbart for ekte kommunale avløpsfnokker. Endringer i fellingsbetingelser og kvalitetsparametere for det innkommende avløpsvannet kunne spores i fnokkbildene. Det ble vist at systemet kunne predikere turbiditetsverdier i utløpet, noe som potensielt kan utnyttes i feilsøking og som en tidlig indikator på svikt i fellingsprosessen.

En prototype på en lav-kostnads fnokksensor ble deretter utviklet og testet ved det samme avløpsrenseanlegget. Prototypen består av et enkelt kretskort og en kameramodul, og en tilpasset programvare ble utviklet for å styre kameraet og justere innstillinger. Undersøkelsene viste at fnokkbildene som ble tatt ved hjelp av prototypen kan benyttes for prediksjon av optimale kjemikaliedoser.

Oppsummert bekrefter resultatene fra dette doktorgradsarbeidet at fnokksensoren som er utviklet har potensiale til å fungere som en selvstendig enhet for billedanalyse i sanntid og kan forbedre nøyaktigheten og kostnadseffektiviteten til eksisterende multi-parameter-systemer for kjemikaliedosering.

Dedicated to my family

Acknowledgements

Over the course of this fairly long journey of my PhD, which presented challenging and interesting work, I have gained a lot of knowledge and experience in different and novel fields. It would never have been possible without the people and organisations to whom I am grateful for my growth as a person and researcher.

First and foremost I would like to express my sincere gratitude to the main supervisor Professor *Harsha Ratnaweera* for his continuous guidance and help during all years of my PhD; not only on research and scientific matters, but often also with solving life issues. I am thankful for his encouragement, advice and constructive discussions, alongside all the support and time kindly provided to help me from the very beginning through to finalising the thesis. I am blessed to be a part of the water team built in RealTek.

Secondly, I would like to thank my co-supervisor Professor *Knut Kvaal*, who has been extensively involved in this work and introduced me to the worlds of image analysis and multivariate statistics. I am very thankful for all his guidance, valuable discussions and feedbacks; for always cheering me up. Knut has been continuously providing advanced programming and statistical support in addition to answering all the arising questions. He developed most plugins and macros used in this work, and I am grateful for having been able to use them.

I would like to thank the *Water Harmony project* funded by the Norwegian Centre for International Cooperation, which provided partial financial assistance. I would like to thank my employer *DOSCON AS* (Norway) for providing exciting and challenging work, and partly financing my PhD studies. I would like to extend my thanks to *Rosim AS* and all my colleagues, who have believed in, encouraged and supported me during the tough last stages of the PhD.

I am grateful to Mr *Jan Erik Andersen* and the colleagues from Skiphelle wastewater treatment plant (Drøbak, Norway) for helping me with the experiments and the practical arrangements at the facility.

I also want to thank Norwegian University of Life Sciences and Faculty of Science and Technology for making this PhD work possible, and for the friendly environment.

I would like to thank all my *colleagues and friends* at NMBU with whom I have spent many hours in the office and laboratory, sharing unforgettable moments from social gatherings and conferences. It is a pleasure to work with you all and be a part of our *WESH team* at NMBU.

My sincere thanks to the closest colleagues – PhD candidates in RealTek *Abhilash Nair* and *Xiaodong Wang* for their support, encouragements and countless talks on research, professional and life aspects. Thanks also to PhD candidate *Aleksander Hykkerud* for help provided in programming part of this work. A special thank you goes to Master students who spent many hours in the laboratory taking part in this research – *Vladyslav Shostak*, *Olha Yanova* and graduated Dr *Yurii Kalashnikov*. I would like to thank Dr *Vegard Nilsen* for his kind assistance in all issues regarding the Norwegian language, culture and laws. Vegard has also kindly agreed to help with the translation of the thesis' abstract to Norwegian.

Table of Contents

List of Figures

List of Tables

List of Acronyms

ACN	Adsorption and charge neutralisation (mechanism)
ADCS	Advanced dosing control systems
AMT	Angle measure technique
ANN	Artificial neural networks
CCD	Charge-coupled device
CNI	Inlet conductivity
DSLR	Digital single-lens reflex (camera)
DWTP	Drinking water treatment plant
FNU	Formazin nephelometric unit
GLCM	Grey level co-occurrence matrix
L/M/HH	Low/medium/high concentration hard water
L/M/HS	Low/medium/high concentration soft water
LED	Light-emitting diode
Me	Metal
MP	Megapixels
NTU	Nephelometric turbidity unit
OPI	Inlet ortho-phosphates
OPO	Outlet ortho-phosphates
Ortho-P	Ortho-phosphates
PC(n)	Principal component
PCA	Principal component analysis
pe	Population equivalent or unit per capita loading
PHI	Inlet pH
PHO	pH after coagulant dosage
PLC	Programmable logic controller
PLSR	Partial least squares regression
QIN	Inlet flow rate
RMSE	Root mean square error
RPM	Revolutions per minute
SCADA	Supervisory control and data acquisition
SS	Suspended solids
Total P	Total Phosphorous
TPI	Inlet total phosphorous
TPO	Outlet total phosphorous
TSS	Total suspended solids
TSSI	Inlet total suspended solids
TSSO	Outlet total suspended solids
TUI	Inlet turbidity
TUO	Outlet turbidity
WWTP	Wastewater treatment plant

1 Introduction and aims

1.1 Status of the coagulant dosing control and current needs

Coagulation is a well-known and widely used water and wastewater treatment method to remove suspended solids, colour, phosphates and other water contaminants. The chemical substances, called coagulants and flocculants, are used for water purification by coagulation process. Thus, the operational costs of the process are highly dependant on coagulants consumption. The worlds' coagulant market predicted for 2017 is over €28 billion (Freedonia 2014). Norway uses about €20 million annually for coagulants (Kemira, personal communication). For instance, the total annual cost for coagulants alone in one Norwegian wastewater treatment plant (WWTP) with a capacity of 130 000 pe is around €0.2 million (VA-Support 2012). Optimisation of the chemical dosage is a way to reduce the operational costs (Ratnaweera 2004; Liu et al. 2013). The coagulant dosages needed for achieving required treatment efficiencies highly depend on inlet water and wastewater parameters. These parameters are flow rate, suspended solids/turbidity, phosphates, pH and some other parameters like particles sizes, streaming current and zeta potential, which are more difficult to measure in full-scale. Usually, they do not highly correlate with each other and should be considered for the optimal coagulant dosage predictions (Rathnaweera 2010; Ratnaweera 1997). At present, a flow-proportional dosing concept is often used for coagulation dosage control, while process optimisation is based on data from jar tests. In other words, the coagulant dosage in WWTPs is determined by the incoming wastewater flow rates. In some cases, the dosage calculated by the flow-proportional method is corrected according to pH in the flocculation chamber (Ratnaweera & Fettig 2015). However, for complicated systems such as coagulation the flow-proportional strategy should be viewed as an oversimplified control, even with the pH correction. This is because the strategy leads to overdosage of chemicals, production of excess amounts of sludge and an overall non-sustainable usage of resources.

Advanced dosing control systems (ADCSs) based on multiple water quality parameters that could be measured online have proven to be successful. Application of such systems enables a reduction of coagulant consumption (i.e. minimise the operational costs) (Manamperuma et al. 2017; VA-Support 2012), reduces the sludge volumes and maintains the desired removal of particles and phosphates (Manamperuma et al. 2013). With the growing need for wastewater treatment processes optimisation, the need for further development of intelligent, accurate and reliable online dosing control systems arises (Ratnaweera & Fettig 2015). However, such smart control systems require high initial investments for implementation because of the online sensors costs. With the rapid growth of the databases and increasing popularity of IoT (internet of things), the question arises about who, when and how should analyse all the amounts of data recorded to maintain and validate sensors (Åmand et al. 2017; Menniti & Schauer 2017). There is also an increasing lack of qualified human resources in the treatment plants, not only to facilitate the collection of data but also to use them sensibly. Hence, decreasing the number of sensors, especially the ones that are expensive and complicated in operation, and reducing the recording of parameters to a minimum while maintaining the ability to perform adequate process control will avoid unnecessary measurements and reduce maintenance expenses.

It must be noted that currently, a reliable and universally accepted model of coagulation process does not exist. Partly because it is difficult to simulate the dynamics of the coagulation-flocculation (Thomas et al. 1999) and sedimentation process (Li & Stenstrom 2014), and also because of a significant number of rapid variations in the inlet wastewater parameters. Hence, the flow-proportional (feed-forward) dosage control has been the most straightforward solution and the most widely used until now. Feed-back control for wastewater coagulation-flocculation system is problematic because the time lag between the dosage point of chemicals (usually, first flocculation chamber) and the actual resulting treatment efficiency (often, outlet of the sedimentation tank) might be in a range from half an hour to several hours depending on the separation process. During this time, the initial conditions of inlet wastewater can (and usually do) change dramatically. Obviously, with the climate change, the changes of wastewater parameters and amounts will be even more rapid and unpredictable.

1.2 Objective of this research

Summarising all the above, there is a strong need to develop an intelligent concept that can predict the optimal dosage accurately, and is capable of substituting many of the expensive and complicated in maintenance online sensors. Such a system may also require the development of new sensors to provide warning systems and decision support systems.

The above reported ADCS is proven to be a very successful concept, though its applicability (affordability and accuracy) can be improved using a novel sensor. Despite several previous efforts by researches to study image analysis in the context of coagulation, the concept has never been used in operational context. This is largely due to various observed weaknesses at the times of the studies. The technological and conceptual development of image analysis techniques and advanced statistical methods have remarkably improved in the last decades, thus providing a mature platform to develop a promising alternative to existing concepts.

Hence, the objective of this PhD work was to develop a prototype of the imaging sensor to be used to improve existing coagulant dosage control strategies. The signal from such a sensor could function as an additional parameter to increase the accuracy or replace other more expensive and complicated sensors in ADCSs. As noted above, this would make ADCSs more efficient and affordable while increasing automation and treatment efficiencies in WWTPs. The ambition for the sensor was that it should acquire images of flocs during the coagulation process and successfully convert the image information to a specified signal, which can be used in ADCSs. The initial assumption was that the flocs formed under different conditions (wastewater parameters and coagulant dosages) would have unique structures and properties.

1.3 Structure of the research and the thesis

The research was organised into four stages, presented by two published papers in peer-reviewed journals, one accepted and one manuscript. The research stages and their respective publications follow the below research roadmap (illustrated in Figure 1):

1. Verify if the flocs formed under the different conditions have unique structures. Evaluate which image analysis methods are applicable to the flocs images (**Paper I**).
2. Establish the relationships between flocs images, coagulant dosages, initial and post-treatment parameters of model wastewaters (**Paper II**).
3. Evaluate the applicability of the alternative image analysis techniques with real wastewater (**Paper III**).
4. Test the performance of the sensor prototype in a full-scale (**Paper IV**).

The justification of research, methods and materials, results and discussions with conclusions and recommendations of the above stages were successfully developed as four research publications, respectively.

Paper I	Paper II	Paper III	Paper IV
Choose image analysis methods	Laboratory experiments with model wastewaters	Approbation of the method in full-scale with wastewater	Testing the sensor prototype in full-scale
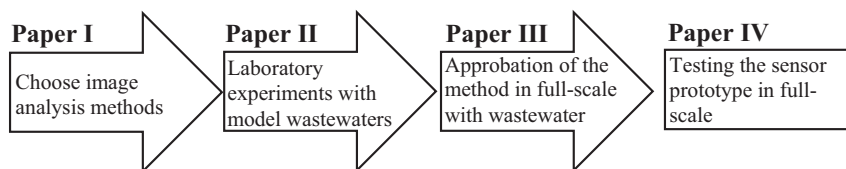
Figure 1. The concept of the PhD thesis

The thesis presents a detailed section on theoretical background and an overview of the research and findings integrating the contents of the four publications. It presents a detailed literature review on the State of the Art of coagulation, dosing control and image analysis techniques, describing the status and the needs justifying the research question. A section on materials and methods is then followed by sections on results and discussions, conclusions, recommendations and bibliography. The four publications are attached in the appendix.

2 Theoretical background

2.1 Coagulation

Surface charge is one of the most significant factors in the stability of particles. In general, surface charge arises due to factors such as the chemical composition of water and the nature of a colloid. Most colloidal particles found in raw water have a negative surface charge. Brownian movement of particles in the water and Van der Waals forces continuously cause collisions of particles with each other. The colloidal particles form a stable dispersed suspension only in cases where those collisions do not result in permanent associations (aggregates). There are two reasons preventing aggregation during a collision:

1. *Electrostatic stabilisation* – all particles have the same charge (positive or negative), so they repel from each other while colliding.
2. *Steric stabilisation* – some materials, for instance, polymers can absorb on the particles' surfaces (coating them), thus preventing any close approaches.

If collisions result in the formation of aggregates, it means that the given solution is unstable. This process is called *coagulation* or *flocculation*. Precisely, the term *coagulation* describes the process and changes in the system that leads to colloidal destabilisation. The term *flocculation* refers to the transport phase – relative motion resulting in collisions of destabilised particles that may cause particle growth (aggregation) (Bache & Gregory 2007; Bratby 2016).

The main aim of coagulation is to destabilise colloidal particles by reduction of repulsion forces, thereby lowering the energy barrier and enabling particles to aggregate. According to the traditional understanding of colloid stability, the destabilisation can be brought about by; an increase in ionic strength, giving some reduction in zeta-potential and a decreased thickness of the diffuse part of the electrical double layer; or specific adsorption of counterions to neutralise the particle charge. In both cases, destabilisation can be achieved by the addition of coagulants. Typical coagulants in wastewater treatment are aluminium and ferric salts (aluminium chloride, aluminium sulphate, ferric sulphate or ferric chloride) and pre-polymerised metal salts.

If conditions of coagulation are good enough and particle aggregates become larger with time and are macroscopically visual, they are called *flocs*. Flocs settle fast if their density is higher than the surrounding medium. In other case, aggregates will float to the top, and this process is called *creaming*.

During wastewater treatment, the coagulation process is often supplemented with additional processes that may consume a significant amount of coagulants. The spontaneous formation of metal hydroxides leading to the so-called "sweep-floc" coagulation and formation of chemical precipitates of phosphates are examples.

Coagulation is a widely used method for particles aggregation with the subsequent solid-liquid separation stage that might be sedimentation, flotation, membrane filtration or precipitation with sand and polymer (ActiFlo® system). Colloidal particles have a size range of roughly 1 nm to 1 μm and can in water be present in the form of viruses, bacteria, natural organic matter (NOM) and other inorganic particles. The coagulation-flocculation process is one of the most commonly used treatment processes for the removal of suspended solids,
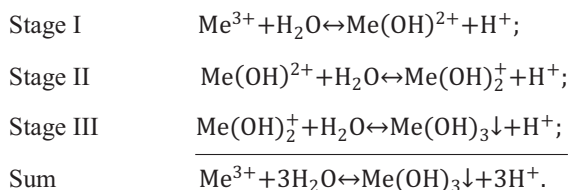
turbidity, organic matter, colour, toxic materials and microorganisms in drinking water supply as well as in industrial and domestic wastewater treatment. The primary purpose of coagulation in municipal wastewater treatment is to remove particles and phosphates.

In water and wastewater treatment the coagulation-flocculation process is a way to solve a significant number of different problems: pulp mill wastewater purification (Wang et al. 2011), kaolin–humic acid solution treatment (Bo et al. 2012), natural organic matter removal (Matilainen et al. 2010; Ratnaweera et al. 1999; Ødegaard et al. 2010), treatment of high alkalinity water (Yan et al. 2008).

The four known mechanisms of coagulation are widely accepted: double-layer compression, adsorption and charge neutralisation (ACN), enmeshment ("sweep floc"), inter-particle bridging. The destabilisation mechanism involved in the coagulation-flocculation process depends on the type and concentration of a coagulant, the nature of colloids and predominant chemistry (Bache & Gregory 2007). For drinking water treatment, the main coagulation mechanism with inorganic coagulants is ACN. In cases of wastewater coagulation-flocculation with inorganic coagulants both ACN and enmeshment may occur, but "sweep floc" is known to be a dominant mechanism (Stumm & O'Melia 1968).

2.1.1 Hydrolysis of metal salts

Two main chemical reactions take place during coagulation of wastewater – hydrolysis of metal salts and phosphate precipitation. Depending on pH, the hydrolysis process runs through many intermediate species (Amirtharajah & Mills 1982; Stumm & O'Melia 1968). The hydrolysis process for the trivalent metals can be represented by three stages of chemical reaction ("Me" represents Al and Fe):

Stage I $\qquad$ $Me^{3+}+H_2O \leftrightarrow Me(OH)^{2+}+H^+$;

Stage II $\qquad$ $Me(OH)^{2+}+H_2O \leftrightarrow Me(OH)_2^+ +H^+$;

Stage III $\qquad$ $Me(OH)_2^+ +H_2O \leftrightarrow Me(OH)_3\downarrow +H^+$;

Sum $\qquad$ $Me^{3+}+3H_2O \leftrightarrow Me(OH)_3\downarrow +3H^+$.

Many researchers extensively studied the hydrolysis process of inorganic metal salts used for coagulation and found different combinations of monomer and polymer hydrolysis products, depending on pH, ionic strength, temperature and solubility of the metal hydroxide precipitate (Van Benschoten & Edzwald 1990a; Van Benschoten & Edzwald 1990b; Stumm & O'Melia 1968). For Al(III) and Fe(III) salts the reaction equation and stages shown above are oversimplified since dimeric, trimeric and polynuclear hydrolysis products are forming (Duan & Gregory 2003). For both Al(III) and Fe(III) in higher range pH ranges the dominant monomer is soluble anionic $Me(OH)_4^-$. Therefore, one more stage can be written as:

$$Me(OH)_3+H_2O \leftrightarrow Me(OH)_4^- +H^+;$$
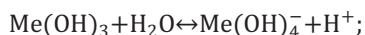
The diagrams of monomeric hydrolysis products and the amorphous hydroxide precipitates are shown for Al(III) and Fe(III) in figure 2a) and 2b), respectively. The "research zone" hashed rectangles are marked in the diagrams to show which pH and which range of dosages were used during this research. The areas fall into the "sweep floc" coagulation mechanism zones (Amirtharajah & Mills 1982).

During the hydrolysis of metal ions, pH is decreasing, and metal hydroxides precipitate as sludge. In some cases, the reduction of pH might be an advantage, for instance in processes where high alkalinity water is treated. Conversely, for cases when biological treatment follows the coagulation process, a decrease in pH may be a problem for the nitrification process (Jones & Hood 1980; Princic et al. 1998). The optimal pH for coagulants containing Al(III) salts is between 6 and 7, for Fe(III) the range is 5.5 – 6.5 (Metcalf & Eddy 2013).

a)   b)
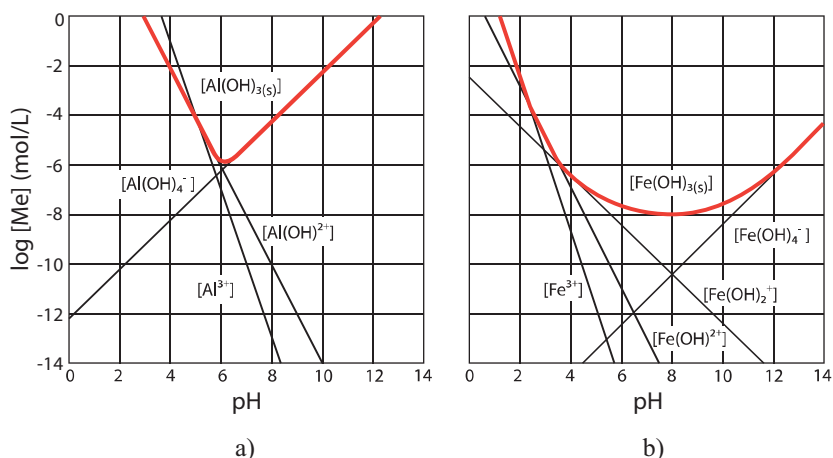
Figure 2. Solubility diagrams of a) Al(III) salts; b) Fe(III) salts (after Amirtharajah and Mills 1982; Bratby 2016).

2.1.2 Phosphate precipitation

Removal of phosphates is crucial in wastewater treatment since its excessive amounts in the recipient water bodies lead to eutrophication (Liu et al. 2009; Smith et al. 1999). Regulations regarding total phosphorus (total P) removal rates during the wastewater treatment process are becoming stricter, especially in sensitive areas of water bodies. The required removal of total P in most Norwegian WWTPs is currently very high – at least 90 %. The changes in the discharge licence practice to include bypass flows at the WWTP into the reporting has forced WWTPs to push for even higher removal rates on the portion of wastewater that goes through the plant. Coagulation-flocculation is an extremely effective method of phosphates removal due to the chemical reactions with metal ions and adsorption or enmeshment of MeP-complexes and/or PO_4^{3-} ions on other particles and/or amorphous $Me(OH)_3$ precipitates, which are present during flocculation and should be further separated.

The reaction of phosphate precipitation with the trivalent metals is:

$$Me^{3+}+PO_4^{3-}\rightarrow MePO_4\downarrow$$
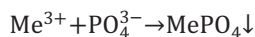
Despite the above equations' stoichiometric proof that 1 mole PO_4^{3-} requires only 1 mole of Me^{3+}, WWTPs use much higher Me:P ratios in practice. The Me:P ratios in Norwegian WWTPs are usually within the range 2.5-4.5:1, but can in some cases reach as high as 8:1 (Ratnaweera 2013). This can be explained by the competition mechanisms of metal ions interacting with OH^- and PO_4^{3-} together with the WWTPs' desire to achieve extremely high P

removals with inefficient dosing control strategies. Some part of the coagulant may also be inert.

Over the last few decades the mechanisms of phosphorus removal in coagulation were thoroughly investigated, but the findings are still disputed. When metal salt is dosed into water, a range of Me-hydrolysed species are formed. Hence, the species that can react with PO_4^{3-} ions are not only Me ions, but also a range of hydrolysis products described in section 2.1.1. In general, coagulation of PO_4^{3-} ions with Me salts forms two major linked mechanisms: 1) adsorption of PO_4^{3-} ions on the Me-hydroxide; 2) formation of Me-hydroxy-phosphate complexes Me(OH)$_{3-x}$(PO$_4$)$_x$ (Ratnaweera 1991).

Regarding phosphorus removal by iron-based coagulants, it was found that the major iron phosphorus compound in sewage treatment plants is Vivianite – ferrous iron phosphate mineral $Fe_3^{2+}[PO_4]_2 \cdot 8H_2O$ (Wilfert et al. 2016). Researchers hypothesise that Vivianite is dominating in most wastewater treatment plants where iron-based coagulants are used, and surplus sludge or anaerobically digested sludge is produced. They also believe this discovery could offer new ways for phosphorous recovery.

Research on mechanisms of phosphate and particles removal (Ratnaweera 1991; Manamperuma et al. 2016) showed that higher concentrations of suspended solids (SS) in wastewater when amounts of orthophosphates (ortho-P) remain constant increases the SS removal efficiency and decreases ortho-P removal efficiency. The hydrolysis of Al(III) ions is a dominant reaction when the initial concentration of SS is high. With higher amounts of ortho-P in wastewater, the phosphate precipitation becomes a dominant reaction (Manamperuma et al. 2016; Manamperuma 2016).

Sometimes, inorganic coagulants are used together with organic polyelectrolytes to improve treatment efficiency and flocs characteristics. The examples of such polyelectrolytes are polyaluminium chloride (PAC) and polydiallyldimethylammonium chloride (PDADMAC). The work conducted by Hatton (1985) showed that the removal efficiencies of orthophosphates with PAC and Al$_2$(SO$_4$)$_3$ were in the same range for equivalent dosages, and the superior performance of PAC was due to higher adsorption of PO_4^{3-} on the floc blanket. The impact of organic polyelectrolytes on coagulation of black water was also studied (Kozminykh et al. 2016). Researchers have found that the combined use of inorganic coagulants and cationic polymers resulted in the highest removal rates of SS, total chemical oxygen demand (COD) and ortho-P, compared to using just coagulants or just polymers.

A range of published studies have shown that the higher the OH/Al ratio of the coagulant, the greater the required dosage for phosphates removal (Jiang & Graham 1998; Fettig et al. 1990; Ødegaard et al. 1990; Ratnaweera et al. 1992; Diamadopoulos & Vlachos 1996; Boisvert et al. 1997).

2.2 Factors that influence the coagulation process

Rapid mixing conditions are required when coagulant chemicals are added to water. Rapid mixing is one of the most critical stages that needs to be correctly operated in a coagulation process, because the destabilisation of the heterogeneous solution takes place and primary floc particles are formed in this stage. The choice of rapid-mix unit type is based on the rate of coagulation reactions – an interaction between the hydrolysis products of Al(III), for instance, and colloidal suspension (Amirtharajah & Mills 1982). Kan et al. (2002) have proved that rapid-mixing time has a significant impact on charge neutralisation and "sweep floc" mechanisms of coagulation. The rapid-mixing time affects the destabilisation of colloid suspension and particle aggregation. According to their results, residual turbidity from the ACN mechanism was lower than that formed from "sweep floc" coagulation. Edzwald (2013) has reviewed different types of impellers and mixing conditions considering sufficient mixing intensities for charge neutralisation and "sweep floc" coagulation mechanisms. He claimed that proper mixing speed and design, depending on the coagulation mechanism, leads to energy and cost savings. The main conclusions of the paper are: a) charge neutralisation mechanism requires high velocity gradients G_S (intense mixing) and short mixing times; b) for sweep-floc coagulation; mixing intensity is not important, while crucial factors are dose (metal concentration) and pH. However, uniform coagulant concentration should be achieved in the mixing volume; c) construction of rapid-mixing reactors or pipes should not be based on the maximum daily flow, but on the average daily flow.

The conditions of the slow mixing stage of the coagulation-flocculation process have an extreme influence on particle aggregation and breakage. Several research efforts have investigated how to construct the appropriate flocculation chambers in treatment plants. The main parameters considered when constructing flocculation compartments are retention time (T) and average velocity gradient (G), as they influence the rate and extent of particle aggregation and the rate and extent of breakage of these aggregates. Flocculation chambers and mixing units might be next types: baffled chambers, granular media beds, diffused air, spiral flow, reciprocating blades, rotating blades. They are well described by Bratby (2016).

Camp & Stein (1943) developed the concept of velocity gradient G. However, other researchers have emphasised the limitations of this concept (McConnachie 1991; Han & Lawler 1992; Kramer & Clark 1997; Pedocchi & Piedra-Cueva 2005). Gregory (2006) stated that the dimensionless term GT ($G \times T$) has been, and still is, of practical importance. Bernhardt & Schell (1993) found that flocculation efficiency could be maintained when high or low G values are exchanged for high or low T values to maintain a particular GT.

Ødegaard (1985) has studied optimisation of floc separation performance. The intensity of mixing together with the residence time, floc volume and residence time distribution were considered. It was established that flow distribution during the flocculation process should preferably be similar to plug flow process. It is better to have three or more compartments: 1 – short time flocculation at a relatively high G-value (where G is mean turbulent velocity gradient); 2 – longer time at a lower constant G-value; 3 – shorter time at lowest possible G-value in order to prevent settling and to maintain the homogeneous suspension. Flocculation time varies from 15 to 35 min.

Parker et al. (1970) studied maximum floc sizes under different mixing intensities. They defined that the flocs breakages in a given mixing intensity depends on an energy cascade effect.

A pre-polymerisation degree of the coagulant influences the coagulation process. It was observed that the higher the OH/Al ratio of the coagulant, the higher dose is required to achieve a comparable orthophosphates removal rate (Manamperuma et al. 2016; Manamperuma 2016; Ratnaweera 1991; Fettig et al. 1990; Ødegaard et al. 1990; Zouboulis et al. 2007).

Water temperature influences the coagulation-flocculation and sedimentation processes. However, what stages it alters is still not very clear. Kang & Cleasby (1995) stated that temperature might affect the hydrolysis process of a metal salt, adsorption, particle motion and particle-particle interactions, fluid motion and precipitation rates. Whether the decrease in coagulation efficiency at low temperatures is caused by chemical factors or particle transportation processes remains disputable. Xiao et al. (2009) observed much slower growth rate of flocs at 2°C compared to 22°C. They suggested it might be a result of impeded Brownian motion. They have compared sedimentation of flocs within different temperatures and observed comparatively fine settling; however, the settling speed was slower at low temperatures. One of the proposed solutions was to enhance the frequency of particles collisions, accelerating the flocs growth rate by increasing the initial turbidity.

As was previously mentioned, the influent wastewater parameters are changing dramatically during the day. Fig. 3 shows one week observation data in a Norwegian WWTP, where the advanced automated dosage control system DOSCON (DOSCON AS, Norway) is implemented. The blue line on the graph corresponds to the inlet wastewater flow. The pattern of wastewater production by population is fairly visible when the rain events do not take place. In the morning, the wastewater inlet to the plant dramatically increases (more than 2 times), declines during the working hours, increases again in the evening and gradually decreases at night. The green line represents the inlet turbidity behaviour. The fluctuations of inlet turbidity values are more rapid and unpredictable. Besides, they do not always follow the wastewater loads. Similar trends can be observed for other wastewater parameters such as pH, conductivity and temperature. All these water characteristics vary independently from each other. Moreover, during the wet-weather conditions and/or snow melt the inlet water qualities and quantities change even faster. Wang et al. (2017) investigated and built a classifier for the daily, weekly and seasonal variations of influent wastewater parameters.

The industry is seeking new methods for tracing and/or predicting the rapid changes in inlet wastewater conditions. The systems, which can adequately react to these changes and prevent failures in the treatment processes or help in troubleshooting, are still in the development phases. There are not many existing solutions at present in the market that effectively addresses the problem. This need for a solution is particularly striking in the coagulation-flocculation optimisation process.
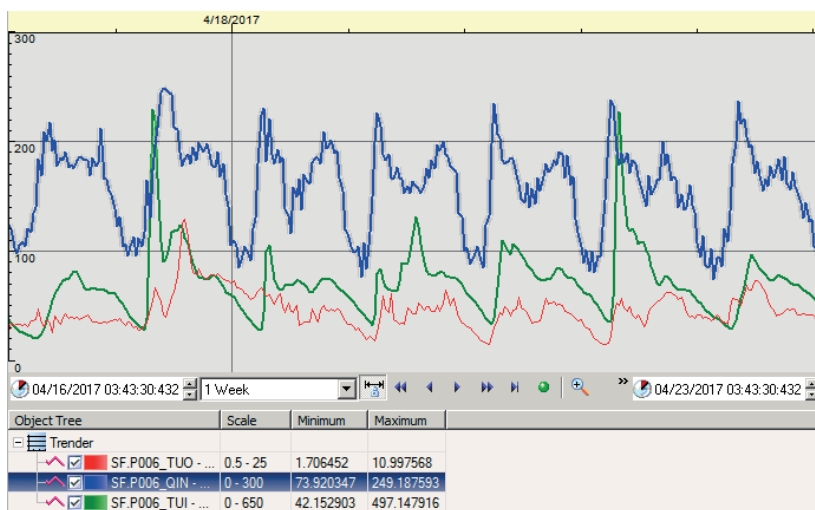
Figure 3. Changes in wastewater load and parameters during the week (image from DOSCON system)

The time lag between a coagulant and/or flocculant dosage point and effluent from the sedimentation basin (or any other subsequent solid-liquid separation process) in addition to rapid and frequent changes of inlet wastewater parameters pose challenges for proper control of the coagulation process.

Optimal dosing of coagulants and flocculants implies the addition of minimum amounts of chemicals to achieve the required treatment efficiencies. The needed amounts of coagulants are dependent on influent wastewater characteristics, such as inlet flow rate, turbidity, SS, pH and total P (Ratnaweera 1997). Many WWTPs operate a coagulation-flocculation process on the flow-proportional basis, considering only inlet wastewater flow rate. Even with the pH overriding function, treatment plants tend to overdose the chemicals and work in non-optimal dosage ranges (Ratnaweera & Fettig 2015). The consequences of operating in non-optimal dosage ranges is the potential for poor treatment efficiencies, alongside increased chemical costs, high sludge volumes leading to increased sludge treatment and management costs, causing challenges to the downstream processes and an overall adverse environmental impact. Thus, the instruments, proper models and procedures are needed for the optimal coagulant dosage control.

All the above passages highlight the complex nature of the coagulation process together with the dynamic nature of many wastewater treatment systems. With the growing global requirement for improved reliability and quality in treatment processes, greater demands are placed on precision in the control and monitoring of automatic wastewater treatment systems than what can be provided by manual control alone. Accurate and affordable online monitoring tools and automatic control units are needed for robust process control and daily process surveillance to achieve the required degrees of wastewater treatment. A wide range of instruments are available for water quality and treatment monitoring, for instance; online sensors measuring flow rate, level, conductivity, pH, colour, turbidity, dissolved oxygen and some other water quality parameters. However, there is a lack of online instruments for some relevant parameters due to price, complexity or frequently needed maintenance, so these

10

parameters currently remain as unresolved challenges in terms of online control. Indirect measurements can be a helpful strategy in such cases. The concept of indirect estimation is also known as virtual sensors or soft sensors.

2.3 Dosage control strategies

Ratnaweera & Fettig (2015) recently reviewed and summarised all attempts made towards advanced monitoring and dosage control of the coagulation process. They state that the flow-proportional dosage control is still the most commonly used method in drinking water and wastewater treatment. However, as discussed in previous sections, such approaches in the control strategy results in an unnecessary overdose of chemicals. Thus, advanced alternative methods of dosage control and online monitoring have been under development for years.

Unlike in biological wastewater treatment, there are no universally accepted conceptual models for the coagulation-flocculation process. This is probably due to the large number of variables affecting the process and complexities of reaction mechanisms. Coagulation-flocculation is a non-linear dynamic system, which is often controlled under non-optimal conditions due to the rapid variations in water qualities, quantities and a lack of tools and techniques for online monitoring and control or their unavailablity due to extremely high prices.

The modelling of the drinking water treatment process, which includes a coagulation stage, had recently been reported by Juntunen et al. (2012). The authors investigated both laboratory and process data. They found the aluminium dose to be one of the most important factors affecting the treated water quality parameters (turbidity and residual aluminium).

There is an increasing availability of measurement and collecting of different types of data during the process development (Giudici & Figini 2009). However, it seems that this has led to a situation where engineers and plant operators experience difficulties in handling the large amounts of process data (Wang 2007). For this reason, new approved software tools are needed to facilitate and automate the monitoring of process performance. Because processing of measurement data can be laborious and time-consuming, new descriptive methods for presenting real-time and other data can provide considerable advantages in the monitoring and control of the process. The increasing amount of data and increasingly complicated process require appropriate software tools, which can be used for extracting valuable information from the massive amount of measurement data.

The process of coagulation involves many complex physical and chemical phenomena that are difficult to evaluate using single measurements. Process dynamics is a typical problem in data-based systems, and successful monitoring of process performance often requires an ability to adapt to changing conditions (Juntunen et al. 2012). On the other hand, manual monitoring of several measurements simultaneously and decision-making based the observed results is both difficult and time-consuming. Therefore, it is useful to adopt a monitoring system that can handle all available measurements and present the available information in a simple, user-friendly and flexible manner (Liukkonen, Juntunen, et al. 2013; Liukkonen, Laakso, et al. 2013).

Streaming current detector (SCD) was evaluated (Dentel et al. 1989a; Dentel et al. 1989b) and tested in the drinking water treatment plants (DWTP) for automatic coagulation

control (Critchley et al. 1990; Yavich & Van De Wege 2013). The coagulant dosage strategy based on zeta potential measurements was documented to be a promising control technique for DWTP (Sharp et al. 2005; Sharp et al. 2016). Advanced soft sensors and coagulation process control models employing artificial neural networks (ANN) have been tested in DWTPs (Baxter et al. 2002; Juntunen et al. 2013; Valentin & Denœux 2001). Yu et al. (2017) applied image analysis with ANN to simultaneously evaluate the removal of SS and colour from textile wastewater.
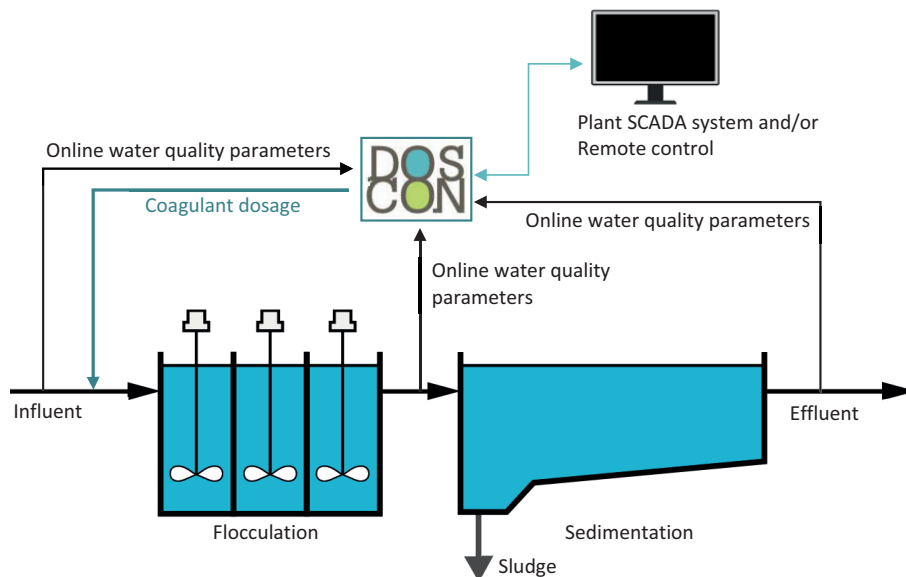
Figure 4. Schematic representation of DOSCON concept

DOSCON® has developed a real-time feed-forward dosage control system, which bases on multi-parameter surveillance of water quality (Ratnaweera 1997; Rathnaweera 2010; Manamperuma et al. 2013; Manamperuma et al. 2017). Fig. 4 illustrates the coagulant and/or flocculant dosage control strategy. Different inlet and outlet water quality parameters measured by online sensors are sent to the programmable logic controller (PLC). PLC performs the calculations and sends the signal to the chemical dosing pump. Multivariate statistical methods are used for the dosage prediction. The input parameters in the model include wastewater flow rate, initial pH, turbidity, conductivity, temperature, total phosphates and/or ortho-phosphates, coagulation pH, effluent turbidity and remaining phosphates. One of the system's advantages is its flexibility to include or exclude different parameters from the predictive model. Thus, the DOSCON system can be adjusted for the treatment plants with different surveillance and control needs and the various online sensors that are installed. In harsh environments such as wastewater, temporary malfunctions among the online sensors are common. DOSCON seems to handle this challenge using a smart concept of online validation of parameters and choosing an algorithm, which includes only the validated parameters from a pre-calibrated set of algorithms. However, a reduction in the number of validated parameters also decreases the efficiency of the system. PLC can be connected to the plant's supervisory control and data acquisition system (SCADA) or be a completely self-standing system. The monitoring of the

parameters is possible both through the PLC's touch screen and by remote access/control via special software. So far, DOSCON has proven its performance in coagulant dosage control for both drinking and wastewater treatment plants (Liu et al. 2013; Manamperuma et al. 2017). The benefit of DOSCON is the significant savings on chemicals and sludge management costs. The pure chemicals savings can be as high as 30 % depending on the size, instrumentation and conventional dosing systems (VA-Support 2012).

Dependency on several online sensors, which can be expensive and complicated to operate, and inadequacies in feed-back control remain as weaknesses in this system.

The time lag between the dosage point and effluent after the separation process remains an unsolved problem. Therefore, the control is mostly based on a feed-forward strategy. A feed-back control will significantly improve the existing dosage control strategy. Hence, the technology is still seeking a verified method to solve the time lag problem (Liu 2016; Liu & Ratnaweera 2017; Su et al. 2017).

2.3 Image analysis in water and wastewater treatment

Computer vision image processing techniques are increasingly more frequently employed to study and characterize complex size, shape and features of particles. Photographic image acquisition techniques make it possible to capture and analyse flocs' images online (Chakraborti et al. 2003). The distinct effect of coagulant dosage on aggregates' features was detected by Lin et al. (2008) in kaolin water suspension. Authors used wet scanning electron microscopy to obtain the morphology of flocs. Jin (2005) used high-resolution digital camera to study the influence of temperature on flocs properties under different coagulant dosages for river water coagulation. The relations between the projected area of particles and coagulant dosages were found. Wang et al. (2011) investigated the changes in flocs' characteristics due to different dosages and coagulation pH in humic acid suspension, obtaining images by digital CCD camera.

Image analysis of particles is a fairly old (Tambo & Watanabe 1979b) and known technique to be used in different water treatment applications, for instance: activated sludge processing (Alves et al. 2000; Amaral & Ferreira 2005; da Motta et al. 2001; Dagot et al. 2001; Jenné et al. 2006; Mesquita et al. 2011), aggregates settling velocity measurements (Vahedi & Gorczyca 2012; 2014), membrane fouling (Mendret et al. 2007), natural organic matter removal (Xiao et al. 2011). Image analysis applied to coagulation and aggregation of particles was also widely studied in different waters: clay (kaolinite) suspension (Tambo & Watanabe 1979a; Xiao et al. 2011; He et al. 2012; Lin et al. 2008), mineral suspensions (Gorczyca & Ganczarczyk 1996), latex particles suspension (Chakraborti et al. 2003), humic acid solution (Wang et al. 2011), water after lime softening (Vahedi & Gorczyca 2011), lake water (Chakraborti et al. 2000; Kim et al. 2001), textile wastewater (Yu et al. 2009; Yu et al. 2012), sewer system wastewater (Zheng et al. 2011). Some attempts were made to characterise alum flocs by image analysis in the drinking water treatment process (Juntunen et al. 2014). Yu et al. (2005) presented the results of online monitoring of industrial wastewater true colour using digital image analysis and ANN.

An extensive literature search showed an inadequacy of studies concerning the floc-dosage or floc-effluent relations in the coagulation process regarding municipal wastewater or model wastewater.

The fractal geometry of nature and the term "fractal dimension" were defined by Mandelbrot (1967; 1982). Li and Ganczarczyk (1989) described the spatial structure of particles appearing in water and wastewater treatment processes by fractal theory, i.e. fractal dimension. Currently, it is one of the most common ways of flocs characterisation in water and wastewater field. The review of Bushell et al. (2002) summarises the techniques available for measuring fractal properties of flocs and aggregates.

For a 2D projected particle image, the fractal dimension, D_{pf} defines how the projected area of the particles rises with the perimeter (He et al. 2012):

$$A \propto P^{2/D_{pf}} , \tag{1}$$

where A – projected area; P – perimeter of the particles. For the 2D projection of an image, the value of fractal dimension varies from $D_{pf}=1$ for the circle shape floc to $D_{pf}=2$ for a chain of particles (a line).

The size of an irregularly shaped particle can be determined as the equivalent diameter, d_p (He et al. 2012):

$$d_p = (4A/\pi)^{1/2} . \tag{2}$$

Even though fractal geometry and particles characterisation by image analysis have gained great popularity among researchers, there were not many identifieable attempts towards applying the method for online coagulant dosage control. One of the few examples include research on flocculation control based on the fractal dimension of flocs in the pilot scale of drinking water coagulation (Chang et al. 2005). However, the technique was not yet applied in full-scale.

The weakness of the particles characterisation method, particularly by image analysis, is that some challenges might arise during the image processing. For instance, the need to remove out-of-focus objects has been documented (Keyvani & Strom 2013). Since the particles characterisation by image analysis is based on objects count algorithm, it is important to ensure that the 3D objects (particles) do not overlap in the 2D images and that the number of particles is not under-estimated. Otherwise, the objects' characteristics could in such cases be over-estimated. Besides, the number of estimated particles on the image depends on the threshold setting and is often a matter of judgment (Bache & Gregory 2007). The problems of image resolution limitations and hardware limitations are gradually being reduced with rapid development in the industry. However, the "ready to use" online solutions with feasible and robust integrated digital image analysis systems are not yet available in the market.

The Grey level co-occurrence matrix (GLCM) method has found broad application in the food industry, for instance, in the determination of meat quality (Shiranita et al. 1998), baking experiments on wheat baguettes (Kvaal et al. 1998) and surface texture characterisation of an Italian pasta (Fongaro & Kvaal 2013). Other applications of texture image analysis for evaluation of food qualities are summarised in a review by Zheng (2006). The studies on colour

texture classifications can also be found in the literature (Palm 2004; Khan et al. 2015; Gui et al. 2013).

Angle measure technique (AMT) is a method of calculating the complexity of an image. It was first introduced by Andrle (1994) for characterising the complexity of geomorphic lines. The purpose was to detect changes in coastline complexity as a function of scale. Esbensen et al. (1996) introduced AMT in chemometrics for general applications and use. Later the technique has been successfully used with image analysis applications (Kvaal et al. 1998; Huang & Esbensen 2000; Huang & Esbensen 2001; Kucheryavski 2007; Mortensen & Esbensen 2005; Dahl & Esbensen 2007; Fongaro & Kvaal 2013; Fongaro et al. 2016).

Numerous computer vision applications use texture analysis to perform automated image recognition, classification and segmentation (Haralick et al. 1973; Zheng et al. 2006; Bharati et al. 2004). Texture is a loosely defined term without an accepted or universally quantitative meaning. A general description is that texture is a combination of repetitive patterns of pixel variations organised in some structural way (Russ 2011). Texture can be defined as a measure of the image's surface roughness defined by parameters of brightness, colour, shape and size variations within some region and its repetitiveness. The texture properties of the materials/objects in the image may be found or correlated in some way. Such properties might be geometric structure, orientation, coarseness, smoothness, roughness and periodicity. Texture is a pattern that can be completely distinct or completely random. Furthermore, texture could be isotropic (without any preferred orientation) or anisotropic (has definite pattern structure) (Levine 1985; Gonzalez & Woods 2010).

Considering the variability of texture properties and their combinations, an infinite number of textures exist and it is difficult to identify and describe all these structures. Nevertheless, different databases have been collected to perform computer vision analyses. One of the most known, classical, databases was collected by Brodatz (1966) and consists of 112 photographs captured under controlled lighting conditions. The data set is commonly used for testing new texture analysis methods, in computer vision and signal processing. Normalized and coloured Brodatz texture data sets are also used (Abdelmounaime & Dong-Chen 2013).

So far, the water and wastewater industry has not employed texture image analysis methods. However, texture analysis has been broadly used for segmentation of the urban scenes (Conners et al. 1984), in situ powder characterisation (Huang & Esbensen 2000; Huang & Esbensen 2001), grain size characterisation (Dahl & Esbensen 2007), classification of the tree barks (Palm 2004) and identification of uranium ore (Fongaro et al. 2016).

3 Experimental procedures and methods

3.1 Laboratory experiments. Jar tests

The laboratory experiments were conducted using an adapted jar test procedure and different types of model wastewater. The model wastewater contained both organic and inorganic components to represent typical domestic wastewater characteristics important for the coagulation processes. Model wastewater was prepared according to procedures from previous studies (Fettig et al. 1990; Ødegaard et al. 1990; Ratnaweera 1991), but the initial recipe table was expanded to better simulate changes in inlet wastewater characteristics. Table 1 represents the composition of model wastewaters used for the research.

Table 1. Composition of model wastewaters

Component	Concentration level								
	Soft water*						Hard water**		
	LS1	*LS2*	*LS3*	*MS1*	*MS2*	*HS*	*LH*	*MH*	*HH*
NaHCO$_3$, mg/l	60	60	60	60	60	60	400	400	400
NaCl, mg/l	400	400	400	400	400	400	0	0	0
NH$_4$Cl, mg/l	50	50	50	100	100	200	50	100	200
K$_2$HPO$_4$, mg/l	12.5	25	30	35	50	100	25	50	100
Humic acid, mg/l	2.5	2.5	2.5	5	5	10	2.5	5	10
Dry milk, mg/l	150	150	150	300	300	600	150	300	600
Potato starch, mg/l	30	30	30	60	60	120	30	60	120
Bentonite, mg/l	20	40	50	70	80	160	40	80	160
CaCl$_2$, mg/l	0	0	0	0	0	0	255	255	255

* *LS* – low concentration soft water, *MS* – medium concentration soft water, *HS* – high concentration soft water;
** *LH* – low concentration hard water, *MH* – medium concentration hard water, *HH* – high concentration hard water.

Initially, three contaminants concentrations were used to prepare model wastewaters of soft type – low (*LS2*), medium (*MS2*) and high (*HS*). However, the concentrations of K$_2$HPO$_4$ and bentonite were varied to simulate changes in inlet wastewater characteristics typical for domestic WWTPs. Hence, the table was expanded to include *LS1*, *LS3* and *MS1*. All 6 model soft water types and 3 hard water types were used in experiments described in Paper II. Paper I is based on the data obtained from the experiments with model wastewater type *MS2*.

The resulting concentrations and characteristics of different model wastewater types are given in table 2.

Flocculator 2000 from Kemira Chemicals AS with programmable mixer units and 1 litre beakers was used for the coagulation experiments in jar test scale. The mixing conditions during coagulation: 1 min rapid mixing (400 RPM), 10 min slow mixing (30 RPM) followed by 20 min of sedimentation without mixing.

It is important to conduct the jar test experiments with a constant coagulation pH to enable the accurate comparison of coagulation performance for various dosages and different coagulants. The pH values were measured by portable pH meter 3110 (WTW GmbH,

Germany) during the slow mixing stage. The constant pH values were maintained by addition of 2 mol/l NaOH.

Table 2. Concentration and parameter measurements of different model wastewaters

Parameter	Concentration level								
	Soft water						Hard water		
	LS1	LS2	LS3	MS1	MS2	HS	LH	MH	HH
pH	8.00± 0.00	7.87± 0.19	8.00± 0.00	8.00± 0.00	7.89± 0.15	7.81± 0.05	7.98± 0.07	8.01± 0.04	7.95± 0.03
TSS, mg/l	98±2	128±22	138 ±2	234±12	255± 64	507±35	110±10	240±30	380±100
Turbidity, NTU	61±2	66±4	79±3	150±0	156±3	352±5	114±11	264±14	558±42
Ortho-P, mg-P/l	2.40± 0.31	4.95± 0.28	6.30± 0.48	7.15± 0.23	10.65± 0.77	21.32± 0.89	5.10± 0.15	10.17± 0.67	17.64± 0.04
Total P, mg-P/l	3.18± 0.37	5.34± 0.47	6.29± 0.69	8.27± 0.23	12.32± 1.52	22.99± 0.98	5.67± 0.05	10.7± 0.05	20.9± 0.05
ζ-potential, mV	-22.2	-19.1	-21.7	-22.0	-21.0	-22.3	-	-	-
Particle size distribution peaks, μm	-	1.7, 45	-	-	1.7, 12	1.7, 55	1.7, 45	12	1.7, 12, 45

The coagulants used during experiments were from Kemira Chemicals AS. The used coagulants and their characteristics are given in table 3. PAX-XL61 was used during the research published in Paper I, the dosage range was 0.29 – 1.08 mmol Al/l, in total 11 dosages were used with two replicates. Three other Kemira coagulants were used during the research described in Paper II. The optimal dosages of each coagulant for each model wastewater type were determined by the preliminary jar tests experiments. Optimal dosages were evaluated regarding total suspended solids (TSS), turbidity, ortho-P and total P reduction. Five dosages, which include under-, near- and over-optimum coagulation conditions for each model wastewater type, estimated during preliminary tests, were used during the experiments with at least two replicates: 0.25, 0.5, 0.75, 1.0, 2.0 mmol Me/l.

Table 3. Properties of used coagulants

Coagulant	Short code	OH:Me	% Me by weight	Density, kg/m^3	Used
Prepolymerised polyaluminium chloride	PAX-XL61	1.9	5.4±0.25	1270±20	Paper I
Aluminium sulphate	ALS	0	4.0±0.3	1280±50	Paper II
Prepolymerised Polyaluminium chloride	PAX-18	1.1	9.0±0.3	1360±20	
Ferric Sulphate	PIX-313	0	11.4±0.3	1550±20	

After sedimentation, the supernatant was analysed for TSS, turbidity, ortho-P, total P, and for some samples zeta-potential (ζ-potential) was measured. TSS measurements were done by filtering water samples, followed by drying at temperature 105°C for 1 hour (ISO 11923:1997). Turbidity was measured by portable turbidimeter 2100Q (Hach®, USA). Ortho-P and total P were measured by EasyChem analyzer (Systea S.p.A, Italy). ζ-potential was measured by Zetasizer Nano (Malvern Instruments Ltd, UK). One of the conventional methods of particle size measurement is the laser diffraction analysis. Particle size distribution in model wastewater was measured by Mastersizer 3000 (Malvern Instruments Limited, UK).

3.2 Image acquisition in laboratory scale

One of the Flocculator 2000 mixers was manually equipped with a 4.5 cm width black curved plastic plate, which served as a dark background for the images of flocs. A schematic representation of the whole installation is shown in figure 5.

The images were acquired during the slow mixing coagulation stage every 20 sec using digital single lens reflex (DSLR) D600 camera (Nikon®, Japan) with 105 mm Nikkor AF-S Micro 1:2.8 G ED lens (Nikon®, Japan). Parameters of the camera: ISO 3200, aperture F/13, shutter speed 1/125 sec. The camera was remotely controlled from a PC via *DigiCamControl 1.2.0* software (Istvan 2012) – images were collected and automatically saved on a separate hard-drive.
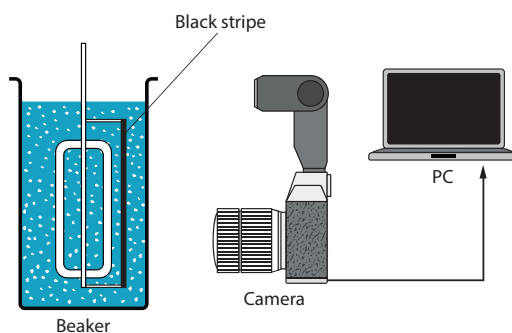
Figure 5. Schematic representation of image acquisition on the laboratory scale

3.3 Image analysis methods

An open source image analysis software *ImageJ v.1.49* (Rasband, 1996/2016) was used. The software is based on plugins and macros and has an extensive amount of already existing image analysis methods for different kinds of applications. It is also possible to write specific scripts and plugins in Java for the certain tasks.

3.3.1 Selection of methods

This PhD work started to study the model wastewater particles' characteristics and encountered several problems, one of which being flocs overlapping on the images. Overlapping of particles in the image results in a non-accurate recognition of objects and their

amounts, which leads to incorrect detection of the flocs' geometrical characteristics. Therefore, alternative methods of flocs images characterisation were explored. Different texture image analysis techniques to correlate images of flocs with coagulant dosages were evaluated.

The texture analysis methods were not previously used for studying the images of coagulated particles in the approach of this PhD work, particularly for improving coagulant dosage control. Hence, this PhD research work is innovative in the perspective of texture image analysis methods applied to images of flocs in the coagulation-flocculation process.

We have tested simple statistical methods of image characterization like Histograms, to see if pure information from the image can be used. Since the GLCM method has documented its broad applicability of texture images characterisation in different fields of studies, we decided that the method might also be successfully applied to wastewater flocs images. GLCM was tried as the fairly old and proven technique. AMT was tested as a new and promising method of the images' complexity characterisation. However, GLCM feature vectors as a result of image analysis are more intuitive to understand and interpret in comparison to AMT, which is a measure of spectrum complexity based on scale.

Our approach is to prove that images of flocs captured by the nonintrusive photographic method are descriptive in terms of coagulant dosage prediction. Thus, the technique could potentially be used for online coagulation dosage control. Advanced dosage control will then be possible by means of determining the lowest coagulant dosage, which leads to required solid-liquid and phosphorous separation.

3.3.2 Object recognition image analysis method

Particles recognition image analysis methods are often used for measuring particles size and dimensions. The sequence of actions for the particles characterisation presented in **Paper I** is illustrated in fig. 6.

All images were collected during the slow mixing period of coagulation. They were gathered in a stack in *ImageJ*. Then 3690 × 3690 pixels area was manually selected and cropped. A stack of images was converted to 8-bit grey scale. Plugin "Subtract Background" was applied to images in order to equalise background pixel values for easier thresholding. If needed, brightness and contrast were adjusted. Onwards, thresholding was used to obtain binary images, with a threshold value estimated manually or by Otsu method (Otsu 1979). Flocs within this transformation had greyscale value 0 – true black, while the background was set to value 255 – true white. Finally, plugin "Analyse particles" was applied to the binary images of flocs. Many geometrical and statistical parameters might be subtracted from the images of aggregates by this plugin. For this research, significant parameters of flocs were: number of particles in the image, mean area, and perimeter of flocs. Knowing that 1230 pixels in the image equal 1 cm, we were able to recalculate flocs' features from pixel values to quantitative values in centimetres.
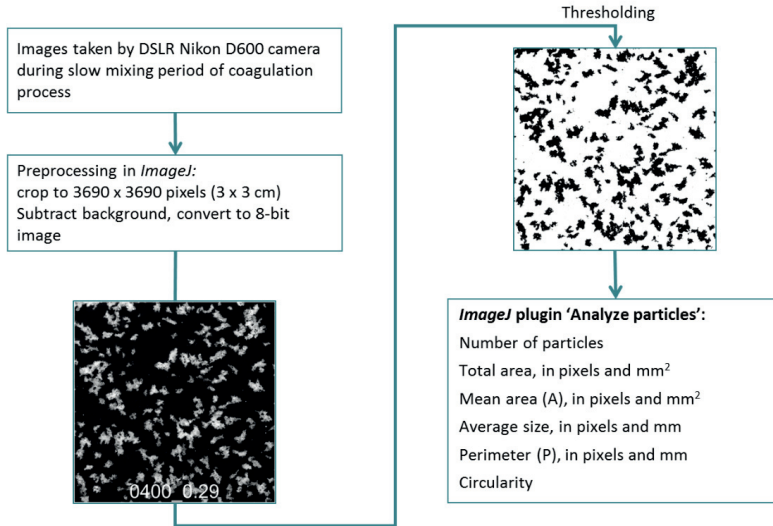
Figure 6. Floc features detection procedure by image analysis. Numbers on the left image indicate 400 s after the start of slow mixing period, dose 0.29 mmol Al/l of PAX-XL61 coagulant

Mean fractal dimension of flocs was calculated in **Paper I** by the equation:

$$D_2 = 2 \times \log(P_{pix}) / \log(A_{pix}), \tag{3}$$

where A_{pix} – mean area of particles, in pixels; P_{pix} – perimeter, a count of pixel edges; 2 is a constant number (Yu et al. 2009).

3.3.3 Image analysis by Grey level co-occurrence matrix (GLCM)

GLCM is one of the statistical methods for measuring texture in the image. Haralick et al. (1973) invented this method, which bases on spatial-dependence grey level co-occurrence matrix of pixels with estimation of image features using second-order statistics.

The size of GLCM is determined by the number of grey levels G in the image. Typically, a 256×256 GLCM matrix is constructed for the 8-bit image (G=256). Figure 7a shows an example of 4×4 pixels image with grey level variations from 1 to 3. The GLCM (fig. 7b) was constructed from the image 7a and resulted in a matrix size 3×3, because the image has only three grey levels. Each value in the matrix is the number $K(i, j|\theta, d)$ of co-occurring grey level pixel-pairs (neighbours), where the first pixel has intensity i and the second has intensity j. Prior to constructing GLCM, two parameters should be decided – θ is the direction of pixel pairs, d is the distance between the pixels. In practice distance d is dependent on the resolution of texture (scale), while direction θ might have a significant influence on GLCM for textures with definite pattern structure. Figure 7c illustrates four unique directions $0°$, $45°$, $90°$ and $135°$ for the two-dimensional image. Blue arrow shows pixel-pair with direction $0°$ and distance 2. The example GLCM (fig. 7b) was constructed using $\theta = 0°$ and $d = 1$.
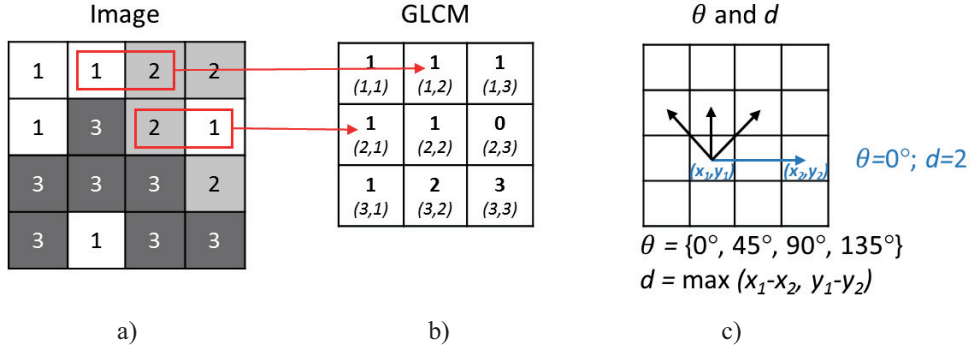
Figure 7. The grey level co-occurrence matrix: a) pixel representation of an image; b) GLCM of an image; c) GLCM construction parameters θ (direction in degrees) and d (distance between the pixels)

The textural features (feature vectors) are calculated from GLCM (Haralick et al. 1973; Zheng et al. 2006). *ImageJ* has a plugin "GLCM Texture" v.0.4, created by Julio E. Cabrera and further updated to "GLCM Texture Too" v. 0.008 by Toby C. Cornish. The output from the program can be given in 11 GLCM feature vectors, calculated for each image: Angular Second Moment (ASM), Contrast, Correlation, Inverse Difference Moment (IDM), Entropy, Energy, Inertia, Homogeneity, Prominence, Variance and Shade. The detailed description, explanation and equations for the above GLCM texture features can be found in the literature (Conners et al. 1984; Haralick et al. 1973; Zheng et al. 2006).

In **Paper I** all 11 textural features of the images were calculated and processed by multivariate analysis. The grey level co-occurrence matrixes were constructed for 4 directions θ (0°, 45°, 90° and 135°) and 5 distances *d* between the pixels (1, 2, 3, 5, 10), and the resulting values of feature vectors were analysed and compared.

It was found that direction θ and distance *d* do not have a significant effect on the textural features of flocs images, so in further research θ = 0° and *d* = 1. Some of the 11 textural features are highly correlated. Therefore, only feature vectors Contrast, Entropy, Homogeneity and Variance are used in **Paper II**, **III** and **IV**.

Contrast *C* is the amount of local grey level variations from one pixel to its neighbour:
$$C = \sum_{i,j}(i-j)^2\, k(i,j)\,, \tag{4}$$
where $k(i, j)$ is the relative frequency or, in other words, a normalised number $K(i, j)$. $k(i, j)$ is found by dividing $K(i, j)$ with the sum of the co-occurrence matrix.

Entropy *E* is the measure of disorder in the image – statistical randomness:
$$E = -\sum_{i,j} k(i,j)\log\big(k(i,j)\big). \tag{5}$$

Homogeneity *H* is the measure of closeness of the values in GLCM to the diagonal. For the inhomogeneous image, *H* will be relatively lower than for homogeneous image, changing in the range from 0 to 1.
$$H = \sum_{i,j}\frac{1}{1+(i-j)^2}k(i,j). \tag{6}$$

Variance *V* is the measure of deviations from the mean value of $k(i, j)$ in the image (sum of squares):
$$V = \sum_{i,j}(i-\mu)^2 k(i,j)\,, \tag{7}$$
where μ is the mean value of $k(i, j)$ in GLCM.

3.4 Pilot scale experiments with municipal wastewater

The concepts developed and evaluated on model wastewater was then verified at the Skiphelle wastewater treatment plant (Drøbak, Norway). Skiphelle WWTP receives municipal wastewater from Drøbak city and the neighbourhood area. Average inlet flow is 4600 m³/day during the days without snowmelt and/or precipitations. The results of studies conducted in Skiphelle WWTP are reported in **Paper III** and **Paper IV**.

Skiphelle WWTP is a mechanical-chemical precipitation plant. The treatment process consists of the following stages; screens, two parallel pre-sedimentation basins, three sequenced flocculation chambers with the different velocity gradients, and two parallel sedimentation chambers. The plant also has a sludge dewatering and thickening system. The inlet and outlet water quality parameters are measured by online sensors and recorded (average values) with 10 min intervals. The data is available for observation in the plant's SCADA system and through the DOSCON® system. The retrieved water parameters include inlet wastewater flow (QIN), inlet pH (PHI), inlet turbidity (TUI), inlet conductivity (CNI), wastewater temperature (TMP), coagulant dosage (Dose), pH after coagulant dosage (PHO) and outlet turbidity (TUO). The plant operators perform daily sampling of inlet and outlet total Phosphorous. Coagulant used in the Skiphelle WWTP is polyaluminium chloride (ECOFLOCK 90, Feralco), 9 ± 0.3 % Al by weight, and density 1356 ± 25 kg/m³.

A special image acquisition cuvette was constructed from 4 mm Plexiglas. The construction and dimensions of the cuvette are shown in fig. 8a. The installation was installed above the second flocculation chamber and consisted of the tube, peristaltic pump, image acquisition cuvette, digital camera and computer (fig. 8b). The schematic representation of the installation is shown in fig. 9a. The chosen tube was 3 cm in diameter to minimise the potential danger of flocs breakage, and the peristaltic pump was placed after the imaging cell. The water flow in the system was upstream of the flocculation chamber and manually adjusted to approximately 40 l/h.
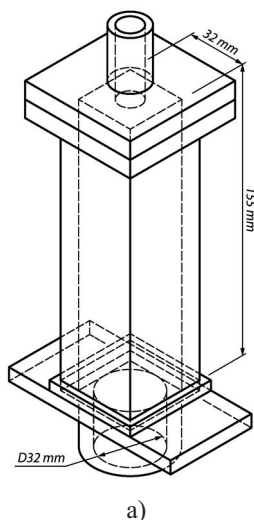

a)   b)

Figure 8. Installation in Skiphelle WWTP: a) drawing of the acrylic cuvette; b) photo of the system

The image acquisition system presented in **Paper III** consisted of DSLR Nikon D600 camera, 105 mm Nikkor AF-S Micro 1:2.8 G ED lens (Nikon®, Japan), SpeedLite YN460 flash (Yongnuo, China). Images of flocs were taken continuously with a pre-set repeatability using free remote camera control software – *DigiCamControl 1.2.0* (Istvan 2012). The size of the image-capturing zone in the cuvette was 3.3 cm × 10.3 cm. In order to obtain flocs with the proper depth of field, the black metal stripe was placed in the centre of the cuvette, which also became a background for the flocs. The contrasting background colour was chosen against the greyish coloured wastewater flocs, making it easier to perform further image analysis. The obtained images had a resolution of 24.3 megapixels each. They were processed in *ImageJ.* For each image 1380 × 3640 pixels (2.4 cm × 6.3 cm) was cropped by manual investigation of the area. Because of slight changes in lighting conditions during image acquisition, all images were pre-processed to mean centre the images' grey-tone values in order to have the same brightness.

Results presented in **Paper IV** were obtained from the same installation. However, an expensive Nikon D600 camera and a computer were replaced by a Raspberry Pi 2 Model B V1.1 single-board computer (price range 35$) and a low-cost (15$) 5 megapixel (OV5647 sensor) camera module with the changeable focal length (fig. 9b). A special program was written in Python to control the camera and make changes in the camera settings. A floodlight LED (anslut® 427-624, 600 lumens) was used as a light source. 3 images were captured every 10 minutes, with 5 second intervals between each images. The size of the image-capturing zone in the cuvette was 3.2 cm × 9.6 cm. The obtained images had a resolution of 5 megapixels each. They were processed in *ImageJ.* For each image 1360 × 1360 pixels (3 cm × 3 cm) was cropped by manual investigation of the area.
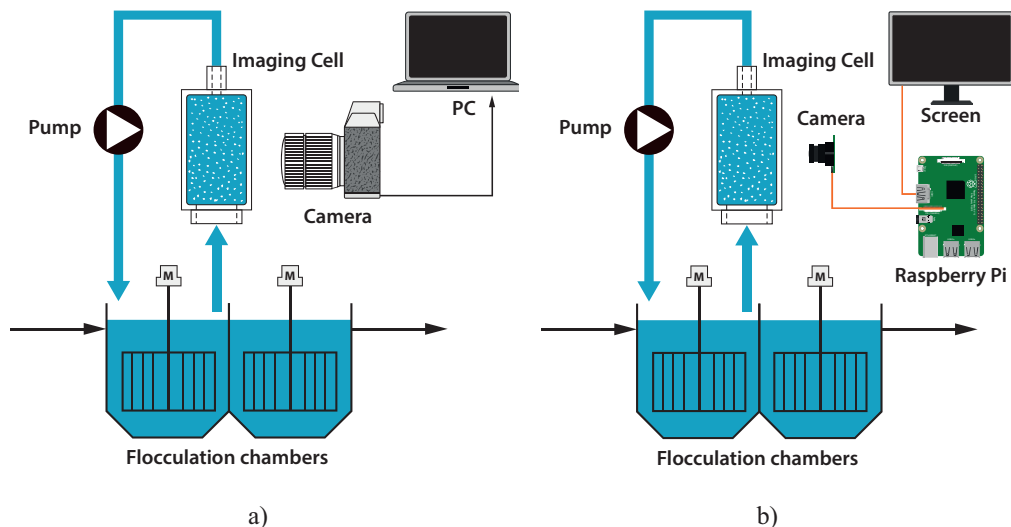

a)   b)

Figure 9. Schematic representation of the installation: a) with DLSR camera; b) with Raspberry Pi board and camera module

4 Results and discussion

4.1 Particle size distribution in model wastewater

The range of initial colloid particle sizes in the model wastewater are important and needs to be defined. Particle size distribution in model wastewater (fig. 10a, b) was measured by Mastersizer 3000. Further tests were also conducted on whether the sizes of coagulated particles could be measured without breaking the flocs (fig. 10c).

Fig. 10a represents the particle size distribution of model wastewater components diluted in deionised water. The concentrations of prepared solutions were corresponding to those in model wastewaters. The humic acid solution has a quite broad particles size distribution with peaks at approximately 9 µm and 190 µm. The bentonite solution has a peak at approximately 25 µm size particles. The potato starch solution has a peak at approximately 45 µm. The dried milk solution has three peaks at approximately 1.7 µm, 12 µm and 75 µm.

Fig. 10b represents the particle size distribution of different model wastewater types according to the level of contaminants and water hardness – medium soft (*MS2*), medium hard (*MH*), low soft (*LS2*), low hard (*LH*), high soft (*HS*) and high hard (*HH*). Both soft and hard model wastewaters with a medium concentration of contaminants contain a higher amount of smaller particles, compared to other model wastewater types. The particle size distribution peaks are at approximately 1.7 µm, 12 µm, corresponding to the dried milk particles' sizes. The other four model wastewater types contain a higher amount of slightly bigger particles (peaks at approximately 1.7 µm and 45 µm), corresponding to the dried milk and potato starch particle sizes peaks.

The medium soft model wastewater (*MS2*) was coagulated with Kemira prepolymerised polyaluminium chloride PAX XL-61, dosage 0.65 mmol Al/l. Flocs appeared during coagulation were transferred to the sampling bottle for further analysis. The breakage of aggregates was visible during transfers. Fig. 10c represents a continuous shift of the particles size distribution peak towards smaller scale during the 15 repetitive laser diffraction measurements. In order to assure that the two-phase system is stable, the average distribution curve should fit certain standard deviation range requirements. The system with coagulated particles has a stable drift with each next measurement, meaning that the flocs are breaking during the circulation in the instrument. Most probably, the breakage is caused by the instrument's peristaltic pump.

Even though Mastersizer 3000 works with the particles in a range from 0.01 µm to 3500 µm, suitable for the studied flocs, the aggregates formed during the coagulation of model wastewater are very fragile. The system was able to measure the particle size distribution for initial model wastewaters and solutions of model wastewater components. However, the estimation of the size of coagulated flocs was not successful due to their fragile nature and breakage during the measurement process.

The conclusion of this part of the studies is that non-intrusive image analysis methods need to be employed to measure the flocs characteristics.
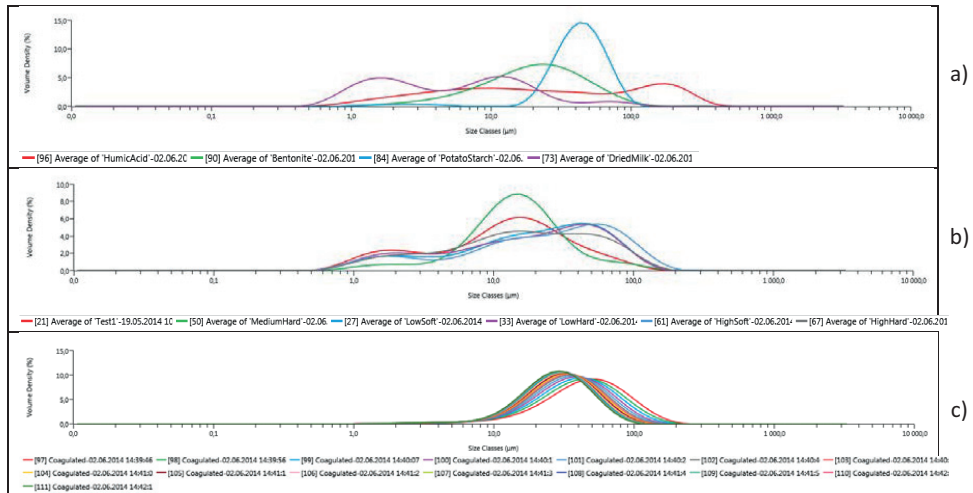
Figure 10. Particles size distribution: a) components of model wastewater; b) different types of model wastewater; c) coagulated model wastewater (MS2, coagulant dose - 0.65 mmol Al/l of PAX XL-61). In collaboration with Reidar Schüller (IKBM, NMBU).

4.2 Image analysis methods and challenges (Paper I)

First of all, it was important to verify that flocs formed under different coagulation conditions have unique structures. Afterwards, we had to test which image analysis techniques was best suited to quantify the images of wastewater flocs.

A non-intrusive image acquisition system described in section 3.2 was used during the slow mixing stage of coagulation in a batch process (jar tests). Fig. 11 shows the relation between coagulant dosage in mmol Al/l and total P efficiency in %. The representative images of flocs are shown for different coagulant dosages. All flocs images in fig. 11 have the same scale and were captured after 400 seconds of slow mixing. Visually, the flocs corresponding to lower coagulant dosages and low total P removal rates are fewer and have bigger size. With the increase in coagulant dose, which results in better treatment efficiency, the formed flocs decreased in size but increased in quantity. These results are consistent with findings Wang et al. (2011).

The first conclusion here is that the flocs aggregated during the addition of different amounts of coagulant have size variations specific to their conditions. Hence, the images of flocs for different dosages are unique and the different structures on images should be possible to quantify.

Researchers often use the object recognition image analysis methods to count particles, estimate particles sizes and fractal dimensions from the images (Chakraborti et al. 2003; Chang et al. 2005; Dukhin et al. 2007; Yu et al. 2009; He et al. 2012; Yu 2014). Such methods were applied to characterise the flocs on images captured during slow mixing of model wastewater coagulation with 11 dosages of the coagulant. The detailed description of the method is given in section 3.3.2.
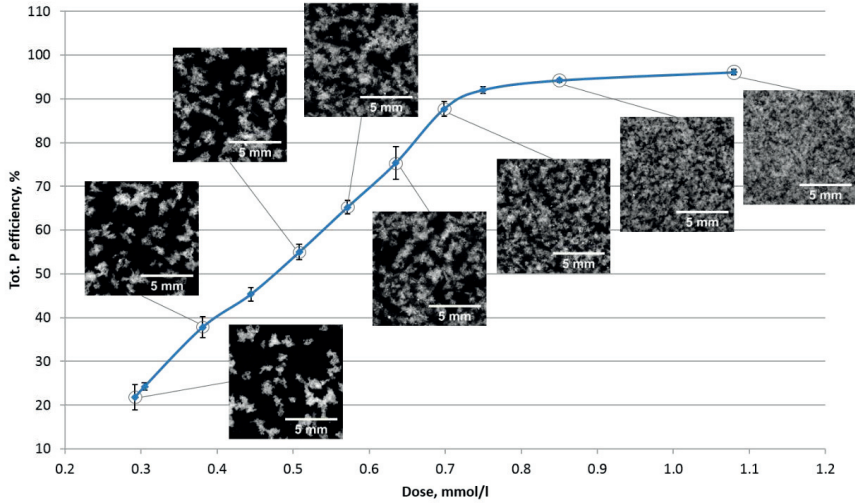
Figure 11. Dependence of total P removal from coagulant dosage and its associated image of flocs during coagulation

With the lower dosages of coagulant, the aggregated particles had bigger size and were easy to determine on the image (fig. 12a). However, with the addition of higher amounts of coagulant, particles were becoming smaller and overlapping in the image. Fig. 12b illustrates the particle detection failure. A higher amount of particles appeared during coagulation with 0.64 mmol Al/l, which resulted in overlapping of the aggregates in a 2D image. The yellow boundary (fig.12b) on the image indicates the structure, which was recognised as one separate particle. This is known to be the wrong estimation by the visual investigation of flocs size during the slow mixing stage. With the dosage 0.85 mmol Al/l it was not possible to detect the separate flocs (fig. 12c) by the particles recognition image analysis method in the given images scale and with such amount of aggregates. Some difficulties have been observed with the application of the particles recognition image analysis method for such highly contaminated water as wastewater.
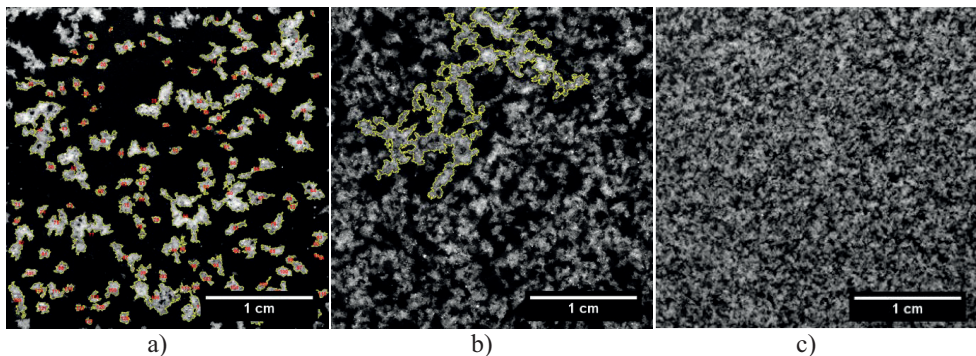

a)   b)   c)

Figure 12. Flocs detection by particle recognition image analysis method: a) coagulant dosage 0.29 mmol Al/l, total P efficiency 22%; b) coagulant dosage 0.64 mmol Al/l, total P efficiency 51%; c) coagulant dosage 0.85 mmol Al/l, total P efficiency 96%.

The challenge on separate particles recognition spurred us to search for the alternative methods of images quantification. Texture image analysis methods were intensively employed in other research areas such as medicine, food industry, slurry and powder production, geographic information systems. However, our literature search did not reveal the application of texture image analysis techniques in wastewater flocs

We have tested three texture recognition methods – analysis of image by its Histogram, GLCM and AMT. Image analysis by Histogram and AMT are described in Paper I, sections 2.5 and 2.7, respectively. GLCM calculation and features extraction are described in section 3.3.3 of this thesis.

Table 4 represents the results of multivariate regression analysis for three different image analysis techniques. The predicted parameter (response Y) for all multivariate models was the coagulant dosage in mmol Al/l, while the predictors matrix X was either the data retrieved from images Histograms, or GLCM feature vectors or AMT spectra. It is important to notice that no information regarding coagulants dosages or wastewater initial and post-treatment parameters were included in the model during prediction. Hence, the prediction applied during coagulation dosages was purely based on information retrieved from the images of flocs by different texture analysis methods.

Table 4. Comparison of multivariate regression analysis results for 3 texture image analysis techniques

Regression method	PCR		PLSR	
Number of components/ factors	PC-2		Factor-2	
Explained Y variance	R^2 cal., %	R^2 val., %	R^2 cal., %	R^2 val., %
Histogram, 256 bins	87.39	83.53	96.35	88.84
Histogram, 128 bins	90.28	88.53	95.07	93.03
Histogram, 64 bins	91.55	89.92	95.51	93.80
GLCM, 4 variables	87.54	87.95	92.89	92.42
GLCM, 6 variables	92.76	92.83	95.58	95.22
GLCM, 11 variables	95.91	95.59	96.89	96.09
AMT, 200 variables	95.84	95.01	96.90	96.19

The explained total validation variance has high values for all tested models, with a minimum of 83 %. It means that the information retrieved from the images of flocs and quantified by different texture methods is correlated with the coagulant dosages. In other words, the images of flocs for particular dosages are unique, and information from the images can be used for dosage prediction.

Partial least squares regression (PLSR) models based on GLCM feature vectors and AMT spectra resulted in 96 % explained validation variance. Hence, both texture analysis techniques can be used for the coagulant dosage prediction. However, from a practical point, it is more convenient to implement GLCM measurements and feature extraction than AMT spectra detection. It also requires less time and computational capacity of the processor to perform GLCM analysis. It was decided to focus on one texture analysis method – GLCM, for further studies and a sensor prototype development.

4.3 Relationships between flocs images, coagulant dosages, initial and post-treatment parameters of model wastewaters (Paper II)

The purpose of this part of the research work was to find the relationships between flocs images, coagulant dosages, initial and post-treatment parameters of model wastewaters. The obtained results, which were presented and discussed in Paper I, needed to be verified in a broader range of initial conditions. We aimed to simulate the different wastewater conditions while concurrently repeating the analysis. Hence, 9 types of model wastewaters with different initial concentrations of particles and phosphates were prepared and coagulation jar tests were performed, together with image analysis.

Fig. 13 represents the coagulation results to compare the outlet removal efficiencies of TSS (13a), turbidity (13b), ortho-P (13c) and total P (13d). The results are shown for model wastewater type *LS2* (low concentration soft water) and 3 coagulants, with similar trends observed in several other water types.

Overall, the results indicate that the particles in terms of TSS and turbidity are better removed with a prepolymerised coagulant. Ortho-P and total P have higher removal rates with non-polymerised ALS and PIX-313. These results are consistent with previous findings (Ratnaweera 1991).
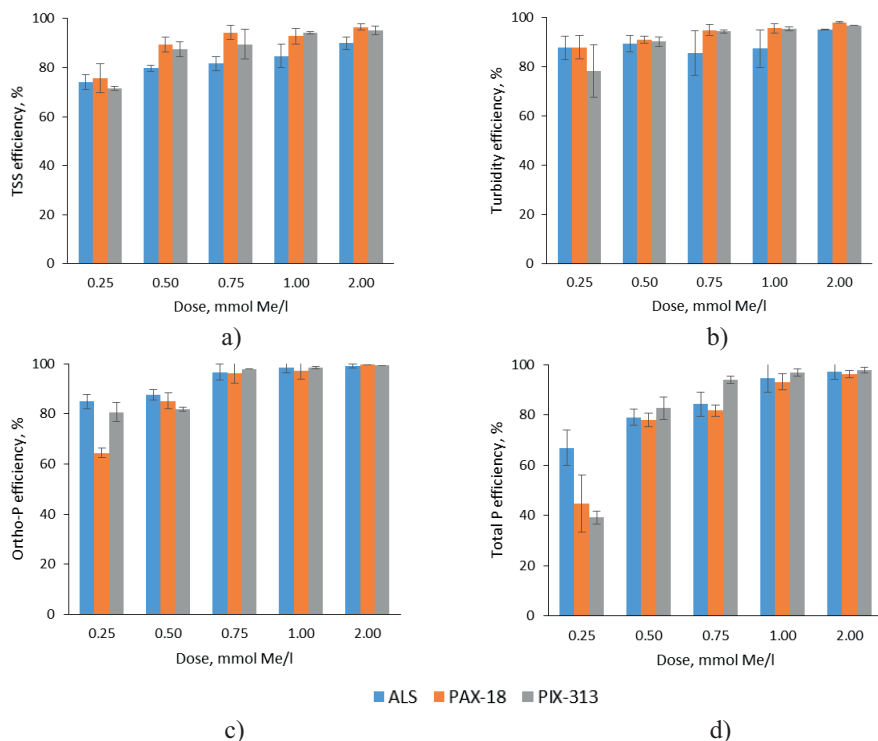

Figure 13. Removal efficiencies after the jar tests with LS2 and 3 coagulants: a) total suspended solids; b) turbidity; c) ortho-phosphates; d) total phosphorus.

Fig. 14 represents the coagulation results for model wastewater type *MS2* (medium concentration soft water) in terms of total suspended solids (14a) and total phosphorous (14b) removal rates. The general trend of coagulants performance is similar to the one described for model wastewater type *LS2*.
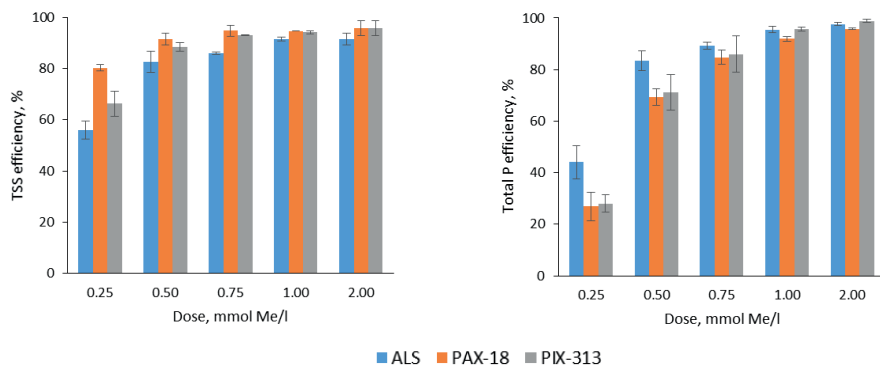
Figure 14. Removal efficiencies after the jar tests with MS2 and 3 coagulants: a) total suspended solids; b) total phosphorus.

The results of the principal component analysis (PCA) based on data from jar tests with all soft model wastewater types and aluminium sulphate as a coagulant are shown in the scores plot (15a) and loadings plot (15b). The data set also included information from the flocs images that were captured during the slow mixing stage of the coagulation process and converted to the GLCM feature vectors Contrast, Entropy, Variance and Homogeneity. According to the loadings plot, all initial model wastewater parameters are correlated, such as inlet turbidity (TUI), inlet total suspended solids (TSSI), inlet total P (TPI), inlet ortho-P (OPI). Outlet model wastewater parameters are also correlated - outlet turbidity (TUO), outlet total suspended solids (TSSO), outlet total P (TPO), outlet ortho-P (OPO). In addition, the outlet values of water qualities are positively correlated with Homogeneity – feature vector calculated from the images of flocs. It means the higher values of outlet water qualities the higher rate of Homogeneity is observed in the associated image of flocs. Contrast and Entropy parameters retrieved from the images are positively correlated with zeta potential and negatively correlated with inlet and outlet model wastewater parameters. The dosage of the coagulant is negatively correlated with Variance on the image.

Figure 15. Results of PCA for ALS coagulant and all types of soft model wastewaters: a) scores plot; b) loadings plot
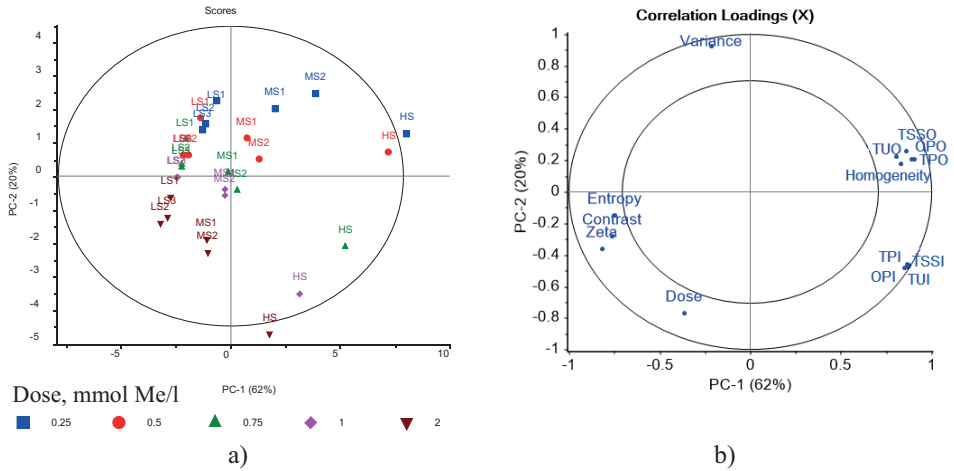
4.4 Approbation of the texture image analysis method in a full-scale (Paper III)

After the successful application of the texture image analysis method for the flocs characterisation in a laboratory batch-scale with model wastewater, the technique needed to be tested in a full-scale with real municipal wastewater. We aimed to determine if the GLCM feature vectors retrieved from the images of flocs during the coagulation process could be used to predict the changes in coagulation conditions and the outlet turbidity values after sedimentation.

During the experimental period in Skiphelle WWTP a long rain event took place (approximately 1 week of rain). Furthermore, this treatment plant regularly receives the septic tank contents by the discharges from special trucks. At these discharge times the inlet wastewater turbidity increases by more than two times, compared to regular municipal discharges. Hence, the data received from the plant can be divided to three major classes concerning the inlet wastewater parameters: normal operational conditions, wet-weather (rain) conditions and events with high inlet turbidity values due to septic tank discharges.

Fig. 16a shows the PCA results obtained from the images of flocs data. In this case, no information about inlet, outlet conditions of the wastewater and coagulant dosages was revealed as the input parameters to the model. Only the GLCM feature vectors were used for conducting the PCA. Nevertheless, the scores on the PCA plot formed the three distinct clusters which correspond to the inlet wastewater conditions, meaning that the information from the images of flocs is correlated with the wastewater quality and quantity parameters. GLCM textural features Contrast, Prominence, Variance and Entropy are positively correlated with the images of flocs during normal operational conditions. Wet-weather conditions have higher values of Shade, Correlation and Energy, while the flocs formed during high inlet turbidity result in higher values of Homogeneity and IDM.

For the comparison, fig. 16b represents results from the conducted PCA including all the wastewater quality parameters, coagulant dosages and GLCM textural features of the flocs images. Observed scores clusters are similar to those shown in fig. 16a. However, the clusters are more distinct because the coagulation operation conditions are included in the model.
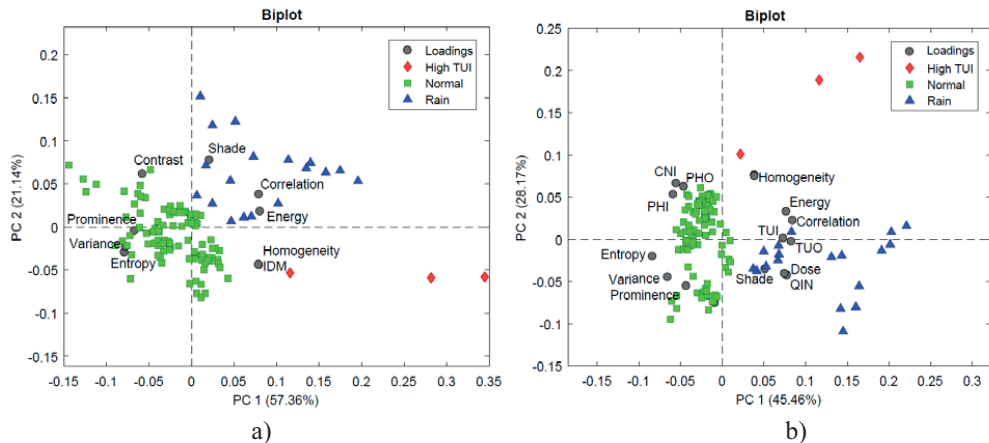

	a)	b)

Figure 16. Biplots of PCA, PC1 vs PC2 for: a) data with 9 GLCM textural features of the images of flocs; b) data containing wastewater quality parameters, GLCM textural features and the coagulant dosage.

In practice, the obtained results mean that the images of flocs quantified by GLCM textural features correspond to inlet municipal wastewater parameters and coagulation conditions. In other words, the flocs images readings can potentially be used to replace some of the sensors. In addition, it might be possible to use the system based on image analysis of flocs for the choice of an action. Because the increase of precipitations in amplitude and frequency are growing due to climate change impacts, WWTPs must take certain actions in the nearest future to sustain their efforts in meeting current requirements for the discharged wastewater qualities. For instance, the different algorithms of coagulant-flocculant dosage predictions can be applied and the choice of such algorithms can be based on the information from the flocs images.

As was mentioned above, climate change is expected to increase its influence on the performance of WWTPs. The amounts of water received by the WWTPs due to more frequent rain events, storms and floods are expected to rise significantly. This will result in not only increased amounts of bypasses and overall reduction of treatment efficiencies in the plants, but will particularly also cause problems for the coagulation process. The liquid-solid separation processes following the flocculation chambers are designed for certain ranges of hydraulic loads. If the amounts of water entering the separation stages are too high, the process might be disturbed. Consequently, the parts of settled or floated sludge might be discharged together with the treated water effluent. In other cases, the retention time of the separation process might be significantly decreased, resulting in a lack of capacity for liquid-solid separation. An example of the sedimentation failure is shown in fig. 17. After four days of rain and increased storm water load to the Skiphelle WWTP, accumulations of sludge were floating on the surface of the sedimentation chamber and had to be removed manually by plant operators. These events

are known to happen during wet-weather conditions. In such cases, the outlet turbidity values increase, indicating poor treatment.

Figure 17. Failure of the sedimentation process – floating sludge

The performance of the treatment plant is evaluated by discharged wastewater quality, which influences the recipient water bodies and the environment in general. Thus, the outlet wastewater quality should be continuously measured and evaluated. More efforts should be made towards troubleshooting during wet-weather conditions in an effort to keep the wastewater effluent quality within the range of limits set by the authorities.

Because of the considerable time lag between a coagulant injection point and the effluent from the sedimentation chamber, it does not make sense to steer the dosing pump based on the outlet turbidity values, since the inlet wastewater qualities and amounts might have changed during that time. Particularly, based on the results of tracer tests performed in Skiphelle WWTP, the time lag was found to be approximately one hour during normal working conditions, while the changes in the inlet conditions were more frequent.

Currently, the development of the soft sensor raises the attention of researchers and instruments providers. We have performed the tests of predicting the outlet turbidity values (TUO) of the municipal WWTP by PLSR. Fig. 18a shows the results of TUO prediction using inlet wastewater quality parameters and coagulant dosages as predictors. The resulting cross-validation R^2 was 0.79. However, most important is the observation that several samples of the data set were considerably underestimated. The maximum TUO limit was set to be 5 NTU. The samples 94 and 95 had TUO values 10.9 and 10.5 NTU, respectively. If the prediction of TUO is underestimated, in this case calculated by PLSR values to be around 4 NTU, the automatic alarm system would not trigger and the action to prevent the increase in outlet turbidity would not be executed.

Fig. 18b shows the results of PLSR that has the highest cross-validation R^2=0.85 among the tested combinations of predictors. The predictors of TUO were as follows: inlet wastewater flow, inlet turbidity and four GLCM textural features – Variance, Prominence, Correlation and Contrast. In comparison with the previous prediction based only on water quality parameters, here the samples 94 and 95 were estimated correctly.

The ambition is to use the image analysis data from the floc sensor to improve the prediction of outlet turbidity values, to enable a preventive action in advance of coagulation

process and/or sedimentation stage failure. The predicted outlet turbidity values can be used to adjust the coagulant dosage well in advance, which is especially important in cases with wet-weather conditions and rapid changes in inlet wastewater qualities.
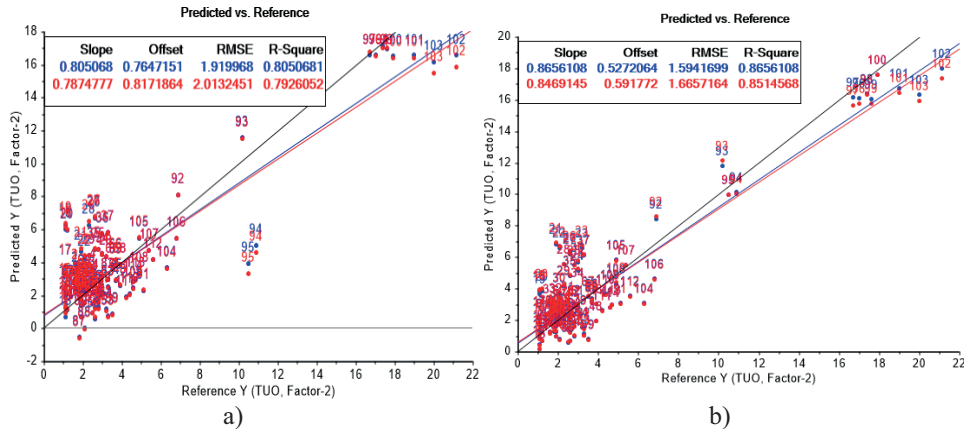

Figure 18. Comparison of two effluent turbidity prediction models: a) TUO=f(QIN, TUI, CNI, PHI, Dose); b) TUO=f(QIN, TUI, Variance, Prominence, Correlation, Contrast)

4.5 Development and testing of the floc sensor prototype (Paper IV)

The sensor prototype was developed using a low-cost computer and camera module. It was tested in the WWTP with municipal wastewater. The ambition and the main focus of this research was to develop a sensor prototype for optimising the coagulant-flocculant dosage prediction. The dosages were predicted by PLSR using inlet wastewater parameters, coagulant dosages and GLCM feature vectors from the images of flocs.

The software developed in Python for the camera module control make it possible to adjust all settings available for that camera. Furthermore, the software has a menu window to set the time laps preferences and even start the sensor at a particular time. These are the other important advantages (in addition to the low price) of the system, compared to having a DLSR camera that is not as flexible with the time lapse control. In addition, because the single-board computer has a Wi-Fi module the captured/processed information can be sent right away to the server, database cloud or remote computer.

For the prediction of coagulant dosages the data matrix was divided into calibration and test data sets, 60 % and 40 % of the data, respectively. The data was divided based on the outlet turbidity values. The X matrix for PLSR included inlet wastewater parameters – QIN, TUI, PHI, TMP; after the dosage measurement PHO; an hour of the day; and GLCM textural features – Contrast, Entropy, Homogeneity and Variance. The response Y was coagulant dosage in ml/s. The PLSR model was calibrated on the data values that correspond to the outlet turbidity measurements between 1.9-5 FNU (desired range of effluent turbidity for the Skiphelle WWTP). The test data set included measurements related to outlet turbidity values higher than 5 FNU.

Fig. 19 shows the results of PLSR – coagulant dosage prediction. The continuous red line is the reference dose (dosages used in the WWTP). Black empty squares represent the dosage prediction of calibration data with corresponding TUO in a range 1.9-5 FNU. Black circles are the dosage predictions that correspond to TUO less than 1.9 FNU. Black diamonds are the dosage predictions that correspond to TUO above 5 FNU. Minimum desired value of effluent turbidity – 1.9 FNU and maximum value 5 FNU are marked by dashed green lines, TUO min and TUO max, respectively.

Overall, the predicted coagulant dosages precisely follow the reference dosages. Prediction R^2 equals 0.92 for calibration and 0.78 for validation with three factors, root mean square error (RMSE) for calibration is 0.182 and 0.297 for validation. The area with high TUO represents a rain event, dosage was manually adjusted by plant operators and tend to be under-estimated. The dosages predicted by the PLSR model (black diamonds, fig. 19) suggest having higher coagulant use for the wet-weather period.
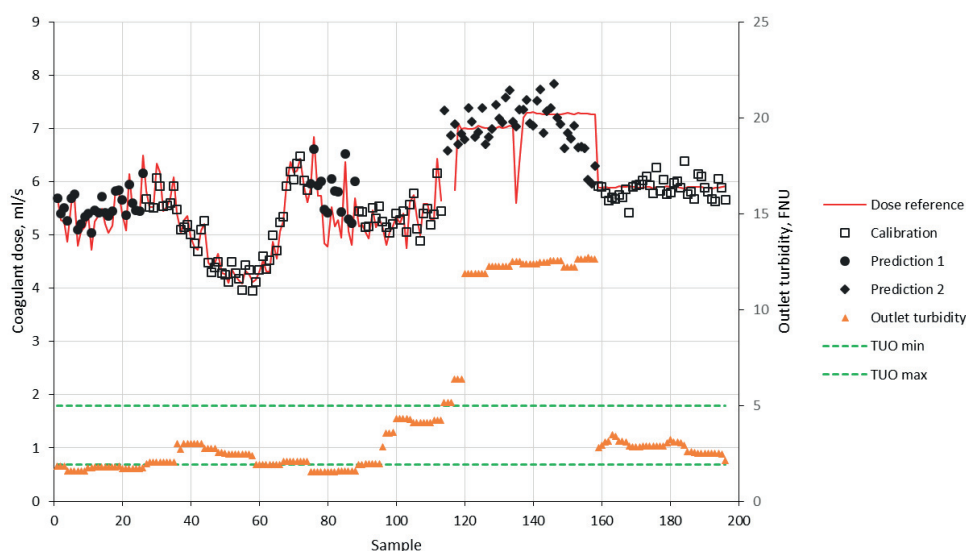
Figure 19. Comparison of the predicted by PLSR coagulant dosages with reference dosages and corresponding effluent turbidity measurements.

Conclusions

Coagulation, being one of the widely used and robust methods in water and wastewater treatment, needs further improvement regarding optimising dosing control. The existing methods are still far from accessible for many plants, thereby highlighting the need for more accurate, simplified and affordable concepts.

Image analysis is identified as a potential method to improve existing multi-parameter based coagulant dosing strategies. However, conventional methods for measuring particles sizes, such as laser diffraction, are not applicable for the flocs appearing during wastewater

coagulation, due to their breakage during the analysis. Conventional methods of particles detection by image analysis, such as calculations of size and fractal dimension using object recognition methods were found to be challenging to apply for wastewater flocs due to the complex image processing techniques needed to be employed. It was found that the flocs overlap in the images due to their small size and massive amount within the scale of non-intrusive image analysis by the digital camera. Hence, the technique was not applicable in the scope of this research. Besides, the main aim of the studies was to construct a floc sensor prototype and use the data from images for coagulant dosage prediction. As a result, the precise characterisation of particles themselves was not a key focus.

Texture image analysis methods proved to satisfy the needs of successful flocs images characterisation during the coagulation-flocculation process. Images of flocs captured during different coagulation conditions resulted in distinct and detectable textural features extraction, for both model wastewater and real municipal wastewater flocs. Thus, images of flocs are unique for different wastewaters, dosages and types of coagulants.

Images of flocs correspond to the coagulation treatment conditions in the WWTP with municipal wastewater. The information retrieved from the images of flocs enabled the detection of normal and wet-weather conditions as well as rapid changes of inlet wastewater parameters. Thus, the flocs' information can be used to trace and react to process changes that are due to variations in inlet wastewater qualities.

The information from the images of flocs significantly improves the prediction of optimal coagulant dosages compared to flow-proportional dosing and models that only include inlet wastewater parameters.

The addition of information retrieved from the flocs images to the regression model improved the outlet turbidity prediction compared to the model containing just inlet wastewater characteristics. The system can be used as an early alarm of coagulation process failure, predicting outlet turbidity values and assessing it to be within the required treatment range.

Recommendations for further studies

A sketch of the floc sensor prototype based on Raspberry Pi single-board computer with a camera module is shown in fig. 20. The sensor prototype should be sealed into a waterproof cabinet and tested for functionality under water, directly in the flocculation chamber. The successful development of this approach will simplify the full-scale installation used in these studies and aggregates sampling. The development would remove the need for the peristaltic pump for wastewater circulation, thus removing a further additional electricity costs. However, it must be expected that the sensor might become dirty quite fast during wastewater coagulation. Hence, an automated cleansing system of the camera lens or regular manual cleanings of the sensor are required.

A proper lighting source for the camera module should be found. Another solution for working in a dark environment might be to employ infrared cameras. The special NoIR camera

module for Raspberry Pi can be found in the market. However, further tests are needed to compare the images of flocs captured by the regular camera and NoIR modules.

Acoustic sensors and acoustic chemometrics is another method that can be tested in its applicability to wastewater coagulant dosage prediction and control.

The floc sensor prototype can be further developed into a self-standing digital image analysis system. Since Raspberry Pi is a computer, it can automatically perform GLCM calculations and textural features extraction when the new images of flocs are captured. A special software should be written for such operations.

Regarding the placement of the sensor, it is important to test the images from each flocculation chamber in the sequence to identify which are best suited for multivariate modelling and coagulant dosage and/or outlet water treatment parameters prediction.

The GLCM features extraction was applied for the grey-scale images in this PhD work. Techniques such as colour GLCM also exist in the current market. It is recommended to apply this technique and test it on the images of flocs. It is likely that the colour information on the image is important and should not be neglected.

The texture image analysis method of data extraction can be an alternative solution for other particles characterisation methods and processes that contain non-fragile particles. The texture features of the particles retrieved from the images may potentially correlate with the real measured values, for instance by the laser diffraction method. However, this should be further tested with the assumption that the calibration is to be done for each particular system.
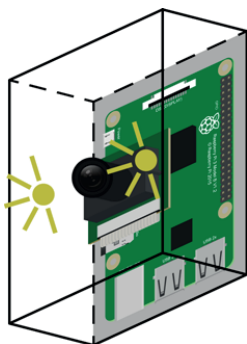

Figure 20. Sketch of the sensor prototype sealed in a waterproof cabinet

Finally, the floc sensor should be integrated into the DOSCON type ADCS system to evaluate its potential to reduce the use of online sensors and its impact on optimal dosage predictions.

References

Abdelmounaime, S. & Dong-Chen, H., 2013. New Brodatz-based image databases for grayscale color and multiband texture analysis. *ISRN Machine Vision*, pp.1–14.

Alves, M. et al., 2000. Characterisation by image analysis of anaerobic sludge under shock conditions. *Water Science and Technology*, 41(12), pp.207–214.

Åmand, L., Andersson, S. & Harding, M., 2017. Instrumentation at Swedish WWTPs – a survey and interview study. In *12th IWA Specialized Conference on Instrumentation, Control and Automation*. Quebec, Canada, 11-14 June 2017, pp. 383–390.

Amaral, A.L. & Ferreira, E.C., 2005. Activated sludge monitoring of a wastewater treatment plant using image analysis and partial least squares regression. *Analytica Chimica Acta*, 544(1–2), pp.246–253.

Amirtharajah, A. & Mills, K.M., 1982. Rapix mix design for mechanism of alum coagulation. *Journal of American Water Works Assoc*, 74(4), pp.210–216.

Andrle, R., 1994. The angle measure technique: A new method for characterizing the complexity of geomorphic lines. *Mathematical Geology*, 26(I), pp.83–97.

Bache, D. & Gregory, R., 2007. *Flocs in water treatment*, IWA Publishing.

Baxter, C.W. et al., 2002. Model-based advanced process control of coagulation. *Water science and technology*, 45(4–5), pp.9–17.

Van Benschoten, J.E. & Edzwald, J.K., 1990a. Chemical aspects of coagulation using aluminium salts - I. Hydrolytic reactions of alum and polyaluminum chloride. *Water research*, 24(12), pp.1519–1526.

Van Benschoten, J.E. & Edzwald, J.K., 1990b. Chemical aspects of coagulation using aluminium salts - II. Coagulation of fulvic acid using alum and polyaluminum chloride. *Water research*, 24(12), pp.1527–1535.

Bernhardt, H. & Schell, H., 1993. Effects of energy input during orthokinetic aggregation on the filterability of generated flocs. In *Water Science and Technology*. pp. 35–65.

Bharati, M.H., Liu, J.J. & MacGregor, J.F., 2004. Image texture analysis: Methods and comparisons. *Chemometrics and Intelligent Laboratory Systems*, 72(1), pp.57–71.

Bo, X. et al., 2012. Effect of dosing sequence and solution pH on floc properties of the compound bioflocculant-aluminum sulfate dual-coagulant in kaolin-humic acid solution treatment. *Bioresource technology*, 113, pp.89–96.

Boisvert, J.P. et al., 1997. Phosphate adsorption in flocculation processes of aluminium sulphate and poly-aluminium-silicate-sulphate. *Water Research*, 31(8), pp.1939–1946.

Bratby, J., 2016. *Coagulation and flocculation in water and wastewater treatment*, IWA Publishing.

Brodatz, P., 1966. *Textures: a photographic album for artists and designers*, Dover, New York, USA.

Bushell, G. et al., 2002. On techniques for the measurement of the mass fractal dimension of aggregates. *Advances in Colloid and Interface Science*, 95, pp.1–50.

Camp, T. & Stein, P., 1943. Velocity gradients and internal work in fluid motion. *Journal Boston Society of Civil Engineers*, 1(4), pp.219–237.

Chakraborti, R.K. et al., 2003. Changes in fractal dimension during aggregation. *Water research*, 37(4), pp.873–883.

Chakraborti, R.K., Atkinson, J.F. & Van Benschoten, J.E., 2000. Characterization of Alum Floc by Image Analysis. *Environmental Science & Technology*, 34(18), pp.3969–3976.

Chang, Y., Liu, Q.-J. & Zhang, J.-S., 2005. Flocculation control study based on fractal theory. *Journal of Zhejiang University Science B*, 6(10), pp.1038–1044.

Conners, R.W., Trivedi, M.M. & Harlow, C.A., 1984. Segmentation of a high-resolution urban scene using texture operators. *Computer Vision, Graphics, and Image Processing*, 25(3), pp.273–310.

Critchley, R., Smith, E. & Pettit, P., 1990. Automatic coagulation control at water-treatment plants in the North-West region of England. *Water and Environment*, 4(6), pp.535–543.

Dagot, C. et al., 2001. Use of image analysis and rheological studies for the control of settleability of filamentous bacteria: application in SBR reactor. *Water Science and Technology*, 43(3), pp.27–33.

Dahl, C.K. & Esbensen, K.H., 2007. Image analytical determination of particle size distribution characteristics of natural and industrial bulk aggregates. *Chemometrics and Intelligent Laboratory Systems*, 89, pp.9–25.

Dentel, S.K., Thomas, A. V. & Kingery, K.M., 1989a. Evaluation of the streaming current detector-I. Use in jar tests. *Water Research*, 23(4), pp.413–421.

Dentel, S.K., Thomas, A. V. & Kingery, K.M., 1989b. Evaluation of the streaming current detector-II. Continuous flow tests. *Water Research*, 23(4), pp.423–430.

Diamadopoulos, E. & Vlachos, C., 1996. Coagulation-filtration of a secondary effluent by means of pre-hydrolyzed coagulants. *Water Science & Technology*, 33(10–11), pp.193–201.

Duan, J. & Gregory, J., 2003. Coagulation by hydrolysing metal salts. *Advances in Colloid and Interface Science*, 100–102, pp.475–502.

Dukhin, A.S. et al., 2007. Characterization of fractal particles using acoustics, electroacoustics, light scattering, image analysis, and conductivity. *Langmuir : the ACS journal of surfaces and colloids*, 23(10), pp.5338–51.

Edzwald, J.K., 2013. Coagulant mixing revisited: theory and practice. *Journal of Water Supply: Research and Technology-AQUA*, 62(2), pp.67–77.

Esbensen, K.H., Hjelme, K.H. & Kvaal, K., 1996. The AMT approach in chemometrics - first forays. *Journal of Chemometrics*, 10, pp.569–590.

Fettig, J., Ratnaweera, H. & Ødegaard, H., 1990. Simultaneous phosphate precipitation and particle destabilization using aluminium coagulants of different basicity H. H. Hahn & R. Klute, eds. *Chemical Water and Wastewater Treatment. Proceedings of the 4th Gothenburg Symposium 1990 October 1–3, 1990 Madrid, Spain*, pp.221–242.

Fongaro, L. et al., 2016. Application of the angle measure technique as image texture analysis method for the identification of uranium ore concentrate samples: New perspective in nuclear forensics. *Talanta*, 152, pp.463–474.

Fongaro, L. & Kvaal, K., 2013. Surface texture characterization of an Italian pasta by means of univariate and multivariate feature extraction from their texture images. *Food Research International*, 51(2), pp.693–705.

Freedonia, 2014. *World Water Treatment Chemicals. Industry study 3122*, The Freedonia Group, Cleveland, USA.

Giudici, P. & Figini, S., 2009. *Applied data mining for business and industry* 2nd ed., Wiley.

Gonzalez, R.C. & Woods, R.E., 2010. *Digital image processing* 3rd ed., Pearson Prentice Hall.

Gorczyca, B. & Ganczarczyk, J., 1996. Image analysis of alum coagulated mineral suspensions. *Environmental Technology*, 17(12), pp.1361–1369.

Gregory, J., 2006. *Particles in Water: properties and processes*, CRC Press, Taylor&Francis Group.

Gui, W. et al., 2013. Color co-occurrence matrix based froth image texture extraction for mineral flotation. *Minerals Engineering*, 46–47, pp.60–67.

Han, M. & Lawler, D.F., 1992. (Relative) insignificance of G in flocculation. *Journal / American Water Works Association*, 84(10), pp.79–91.

Haralick, R.M., Shanmugam, K. & Dinstein, I., 1973. Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics*, 3(6), pp.610–621.

He, W. et al., 2012. Characteristic analysis on temporal evolution of floc size and structure in low-shear flow. *Water research*, 46(2), pp.509–20.

Huang, J. & Esbensen, K.H., 2001. Applications of AMT (Angle Measure Technique) in image analysis: Part II: Prediction of powder functional properties and mixing components using Multivariate AMT Regression (MAR). *Chemometrics and Intelligent Laboratory Systems*, 57(1), pp.37–56.

Huang, J. & Esbensen, K.H., 2000. Applications of Angle Measure Technique (AMT) in image analysis: Part I. A new methodology for in situ powder characterization. *Chemometrics and Intelligent Laboratory Systems*, 54(1), pp.1–19.

Istvan, D., 2012. DigiCamControl. Available at: http://digicamcontrol.com/.

Jenné, R. et al., 2006. Use of image analysis for sludge characterisation: studying the relation between floc shape and sludge settleability. *Water Science & Technology*, 54(1), pp.167–174.

Jiang, J. & Graham, N.J.D., 1998. Pre-polymerised inorganic coagulants and phosphorus removal by coagulation - A review. , 24(3), pp.237–244.

Jin, Y., 2005. *Use of a high resolution photographic technique for studying coagulation/flocculation in water treatment*. University of Saskatchewan.

Jones, R.D. & Hood, M.A., 1980. Effects of temperature, ph, salinity, and inorganic nitrogen on the rate of ammonium oxidation by nitrifiers isolated from wetland environments. *Microbial Ecology*, 6(4), pp.339–347.

Juntunen, P. et al., 2014. Characterization of alum floc in water treatment by image analysis and modeling. *Cogent Engineering*, 1, pp.1–13.

Juntunen, P. et al., 2013. Dynamic soft sensors for detecting factors affecting turbidity in drinking water. *Journal of Hydroinformatics*, 15(2), pp.416–426.

Juntunen, P. et al., 2012. Modelling of water quality: an application to a water treatment process. *Applied Computational Intelligence and Soft Computing*, 2012, pp.1–9.

Kan, C., Huang, C. & Pan, J.R., 2002. Time requirement for rapid-mixing in coagulation. *Colloids and Surfaces A: Physicochemical and Engineering Aspects*, 203(1–3), pp.1–9.

Kang, L.-S. & Cleasby, J.L., 1995. Temperature effects on flocculation kinetics using Fe(III) coagulant. *Journal of Environmental Engineering*, 121(12), pp.893–901.

Keyvani, A. & Strom, K., 2013. A fully-automated image processing technique to improve measurement of suspended particles and flocs by removing out-of-focus objects. *Computers & Geosciences*, 52, pp.189–198.

Khan, F.S. et al., 2015. Compact color-texture description for texture classification. *Pattern Recognition Letters*, 51, pp.16–22.

Kim, S.-H., Moon, B.-H. & Lee, H.-I., 2001. Effects of pH and dosage on pollutant removal and floc structure during coagulation. *Microchemical Journal*, 68(2–3), pp.197–203.

Kozminykh, P., Heistad, A. & Ratnaweera, H.C., 2016. Impact of organic polyelectrolytes on coagulation of source-separated black water. *Environmental Technology*, 37(14), pp.0–10.

Kramer, T.A. & Clark, M.M., 1997. Influence of strain-rate on coagulation kinetics. *Journal of Environmental Engineering*, 123(5), pp.444–452.

Kucheryavski, S., 2007. Using hard and soft models for classification of medical images. *Chemometrics and Intelligent Laboratory Systems*, 88(1), pp.100–106.

Kvaal, K. et al., 1998. Multivariate feature extraction from textural images of bread. *Chemometrics and Intelligent Laboratory Systems*, 42(1–2), pp.141–158.

Levine, M., 1985. *Vision in Man and Machine*,

Li, B. & Stenstrom, M.K., 2014. Research advances and challenges in one-dimensional modeling of secondary settling tanks – A critical review. *Water Research*, 65, pp.40–63.

Li, D.H. & Ganczarczyk, J., 1989. Fractal geometry of particle aggregates generated in water and wastewater treatment processes. *Environmental Science & Technology*, 23(11), pp.1385–1389.

Lin, J.L. et al., 2008. Coagulation dynamics of fractal flocs induced by enmeshment and electrostatic patch mechanisms. *Water Research*, 42(17), pp.4457–4466.

Liu, D., Li, F. & Zhang, B., 2009. Removal of algal blooms in freshwater using magnetic polymer. *Water Science & Technology*, 59(6), pp.1085–1091.

Liu, W., 2016. *Enhancement of coagulant dosing control in water and wastewater treatment process*. (Doctoral dissertation). Norwegian University of Life Sciences.

Liu, W. & Ratnaweera, H., 2017. Feed-forward-based software sensor for outlet turbidity of coagulation process considering plug flow condition. *International Journal of Environmental Science and Technology*, 14(8), pp.1689–1696.

Liu, W., Ratnaweera, H. & Song, H., 2013. Better treatment efficiencies and process economies with real-time coagulant dosing control. In *11th IWA Conference on Instrumentation Control and Automation*. Narbonne, France, 18–20 September 2013.

Liukkonen, M., Juntunen, P., et al., 2013. A software platform for process monitoring: Applications to water treatment. *Expert Systems with Applications*, 40(7), pp.2631–2639.

Liukkonen, M., Laakso, I. & Hiltunen, Y., 2013. Advanced monitoring platform for industrial wastewater treatment: Multivariable approach using the self-organizing map. *Environmental Modelling and Software*, 48, pp.193–201.

Manamperuma, L., 2016. *Optimisation of the coagulation process to improve plant availability of phosphorus in wastewater sludge*. PhD. Norwegian University of Life Sciences, Ås.

Manamperuma, L., Ratnaweera, H.C. & Rathnaweera, S.S., 2013. Retrofitting coagulant dosing control using real-time water quality measurements to reduce coagulant consumption. In *11th IWA Conference on Instrumentation Control and Automation*. Narbonne, France, 18–20 September 2013.

Manamperuma, L., Wei, L. & Ratnaweera, H., 2017. Multi-parameter based coagulant dosing control. *Water Science and Technology*, 75(9), pp.2157–2162.

Manamperuma, L.D., Ratnaweera, H.C. & Martsul, A., 2016. Mechanisms during suspended solids and phosphate concentration variations in wastewater coagulation process. *Environmental Technology (United Kingdom)*, 37(19), pp.2405–2413.

Mandelbrot, B., 1967. How long is the coast of Britain? Statistical self-similarity and fractional dimension. *Science*, 156(3775), pp.636–638.

Mandelbrot, B., 1982. *The fractal geometry of nature*, W. H. Freeman and Company.

Matilainen, A., Vepsäläinen, M. & Sillanpää, M., 2010. Natural organic matter removal by coagulation during drinking water treatment: a review. *Advances in colloid and interface science*, 159(2), pp.189–97.

McConnachie, G.L., 1991. Turbulence intensity of mixing in relation to flocculation. *Journal of Environmental Engineering*, 117(6), pp.731–750.

Mendret, J. et al., 2007. An optical method for in situ characterization of fouling during filtration. *AIChE Journal*, 53(9), pp.2265–2274.

Menniti, A. & Schauer, P., 2017. One utility's approach to evaluating new instrumentation and ongoing maintenance and validation. In *12th IWA Specialized Conference on Instrumentation, Control and Automation*. Quebec, Canada, 11-14 June 2017, pp. 391–400.

Mesquita, D.P., Amaral, A.L. & Ferreira, E.C., 2011. Characterization of activated sludge abnormalities by image analysis and chemometric techniques. *Analytica Chimica Acta*, 705(1–2), pp.235–42.

Metcalf, E. & Eddy, H., 2013. *Wastewater engineering: treatment and resource recovery* 5th ed. G. Tchobanoglous et al., eds., McGraw-Hill Education.

Mortensen, P.P. & Esbensen, K.H., 2005. Optimization of the Angle Measure Technique for image analytical sampling of particulate matter. *Chemometrics and Intelligent Laboratory Systems*, 75(2), pp.219–229.

da Motta, M. et al., 2001. Characterisation of activated sludge by automated image analysis. *Biochemical Engineering Journal*, 9(3), pp.165–173.

Ødegaard, H., 1985. Engineering aspects of flocculation. In G. F. Verlag, ed. *Chemical water and wastewater treatment*. pp. 81–102.

Ødegaard, H. et al., 2010. NOM removal technologies – Norwegian experiences. *Drinking Water Engineering and Science*, 3, pp.1–9.

Ødegaard, H., Fettig, J. & Ratnaweera, H., 1990. Coagulation with prepolymerized metal salts H. H. Hahn & R. Klute, eds. *Chemical water and wastewater treatment. Proceedings of the 4th Gothenburg Symposium 1990 October 1–3, 1990 Madrid, Spain*, 304, pp.189–220.

Otsu, N., 1979. A threshold selection method from gray-level histograms. *IEEE Trans. Systems Man Cybernet.*, 9(1), pp.62–66.

Palm, C., 2004. Color texture classification by integrative Co-occurrence matrices. *Pattern Recognition*, 37(5), pp.965–976.

Parker, D.S., Kaufman, W.J. & Jenkins, D., 1970. *Characteristics of biological flocs in turbulent regimes*, Sanitary Engineering Research Laboratory, University of California.

Pedocchi, F. & Piedra-Cueva, I., 2005. Camp and Stein's velocity gradient formalization. *Journal of Environmental Engineering*, 131(10), pp.1369–1376.

Princic, A. et al., 1998. Effects of pH and Oxygen and Ammonium concentrations on the community structure of nitrifying bacteria from wastewater. *Applied and Environmental Microbiology*, 64(10), pp.3584–3590.

Rasband, W.S., ImageJ. *U. S. National Institutes of Health, Bethesda, Maryland, USA, 1997-2016*. Available at: http://imagej.nih.gov/ij/.

Rathnaweera, S.S., 2010. *Modelling and optimisation of wastewater coagulation process*. PhD. Norwegian University of Life Sciences, Ås.

Ratnaweera, H., 1997. *Chemical wastewater treatment: a concept for optimal dosing of coagulants*, Norwegian Institute for Water Research, Report SNO 3713-97.

Ratnaweera, H., 2004. Coagulant dosing control - a review. *Chemical water and wastewater treatment, IWA publishing, London*, 7, pp.3–13.

Ratnaweera, H., 1991. *Influence of the degree of coagulant prepolymerization on wastewater coagulation mechanisms*. (Doctoral dissertation). Norwegian Institute of Technology, Trondheim.

Ratnaweera, H. & Fettig, J., 2015. State of the art of online monitoring and control of the coagulation process. *Water*, 7(11), pp.6574–6597.

Ratnaweera, H., Hiller, N. & Bunse, U., 1999. Comparison of the coagulation behavior of different Norwegian aquatic NOM sources. *Environment International*, 25(2), pp.347–355.

Ratnaweera, H., Ødegaard, H. & Fettig, J., 1992. Coagulation with prepolymerized aluminium salts and their influence on particle and phosphate removal. *Wat. Sci. Tech.*, 26(5–6), pp.1229–1237.

Ratnaweera, H.C., 2013. Phosphorus recovery from wastewater — should we rebuild our wastewater treatment plants? *VANN, no. 04*, pp.551–556.

Russ, J.C., 2011. *The image processing handbook* 6th ed., CRC Press.

Sharp, E. et al., 2005. Application of zeta potential measurements for coagulation control: pilot-plant experiences from UK and US waters with elevated organics. *Water Science and Technology: Water Supply*, 5(5), pp.49–56.

Sharp, E. et al., 2016. Using online zeta potential measurements for full-scale coagulation control in drinking water treatment. In *IWA Specialist Conference, Advances in particle science and separation: meeting tomorrow's challenges*. Oslo, Norway, 22-24 June 2016.

Shiranita, K., Miyajima, T. & Takiyama, R., 1998. Determination of meat quality by texture analysis. *Pattern Recognition Letters*, 19, pp.1319–1324.

Smith, V.H., Tilman, G.D. & Nekola, J.C., 1999. Eutrophication: Impacts of excess nutrient inputs on freshwater, marine, and terrestrial ecosystems. *Environmental Pollution*, 100, pp.179–196.

Stumm, W. & O'Melia, C., 1968. Stoichiometry of coagulation. *American Water Works*

Association, 60(5), pp.514–539.

Su, X., Xu, S. & Xu, S., 2017. Compound control system for coagulant dosing process based on a fuzzy cerebellar model articulation controller. In *Proceedings of the 36th Chinese Control Conference*. Dalian, China, 26-28 July 2017, pp. 3931–3936.

Tambo, N. & Watanabe, Y., 1979a. Physical aspect of flocculation process—I: Fundamental treatise. *Water Research*, 13(5), pp.429–439.

Tambo, N. & Watanabe, Y., 1979b. Physical characteristics of flocs—I. The floc density function and aluminium floc. *Water Research*, 13(5), pp.409–419.

Thomas, D.N., Judd, S.J. & Fawcett, N., 1999. Flocculation modelling: a review. *Water Research*, 33(7), pp.1579–1592.

VA-Support, 2012. *DOSCON – Experience at NRA. (03.07.2012)*

Vahedi, A. & Gorczyca, B., 2011. Application of fractal dimensions to study the structure of flocs formed in lime softening process. *Water research*, 45(2), pp.545–56.

Vahedi, A. & Gorczyca, B., 2012. Predicting the settling velocity of flocs formed in water treatment using multiple fractal dimensions. *Water research*, 46(13), pp.4188–4194.

Vahedi, A. & Gorczyca, B., 2014. Settling velocities of multifractal flocs formed in chemical coagulation process. *Water Research*, 53, pp.322–328.

Valentin, N. & Denœux, T., 2001. A neural network-based software sensor for coagulation control in a water treatment plant. *Intelligent Data Analysis*, 5, pp.23–39.

Wang, J.-P. et al., 2011. Optimization of the coagulation-flocculation process for pulp mill wastewater treatment using a combination of uniform design and response surface methodology. *Water research*, 45(17), pp.5633–5640.

Wang, K., 2007. Applying data mining to manufacturing: the nature and implications. *Journal of Intelligent Manufacturing*, 18(4), pp.487–495.

Wang, X., Kvaal, K. & Ratnaweera, H., 2017. Characterization of influent wastewater with periodic variation and snow melting effect in cold climate area. *Computers and Chemical Engineering*, 106, pp.202–211.

Wang, Y.L. et al., 2011. Effect of polyferric chloride (PFC) doses and pH on the fractal characteristics of PFC–HA flocs. *Colloids and Surfaces A: Physicochemical and Engineering Aspects*, 379(1–3), pp.51–61.

Wilfert, P. et al., 2016. Vivianite as an important iron phosphate precipitate in sewage treatment plants. *Water Research*, 104, pp.449–460.

Xiao, F. et al., 2009. Effects of low temperature on coagulation kinetics and floc surface morphology using alum. *Desalination*, 237(1–3), pp.201–213.

Xiao, F. et al., 2011. PIV characterisation of flocculation dynamics and floc structure in water treatment. *Colloids and Surfaces A: Physicochemical and Engineering Aspects*, 379(1–3), pp.27–35.

Yan, M. et al., 2008. Enhanced coagulation for high alkalinity and micro-polluted water: the third way through coagulant optimization. *Water research*, 42(8–9), pp.2278–2286.

Yavich, A.A. & Van De Wege, J., 2013. Chemical feed control using coagulation computer models and a streaming current detector. *Water Science & Technology*, 67(12), p.2814.

Yu, R.-F. et al., 2009. Simultaneously monitoring the particle size distribution, morphology and suspended solids concentration in wastewater applying digital image analysis (DIA). *Environmental monitoring and assessment*, 148(1–4), pp.19–26.

Yu, R.-F., Cheng, W.-P. & Chu, M.-L., 2005. On-line monitoring of wastewater true color using digital image analysis and artificial neural network. *Journal of Environmental Engineering*, 131(1), pp.71–79.

Yu, R.-F., Cheng, W.-P. & Huang, H.-D., 2012. On-line assessment of the particle separation in chemical flocculent suspension by image analysis. In Martin Jekel, ed. *International Conference on Particle Separation*. Berlin, Germany, 18-20th June 2012: International Water Association, pp. 53–57.

Yu, R., Chen, H. & Cheng, W., 2017. Applying online image analysis to simultaneously evaluate the removals of suspended solids and color from textile wastewater in chemical flocculated sedimentation. *Journal of Environmental Informatics*, 29(1), pp.29–38.

Yu, R.F., 2014. On-line evaluating the SS removals for chemical coagulation using digital image analysis and artificial neural networks. *International Journal of Environmental Science and Technology*, 11(7), pp.1817–1826.

Zheng, C., Sun, D.-W. & Zheng, L., 2006. Recent applications of image texture for evaluation of food qualities—a review. *Trends in Food Science & Technology*, 17(3), pp.113–128.

Zheng, H. et al., 2011. Investigations of coagulation–flocculation process by performance optimization, model prediction and fractal structure of flocs. *Desalination*, 269(1–3), pp.148–156.

Zouboulis, A.I., Tzoupanos, N.D. & Moussas, P.A., 2007. Inorganic pre-polymerized coagulants : current status and future trends. , (c), pp.292–300.

Appendix – List of Publications

Paper I

Sivchenko, N., Kvaal, K., & Ratnaweera, H. (2016). Evaluation of image texture recognition techniques in application to wastewater coagulation. *Cogent Engineering*, *3*(1), 1206679. http://doi.org/10.1080/23311916.2016.1206679

Paper II

Sivchenko, N., Shostak, V., Kalashnikov, Y., Ratnaweera, H, & Kvaal, K. Relationships between floc texture features and model wastewater qualities. Manuscript

Paper III

Sivchenko, N., Ratnaweera, H., & Kvaal, K. (2017). Approbation of the texture analysis imaging technique in the wastewater treatment plant. *Cogent Engineering*, *4*(1), 1373416. http://doi.org/10.1080/23311916.2017.1373416

Paper IV

Sivchenko, N., Kvaal, K., & Ratnaweera, H. Floc sensor prototype tested in the municipal wastewater treatment plant. In press, *Cogent Engineering*.

Additional scientific work performed during the PhD program:

Sivchenko, N., Kvaal, K., & Ratnaweera, H. (2013). Characterization of flocs in coagulation-flocculation process by image analysis and mathematical modelling. In *Poster, NORDIWA 13th Nordic Wastewater Conference*. Malmo, Sweden, 8-10 October 2013.

Sivchenko, N., Kvaal, K., & Ratnaweera, H. (2014). Image analysis of flocs and mathematical modelling applied to coagulation-flocculation process. In *Poster, IWA Specialist Conference, Advances in particle science and separation: from mm to nm scale and beyond.* Sapporo, Japan, 15-18 June 2014.

Paper I

Sivchenko, N., Kvaal, K., & Ratnaweera, H. (2016). Evaluation of image texture recognition techniques in application to wastewater coagulation. *Cogent Engineering*, *3*(1), 1206679. http://doi.org/10.1080/23311916.2016.1206679
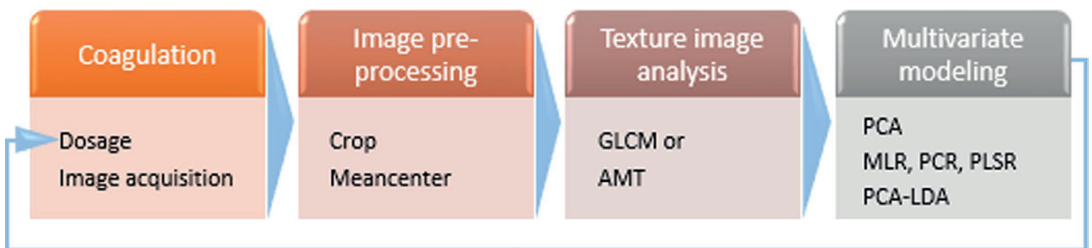
CIVIL & ENVIRONMENTAL ENGINEERING | RESEARCH ARTICLE

Evaluation of image texture recognition techniques in application to wastewater coagulation

Nataliia Sivchenko, Knut Kvaal and Harsha Ratnaweera

� **cogent**
engineering

CrossMark

*Corresponding author: Nataliia
Sivchenko, Department of Mathematical
Sciences and Technology, Norwegian
University of Life Sciences, P.O. Box
5003-IMT, Aas 1432, Norway
E-mails: nataliia.sivchenko@gmail.com,
nataliia.sivchenko@nmbu.no

Reviewing editor:
Yiu Fai Tsang, The Hong Kong Institute
of Education, Hong Kong

Additional information is available at
the end of the article

CIVIL & ENVIRONMENTAL ENGINEERING | RESEARCH ARTICLE

Evaluation of image texture recognition techniques in application to wastewater coagulation

Nataliia Sivchenko[1]*, Knut Kvaal[1] and Harsha Ratnaweera[1]

Abstract: Flocs formation and growth are important characteristics in wastewater coagulation process. The shape and size of flocs highly affect further separation processes, therefore resulting treatment efficiency of wastewater after coagulation. Observed images of flocs tend to show strong relations to coagulation parameters: dose and coagulation time. In this article, three texture recognition techniques were evaluated for the ability to mathematically describe the relationship between the images of flocs and coagulant dosages. The easily computable texture analysis methods were found to be potential techniques for the characterization of the particles images. Ten out of eleven co-occurrence matrix-based grey level co-occurrence matrix (GLCM) texture features were found to be significant for the dosage prediction by a principal component regression model with only one principal component. Two features (Inverse difference moment and Variance) were selected for the multiple linear regression model. Test set prediction accuracy varied from 83 to 96% depending on texture analysis method and multivariate model. Best dosage prediction and image classification results were achieved by GLCM and angle measure technique. The results of image texture analysis coupled with multivariate modelling techniques indicate that it is possible to characterize and relate flocs images, captured during coagulation, with different coagulant dosages, as well as predict those dosages.

ABOUT THE AUTHORS

The authors are members of Water, Environment, Sanitation and Health (WESH) research group at the Norwegian University of Life Sciences (NMBU), Department of Mathematical Sciences and Technology (IMT). WESH group focuses on water and wastewater related issues and is heavily involved in teaching and supervision of MSc and PhD programmes in water engineering and technology. The group also has one of the largest externally funded research and educational project portfolios at IMT, and collaborates with partners from EU, North America, Eurasia, Asia and Africa. It also has a number of Research, Development and Innovation projects with Norwegian partners. The main research and development areas of the WESH group are: process control and optimization of coagulation and biological treatment processes; membrane fouling and filtration processes; microbial water quality and risk assessment; decentralized wastewater systems; modelling of sewer systems.

PUBLIC INTEREST STATEMENT

The performance of coagulation process is crucial in wastewater treatment. By observing changes in flocs' aggregation, it is potentially possible to predict treatment efficiency and optimize the coagulation process – on-line dosage control. In this article, we present the new approach in flocs images description. It is based on the characterization of the whole image (texture image analysis) instead of characterizing each floc separately. Texture image analysis methods applied to images of flocs and coupled with multivariate modelling seem to be a potential tool for on-line prediction of the optimal coagulant dosage and resulting treatment efficiency.

�֍ **cogent** ⚬ **oa**

cogent ·· engineering

1. Introduction

Coagulation is a well known and essential water purification process for drinking water, wastewater, and industrial water treatment. Added to the colloidal suspension, organic or inorganic coagulant destabilizes system which leads to particles' aggregation and subsequent separation. Intelligent dosage control systems are required to optimize and control the removal of particles and phosphates, minimize the use of a coagulant, achieve better sludge management (by reduction of the sludge volumes) and enhance the plant availability of phosphates. Nowadays optimization and process control of coagulation are mainly done by simple flow-proportional dosing concept or based on data from jar tests (Ratnaweera & Fettig, 2015). Nevertheless, with the recent development of a big range of on-line sensors the different coagulation control systems appear (Annadurai, Sung, & Lee, 2004; Juntunen, Liukkonen, Lehtola, & Hiltunen, 2013; Ratnaweera, Lei, & Lindholm, 2002). However, the initial costs of such control systems are usually quite high. Furthermore, the control systems based on outlet water parameters have a perceptible time lag between dosing point and effluent, varying from 30 min to several hours depending on the separation method. This peculiarity makes such systems inappropriate for wastewater coagulation dosing control, where influent water parameters and its flow rate changes dramatically within a short time. Hence, water and wastewater treatment industry are seeking for simple, cheap, robust and precise real-time dosing control systems.

It has been relatively long since researchers started to model the mechanisms of flocculation (Vold, 1963). The developments in this area are summarized in the review by Thomas, Judd, and Fawcett (1999). Visualization, modelling, and understanding of flocs formation mechanisms reveals the ability to determine relationships between particles' properties and coagulation efficiency thus optimize the coagulation process. Many researchers have made their contribution to the understanding of the particles aggregation, breakage and regrowth mechanisms. Jarvis, Jefferson, Gregory, and Parsons (2005) summarized the development of floc strength and breakage mechanisms in their review. Influence of coagulation mixing conditions (e.g. share rates, impeller construction) into flocs formation, breakage and regrowth were also studied (Cao et al., 2011; He, Nan, Li, & S., 2012). However, due to complicated nature of flocs and the huge number of parameters influencing the coagulation process there is still no comprehensive or universally accepted mathematical description of floc formation mechanisms.

The modelling of drinking water treatment process, which included coagulation stage, had been recently performed by Juntunen, Liukkonen, Pelo, Lehtola, and Hiltunen (2012). The authors investigated both laboratory and process data. They found aluminium dose to be one of the most important factors affecting the treated water quality parameters (turbidity and residual aluminium).

Recently, computer vision image processing techniques started to be employed more frequently in order to study and characterize complex size, shape and features of particles. Photographic image acquisition techniques make it possible to capture and analyse flocs' images *in situ* (Chakraborti, Gardner, Atkinson, & Van Benschoten, 2003). Our hypothesis is that it should be possible to correlate certain floc properties with coagulant dosages and treatment efficiencies using images of flocs. The distinct effect of coagulant dosage on aggregates' features was detected by Lin, Huang, Chin, and Pan (2008) in kaolin water suspension. Authors used wet scanning electron microscopy in order to obtain morphology of flocs. Jin (2005) used the high-resolution digital camera to study the influence of temperature on flocs properties under different coagulant dosages for river water coagulation. The relation between projected area of particles and coagulant dosages were found. Wang et al. (2011) investigated the changes in flocs' characteristics due to different dosages and coagulation pH

cogent · engineering

in humic acid suspension, obtaining images by digital CCD camera. Our approach in this article is to prove that images of model wastewater flocs captured *in situ* by the nonintrusive photographic method are descriptive in terms of coagulant dosage prediction. Thus, the technique could potentially be used for on-line coagulation dosage control. Advanced dosage control then will be possible in means of determining lowest coagulant dosage which leads to required solid, liquid and phosphorous separation.

Image analysis of particles is a fairly old (Tambo & Watanabe, 1979b) and known technique to be used in different water treatment applications, for instance: activated sludge processing (Alves et al., 2000; Amaral & Ferreira, 2005; da Motta, Pons, Roche, & Vivier, 2001; Dagot et al., 2001; Gorczyca & Ganczarczyk, 2002; Jenné, Banadda, Smets, Deurinck, & Van Impe, 2007; Jenné et al., 2006; Mesquita, Amaral, & Ferreira, 2011), aggregates settling velocity measurements (Gorczyca & Ganczarczyk, 1996; Vahedi & Gorczyca, 2012, 2014), membrane fouling (Mendret, Guigui, Schmitz, Cabassud, & Duru, 2007), natural organic matter removal (Xiao, Lam, Li, Zhong, & Zhang, 2011). Image analysis applied to coagulation and aggregation of particles was also widely studied in different waters: clay (kaolinite) suspension (He et al., 2012; Lin et al., 2008; Tambo & Hozumi, 1979; Tambo & Watanabe, 1979a; Xiao et al., 2011), mineral suspensions (Gorczyca & Ganczarczyk, 1996), latex particles suspension (Chakraborti et al., 2003), humic acid solution (Wang et al., 2011), water after lime softening (Vahedi & Gorczyca, 2011, 2012, 2014), lake water (Chakraborti, Atkinson, & Van Benschoten, 2000; Kim, Moon, & Lee, 2001), textile wastewater (Yu, Chen, Cheng, & Chu, 2009; Yu, Cheng, & Huang, 2012), sewer system wastewater (Zheng et al., 2011). However, not many studies concerned the floc-dosage relation in the coagulation process regarding wastewater or model wastewater.

Li and Ganczarczyk (1989) described the spatial structure of particles appearing during water and wastewater treatment processes by fractal theory, i.e. fractal dimension. Since then, it has become the most common way of flocs characterization. In this research, we were not much concerned with the study of the particles' characteristics itself, instead, we were testing the different image analysis techniques to correlate images of flocs with coagulant dosages. The technique presented in this article analyses the whole captured image, not its particular parts or regions of interest, and is known as image texture analysis. The advantage of such method is that the flocs do not have to be segmented from the image, which is quite difficult and challenging in some cases, e.g. in a high turbidity waters and with a high concentration of particles when they are overlapping in the image.

Texture analysis is a well-known method of image recognition and especially classification in numerous computer vision applications. However, as far as we know, it has not been previously reported to be used in application to images of coagulated particles. Texture is a loosely defined term without an accepted or universal quantitative meaning. Texture can be defined as a measure of the image's surface roughness in a meaning of brightness, colour, shape and size variations within some region and its repetitiveness. Texture is a pattern which can be completely distinct or completely random. Furthermore, texture could be isotropic (without any preferred orientation) or anisotropic (has definite pattern structure). Images of flocs presented in this study were described by texture recognition techniques.

There have been many attempts in combining the digital image analysis with coagulation/flocculation process. The clear correlations between floc properties and the flocculation process parameters were found by Juntunen, Liukkonen, Lehtola, and Hiltunen (2014). The authors documented that, for instance, the surface area of floc and the number of floc particles highly correlate with the process data, such as lime feed, pH, turbidity etc. However, in contrast to our research, the authors performed the analysis for drinking water treatment plant with the characterization of flocs by object recognition methods in image analysis. In this paper, we evaluate the texture image analysis methods to be used for coagulant dose prediction by the images of flocs. It seems to be a promising tool for further coagulation–flocculation dosage control studies. Potentially, the texture analysis methods in application to particles separation processes may simplify the image characterization part.

cogent ·· engineering

2. Materials and methods

2.1. Raw water

Model wastewater was used to perform all the experiments in this research. In order to have a better representation of the typical domestic wastewater characteristics, the synthetic wastewater was prepared according to previous studies (Fettig, Ratnaweera, & Ødegaard, 1990; Ødegaard, Fettig, & Ratnaweera, 1990; Ratnaweera, 1991). The model wastewater contains both the organic and inorganic components. Dried milk, potato starch, and humic acid represent the organic part of the suspension. Different inorganic salts are used to well define amounts of orthophosphate and ammonium. Bentonite is used to imitate the inorganic particles. For this particular research, model wastewater was prepared with a medium concentration level of salts and particles and represents typical soft Norwegian wastewater. Components and their concentrations were as follows: dried milk (Nestle, Norway)—300 mg/l, potato starch (Hoff, Norway)—60 mg/l, bentonite (Alfa Aesar, USA)—80 mg/l, NaCl (Merck, Germany)—400 mg/l, K_2HPO_4 (Merk, Germany)—50 mg/l, NH_4Cl (Kebo, Germany)—100 mg/l, Humic acid sodium salt (Merck, Germany)—5 mg/l, $NaHCO_3$ (Merck, Germany)—60 mg/l. Ås city tap water was used as a solvent, which is typical soft Norwegian lake water with low concentrations of most components. The model wastewater was prepared for each experimental series and used within 1.5–4 h after preparation. Average model wastewater parameters were next: 8.03 ± 0.13 pH, 187 ± 10.7 mg/l suspended solids, 256 ± 12 FNU turbidity, 11.2 ± 0.24 mg/l total Phosphorus, 9.5 ± 0.07 mg/l Orthophosphates. The initial model wastewater particles had size distribution 0.52–150 μm with two pics at 1,88 and 14.5 μm, determined by Malvern Mastersizer 3000 (Malvern Instruments Ltd, UK).

2.2. Coagulation procedures

All investigations were performed in jar-test scale, Flocculator 2000 (Kemira, Sweden) with programmable mixer units and 1 l beakers. The mixing conditions during coagulation were: 1 min rapid mixing (400 revolutions per min—RPM), 10 min slow mixing (30 RPM) and 20 min sedimentation without mixing. Prepolymerized aluminium chloride coagulant PAX XL61 (Kemira, Sweden) was used in this research. All tests were performed with wastewater temperature 16–17 C and at a constant pH of 7.5 adjusted by required acid or base addition. Coagulant doses varied from 0.29 to 1.08 mmol Al/l, 11 dosages in total. After sedimentation stage, approximately 200 ml of treated water samples were taken from 5 cm depth by a peristaltic pump.

Each sample of treated water taken after sedimentation step was analysed for turbidity, total suspended solids, orthophosphates and total phosphorous concentrations.

2.3. Image acquisition and pre-processing

A nonintrusive optical sampling technique was used to capture images of appearing aggregates. The principal scheme is shown in Figure 1. The images of flocs were obtained during the whole slow mixing period of coagulation with repeatability 1 image per 20 seconds provided by free remote camera control software—DigiCamControl 1.2.0. Image capturing equipment used during the investigations was as follows: digital single lens reflex (DSLR) Nikon D600 camera, 105 mm Nikkor AF-S Micro 1:2.8 G ED lens (Nikkor, China), SpeedLite YN460 flash (Yongnuo, China). The size of the image-capturing zone in the beaker was 4.8 × 3.3 cm. In order to obtain flocs with the proper depth of field, the mixer unit was modified with an attachment of 4.5 cm width black curved plastic stripe, which also become a background for the images of flocs. The choice of the background colour was based on a fact that the particles, present in model wastewater, are white in colour. Thus, using a contrasting background it is easier to perform the further image analysis.

Images obtained during coagulation have a resolution of 24.3 megapixels each. They were processed in the open source image analysis software *ImageJ v.1.49* (Rasband, 1997/2016) that bases on plugins and macros. For each image 3,690 × 3,690 pixels (3 × 3 cm) area was cropped by manual investigation of the area. Because of slight changes in lighting conditions during image acquisition

Figure 1. Schematic representation of image acquisition.
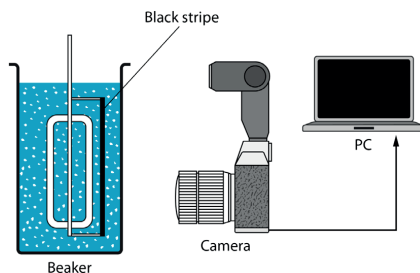
for different dosages, all images had been pre-processed in order to have the same brightness intensity. Knut Kvaal wrote *ImageJ* plug-in to mean centre the images' grey-tone values.

In total, coagulation jar tests were performed using 11 coagulant dosages, 10 images from each experiment after 6 min of coagulation were selected as representative samples, resulting in 110 samples. Coagulant dose was set as the response Y parameter in multivariate models.

2.4. Fractal dimension of aggregates

Fractal dimension is a very common and widely used parameter to characterize the geometrical features of fractal particles, such as flocs. For a 2-dimensional (2D) projected particle image, the fractal dimension, D_{pf} defines how the projected area of the particles rises with the perimeter (He et al., 2012):

$$A \propto P^{2/D_{pf}} \tag{1}$$

where A—projected area; P—perimeter of the particles. For the 2D projection of an image, the value of fractal dimension varies from $D_{pf} = 1$ for the circle shape floc to $D_{pf} = 2$ for a chain of particles (a line).

The size of irregular shape particle can be determined as the equivalent diameter, d_p (He et al., 2012):

$$d_p = (4A/\pi)^{1/2}. \tag{2}$$

Flocs' morphological characteristics were obtained using *ImageJ* and plug-in "Analyse particles". First, all images collected during the slow mixing period of coagulation were gathered in a stack. Then 3,690 × 3,690 pixels area was manually estimated and cropped. A stack of images was converted to 8-bit greyscale. Plug-in "Subtract Background" was applied on images in order to equalize background pixel values for easier thresholding. If needed, brightness and contrast were adjusted. Onwards, thresholding was used to obtain binary images, with a threshold value estimated manually or by Otsu method (Otsu, 1979). Flocs within this transformation had greyscale value 0—true black, while the background was set to value 255—true white. Finally, plug-in "Analyse particles" was applied to the binary images of flocs. Many geometrical and statistical parameters might be subtracted from the images of aggregates by this plug-in. For this research, significant parameters of flocs were: number of particles in the image, mean area, and perimeter of flocs. Knowing that 1,230 pixels in the image equal 1 cm, we were able to recalculate flocs' features from pixel values to quantitative values in centimetres.

Mean fractal dimension of flocs was calculated by equation:

$$D_f = 2 \times \log\left(P_{pix}\right)/\log(A_{pix}) \tag{3}$$

where A_{pix}—mean area of particles, in pixels; P_{pix}—perimeter, a count of pixel edges; 2 is a constant number (Yu et al., 2009).

2.5. Image analysis by histogram

The histogram is a frequency distribution of the intensity values that occur in an image. The typical 8-bit greyscale image has 256 intensity values (grey shades). Histogram of a greyscale 8-bit image is a one-dimensional (1D) vector of length 256, which contain information about the number of pixels of particular intensity in the image. However, histograms do not contain the spatial information about pixels in the image. The histogram is a fast and simple way to illustrate statistical information of the image—distribution of intensity values, thus, it is a popular tool for real-time image processing. Histogram analysis had also been illustrated as a tool for elimination of poor quality images (Liukkonen, Hiltunen, & Hiltunen, 2015). The information from histograms could be enough for images comparison.

In order to decrease the number of variables (intensity values) the 3 sizes of histogram bins were tried in the further multivariate analysis: 256, 128 and 64 bins. It means that initial histogram was scaled 2 and 4 times using the plug-in in *ImageJ* created by Knut Kvaal "Compute hist stack bins". The resulting outputs were given as vectors of 256, 128 and 64 intensity values for each image in the stack. Hence, the resulting matrixes were: 110×256, 110×128 and 110×64.

2.6. Image analysis by GLCM

GLCM is a typical statistical method of measuring texture. Haralick, Shanmugam, and Dinstein (1973) invented this method, which bases on spatial-dependence GLCM of pixels with estimation of image features using second-order statistics.

ImageJ has a plug-in "GLCM Texture" v.0.4, created by Julio E. Cabrera and further updated to "GLCM Texture Too" v. 0.008 by Toby C. Cornish. Knut Kvaal has done some in-house revision of the plug-in and called the version 0.009, which was used in this analysis. GLCM algorithm was applied considering the distance between the pixel pairs: 1, 2, 3, 5, 10 pixels and the angles were 0°, 45°, 90° and 135°. The resulting output was given as a vector of the next 11 parameters per each image in the stack: angular second moment (ASM), contrast, correlation, inverse difference moment (IDM), entropy, energy, inertia, homogeneity, prominence, variance, and shade. Hence, 20 matrixes were obtained with the size 110×11 each. The detailed description, explanation, and equations for above GLCM texture features are given elsewhere (Conners, Trivedi, & Harlow, 1984; Haralick et al., 1973; Zheng, Sun, & Zheng, 2006).

2.7. Image analysis by AMT

Andrle (1994) introduced angle measure technique (AMT) as a method to characterize the complexity of geomorphic lines. The purpose was to detect changes in coastlines complexity as a function of scale, so firstly it was an approach to analyse 1D data. Later Esbensen, Hjelmen, and Kvaal (1996) have inducted AMT into chemometrics for textural analysis of generic "measurement series" and have shown that this technique is also applicable for the images (2D data). Further successful implementation of AMT (Dahl & Esbensen, 2007; Fongaro & Kvaal, 2013; Fongaro, Lin Ho, Kvaal, Mayer, & Rondinella, 2016; Huang & Esbensen, 2000, 2001; Kucheryavski, 2007; Kvaal, Wold, Indahl, Baardseth, & Næs, 1998) proved that the method could be used for images' texture description. AMT spectrums can also be used for interpretative purposes. The AMT is implemented as an in-house made *ImageJ* plug-in (Kvaal, 2014).

AMT is a transform-based texture recognition method, which in effect creates a completely new domain – scale domain. The AMT algorithm, regarding image analysis, can be described by five steps.

(1) Unfold the image (Kvaal et al., 2008; Mortensen & Esbensen, 2005) into a 1D vector of grey level values, also called generic "measurement series". Row by row or spiral unfold are most commonly used.

Table 1. Results of treated wastewater after sedimentation										
Class	Dose (mmol/l)	pH	TSS (mg/l)	Turbidity (FNU)	Ortho-P (mg/l)	Total P (mg/l)	Repl.[1]	Number of images		
								Total	Cal.[2]	Val.[3]
Raw water		8.03 ± 0.13	187 ± 10.7	256 ± 12.22	9.51 ± 0.07	11.2 ± 0.24	8			
L	0.29	7.53 ± 0.01	26 ± 0.5	14.3 ± 1.90	7.38 ± 0.36	8.84 ± 0.33	2	60	5	5
L	0.3	7.56 ± 0.05	12 ± 1.0	5.98 ± 0.13	7.07 ± 0.08	8.29 ± 0.10	2	96	5	5
L	0.38	7.51 ± 0.02	16 ± 2.0	4.39 ± 0.16	6.29 ± 0.22	6.97 ± 0.27	2	60	5	5
L	0.44	7.54 ± 0.00	12 ± 1.5	2.52 ± 0.28	5.87 ± 0.11	6.12 ± 0.17	2	60	5	5
M	0.51	7.50 ± 0.02	10 ± 1.0	1.42 ± 0.36	4.61 ± 0.09	5.05 ± 0.20	2	54	5	5
M	0.57	7.53 ± 0.01	10 ± 0.5	1.15 ± 0.02	3.71 ± 0.13	3.89 ± 0.18	2	60	5	5
M	0.64	7.52 ± 0.04	8 ± 2.0	1.34 ± 0.18	2.62 ± 0.38	2.76 ± 0.42	2	59	5	5
H	0.7	7.55 ± 0.01	6 ± 2.0	1.48 ± 0.08	1.24 ± 0.19	1.38 ± 0.19	2	60	5	5
H	0.75	7.48 ± 0.01	5 ± 1.0	2.71 ± 0.22	0.63 ± 0.02	0.89 ± 0.09	2	60	5	5
H	0.85	7.52 ± 0.01	7 ± 1.0	2.85 ± 0.58	0.54 ± 0.01	0.65 ± 0.05	2	60	5	5
H	1.08	7.49 ± 0.005	32 ± 2.5	3.86 ± 0.91	0.26 ± 0.06	0.44 ± 0.07	2	60	5	5

Notes: 1—number of dose replicates, 2—number of images in calibration data-set, 3—number of images in validation data-set.

(2) Generate a set of random initial starting points, the number of these points should be chosen according to the image resolution.

(3) Develop the AMT complexity measures depending on a current value of the local scale (radius). The average value for a given number of starting points is calculated.

(4) Increase the local scale by 1 and repeat the last step specified number of times—manually selected value of scale.

(5) Transform the image into 1D spectrum, plotted as scale (radius) vs. measured mean angle (MA). Thus, the resulting complexity spectrum contains all the angle measures pertaining to all scales of the measurement series (unfolded image).

A detailed description of the AMT algorithm is given elsewhere (Esbensen et al., 1996; Fongaro et al., 2016; Kvaal et al., 2008).

Further, in the analysis the AMT spectrums data were used as the X-input for multivariate statistical analysis. Obtained matrix consisted of 110 samples and 200 variables (scale value).

2.8. Mathematical modelling

Advanced mathematical and statistical tools are often used to explore, evaluate, model, calibrate and predict data from both laboratory-scale and full-scale experiments. Certain process can be optimized with the help of mathematical modelling. Multivariate regression analysis is a powerful statistical tool, which enables the ability to derive relationships between several input parameters and one or several output parameters. It is especially helpful when the initial data are broad, i.e. massive influencing *X*-parameters is huge and it is not easy to establish the relationships between different parameters by simple linear regression. Furthermore, in many cases, the input parameters are highly correlated.

The Unscrambler® *X* 10.3 (CAMO Software AS, Norway) was used for the multivariate data analysis of 110 image samples.

Principal component analysis (PCA) was primarily used to explore the data, while principal component regression (PCR) and partial least squares regression (PLSR) were used to model the relations between input *X* parameters and corresponding *Y* responses. Models were first validated by

cross-validation and then using a half of the data for calibration and half of the data for test set vali-dation. The number of principal components (PCs)/factors was chosen according to the explained variance. Multiple linear regression (MLR) was used to predict the dosages by some GLCM feature vectors.

Linear discriminant analysis (LDA) with PCs was applied to obtain confusion matrixes of the histo-gram, GLCM, and AMT spectral data classification. Each data-set was divided into 3 classes accord-ing to total *P* efficiency (coded *L*, *M*, and *H*): *L* (low)—coagulant dosages which lead to low (20–50%) treatment efficiency after sedimentation; *M* (medium)—50–85% of total *P* removal; *H* (high)—above 85% of total *P* removal (Table 1). As a result, class *L* consisted of the data for 4 dosages: 0.29, 0.30, 0.38 and 0.44 mmol Al/l with 10 images from each experiment, 40 samples in total. Class *M*—3 dos-ages: 0.51, 0.57, 0.64 mmol Al/l, 30 samples in total. Class *H*—4 dosages: 0.70, 0.75, 0.85, 1.08 mmol Al/l, 40 samples in total.

3. Results and discussion

3.1. Images of flocs
The presented imaging technique shows the ability to investigate images of flocs during the process of coagulation–flocculation. Capturing images each 20 s gives an opportunity to observe and de-scribe changes in aggregates formation. Figure 2 shows the evolution of flocs during coagulation. Even by human eye, there is a very noticeable difference between the images. At the very first mo-ment after fast mixing, when mixing unit stops to change the rotation speed, first image was taken and it was called "zero seconds" of coagulation. Particles in water are quite small at this moment, initial flocs. However, within the next 20 s these initial particles rapidly form larger aggregates. The growth of flocs is distinctly visible until approximately 3–5 min of coagulation. After that, the flocs visually seem to be quite stable on their geometrical characteristics. However, according to the re-sults of further image analysis, flocs are continuing to grow slightly with time (Table 2).

3.2. Coagulant doses and images of flocs
Figure 3 shows the differences in flocs' structure due to coagulant dose (given in mmol Al/l). The figure describes relationships between dose and total *P* treatment efficiency. Images on the graph show how flocs looked during coagulation at some given time (400 s after the start of slow mixing). From Figure 3, the increase of coagulant dose results in a decrease of flocs' size, which is consistent with findings by Wang et al. (2011).

It is noticeable from the images that detection of physical characteristics (thereby fractal dimen-sion) of aggregates in this particular case (image scale) is possible when coagulant dosages are

Figure 2. Eight representative samples of pre-processed images, which illustrate the flocs aggregation process. The number on image shows time of coagulation in seconds. Coagulant dose—0.3 mmol Al/l.
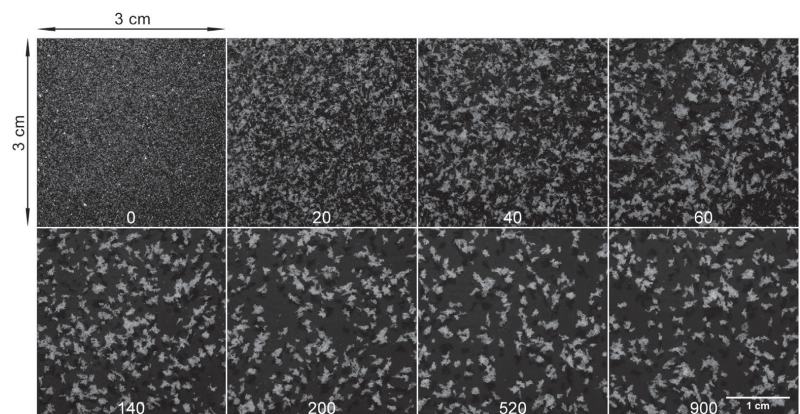
cogent · engineering

Table 2. Changes in mean fractal dimension and equivalent diameter of flocs with coagulation time			
Time (s)	Flocs count	Mean D_f*	d_p** (mm)
0	18,733	1.87 ± 0.08	0.09 ± 0.03
20	908	1.65 ± 0.07	0.29 ± 0.20
40	587	1.57 ± 0.07	0.36 ± 0.30
60	412	1.55 ± 0.07	0.47 ± 0.39
100	195	1.53 ± 0.07	0.65 ± 0.57
200	186	1.50 ± 0.07	0.72 ± 0.48
400	144	1.48 ± 0.06	0.85 ± 0.54
1,180	127	1.46 ± 0.05	0.90 ± 0.49

*Fractal dimension.
**Mean equivalent diameter.

quite low, but inapplicable for dosages with adequate treatment efficiencies. Closer to optimal coagulant dosages, flocs appear to be smaller, hence, they are overlapping in the image and accurate floc feature detection becomes inapplicable in this scale.

The results of coagulation are presented in Table 1.

3.3. Fractal dimension
Flocs' features as characteristics of particles were detected from the images of flocs for low coagulant dosages.

The results of one time series of experiments, dose—0.29 mmol Al/l, are shown in Table 2. Expectedly, the size (equivalent diameter) of particles is growing with coagulation time. The mean fractal dimension of initial particles is close to 2. As particles aggregate, fractal dimension rapidly decreases during the first minute of coagulation and then continues slowly decreasing. These results are consistent with findings of Chakraborti et al. (2003). Figure 4 shows the examples of 4 flocs (from 23,000 particles analysed for one dose) with different fractal dimensions.

Figure 3. Dependence of total P removal from coagulant dose and its associated image of flocs during coagulation.
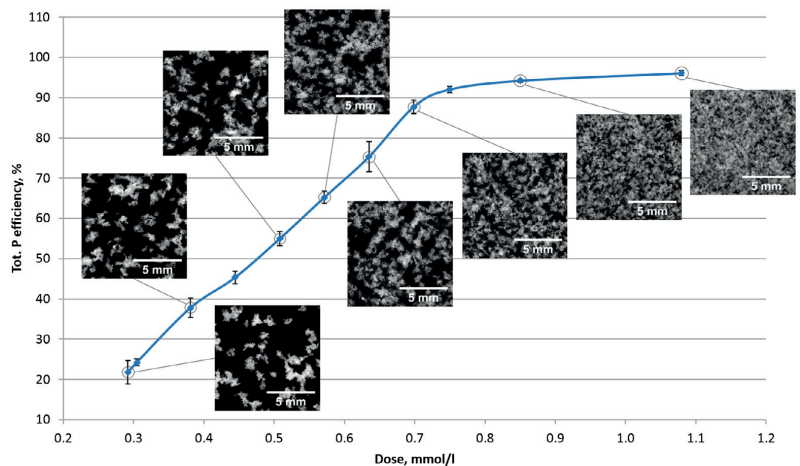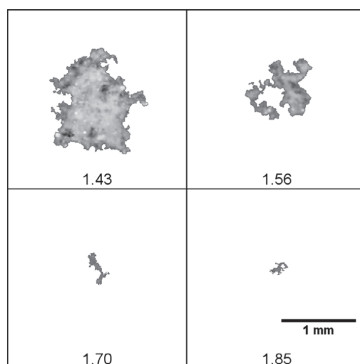
Figure 4. Examples of flocs detected for the coagulant dose 0.29 mmol Al/l, the number on the image equals fractal dimension.

3.4. Histogram analysis

Histograms provide the grey level distribution of pixels in the image. Images of flocs for each coagulant dose had their own typical histograms, so it was possible to distinguish which image's histogram corresponds to which coagulant dose. The results of classification are shown in Table 3.

PCA showed a poor explanation of *X* variance of 256 grey level histogram data. With 2 PCs 21.59% of calibration *X* variance was explained and 18.11% of cross-validation. Each next PC increased the explained variance by 1–3%.

The number of grey level values on histogram was decreased by 2 and 4 times, resulting in a better prediction of coagulant dosage. Thus, regression analysis of 64 bins histograms gave the best PCR and PLSR models, listed in Table 4.

3.5. GLCM analysis

PCR analysis was applied to all 20 GLCMs (5 different pixel steps and 4 directions) to find the best combination of pixel step and direction. *X* was set as 11 GLCM variables for each image, *Y*—coagulant dose. However, it was found that neither differences in pixel step nor direction give a significant change in coagulant dosage prediction by PCR. Cross-validation R^2 fluctuated between 95 and 96%, meaning that images of flocs are invariant and independent of the difference in pixel step. Hence, all further analyses were done for the GLCM obtained with the pixel step 1 and direction 0°.

The results of PCR model are shown in Figure 5. Coagulant dosages were quite well separated in the score plot with the first two PCs, where the total explained *X* variance equals 87.69% (PC1 = 73.82%, PC2 = 13.87%) for calibration and 84.17% (PC1 = 70.33%, PC2 = 13.84%) for cross-validation. Validation method was based on systematic block cross-validation, which created 11 submodels where 5 samples (replicates of one dose) had been left out in each of the submodels. The total explained *Y* variance hence, the accuracy of prediction equals 95.50% (calibration) and 94.98% (cross-validation) with only one PC.

All variables in the correlation loadings plot (Figure 5) fall within the circles, which depict 50–100% of explained variance. The uncertainties in various model parameters were estimated by jack-knifing for the model with 1 PC. The variables found significant are marked with circles in the correlation loadings plot and with stripes pattern in the regression coefficients plot. Ten out of eleven *X* variables are found to be significant at the 95% confidence interval. Variable "Entropy" has uncertainty limits crossing the zero line, so it is not significant at the 5% level for the model with 1 PC.

For the real on-line applications, it is preferred to have a linear regression model as it is simpler to implement. The selection of variables and estimation of a relevant number of model inputs are important parts of the data analysis. In order to perform the correct variable selection, *R console* with the "NMBU" plug-in (Liland & Sæbø, 2016) was used. Two results of the best subsets for a different amount of variables are shown in Appendix A. For instance, the detected best model with two variables "IDM" and "Variance" has the prediction $R^2 = 0.967$ and $R^2_{adj} = 0.965$. Since R^2 and R^2_{adj} are close to 1, the regression model shows the high agreement with the experimental results. With the increase in number of variables, R^2 and R^2_{adj} tend to slightly increase. However, the loadings plot (Figure 5) shows that *X* variables are highly correlated and it is worth to keep as low amount of variables as possible. In this particular case, we would suggest choosing a model with two variables. Nevertheless, the best way of choosing the correct number of variables would be to make prediction analysis with the completely new test data-set.

The linear equation for the model with two variables is:

$$Dose = \beta_0 + \beta_1 \times IDM + \beta_2 \times Variance + \varepsilon, \tag{4}$$

Table 3. Confusion matrixes of PCA-LDA for three image analysis methods

Predicted class		Actual class		
		L	*M*	*H*
(a) Histogram, 256 bins, 3 PCs				
	L	40	1	0
	M	0	29	8
	H	0	0	32
(b) GLCM, 11 variables, 3PCs				
	L	40	3	0
	M	0	27	8
	H	0	0	32
(c) AMT, 200 variables, 2 PCs				
	L	40	4	0
	M	0	26	0
	H	0	0	40

Table 4. Results of multiple regression analysis for 3 different image analysis techniques

Regression method	PCR		PLSR	
Number of components/factors	PC-2		Factor-2	
Explained *Y* variance	R^2 cal. (%)	R^2 val. (%)	R^2 cal. (%)	R^2 val. (%)
Histogram, 256 bins	87.39	83.53	96.35	88.84
Histogram, 128 bins	90.28	88.53	95.07	93.03
Histogram, 64 bins	91.55	89.92	95.51	93.80
GLCM, 4 variables	87.54	87.95	92.89	92.42
GLCM, 6 variables	92.76	92.83	95.58	95.22
GLCM, 11 variables	95.91	95.59	96.89	96.09
AMT, 200 variables	95.84	95.01	96.90	96.19

Figure 5. Results of PCR of 11 GLCM textural features obtained from the images of flocs for different coagulant dosages.
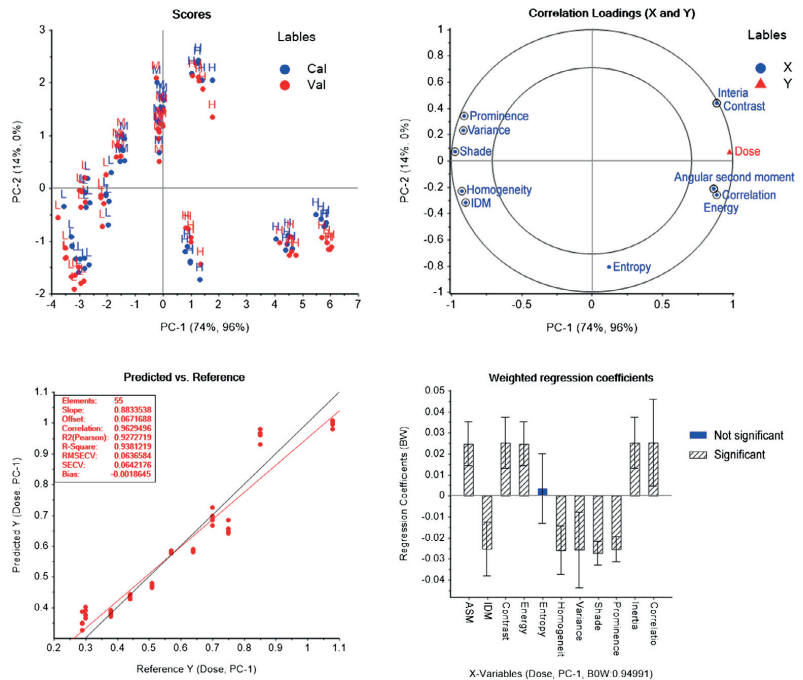
where β_0—intercept, β_1, β_2—coefficients which determine the contribution of each independent variable to the estimation of the dependent variable (dose), ε random error. The coefficients were found by MLR with $R = 0.983$, R^2(cal) = 0.967 and R^2(cross-val) = 0.947, $p < 0.001$ and the final model equation was as follow:

$$\widehat{Dose} = 3.09 - 2.44 \times IDM - 9 \times 10^{-4} \times Variance, \tag{5}$$

It should be noted that the big difference in the values of parameters $\widehat{\beta}_1$ and $\widehat{\beta}_2$ is caused by the difference in scale of variables "IDM" and "Variance" (Figure 6(a)). The comparison and equal significance of both variables can be seen from the plot "weighted regression coefficients" (Figure 5).

It is not easy to interpret the physical sense of this linear equation because the GLCM feature vectors are the second-order statistical parameters of the original images. When IDM = Variance = 0, the Dose is 3.09 mmol Al/l. However, the value $\widehat{\beta}_0$ does not have a practical interpretation since under the normal conditions images of flocs have grey-tone pixel variations, thus, IDM and Variance differ from 0. When the IDM increases by 1, while Variance remains constant, the estimated mean change of Dose decreases by 2.44. When the Variance increases by 1,000, while IDM remains constant, the estimated mean change of Dose decreases by 0.9. The typical images with high IDM and Variance refer to low dosages (Figure 6(a)), which can be visually seen from Figure 3.

The results of MLR prediction are shown in Figure 6(b). The total explained *Y* variance equals 96.66% (calibration) and 94.74% (cross-validation) with 2 GLCM feature vectors—"IDM" and "Variance". It should be emphasized, that this research aims at showing that the texture analysis methods could be applied for determining the coagulant dose. The good repeatability of the results from the images of dynamic coagulation process should also be noted. Despite the high prediction

Figure 6. (a) mean values of variables IDM and Variance vs. dose; (b) scatter plot of predicted vs. observed coagulant dosages for MLR model with two GLCM feature vectors.
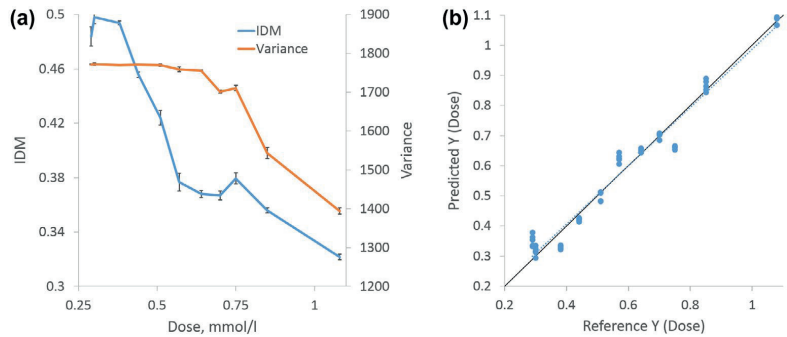

results by PCR and MLR, the models have to be further verified and tested on the bigger data-sets, best—on the data from a wastewater treatment plant.

3.6. AMT

AMT is a potential tool for observing and describing images of flocs both for high and low coagulant dosages. It describes the texture complexity of an image, converting information of grey level values to descriptive spectrums. The spectrum shows the dependence of the MA (calculated for given number of random sample points, in our research it was set to 10,000 samples) in an unfolded image from the scale (radius). Figure 7(a) shows the AMT spectra (complexity) results of flocs images captured at the same coagulation time (400 s) for eleven coagulant dosages from 0.29 to 1.08 mmol Al/l. In a range of low coagulant dosages, the flocs tend to be bigger and detached from each other (see example images in Figure 3). AMT spectrums are not always easy to interpret. However, in application to coagulation, there are visible trends in images' complexities. For the flocs images related to low coagulant dosages the picks of AMT spectrums are lying within larger scale values and have lower measured angles (Figure 7(a)). In contrast, images of flocs associated with higher coagulant dosages have AMT spectrums picks lying in lower scale range and having higher measured angles. This trend– spectrums' vertexes shift to the left, will be easier to explain if we look into the graphs of images' unfold. Figure 8(a) illustrates the unfolding of the first row of an image related to low coagulant dose (0.29 mmol/l). The neighbourhood pixel grey level variations are not dramatic and the pictures of grey levels (which associate with the flocs on the image) are not very frequent. The biggest grey level pixel variations, thus high measured angles, between the neighbouring areas within the whole image occurs in the distance (scale) of approximately 90 pixels, which is visible from Figure 7(a). In contrast, the unfolding of the image associated with higher coagulant dose shows huge and frequent neighbourhood grey-tone variations (Figure 8(b)). Thus, the biggest grey level changes on the image occur in a distance (scale) of approximately 50 pixels, resulting in highest measured angles.

Figure 7(b) shows the AMT spectra results for a time series of flocs images for one certain coagulant dose—0.3 mmol Al/l. For the first seconds of coagulation, AMT spectrums of flocs images are quite sharp, with the maximum angle values within low scales. This can be explained by significant grey level changes in neighbourhood pixel areas in the images due to the high amount of small particles. With time, while aggregates are merging and growing the differences in grey-tone pixel values of close areas are decreasing, resulting in a shift of AMT spectra's vertex to the right, thus to higher scales (bigger distances). From some time, changes in flocs images' textures are becoming almost invisible. The aggregates and their spatial locations change from image to image, because of continuing coagulation, but overall textural characteristics, in other words—images' complexities, remain quite stable, what is also visible from AMT spectrums. Spectrums related to coagulation time 160–1,180 s are overlapping, indicating the stability of images' textures, thus certain stability of flocs' geometrical characteristics. The tendencies described above were observed for all coagulant doses.

Figure 7. AMT spectrums for: (a) different doses at the same time—400 s after the start of coagulation; (b) the time series images of flocs, dose—0.3 mmol Al/l.
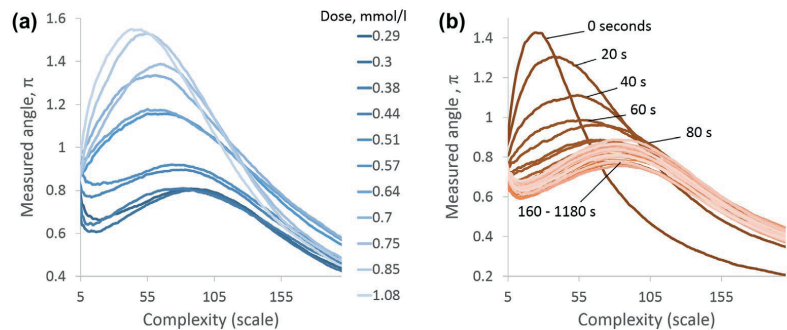
Summarizing all above, the AMT spectrums regarding images of flocs can be interpreted as follows: the bigger the flocs and the better they are separated from each other, the more gentle is AMT slope of the spectra; the smaller the flocs and the higher their amount on the image (overlapping), the sharper the resulting AMT spectra, which situates in the lower scales.

The spectrums' data from time series images were processed in *The Unscrambler*® by PCA. The results of PCA showed a good separation of "stable" images from those taken during the first seconds of coagulation. Three principle components resulted in 97.5% of explained calibration variance and 96.3% of cross-validation variance.

Figure 9 shows the scores on PC1 of AMT spectrum data for time series of flocs images. Results indicate that after approximately 200 s of coagulation images of flocs become quite stable.
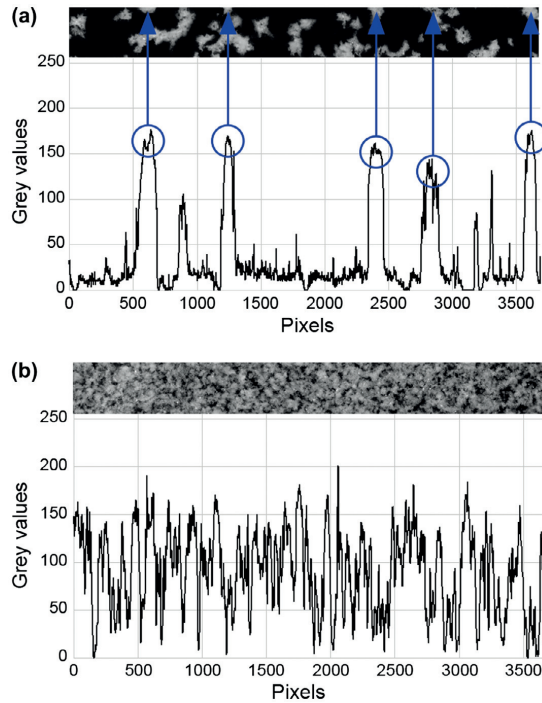
3.7. Mathematical modelling

LDA has shown that all three methods of texture image processing give quite good results of data classification. The information about the coagulant dosages and treatment efficiencies was not revealed during LDA, only information from the images was used for classification. LDA was based on score vectors of PCA for classification with estimation of prior probabilities. Confusion matrix results for three image analysis techniques are shown in Table 3.

Data based on histogram analysis of flocs images were classified with 91.82% accuracy (Table 3(a)) using three PCs. One out of ten samples of dose 0.51 mmol Al/l was classified as a low tot. *P* efficiency (class *L*) instead of medium efficiency (class *M*). Also, 8 samples of dose 0.70 mmol Al/l that belongs to class *H* (high tot. *P* efficiency) were classified to *M*.

Data from GLCM image analysis gave the classification accuracy of 90.00% using 3 PCs (Table 3(b)). The classification results are very much similar to the results described above for histogram analysis. However, two more sample images of dose 0.51 mmol Al/l were misclassified.

The highest PCA–LDA classification accuracy of 96.36% was determined for the AMT spectrums data of images of flocs using 2 PCs (Table 3(c)). Only 4 samples of dose 0.51 mmol Al/l from class *M* were misclassified to *L*. All 40 samples of high treatment efficiency (class *H*) were classified correctly. This is a very good result, showing that AMT is a perspective tool for image analysis of flocs—AMT spectrums contain necessary data for modelling and precise classification. AMT has the potential to be used for further image analysis of flocs and development of coagulation dosage control system.

Figure 8. Unfold of the first photograph's line of pixels for the coagulant dosages: (a) 0.29 mmol/l; (b) 1.08 mmol/l.

The results of multivariate regression analysis are shown in Table 4. PCR and PLSR were used for mathematical modelling. Response *Y* is the coagulant dosages in mmol Al/l, *X*s are matrixes of histogram data, GLCM feature vectors, and AMT spectrums. Data-sets were divided into two equal parts for calibration and validation of the model. Hence, calibration and validation sets consisted of 11 values of coagulant dosages and 5 replicates of flocs' images per each dose resulting in a total of 55 samples. The obtained matrices of histogram image analysis consisted of 256, 128 and 64 variables (bins). GLCM data was tested with 4 (ASM, contrast, correlation, and entropy), 6 (ASM, contrast, variance, correlation, IDM, entropy) and all 11 variables. AMT data consisted of spectrums with 200

Figure 9. Scores on PC1 of AMT spectrums, calculated from images of flocs for 2 doses—0.3 and 0.75 mmol Al/l.
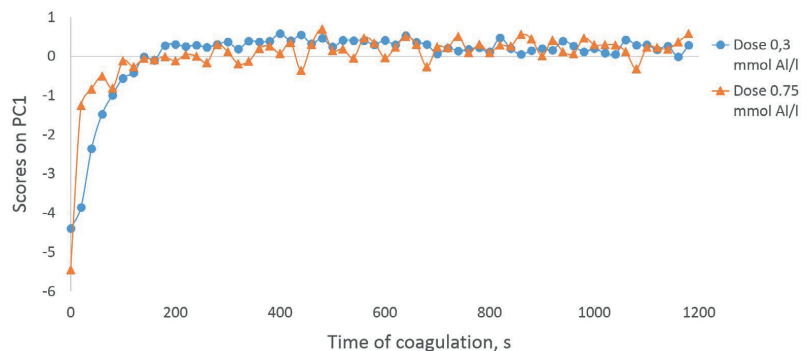
points, which were set as variables. R^2 of calibration and validation data-sets was used to compare the obtained regression models.

PLSR model based on AMT spectrums showed the best result—96.19% variance explained by 2 factors for validation data. PLSR of GLCM data with 11 variables gave the same result—96.09% Y variance explained. Slightly lower validation R^2 (93.80%) was calculated for PLSR based on 64 bins histogram of flocs' images. All PCR models showed from 1 to 10% lower R^2 compared to PLSR models. This can be explained by the model calculation procedure since PLSR technique models relations between X and Y considering Y during the modelling process, while PCR bases only on X matrix.

All validated models showed quite good results with a minimum of 83% variance explained. It means that images of flocs correlate with coagulant dosages and that the dosages could be predicted by multiple statistical methods. Images of flocs for particular dosage are unique, thus, potentially could be used for dosage control. However, it should be further tested and proved through bigger data with the wider range of coagulant dosages and more sample points. Further analysis should be also performed for other initial parameters of model wastewater and/or industrial wastewater. Among the limitations of the texture recognition methods, the dependence on the lighting conditions should be stated. The described texture analysis methods are dependent on the grey-tone pixel values in the images, thus, to be able to compare the images it is important to have stable lighting conditions or perform mean centring of the images during pre-processing stage. Despite lighting conditions which might vary during laboratory scale experiments, with the full-scale installation it is most likely all conditions are set to a steady state, so this problem should not arise.

4. Conclusions

Changes in flocs' structure depending on the coagulant dosage and treatment efficiency were observed.

According to the treatment efficiency, the 11 coagulant dosages were categorized into three classes for further PCA-LDA. Image analysis by AMT showed the best classification results.

The image analysis method which bases on fractal dimension of flocs has been widely studied. We have observed that there are some difficulties with the application of this method for highly contaminated water—such as wastewater.

Various texture image analysis methods, applied in the other scientific fields were tested for wastewater coagulation. GLCM and AMT were found to be the most promising methods of flocs images identification and correlation with coagulant dosages.

Ten out of eleven co-occurrence matrix-based texture features were found to be significant for dosage prediction. The linear regression model was built based on two of these feature vectors. IDM and Variance were found to adequately describe the flocs' images and their respective dosages.

Texture image analysis methods are easily computable and potentially could be used together or instead of conventional object-recognition image analysis techniques in order to simplify the image analysis methodology, especially in cases when the particles detection is complicated by the initial environment and could lead to under- or overestimation of the flocs' parameters.

Concluding, texture image analysis methods and regression models are the way to correlate images of flocs with dosages and treatment efficiencies, which enables optimal coagulant dosage control.

cogent ·· engineering

Author details
Nataliia Sivchenko[1]
E-mail: nataliia.sivchenko@gmail.com
ORCID ID: http://orcid.org/0000-0002-1054-0794
Knut Kvaal[1]
E-mail: knut.kvaal@nmbu.no
ORCID ID: http://orcid.org/0000-0002-0579-7448
Harsha Ratnaweera[1]
E-mail: harsha.ratnaweera@nmbu.no
ORCID ID: http://orcid.org/0000-0003-1456-2541
[1] Department of Mathematical Sciences and Technology, Norwegian University of Life Sciences, P.O. Box 5003-IMT, Aas 1432, Norway.

Cover image
Source: Authors.

Citation information
Cite this article as: Evaluation of image texture recognition techniques in application to wastewater coagulation, Nataliia Sivchenko, Knut Kvaal & Harsha Ratnaweera, *Cogent Engineering* (2016), 3: 1206679.

References
Alves, M., Cavaleiro, A. J., Ferreira, E. C., Amaral, A. L., Mota, M., da Motta, M., & Pons, M. N. (2000). Characterisation by image analysis of anaerobic sludge under shock conditions. *Water Science and Technology, 41*, 207–214.
Amaral, A. L., & Ferreira, E. C. (2005). Activated sludge monitoring of a wastewater treatment plant using image analysis and partial least squares regression. *Analytica Chimica Acta, 544*, 246–253. doi:10.1016/j.aca.2004.12.061
Andrle, R. (1994). The angle measure technique: A new method for characterizing the complexity of geomorphic lines. *Mathematical Geology, 26*, 83–97. doi:10.1007/BF02065877
Annadurai, G., Sung, S. S., & Lee, D. J. (2004). Simultaneous removal of turbidity and humic acid from high turbidity stormwater. *Advances in Environmental Research, 8*, 713–725. http://dx.doi.org/10.1016/S1093-0191(03)00043-1
Cao, B., Gao, B., Liu, X., Wang, M., Yang, Z., & Yue, Q. (2011). The impact of pH on floc structure characteristic of polyferric chloride in a low DOC and high alkalinity surface water treatment. *Water Research, 45*, 6181–6188. doi:10.1016/j.watres.2011.09.019
Chakraborti, R. K., Atkinson, J. F., & Van Benschoten, J. E. (2000). Characterization of alum floc by image analysis. *Environmental Science & Technology, 34*, 3969–3976. http://doi.org/10.1021/es990818o
Chakraborti, R. K., Gardner, K. H., Atkinson, J. F., & Van Benschoten, J. E. (2003). Changes in fractal dimension during aggregation. *Water Research, 37*, 873–883. http://dx.doi.org/10.1016/S0043-1354(02)00379-2
Conners, R. W., Trivedi, M. M., & Harlow, C. A. (1984). Segmentation of a high-resolution urban scene using texture operators. *Computer Vision, Graphics, and Image Processing, 25*, 273–310. doi:10.1016/0734-189X(84)90197-X
da Motta, M., Pons, M.-N., Roche, N., & Vivier, H. (2001). Characterisation of activated sludge by automated image analysis. *Biochemical Engineering Journal, 9*, 165–173. doi:10.1016/S1369-703X(01)00138-3
Dagot, C., Pons, M., Casellas, M., Guibaud, G., Dollet, P., & Baudu, M. (2001). Use of image analysis and rheological studies

for the control of settleability of filamentous bacteria: Application in SBR reactor. *Water Science and Technology, 43*, 27–33.
Dahl, C. K., & Esbensen, K. H. (2007). Image analytical determination of particle size distribution characteristics of natural and industrial bulk aggregates. *Chemometrics and Intelligent Laboratory Systems, 89*, 9–25. doi:10.1016/j.chemolab.2007.05.005
Esbensen, K. H., Hjelmen, K. H., & Kvaal, K. (1996). The AMT approach in chemometrics–first forays. *Journal of Chemometrics, 10*, 569–590. http://dx.doi.org/10.1002/(ISSN)1099-128X
Fettig, J., Ratnaweera, H., & Ødegaard, H. (1990, october 1–3). Simultaneous phosphate precipitation and particle destabilization using aluminium coagulants of different basicity. *Chemical Water and Wastewater Treatment: Proceedings of the 4th Gothenburg Symposium Madrid, Spain, 221*–242. doi:10.1007/978-3-642-76093-8
Fongaro, L., & Kvaal, K. (2013). Surface texture characterization of an Italian pasta by means of univariate and multivariate feature extraction from their texture images. *Food Research International, 51*, 693–705. doi:10.1016/j.foodres.2013.01.044
Fongaro, L., Lin Ho, D. M., Kvaal, K., Mayer, K., & Rondinella, V. V. (2016). Application of the angle measure technique as image texture analysis method for the identification of uranium ore concentrate samples: New perspective in nuclear forensics. *Talanta, 152*, 463–474. doi:10.1016/j.talanta.2016.02.027
Gorczyca, B., & Ganczarczyk, J. (1996). Image Analysis of Alum Coagulated Mineral Suspensions. *Environmental Technology, 17*, 1361–1369. doi:10.1080/09593331708616505
Gorczyca, B., & Ganczarczyk, J. (2002). Flow rates through alum coagulation and activated sludge flocs. *Water Quality Research Journal of Canada, 37*, 389–398.
Haralick, R. M., Shanmugam, K., & Dinstein, I. (1973). Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics, 3*, 610–621. doi:10.1109/TSMC.1973.4309314
He, W., Nan, J., Li, H., & Li, S. (2012). Characteristic analysis on temporal evolution of floc size and structure in low-shear flow. *Water Research, 46*, 509–520. doi:10.1016/j.watres.2011.11.040
Huang, J., & Esbensen, K. H. (2000). Applications of angle measure technique (AMT) in image analysis: Part I. A new methodology for *in situ* powder characterization. *Chemometrics and Intelligent Laboratory Systems, 54*, 1–19. doi:10.1016/s0169-7439(00)00100-3
Huang, J., & Esbensen, K. H. (2001). Applications of AMT (angle measure technique) in image analysis: Part II: Prediction of powder functional properties and mixing components using Multivariate AMT Regression (MAR). *Chemometrics and Intelligent Laboratory Systems, 57*, 37–56. doi:10.1016/s0169-7439(01)00120-4
Jarvis, P., Jefferson, B., Gregory, J., & Parsons, S. A. (2005). A review of floc strength and breakage. *Water Research, 39*, 3121–3137. http://dx.doi.org/10.1016/j.watres.2005.05.022
Jenné, R., Banadda, E., Gins, G., Deurinck, J., Smets, I. Y., Geeraerd, A. H., & Van Impe, J. F. (2006). Use of image analysis for sludge characterisation: Studying the relation between floc shape and sludge settleability. *Water Science & Technology, 54*, 167–174.
Jenné, R., Banadda, E. N., Smets, I., Deurinck, J., & Van Impe, J. (2007). Detection of filamentous bulking problems: Developing an image analysis system for sludge composition monitoring. *Microscopy and Microanalysis, 13*, 36–41. doi:10.1017/S1431927607070092
Jin, Y. (2005). *Use of a high resolution photographic technique for studying coagulation/flocculation in water treatment*. Saskatoon: University of Saskatchewan.

Juntunen, P., Liukkonen, M., Lehtola, M., & Hiltunen, Y. (2013). Dynamic soft sensors for detecting factors affecting turbidity in drinking water. *Journal of Hydroinformatics, 15*, 416–426. http://dx.doi.org/10.2166/hydro.2012.052

Juntunen, P., Liukkonen, M., Lehtola, M., & Hiltunen, Y. (2014). Characterization of alum floc in water treatment by image analysis and modeling. *Cogent Engineering, 1*, 1–13. http://doi.org/10.1080/23311916.2014.944767

Juntunen, P., Liukkonen, M., Pelo, M., Lehtola, M. J., & Hiltunen, Y. (2012). Modelling of water quality: An application to a water treatment process. *Applied Computational Intelligence and Soft Computing, 2012*, 1–9. doi:10.1155/2012/846321

Kim, S.-H., Moon, B.-H., & Lee, H.-I. (2001). Effects of pH and dosage on pollutant removal and floc structure during coagulation. Microchemical Journal, *68*, 197–203. doi:10.1016/S0026-265X(00)00146-6

Kucheryavski, S. (2007). Using hard and soft models for classification of medical images. *Chemometrics and Intelligent Laboratory Systems, 88*, 100–106. doi:10.1016/j.chemolab.2006.08.012

Kvaal, K. (2014). *AMT plugin for ImageJ*. Retrieved from http://arken.nmbu.no/~kkvaal/eamtexplorer/imagej_plugins.html

Kvaal, K., Kucheryavski, S. V., Halstensen, M., Kvaal, S., Flø, A. S., Minkkinen, P., & Esbensen, K. H. (2008). eAMTexplorer: A software package for texture and signal characterization using angle measure technique. *Journal of Chemometrics, 22*, 717–721. doi:10.1002/cem.1160

Kvaal, K., Wold, J. P., Indahl, U. G., Baardseth, P., & Næs, T. (1998). Multivariate feature extraction from textural images of bread. *Chemometrics and Intelligent Laboratory Systems, 42*, 141–158. doi:10.1016/S0169-7439(98)00017-3

Li, D. H., & Ganczarczyk, J. (1989). Fractal geometry of particle aggregates generated in water and wastewater treatment processes. *Environmental Science & Technology, 23*, 1385–1389. http://doi.org/10.1021/es00069a009

Liland, K. H., & Sæbø, S. (2016). *R Commander plug-in for university level applied statistics*. Retrieved from https://cran.r-project.org/web/packages/RcmdrPlugin.NMBU/index.html

Lin, J. L., Huang, C., Chin, C. J. M., & Pan, J. R. (2008). Coagulation dynamics of fractal flocs induced by enmeshment and electrostatic patch mechanisms. *Water Research, 42*, 4457–4466. http://dx.doi.org/10.1016/j.watres.2008.07.043

Liukkonen, M., Hiltunen, T., & Hiltunen, Y. (2015). Evaluation of fluidized bed condition by image analysis and modelling. *IFAC Proceedings Volumes (IFAC-PapersOnline), 48*, 892–897. doi:10.1016/j.ifacol.2015.05.191

Mendret, J., Guigui, C., Schmitz, P., Cabassud, C., & Duru, P. (2007). An optical method for *in situ* characterization of fouling during filtration. *AIChE Journal, 53*, 2265–2274. doi:10.1002/aic.11257 http://dx.doi.org/10.1002/(ISSN)1547-5905

Mesquita, D. P., Amaral, A. L., & Ferreira, E. C. (2011). Characterization of activated sludge abnormalities by image analysis and chemometric techniques. *Analytica Chimica Acta, 705*, 235–242. doi:10.1016/j.aca.2011.05.050

Mortensen, P. P., & Esbensen, K. H. (2005). Optimization of the angle measure technique for image analytical sampling of particulate matter. *Chemometrics and Intelligent Laboratory Systems, 75*, 219–229. doi:10.1016/j.chemolab.2004.08.004

Ødegaard, H., Fettig, J., & Ratnaweera, H. (1990, october 1–3). Coagulation with prepolymerized metal salts. *Chemical water and wastewater treatment: Proceedings of the 4th Gothenburg symposium Madrid, Spain, 304*, 189–220. http://dx.doi.org/10.1007/978-3-642-76093-8

Otsu, N. (1979). A threshold selection method from gray-level histograms. *IEEE Trans. Systems Man Cybernet., 9*, 62–66.

Rasband, W. S. (1997/2016). *ImageJ, US National Institutes of Health, Bethesda, MD*, http://imagej.nih.gov/ij/

Ratnaweera, H. (1991). *Influence of the degree of coagulant prepolymerization on wastewater coagulation mechanisms*. Trondheim: Norwegian Institute of Technology.

Ratnaweera, H., & Fettig, J. (2015). State of the art of online monitoring and control of the coagulation process. *Water, 7*, 6574–6597. doi:10.3390/w7116574

Ratnaweera, H., Lei, L., & Lindholm, O. (2002). Simulation program for wastewater coagulation. *Water Science & Technology, 46*, 27–33.

Tambo, N., & Hozumi, H. (1979). Physical characteristics of flocs—II: Strength of floc. *Water Research, 13*, 421–427. doi:10.1016/0043-1354(79)90034-4

Tambo, N., & Watanabe, Y. (1979a). Physical aspect of flocculation process—I: Fundamental treatise. *Water Research, 13*, 429–439. doi:10.1016/0043-1354(79)90035-6

Tambo, N., & Watanabe, Y. (1979b). Physical characteristics of flocs—I. The floc density function and aluminium floc. *Water Research, 13*, 409–419. doi:10.1016/0043-1354(79)90033-2

Thomas, D. N., Judd, S. J., & Fawcett, N. (1999). Flocculation modelling: A review. *Water Research, 33*, 1579–1592. http://dx.doi.org/10.1016/S0043-1354(98)00392-3

Vahedi, A., & Gorczyca, B. (2011). Application of fractal dimensions to study the structure of flocs formed in lime softening process. *Water Research, 45*, 545–556. doi:10.1016/j.watres.2010.09.014

Vahedi, A., & Gorczyca, B. (2012). Predicting the settling velocity of flocs formed in water treatment using multiple fractal dimensions. *Water Research, 46*, 4188–4194. doi:10.1016/j.watres.2012.04.031

Vahedi, A., & Gorczyca, B. (2014). Settling velocities of multifractal flocs formed in chemical coagulation process. *Water Research, 53*, 322–328. doi:10.1016/j.watres.2014.01.008

Vold, M. J. (1963). Computer simulation of floc formation in a colloidal suspension. *Journal of Colloid Science, 18*, 684–695. doi:10.1016/0095-8522(63)90061-8

Wang, Y. L., Feng, J., Dentel, S. K., Lu, J., Shi, B. Y., & Wang, D. S. (2011). Effect of polyferric chloride (PFC) doses and pH on the fractal characteristics of PFC–HA flocs. *Colloids and Surfaces A: Physicochemical and Engineering Aspects, 379*, 51–61. doi:10.1016/j.colsurfa.2010.11.044

Xiao, F., Lam, K. M., Li, X. Y., Zhong, R. S., & Zhang, X. H. (2011). PIV characterisation of flocculation dynamics and floc structure in water treatment. *Colloids and surfaces a: Physicochemical and engineering aspects, 379*, 27–35. doi:10.1016/j.colsurfa.2010.11.053

Yu, R.-F., Chen, H.-W., Cheng, W.-P., & Chu, M.-L. (2009). Simultaneously monitoring the particle size distribution, morphology and suspended solids concentration in wastewater applying digital image analysis (DIA). *Environmental Monitoring and Assessment, 148*, 19–26. doi:10.1007/s10661-007-0135-z

Yu, R.-F., Cheng, W.-P., & Huang, H.-D. (2012, June 18–20). On-line assessment of the particle separation in chemical flocculent suspension by image analysis. In M. Jekel (Ed.), *International conference on particle separation* (pp. 53–57). Berlin: International Water Association.

Zheng, C., Sun, D.-W., & Zheng, L. (2006). Recent applications of image texture for evaluation of food qualities—A review. *Trends in Food Science & Technology, 17*, 113–128. http://doi.org/10.1016/j.tifs.2005.11.006

Zheng, H., Zhu, G., Jiang, S., Tshukudu, T., Xiang, X., Zhang, P., & He, Q. (2011). Investigations of coagulation–flocculation process by performance optimization, model prediction and fractal structure of flocs. *Desalination, 269*, 148–156. http://doi.org/10.1016/j.desal.2010.10.054

Appendix A

The output from *R*: results of the variable selection for MLR model by best subsets.

>best subsets (RegModel.3, *n*best = 2, nvmax = 11, force.in = NULL)

Reordering variables and trying again:

		ASM	Cont	Corr	Ener	Entr	Homog	IDM	Iner	Prom	Shade	Var	RSS	R^2	$R^2_{adj.}$	Cp
1	(1)										*		0.2097	0.9296	0.9282	439.5
1	(2)						*						0.3339	0.8878	0.8857	730.1
2	(1)							*				*	0.0994	0.9666	0.9653	183.6
2	(2)			*				*					0.1029	0.9654	0.9641	191.6
3	(1)			*				*			*		0.0574	0.9807	0.9796	87.3
3	(2)							*			*	*	0.0605	0.9797	0.9785	94.6
4	(1)			*			*	*			*		0.0432	0.9855	0.9843	56.1
4	(2)						*	*			*	*	0.0469	0.9843	0.9829	64.6
5	(1)			*		*		*	*		*		0.0283	0.9905	0.9895	23.2
5	(2)		*	*		*		*			*		0.0283	0.9905	0.9895	23.2
6	(1)			*		*	*		*		*	*	0.0209	0.9929	0.9921	7.9
6	(2)		*	*		*	*				*	*	0.0209	0.9929	0.9921	7.9
7	(1)			*		*	*		*	*	*	*	0.0187	0.9937	0.9928	4.7
7	(2)		*	*		*	*			*	*	*	0.0187	0.9937	0.9928	4.7
8	(1)		*	*		*	*	*		*	*	*	0.0185	0.9938	0.9927	6.2
8	(2)			*		*	*	*	*	*	*	*	0.0185	0.9938	0.9927	6.2
9	(1)	*		*		*	*	*	*	*	*	*	0.0184	0.9938	0.9926	8.0
9	(2)	*	*	*		*	*	*		*	*	*	0.0184	0.9938	0.9926	8.0

Notes: ASM—angular second moment, Cont—contrast, Corr—correlation, Ener—energy, Entr—entropy, Homog—homogeneity, IDM—inverse difference moment, Iner—inertia, Prom—prominence, Var—variance.

Paper II

Sivchenko, N., Shostak, V., Kalashnikov, Y., Ratnaweera, H, & Kvaal, K.
Relationships between floc texture features and model wastewater qualities. Manuscript

Relationships between model wastewater parameters and texture features of flocs images

N. Sivchenko*, V. Shostak, Y. Kalashnikov, H. Ratnaweera, K. Kvaal

Norwegian University of Life Sciences, PO Box 5003-RealTek, 1432 Aas, Norway

* corresponding author: nataliia.sivchenko@nmbu.no

Abstract: Coagulation is a well-known and broadly used method of water and wastewater treatment. The conceptual and universally accepted model of the coagulation-flocculation process still does not exist due to a complex nature of the process. The industry requires advanced, robust and at the same time cost-efficient dosage control strategies as well as new monitoring tools to achieve the required treatment efficiencies. The texture image analysis technique – grey level co-occurrence matrix (GLCM) is applied in this research to evaluate the images of flocs acquired during different coagulation conditions. The varied parameters were coagulant dosages and the initial model wastewater characteristics in terms of concentrations of particles and phosphates. Three different coagulants were tested. The data obtained by GLCM and jar tests analysis were processed by principal component analysis (PCA) to find the relationships between model wastewater parameters and texture features of flocs images. GLCM feature vector Homogeneity is found to correlate with outlet model wastewater parameters. Variance on the image is negatively associated with increase of coagulant dosages. Contrast and Entropy texture features of the image are positively associated with zeta potential and coagulant dosages and negatively associated with the increase of particles and phosphates in the initial model wastewater.

Keywords: coagulation; model wastewater; image analysis; texture analysis; PCA

1. Introduction

Coagulation-flocculation is a fairly old and extensively used process in drinking water, domestic and industrial wastewater treatment to remove suspended solids, colour, phosphates and other water contaminants. Destabilisation of colloidal system with further aggregation of particles is enabled by addition of coagulants, typically aluminium or iron salts. The aggregated particles are called flocs (Bratby 2016). Flocs are removed after or during coagulation by different solid-liquid separation methods, for instance, sedimentation, flotation, contact filtration, ActiFlo® etc. Currently, a reliable and universally accepted model of coagulation process does not exist. Partly because it is difficult to simulate the dynamics of the coagulation-flocculation (Thomas et al. 1999) and sedimentation process (Li & Stenstrom 2014), and also because of a significant number of rapid variations in the inlet wastewater parameters.

Coagulation is a complex non-linear process that is challenging to operate and control. Optimisation of the process is often based on the results from jar tests analysis. It was documented that a flow-proportional feed-forward dosing concept is most often used for coagulation dosage control, sometimes with corrections based on pH in the flocculation chamber (Ratnaweera & Fettig 2015). However, the optimal dosage of coagulant is a function of some other inlet wastewater characteristics such as suspended solids/turbidity and phosphates. The multi-parameter dosage control strategies that include these water qualities were documented to be successful both in drinking water (Liu et al. 2013; Liu & Ratnaweera 2016) and wastewater treatment plants (VA-Support 2012; Manamperuma et al. 2013; Manamperuma et al. 2017). Some other advanced dosage control strategies are found in literature. Those are coagulant dosage control systems based on streaming current detector (Dentel et al. 1989a; 1989b; Critchley et al. 1990; Yavich & Van De Wege 2013), nephelometric turbidity measurements (Cheng et al. 2008; Yu et al. 2017), zeta potential

measurements (Sharp et al. 2005; Sharp et al. 2016) and fuzzy logic control (Su et al. 2017). Advanced soft sensors and coagulation process control models employing artificial neural networks (ANN) have been tested in drinking water treatment plants (Baxter et al. 2002; Juntunen et al. 2013; Valentin & Denœux 2001). Despite several attempts were made to optimise the coagulant dosage control, the industry is still seeking robust, reliable and cost effective dosage control strategies, advanced monitoring and processing tools and new sensors (Ratnaweera 2004; Ratnaweera & Fettig 2015).

Among the different methods of flocs characterisation image analysis is a promising tool. Image analysis was extensively used for particles characterisation such as fractal dimension, floc size, strength and breakage (Jarvis et al. 2005). Fractal dimension and floc size appear to be the most known and conventional parameters used to monitor the coagulation-flocculation process (Chakraborti et al. 2000; Chakraborti et al. 2003; Bouyer et al. 2005; He et al. 2012; Vlieghe et al. 2014; Xu et al. 2016; Yu 2014). Even though fractal geometry and particles characterisation by image analysis have gained great popularity among researchers, there were not many identifieable attempts towards applying the method for online coagulant dosage control. One of a few examples include research on flocculation control based on the fractal dimension of flocs in the pilot scale of drinking water coagulation (Chang et al. 2005). However, the technique was not yet applied in full-scale.

We have found some challenges applying the particles characterisation methods of image analysis during our previous research (Sivchenko et al. 2014). In cases of highly contaminated waters, such as wastewater, it was found that the flocs overlap in the images making it inappropriate to apply particles' detection methods. It was documented that the texture image analysis techniques, that are successfully used in other research fields, are applicable to evaluate and characterise the images of flocs (Sivchenko et al. 2016). Particularly, grey level co-occurrence matrix (GLCM) (Haralick 1979) texture analysis method was found to be appropriate for the characterisation of images of wastewater flocs and a potential tool for the coagulant dosages and treatment efficiencies predictions. The technique has been successfully tested in a municipal wastewater treatment plant (Sivchenko et al. 2017). The low-cost floc sensor prototype has been developed and tested at the same wastewater plant (Sivchenko et al. n.d.).

The aim of this research was to study the relationships between coagulation conditions (initial wastewater parameters, coagulant type and dosages) and the textural characteristics of flocs images associated with those conditions. Particularly, the goal was to find the relationships between the coagulation conditions and GLCM textural features depending on initial model wastewater parameters. In addition, the relationships between the initial model wastewater conditions and coagulant dosages depending on the coagulant type were studied.

2 Materials and methods

2.1 Model wastewater

The laboratory experiments were conducted using an adapted jar test procedure and different types of model wastewater. The model wastewater contained both organic and inorganic components to represent typical domestic wastewater characteristics important for the coagulation processes. Model wastewater was prepared according to procedures from previous studies (Fettig et al. 1990; Ødegaard et al. 1990; Ratnaweera 1991). Model soft waters were prepared to simulate Norwegian wastewaters, while model hard waters represent the European wastewaters. The initial recipe table for soft waters was expanded to better simulate changes in inlet wastewater characteristics. Table 1 represents the composition of model wastewaters used for the research.

Table 1. Composition of model wastewaters

Component	Concentration level								
	Soft water*						Hard water**		
	LS1	LS2	LS3	MS1	MS2	HS	LH	MH	HH
NaHCO$_3$, mg/l	60	60	60	60	60	60	400	400	400
NaCl, mg/l	400	400	400	400	400	400	0	0	0
NH$_4$Cl, mg/l	50	50	50	100	100	200	50	100	200
K$_2$HPO$_4$, mg/l	12.5	25	30	35	50	100	25	50	100
Humic acid, mg/l	2.5	2.5	2.5	5	5	10	2.5	5	10
Dry milk, mg/l	150	150	150	300	300	600	150	300	600
Potato starch, mg/l	30	30	30	60	60	120	30	60	120
Bentonite, mg/l	20	40	50	70	80	160	40	80	160
CaCl$_2$, mg/l	0	0	0	0	0	0	255	255	255

* *LS* – low concentration soft water, *MS* – medium concentration soft water, *HS* – high concentration soft water;
** *LH* – low concentration hard water, *MH* – medium concentration hard water, *HH* – high concentration hard water.

Initially, three contaminants concentrations were used to prepare model wastewaters of soft type – low (*LS2*), medium (*MS2*) and high (*HS*). However, the concentrations of K$_2$HPO$_4$ and Bentonite were varied to simulate changes in inlet wastewater characteristics typical for domestic WWTPs. Hence, the table was expanded to include *LS1*, *LS3* and *MS1*.

The resulting concentrations and characteristics of different model wastewater types are given in table 2.

Table 2. Concentration and parameter measurements of different model wastewaters

Parameter	Concentration level								
	Soft water						Hard water		
	LS1	LS2	LS3	MS1	MS2	HS	LH	MH	HH
pH	8.00± 0.00	7.87± 0.19	8.00± 0.00	8.00± 0.00	7.89± 0.15	7.81± 0.05	7.98± 0.07	8.01± 0.04	7.95± 0.03
TSS, mg/l	98±2	128±22	138 ±2	234±12	255± 64	507±35	110±10	240±30	380±100
Turbidity, NTU	61±2	66±4	79±3	150±0	156±3	352±5	114±11	264±14	558±42
Ortho-P, mg-P/l	2.40± 0.31	4.95± 0.28	6.30± 0.48	7.15± 0.23	10.65± 0.77	21.32± 0.89	5.10± 0.15	10.17± 0.67	17.64± 0.04
Total P, mg-P/l	3.18± 0.37	5.34± 0.47	6.29± 0.69	8.27± 0.23	12.32± 1.52	22.99± 0.98	5.67± 0.05	10.7± 0.05	20.9± 0.05

The coagulants used during experiments were from Kemira Chemicals AS. The used coagulants and their characteristics are given in table 3. The optimal dosages of each coagulant for each model wastewater type were determined by the preliminary jar tests experiments. Optimal dosages were evaluated regarding total suspended solids (TSS), turbidity, ortho-phosphates (ortho-P) and total phosphorous (total P) removal. Five dosages, which include under-, near- and over-optimum coagulation conditions for each model wastewater type, estimated during preliminary tests, were used during the experiments with at least two replicates: 0.25, 0.5, 0.75, 1.0, 2.0 mmol Me/l.

Table 3. Properties of used coagulants

Coagulant Kemira Chemicals AS	Short code	OH:Me	% Me by weight	Density, kg/m³
Aluminium sulphate	ALS	0	4.0±0.3	1280±50
Prepolymerised Polyaluminium chloride	PAX-18	1.1	9.0±0.3	1360±20
Ferric Sulphate	PIX-313	0	11.4±0.3	1550±20

Flocculator 2000 from Kemira Chemicals AS with programmable mixer units and 1 litre beakers were used for the coagulation experiments in jar test scale. The mixing conditions during coagulation: 1 min rapid mixing (400 RPM), 10 min slow mixing (30 RPM) followed by 20 min of sedimentation without mixing.

The constant pH values 7±0.05 during the slow mixing stage were maintained by addition of 2 mol/l NaOH.

After sedimentation, the supernatant was analysed for TSS, turbidity, ortho-P, total P, and for some samples zeta-potential (ζ-potential) was measured. TSS measurements were done by filtering water samples, followed by drying at temperature 105 ºC for 1 hour (ISO 11923:1997). Turbidity was measured by portable turbidimeter 2100Q (Hach®, USA). Ortho-P and total P were measured by EasyChem analyzer (Systea S.p.A, Italy). ζ-potential was measured by Zetasizer Nano (Malvern Instruments Ltd, UK).

2.2 Image acquisition and analysis

One of the Flocculator 2000 mixers was manually equipped with a 4.5 cm width black curved plastic plate, which served as a dark background for the images of flocs. A schematic representation of the whole installation is shown in figure 1.

The images were acquired during the slow mixing coagulation stage every 20 sec using digital single lens reflex (DSLR) D600 camera (Nikon®, Japan) with 105 mm Nikkor AF-S Micro 1:2.8 G ED lens (Nikon®, Japan). Settings of the camera: ISO 3200, aperture F/13, shutter speed 1/125 sec. The camera was remotely controlled from a PC via *DigiCamControl 1.2.0* software – images were collected and automatically saved on a separate hard-drive.
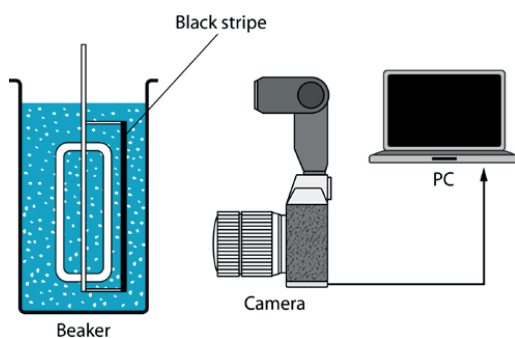

Figure 1. Schematic representation of image acquisition system

The resolution of images obtained during coagulation is 24.3 megapixels each. All images were processed by the open source image analysis software *ImageJ v.1.51* (Rasband, 1997/2016) that bases on plugins and macros. The images were combined in stacks and converted from RGB format to 8-bit grey-scale images. For each image 2610×2610 pixels

(3×3 cm) area was cropped by manual investigation of the area. All images in stacks were preprocessed by mean center of the grey-tone values in order to have the same brightness intensities.

ImageJ plugin "GLCM Texture Too" v. 0.008 by Toby C. Cornish, was used to calculate the GLCM feature vectors Contrast, Entropy, Homogeneity and Variance. Based on the previous studies (Sivchenko, Kvaal, & Ratnaweera, 2016) the distance between the pixel pairs was 1 pixel and the angle was 0°. The detailed description and calculation of the GLCM feature vectors is available in literature (Conners, Trivedi, & Harlow, 1984; Haralick, 1979; Zheng, Sun, & Zheng, 2006).

2.3 Statistical analysis

Principal component analysis (PCA) was performed to find the relationships between water quality parameters and GLCM feature vectors. PCA is a statistical data analysis technique to reduce the dimensionality of the data set, overview and describe the interrelationships among variables and to find so-called hidden structures in the data. The statistical software *The Unscrambler® X 10.4.1* (CAMO Software AS, Norway) was used to perform PCA.

3 Results and discussion

Fig. 2 represents the comparison between treatment efficiencies of different model wastewater types depending on dosages of three coagulants. Fig. 2a, 2c and 2e show the differences of TSS removal efficiencies depending on coagulant and its dosage for model wastewater types *LS2*, *MS2* and *HS*, respectively. Fig. 2b, 2d and 2f show the differences of total P removal efficiencies depending on coagulant and its dosage for model wastewater types *LS2*, *MS2* and *HS*, respectively. The presented bar charts include the removal efficiencies of TSS and total P, however, the removal efficiencies for turbidity and ortho-P were also calculated. Turbidity follows the TSS removal pattern while ortho-P is proportional to total P.

The higher removal rates were usually observed for particles than phosphates. Overall, the results indicate that the particles in terms of TSS and turbidity are better removed with a pre-polymerised coagulant PAX-18. Ortho-P and total P have higher removal rates with non-polymerised ALS and PIX-313. These results are consistent with previous findings (Ratnaweera 1991; Manamperuma et al. 2016).

The removal efficiencies of TSS and total P are decreasing with the increase of particles and phosphates in model wastewater, it is especially visible for the low coagulant dosages.

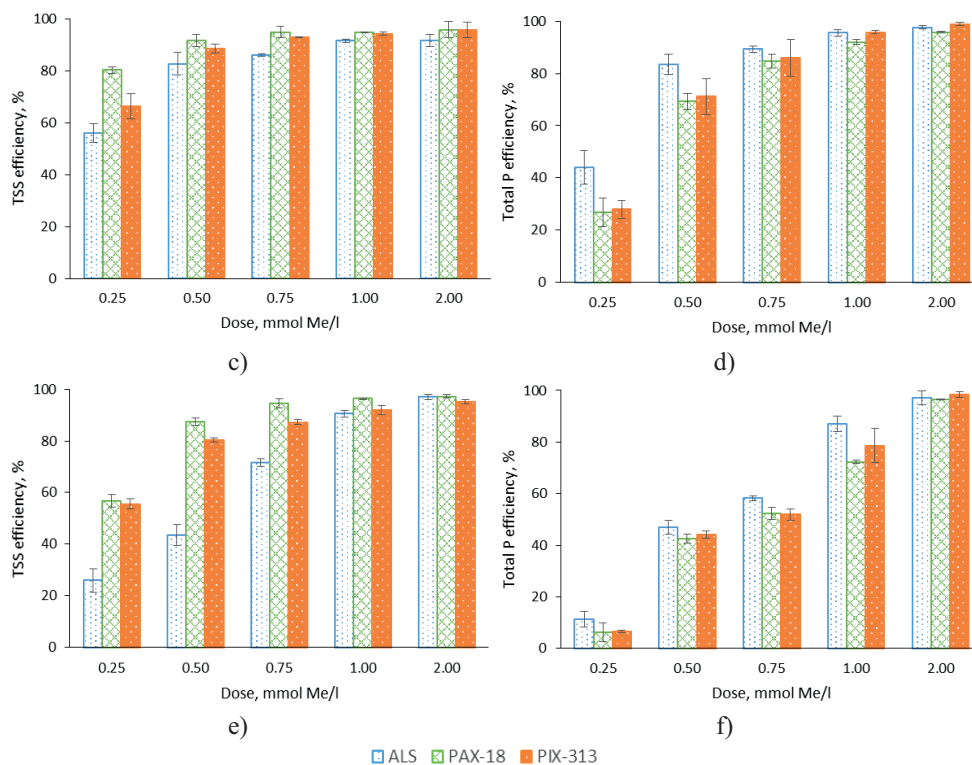a)   b)
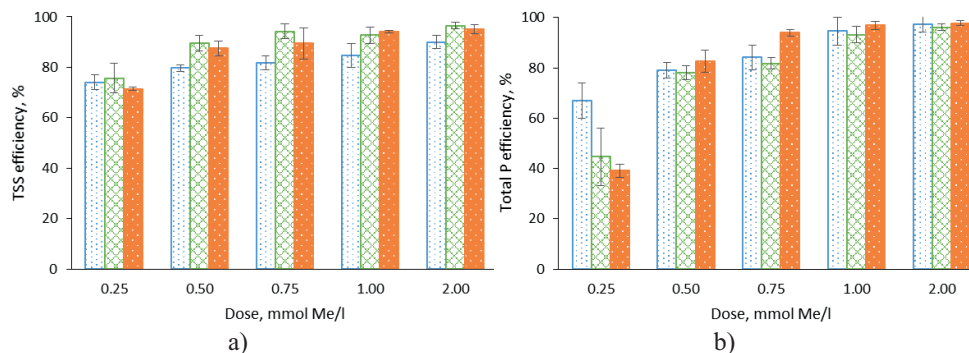
Figure 2. Comparison of treatment efficiencies by three coagulants: a) TSS for LS2; b) total P for LS2; c) TSS for MS2; d) total P for MS2; e) TSS for HS; f) total P for HS.

Fig. 3 represents the comparison of TSS (fig. 3a) and total P (fig. 3b) removal efficiencies by 3 coagulants for hard model wastewater with medium concentration of contaminants (*MH*). In comparison with soft water type *MS2* the higher removal rates are achieved with lower dosages of all coagulants. This can be explained by the Ca-salts present in hard water types. The effect of pre-polymerisation of the coagulant is especially visible with the low dosages of coagulants. The general trend of coagulants performance of the other two hard model wastewater types are similar to the one described for model wastewater type *MH*.
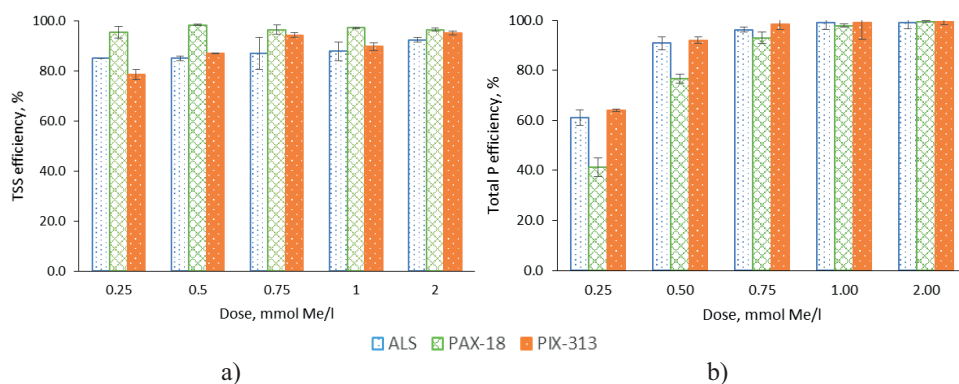
Figure 3. Comparison of treatment efficiencies by three coagulants: a) TSS for MH; b) total P for MH.

Results of the principal component analysis (PCA) based on data from jar tests with all soft model wastewater types and aluminium sulphate as a coagulant are shown in the scores plot (4a) and loadings plot (4b). The data set also included information from the flocs images that were captured during the slow mixing stage of the coagulation process and converted to the GLCM feature vectors Contrast, Entropy, Variance and Homogeneity. Two principal components (PCs) explain the data variation by 82%. PC1 is positively associated with inlet and outlet model wastewater parameters and textural feature Homogeneity; and negatively associated with zeta potential and two textural features from the images of flocs – Contrast and Entropy. PC2 is positively associated with textural feature Variance and negatively associated with the coagulant dosage.

According to the loadings plot, all initial model wastewater parameters are correlated, such as inlet turbidity (TUI), inlet total suspended solids (TSSI), inlet total P (TPI), inlet ortho-P (OPI). Outlet model wastewater parameters are also correlated - outlet turbidity (TUO), outlet total suspended solids (TSSO), outlet total P (TPO), outlet ortho-P (OPO). In addition, the outlet values of water qualities are positively correlated with Homogeneity – feature vector calculated from the images of flocs. It means the higher values of outlet water qualities the higher rate of Homogeneity is observed in the associated image of flocs. Contrast and Entropy parameters retrieved from the images are positively correlated with zeta potential and negatively correlated with inlet and outlet model wastewater parameters. The dosage of the coagulant is negatively correlated with Variance on the image.

Model wastewaters with high levels of contaminants (*HS*) have high Homogeneity values in the images of flocs. The flocs formed under such conditions are visually smaller, have higher quantity in the beaker, thus overlap in the images. Higher values of Variance on the images are observed for the lower coagulant dosages. Model wastewaters with low levels of contaminants (*LS*) are positively correlated with Entropy and Contrast in the associated images of flocs.
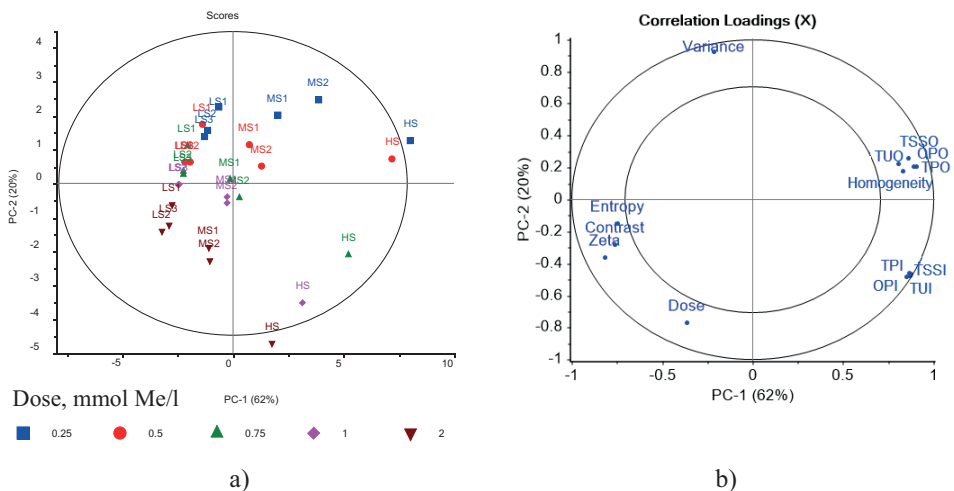

Figure 4. Results of PCA for ALS coagulant and all types of soft model wastewaters: a) scores plot; b) loadings plot.

Conclusions

The textural structures on images of flocs change depending on coagulant type, dosage and corresponding treatment efficiency.

Texture image analysis method GLCM adequately quantifies the textural characteristics of flocs images by feature vectors Contrast, Entropy, Homogeneity and Variance.

Contrast and Entropy of the flocs images are positively correlated with zeta potential measurements. Homogeneity of the image is positively correlated with outlet (after sedimentation) model wastewater qualities. Feature vector Variance is negatively associated with the coagulant dosage.

The images of flocs, captured during coagulation of model wastewaters with high initial levels of particles and phosphates, are more homogeneous compare to the images of other model wastewater types. Thus, have higher calculated values of feature vector Homogeneity. The images of flocs associated with low levels of contaminants have higher values of Contrast and Entropy.

The images of flocs associated with low (under-optimum) dosages of coagulant have higher values of feature vector Variance retrieved from the images.

Pre-polymerised coagulant has higher removal efficiencies of particles for all dosages and model wastewater types compared to inorganic Al- and Fe-based coagulants. However, non-polymerised metal salts achieve better treatment efficiencies of total and ortho-phosphates.

Concluding, GLCM texture image analysis method and statistical methods are the way to correlate images of flocs with coagulant dosages and treatment efficiencies, which enables optimal dosage prediction.

Acknowledgment

References

Baxter, C.W. et al., 2002. Model-based advanced process control of coagulation. *Water science and technology*, 45(4–5), pp.9–17.

Bouyer, D. et al., 2005. Experimental analysis of floc size distributions in a 1-L jar under different hydrodynamics and physicochemical conditions. *Journal of colloid and interface science*, 292(2), pp.413–28.

Bratby, J., 2016. *Coagulation and flocculation in water and wastewater treatment*, IWA Publishing.

Chakraborti, R.K. et al., 2003. Changes in fractal dimension during aggregation. *Water research*, 37(4), pp.873–883.

Chakraborti, R.K., Atkinson, J.F. & Van Benschoten, J.E., 2000. Characterization of Alum Floc by Image Analysis. *Environmental Science & Technology*, 34(18), pp.3969–3976.

Chang, Y., Liu, Q.-J. & Zhang, J.-S., 2005. Flocculation control study based on fractal theory. *Journal of Zhejiang University Science B*, 6(10), pp.1038–1044.

Cheng, W.P., Kao, Y.P. & Yu, R.F., 2008. A novel method for on-line evaluation of floc size in coagulation process. *Water research*, 42(10–11), pp.2691–7.

Conners, R.W., Trivedi, M.M. & Harlow, C.A., 1984. Segmentation of a high-resolution urban scene using texture operators. *Computer Vision, Graphics, and Image Processing*, 25(3), pp.273–310.

Critchley, R., Smith, E. & Pettit, P., 1990. Automatic Coagulation Control at Water-Treatment Plants in the North-West Region of England. *Water and Environment*, 4(6), pp.535–543.

Dentel, S.K., Thomas, A. V. & Kingery, K.M., 1989a. Evaluation of the streaming current detector-I. Use in jar tests. *Water Research*, 23(4), pp.413–421.

Dentel, S.K., Thomas, A. V. & Kingery, K.M., 1989b. Evaluation of the streaming current detector-II. Continuous flow tests. *Water Research*, 23(4), pp.423–430.

Fettig, J., Ratnaweera, H. & Ødegaard, H., 1990. Simultaneous phosphate precipitation and particle destabilization using aluminium coagulants of different basicity H. H. Hahn & R. Klute, eds. *Chemical Water and Wastewater Treatment. Proceedings of the 4th Gothenburg Symposium 1990 October 1–3, 1990 Madrid, Spain*, pp.221–242.

Haralick, R.M., 1979. Statistical and structural approaches to texture. *Proceedings of the IEEE*, 67(5), pp.786–804.

He, W. et al., 2012. Characteristic analysis on temporal evolution of floc size and structure in low-shear flow. *Water research*, 46(2), pp.509–20.

Jarvis, P. et al., 2005. A review of floc strength and breackage. *Water Research*, 39(14), pp.3121–3137.

Juntunen, P. et al., 2013. Dynamic soft sensors for detecting factors affecting turbidity in drinking water. *Journal of Hydroinformatics*, 15(2), pp.416–426.

Li, B. & Stenstrom, M.K., 2014. Research advances and challenges in one-dimensional modeling of secondary settling Tanks – A critical review. *Water Research*, 65, pp.40–63.

Liu, W. & Ratnaweera, H., 2016. Improvement of multi-parameter-based feed-forward coagulant dosing control systems with feed-back functionalities. *Water Science and Technology*, 74(2), pp.491–499.

Liu, W., Ratnaweera, H. & Song, H., 2013. Better treatment efficiencies and process economies with real-time coagulant dosing control. In *11th IWA Conference on Instrumentation Control and Automation*. Narbonne, France, 18–20 September 2013.

Manamperuma, L., Ratnaweera, H.C. & Rathnaweera, S.S., 2013. Retrofitting coagulant dosing control using real-time water quality measurements to reduce coagulant consumption. In *11th IWA Conference on Instrumentation Control and Automation*. Narbonne, France, 18–20 September 2013.

Manamperuma, L., Wei, L. & Ratnaweera, H., 2017. Multi-parameter based coagulant dosing control. *Water Science and Technology*, 75(9), pp.2157–2162.

Manamperuma, L.D., Ratnaweera, H.C. & Martsul, A., 2016. Mechanisms during suspended solids and phosphate concentration variations in wastewater coagulation process.

Environmental Technology (United Kingdom), 37(19), pp.2405–2413.

Ødegaard, H., Fettig, J. & Ratnaweera, H., 1990. Coagulation with prepolymerized metal salts H. H. Hahn & R. Klute, eds. *Chemical water and wastewater treatment. Proceedings of the 4th Gothenburg Symposium 1990 October 1–3, 1990 Madrid, Spain*, 304, pp.189–220.

Rasband, W.S., ImageJ. *U. S. National Institutes of Health, Bethesda, Maryland, USA, 1997-2016*. Available at: http://imagej.nih.gov/ij/.

Ratnaweera, H., 2004. Coagulant dosing control - a review. *Chemical water and wastewater treatment, IWA publishing, London*, 7, pp.3–13.

Ratnaweera, H., 1991. *Influence of the degree of coagulant prepolymerization on wastewater coagulation mechanisms*. PhD. Norwegian Institute of Technology, Trondheim.

Ratnaweera, H. & Fettig, J., 2015. State of the art of online monitoring and control of the coagulation process. *Water*, 7(11), pp.6574–6597.

Sharp, E. et al., 2005. Application of zeta potential measurements for coagulation control: pilot-plant experiences from UK and US waters with elevated organics. *Water Science and Technology: Water Supply*, 5(5), pp.49–56.

Sharp, E. et al., 2016. Using online zeta potential measurements for full-scale coagulation control in drinking water treatment. In *IWA Specialist Conference, Advances in particle science and separation: meeting tomorrow's challenges*. Oslo, Norway, 22-24 June 2016.

Sivchenko, N., Kvaal, K. & Ratnaweera, H., 2016. Evaluation of image texture recognition techniques in application to wastewater coagulation Y. F. Tsang, ed. *Cogent Engineering*, 3(1).

Sivchenko, N., Kvaal, K. & Ratnaweera, H., Floc sensor prototype tested in the municipal wastewater treatment plant. *Cogent Engineering*.

Sivchenko, N., Kvaal, K. & Ratnaweera, H., 2014. Image analysis of flocs and mathematical modelling applied to coagulation-flocculation process. In *Poster, IWA Specialist Conference, Advances in particle science and separation: from mm to nm scale and beyond.* Sapporo, Japan, 15-18 June 2014.

Sivchenko, N., Ratnaweera, H. & Kvaal, K., 2017. Approbation of the texture analysis imaging technique in the wastewater treatment plant. *Cogent Engineering*, 4(1).

Su, X., Xu, S. & Xu, S., 2017. Compound control system for coagulant dosing process based on a fuzzy cerebellar model articulation controller. In *Proceedings of the 36th Chinese Control Conference*. Dalian, China, 26-28 July 2017, pp. 3931–3936.

Thomas, D.N., Judd, S.J. & Fawcett, N., 1999. Flocculation modelling: a review. *Water Research*, 33(7), pp.1579–1592.

VA-Support, 2012. *DOSCON - Experience at NRA. (03.07.2012)*

Valentin, N. & Denœux, T., 2001. A neural network-based software sensor for coagulation control in a water treatment plant. *Intelligent Data Analysis*, 5, pp.23–39.

Vlieghe, M. et al., 2014. In situ characterization of floc morphology by image analysis in a turbulent Taylor-couette reactor. *American Institute of Chemical Engineers*, 60(7), pp.2389–2403.

Xu, Y. et al., 2016. Effect of reused alum-humic-flocs on coagulation performance and floc characteristics formed by aluminum salt coagulants in humic-acid water. *Chemical Engineering Journal*, 287, pp.225–232.

Yavich, A.A. & Van De Wege, J., 2013. Chemical feed control using coagulation computer models and a streaming current detector. *Water Science & Technology*, 67(12), p.2814.

Yu, R., Chen, H. & Cheng, W., 2017. Applying online image analysis to simultaneously evaluate the removals of suspended solids and color from textile wastewater in chemical flocculated sedimentation. *Journal of Environmental Informatics*, 29(1), pp.29–38.

Yu, R.F., 2014. On-line evaluating the SS removals for chemical coagulation using digital image analysis and artificial neural networks. *International Journal of Environmental Science and Technology*, 11(7), pp.1817–1826.

Zheng, C., Sun, D.-W. & Zheng, L., 2006. Recent applications of image texture for evaluation of food qualities—a review. *Trends in Food Science & Technology*, 17(3), pp.113–128.

Paper III

Sivchenko, N., Ratnaweera, H., & Kvaal, K. (2017). Approbation of the texture analysis imaging technique in the wastewater treatment plant. *Cogent Engineering*, *4*(1), 1373416. http://doi.org/10.1080/23311916.2017.1373416

:: cogent
engineering

CrossMark

*Corresponding author: N. Sivchenko, Department of Mathematical Sciences and Technology, Norwegian University of Life Sciences, P.O. Box 5003-IMT, 1432 Aas, Norway
E-mails: nataliia.sivchenko@nmbu.no, nataliia.sivchenko@gmail.com

CIVIL & ENVIRONMENTAL ENGINEERING | RESEARCH ARTICLE

Approbation of the texture analysis imaging technique in the wastewater treatment plant

N. Sivchenko[1]*, H. Ratnaweera[1] and K. Kvaal[1]

Abstract: Monitoring of effluent turbidity is essential to evaluate the coagulation process in wastewater treatment plants (WWTPs). A digital imaging system based on the texture analysis of flocs has been tested in a Norwegian municipal WWTP to predict changes in coagulation conditions and outlet turbidity. Principal component analysis (PCA) was applied to prove that the textural features of flocs' images depend on the inlet wastewater parameters and coagulation conditions. Partial least squares regression (PLSR) was performed for the outlet turbidity prediction. The best model resulted in 86.6% prediction accuracy using two wastewater quality parameters (inlet flow and inlet turbidity) and 4 textural feature vectors retrieved from the images of flocs. Furthermore, the outlet turbidity predicted by this model resulted in a lower amount of underestimated values compare to the model, which contained only wastewater quality parameters. A new term—floc texture index (FTI) summarizes the textural features of flocs' images resulting in a single variable (thought linear combination). This further simplifies the current multivariate dosage control system. Analysing the plant data indicates that an FTI value below 6 would result in the outlet turbidity values above 5. This can be used as an early warning system of coagulation failure. The results of these studies demonstrate the potential of the digital imaging system to improve an existing online coagulant dosing control strategy.

ABOUT THE AUTHORS

The authors are members of Water, Environment, Sanitation and Health (WESH) research group at the Norwegian University of Life Sciences (NMBU), Faculty of Sciences and Technology (RealTek). WESH group focuses on water and wastewater related issues and is heavily involved in teaching and supervision of MSc and PhD programs in water engineering and technology. The group also has one of the largest externally funded research and educational project portfolios at RealTek, and collaborates with partners from EU, North America, Eurasia, Asia and Africa. It also has a number of Research, Development and Innovation projects with Norwegian partners. The main research and development areas of the WESH group are: process control and optimization of coagulation and biological treatment processes; membrane fouling and filtration processes; microbial water quality and risk assessment; decentralized wastewater systems; modelling of sewer systems.

PUBLIC INTEREST STATEMENT

Last decades researchers working with water and wastewater treatment often try to benefit from applying image analysis based systems to the different processes. Doing this we are able to monitor or even properly control these processes, for instance wastewater coagulation, which is often difficult otherwise. In this paper, we are testing how the digital imaging system based on texture analysis of particles (flocs) performs in a full-scale Norwegian wastewater treatment plant. Retrieving the textural parameters from the images of flocs obtained during different coagulation conditions in the plant, we were able to indicate the changes in wastewater inlet and outlet quality parameters. The prediction of outlet wastewater quality parameters (turbidity in this case) is necessary to provide efficient coagulant dosage control and/or develop the coagulation failure alarm system. The results of these studies are to be used to improve the existing online coagulant dosing control system.

:: cogent • oa

Subjects: Process Control - Chemical Engineering; Waste & Recycling; Water Engineering; Water Science; Image Processing

Keywords: coagulation; wastewater treatment plant; image analysis; texture analysis; effluent turbidity prediction

1. Introduction

Coagulation is a well-known and widely used water and wastewater treatment method to remove suspended solids, phosphates and other water impurities. Particles which aggregate during the coagulation/flocculation process are called flocs. The amount of coagulant (chemical material, addition of which results in suspension's destabilisation), its nature, concentration and mixing conditions are the main factors influencing flocs aggregation and breakage. The size and surface properties of flocs formed under different coagulation mechanisms strongly influence their behaviour during further solid-liquid separation process (Bache & Gregory, 2007). Many researchers have studied different floc features, such as size distribution and fractal dimension (Chakraborti, Atkinson, & Van Benschoten, 2000; Li, Zhu, Wang, Yao, & Tang, 2006; Vahedi & Gorczyca, 2011). Some attempts were done to create online systems for flocs characterisation and monitoring. A method of online floc size evaluation based on nephelometric turbidity measurements was presented (Cheng, Kao, & Yu, 2008) and further developed into nephelometric turbidimeter monitoring system (NTMS) (Cheng, Chang, Chen, Yu, & Huang, 2011; Yu, Chen, & Cheng, 2017). Photometric dispersion analyser (PDA) is an optical instrument often used to study the aggregates' characteristics (Burgess, Curley, Wiseman, & Xiao, 2002; Chou, Lin, & Huang, 1998; Wu, Wang, Hu, & Ye, 2013). Nevertheless, many attempts were done so far to study, characterise and control particles aggregating during coagulation, there is still a gap in the application knowledge of how to use the floc features to optimize the coagulation process, provide cheap and robust dosage control system.

Nowadays the flow-proportional dosing concept is usually used for coagulation dosage control, while process optimisation often bases on results from the jar tests and the operator's experience (Ratnaweera & Fettig, 2015). Streaming current detector (SCD) was evaluated (Dentel, Thomas, & Kingery, 1989a, 1989b) and tested in the drinking water treatment plants (DWTP) for automatic coagulation control (Critchley, Smith, & Pettit, 1990; Yavich & Van De Wege, 2013). The coagulant dosage strategy based on zeta potential measurements documented to be a promising control technique (Sharp et al., 2005, 2016). Advanced soft sensors and coagulation process control models employing artificial neural networks (ANN) have been tested in DWTPs (Baxter et al., 2002; Juntunen, Liukkonen, Lehtola, & Hiltunen, 2013; Valentin & Denœux, 2001). Advanced dosing control systems based on multiple water quality parameters that could be measured online have confirmed to be successful both in DWTPs (Liu, 2016; Liu, Ratnaweera, & Song, 2013) and wastewater treatment plants (WWTPs) (VA-Support, 2012). Application of such systems enables a reduction of coagulant consumption (i.e. minimise the operational costs), reduces the sludge volumes and maintains the desired removal of particles and phosphates (Manamperuma, Ratnaweera, & Rathnaweera, 2013; Manamperuma, Wei, & Ratnaweera, 2017). With the growing need of wastewater treatment processes optimisation, the need for further development of intelligent, accurate and reliable online dosing control systems arises (Ratnaweera & Fettig, 2015).

Last decades digital image analysis (DIA) techniques have often been employed to determine particles characteristics, such as floc size, fractal dimension, strength and breakage (Jarvis, Jefferson, Gregory, & Parsons, 2005). Fractal dimension and floc size appear to be the most known and conventional parameters used in numerous attempts to monitor and/or control the coagulation process by DIA (Bouyer, Coufort, Liné, & Do-Quang, 2005; Chang, Liu, & Zhang, 2005; He, Nan, Li, & Li, 2012; Vlieghe, Coufort-Saudejaud, Frances, & Liné, 2014; Xu, Chen, Cui, & Shi, 2016; Yu, 2014). Other floc parameters retrieved from the images (average equivalent diameter, eccentricity and colour index) are also essential, for instance, in combination with self-organizing maps to monitor changes during the coagulation process and estimate flocs quality (Juntunen, Liukkonen, Lehtola, & Hiltunen, 2014).

In most cases the applicability of particles detection methods in wastewater coagulation process has been limited to lab scale due to complicated and inaccurate measurements in the field, hardware and software limitations. We propose a new approach to image analysis of flocs, which was previously tested on the laboratory scale batch tests (jar tests) (Sivchenko, Kvaal, & Ratnaweera, 2014, 2016). It is a comparatively easy method of image characterization, which bases on analysis of the whole image texture instead of concerning the shape characteristics of each particle in the image. Such approach simplifies the image analysis stage, e.g. no need of particles extraction/segmentation from the image and its count which found to be problematic in wastewater applications (Sivchenko et al., 2016). Hence, it gives significant simplification of the software to be developed. Furthermore, out of focus flocs, which are often present in the images, are not a problem for this texture complexity recognition method. Thus, potentially the cheap cameras could be employed for the floc sensor development. This paper presents the applicability of the concept in continuous mode with real wastewater in the context to use it as a dosage control technique for optimizing the coagulation process.

2. Experimental setup and methods

2.1. Wastewater treatment plant

Full-scale tests were conducted in the Frogn wastewater treatment plant (Drøbak, Norway) in September and October 2015. Frogn WWTP receives municipal wastewater from Drøbak city and the neighbourhood area. Average inlet flow is 4,600 m³/day during the days without snowmelt and/or precipitations. The tests period include days when the long precipitation period took place. The maximum flow rate to the plant reached 18,845 m³/day on the 18 September 2015.

Frogn WWTP is a mechanical-chemical precipitation plant. The treatment process consists of the next stages: screens, two parallel pre-sedimentation basins, three sequenced coagulation chambers with the different velocity gradients, and two parallel sedimentation chambers. The plant also has the sludge dewatering and thickening system. The inlet and outlet water quality parameters are measured by online sensors and recorded (average values) with the 15 min interval. The data is available for observation in the plant's SCADA system. The retrieved water parameters included inlet wastewater flow (QIN), inlet pH (PHI), inlet turbidity (TUI), inlet conductivity (CNI), coagulant dosage (Dose), pH after coagulant dosage (PHO) and outlet turbidity (TUO). The plant operators perform daily sampling of inlet and outlet total Phosphorous (total *P*). The summarised data of water quality parameters for the tests period is given in Table 1. Coagulant used in the Frogn WWTP is polyaluminium chloride (ECOFLOCK 90, Feralco), 9 ± 0.3% Al by weight, and density 1,356 ± 25 kg/m³.

Table 1. Inlet and outlet wastewater quality parameters of Frogn WWTP				
Variables	**Mean**	**Min**	**Max**	**Standard deviation**
Inlet flow, m³/h	310.4	102.8	789.7	164.9
Inlet pH	7.4	6.7	8.0	0.3
Inlet turbidity, FNU	211.6	75.6	500.0	93.4
Inlet conductivity, µS/cm	758.9	400.3	1,154.5	174.7
pH after coagulant dosage	6.9	6.5	7.3	0.2
Outlet turbidity, FNU	3.8	1.1	21.1	4.3
Inlet total *P*, mg P/l	3.86	0.91	6.29	1.75
Outlet total *P*, mg P/l	0.18	0.10	0.29	0.05
Inlet total *P*, mmol P/l	0.12	0.03	0.20	0.06
Coagulant dose, ml/m³	83.0	49.5	196.3	24.6
Coagulant dose, mmol Al/l	0.38	0.23	0.89	0.11
Ratio mol Al: mol P	3.65	2.07	8.38	1.92

cogent ·· engineering

2.2. Image acquisition and pre-processing

A special installation was designed to observe changes in flocs' structure *in situ*. The installation was set above the second flocculation chamber and consisted of the tube, peristaltic pump, acrylic cell for image acquisition, digital camera and computer (Figure 1). To minimise the potential danger of flocs breakage, the chosen tube was 3 cm in diameter, and the peristaltic pump was placed after the imaging cell. The water flow in the system was upstream and manually adjusted to approximately 40 l/h.

Images of flocs were constantly taken with a pre-set repeatability using free remote camera control software—*DigiCamControl 1.2.0*. Image capturing equipment used during the investigations was as follows: digital single lens reflex (DSLR) Nikon D600 camera, 105 mm Nikkor AF-S Micro 1:2.8 G ED lens (Nikkor, China), SpeedLite YN460 flash (Yongnuo, China). The size of the image-capturing zone in the cuvette was 3.3 × 10.3 cm. In order to obtain flocs with the proper depth of field, the black metal stripe was placed in the centre of the cuvette, which also became a background for the flocs. The choice of the background colour was based on the fact that the wastewater flocs are greyish coloured. Thus, using a contrasting background, it is easier to perform the further image analysis.

The obtained images have a resolution of 24.3 megapixels each. They were processed in the open source image analysis software *ImageJ v.1.49* (Rasband, 1997/2016) that bases on plugins and macros. For each image 1380 × 3640 pixels (2.4 × 6.3 cm) area was cropped by manual investigation of the area. Because of slight changes in lighting conditions during image acquisition, all images were pre-processed in order to have the same brightness intensity. We wrote *ImageJ* plugin to mean centre the images' grey-tone values.

2.3. Image analysis by Grey level co-occurrence matrix (GLCM)

GLCM is a common method for the image texture measurement. Previously it was successfully applied in the laboratory scale data (Sivchenko et al., 2016). The GLCM method was chosen among the other texture analysis methods because it is quite simply computable and does not require heavy programming for the sensor prototype to be developed.

ImageJ plugin "GLCM Texture Too" v. 0.009 was used to obtain the GLCM feature vectors. The resulting output was given as a vector of the next 9 parameters per each image: contrast, correlation, inverse difference moment (IDM), entropy, energy, homogeneity, prominence, variance, and shade. Hence, the data matrix was obtained with the size 342 × 9. The detailed description, explanation,

Figure 1. (a) Schematic representation of the installation and (b) a photograph from the Frogn WWTP.
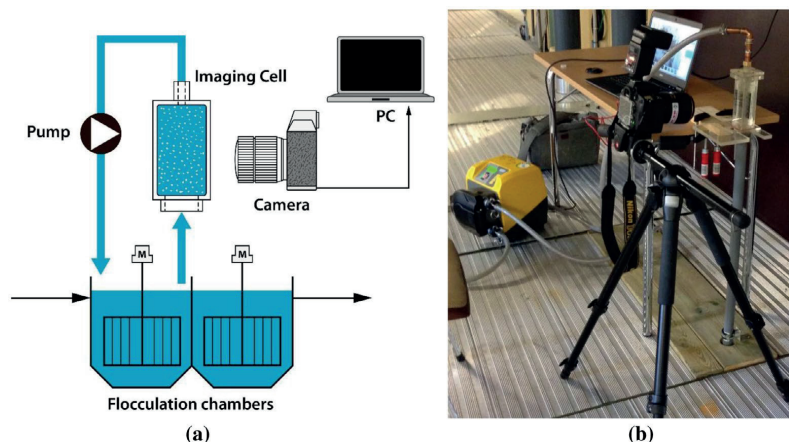
(a)　　　　　　　　　　(b)

and equations for above GLCM texture features are given elsewhere (Conners, Trivedi, & Harlow, 1984; Haralick, Shanmugam, & Dinstein, 1973; Zheng, Sun, & Zheng, 2006).

2.4. Conjugation of two data sets

Three images for each 15 min were chosen to be representative. The measured GLCM feature vectors were averaged for each 3 images and matched with the retrieved water quality parameters. According to the tracer tests conducted in Frogn WWTP, the outlet turbidity values were 45 min shifted to meet the response lag between the coagulant injection point and outlet from the sedimentation tank. After the removal of missing values and outliers, the resulting data-set contains 114 samples. The data include 81.6% of samples under the normal operation conditions, 15.8% of samples under rainy weather conditions and 2.6% of samples which had high inlet turbidity values due to the periodic discharge from the septic tanks to the inlet.

2.5. Multivariate statistical analysis and modelling

The resulted data matrix was processed in statistical software *The Unscrambler® X 10.3* (CAMO Software AS, Norway) and in *MATLAB®* using *PLS toolbox 8.2* (Eigenvector Research, Inc., USA). Principal component analysis (PCA) was performed to find the relationships between water quality parameters and GLCM feature vectors. PCA is a statistical data analysis technique to reduce the dimensionality of the data-set, overview and describe the interrelationships among variables and to find so-called hidden structures in the data. Partial least squares regression (PLSR) was performed to predict outlet turbidity based on different combinations of water quality parameters and GLCM texture features.

3. Results and discussion

3.1. Data overview

At the time when the tests were conducted, Frogn WWTP used the flow-proportional concept of coagulation dosage control with the ability of manual dose adjustment. However, now the plant has an advanced dosage control system—DOSCON® (DOSCON AS, Oslo, Norway), which uses the inlet wastewater quality parameters to calculate the optimal dose. A detailed description of the control strategy can be found elsewhere (Liu & Ratnaweera, 2017; Manamperuma et al., 2017). Although the coagulant consumption and the plant performance were significantly improved, the system still depends on the reliability of the online equipment functioning in a harsh environment.

Figure 2 shows an example of the load changes during the 9 days observation period in September 2015. The first 3 days represent the typical wastewater flow variations during the day under the normal (dry) weather conditions. On the fourth day, the rainy week started, and the wastewater load had significantly increased reaching a peak of 900 m³/h at the beginning of the ninth day. Corresponding flow-proportional coagulant dosages in µl/l are also marked in the figure. The images of exemplary floc structures appeared under different coagulation conditions are shown for the days when image analysis observations took place. Visual investigation of flocs can be described as next: the flocs formed during normal operation conditions tend to be bigger compared to those formed during the rainy days. This can be explained by the change in inlet wastewater composition, when domestic wastewater was highly diluted by stormwater. During the rain events avarage inlet flow to the plant increased 3 times, avarage inlet turbidity decreased 2–2.5 times, the average inlet conductivity of wastewaer decreased 1.3 times and average inlet pH decreased from 7.2 to 6.9. These changes in wastewater inlet parameters could result in a change of aggregates properties and size (Bache & Gregory, 2007). Even though the coagulant dosing control in the plant bases on the flow-proportional concept, the dosage of coagulant was lowered manually during the rainy days period. Relatively high turbidity values of outlet water from the sedimentation tank during the last two rainy days point out on poor coagulation conditions with the non-optimal dosages. The multivariate statistical analysis was employed to test if the images of appearing during coagulation flocs reveal the information about coagulation conditions and have relations with the wastewater characteristics, doses and effluent water quality.

cogent · engineering

Figure 2. Graph showing the load changes in the Frogn WWTP during normal days and in rainy days. The images represent the typical floc structures for those particular days when the image analysis installation was functioning. Numbers within the each image mean the corresponding coagulant dosages in μl/l.
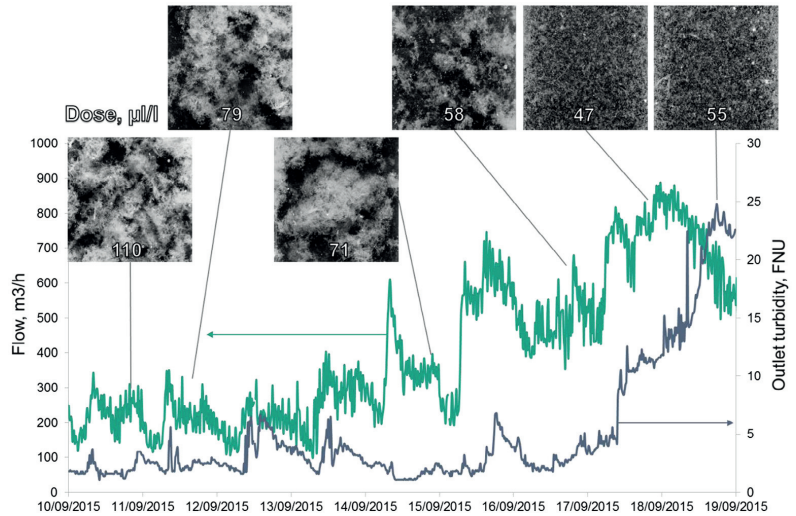
3.2. Principal component analysis

Previously obtained in the laboratory scale results, proved that the images of flocs are unique for the different water conditions and the texture analysis methods have a potential to be used for the further floc sensor development. Principal component analysis (PCA) applied to the full-scale data showed that the images, obtained from different inlet wastewater parameters and coagulation conditions, also contain unique information.

Figure 3(a) shows the results of PCA based only on 9 GLCM textural features of the images of flocs. With four principal components (PCs) the total explained variance equals 96.3% (PC1 = 57.4%, PC2 = 78.5%, PC3 = 91.5%) for calibration and 91.2% (PC1 = 49.4%, PC2 = 68.2%, PC3 = 85.2%) for cross-validation. The images of flocs by means of GLCM feature vectors separated samples on PC1 versus PC2 scores plot into three distinct classes. These classes match the wastewater coagulation conditions: normal, during the rain events and the high inlet turbidity. Thus, the images of flocs can be used as an early indication of the changes in the wastewater parameters to take further manual (or automated) troubleshooting action.

Figure 3. Scores plot of PCA, PC1 vs. PC2 for: (a) data with 9 GLCM textural features of the images of flocs; (b) data containing wastewater quality parameters, GLCM textural features and the coagulant dosage.
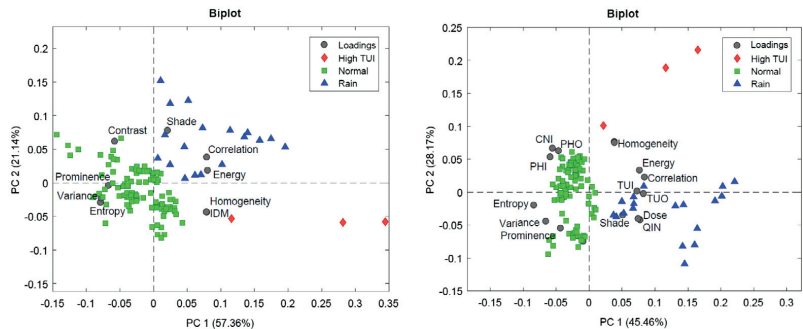
cogent · · engineering

Figure 3(b) shows the results of PCA performed for all inlet and outlet wastewater quality parameters (QIN, TUI, PHI, CNI, PHO, TUO), GLCM textural features (9 variables) and the coagulant dosage in ml/s. It was done in order to compare how different would be the scores plot from the one described above. However, no huge differences can be observed in means of samples' classes. Samples corresponding to the rain events are more stretched by PC2, which associates with high turbidity inlet and outlet, high wastewater flow and coagulant dosage. Total explained variance for calibration: PC1 = 45.5%, PC2 = 73.6%, PC3 = 83.4%, PC4 = 89.6%; for cross-validation: PC1 = 38.3%, PC2 = 66.9%, PC3 = 76.9%, PC4 = 83.3%.

3.3. Floc texture index (FTI)

Figure 3 shows the potential for the sensor prototype to be developed. However, in order to have the instrument functioning as a sensor, preferably, there should be only one signal coming out from the sensor. Thus, 9 GLCM feature vectors should be reduced to one variable. We are introducing the entirely new term—floc texture index (FTI). FTI is a sum of four GLCM feature vectors:

$$FTI = (Contrast + Entropy + Homogeneity + Variance) \times 10^{-2} \tag{1}$$

The other GLCM feature vectors were excluded from the equation, because they have very high (near 1) correlation coefficients to some of the included variables, what can be seen from the loadings in Figure 3. Since Variance has values in hundred scale, FTI was divided by 100 for simplification. The results of FTI calculation for 3 sample classes are presented in Table 2.

The calculated FTIs and corresponding observed outlet turbidity values are presented in Figure 4. Frogn WWTP aims to keep the effluent turbidity (TUO) below 5 FNU. Hence, TUO values below 5 FNU are marked as orange triangles, while the values above 5 FNU are marked as orange open triangles. The lower limit for FTI was chosen to be 6. While FTI is above 6, the corresponding TUO in most cases is below 5 FNU, and vice versa. Three events are highlighted in the figure. The first red dashed box highlights the rain event. The images of flocs quantified as FTIs showed an early indication that the outlet turbidity would exceed the maximum desired level. Even though the TUO values were shifted in the data-set to correspond the inlet wastewater quality and dosed amount of coagulant, the flow through the treatment plant is a dynamic system and not an ideal plug flow, so sometimes the time lag between a flocculation chamber and an outlet from the sedimentation tank is higher than 45 min. In such cases, the early indication of the changes in coagulation conditions by the images of flocs (FTI) is desired and a significant advantage of the planned dosage control system. The 2a and 2b red dashed boxes highlight the events when coagulant dosages were not optimal, which resulted

Table 2. FTI ranges for 3 classes in the data-set				
FTIClass	**Mean**	**Min**	**Max**	**Standard deviation**
Normal conditions	7.97	4.63	11.89	2.25
Rain events	3.09	1.23	6.05	1.47
High inlet turbidity	2.48	1.57	3.90	1.24

Figure 4. Floc texture indexes and corresponding outlet turbidity for the obtained data-set

cogent ·· engineering
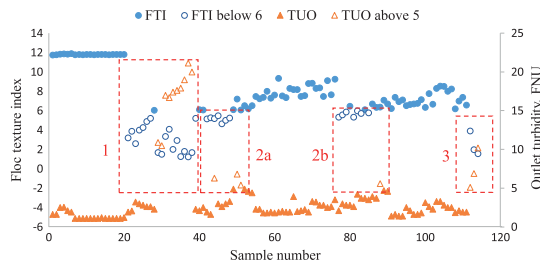
in the increase of effluent turbidity. The third event highlighted by the red dashed box corresponds to the high inlet turbidity. Overall, FTI was able to indicate all non-optimal conditions of the coagulation process in this particular data-set.

With a bigger training data-set, it is potentially possible that FTI will have lower and higher limits to indicate changes in coagulation conditions and/or take action in raising or lowering the coagulant dosage.

3.4. Effluent turbidity prediction

Different PLSR models were tested in order to get the best prediction of the effluent turbidity based on inlet wastewater quality parameters and images of flocs. Nowadays DOSCON® is using a multi-parameter based feed-forward control strategy (Manamperuma et al., 2017). To strengthen the robustness and wider applicability of the dosage control system, the soft sensor should be developed to predict the effluent turbidity. If the operator of the plant knows few hours in advance that there is a potential danger the outlet turbidity will exceed the maximum desired value, he has enough time to take action and justify the coagulant dosage. Ideally, such system is to be developed to the self-standing dosage control strategy.

The models with highest explained variances are presented in Table 3. The response (Y) for all models was the effluent turbidity (TUO), while the X-matrix consisted of different variables. The prediction by only wastewater inlet parameters (QIN, TUI, CNI, PHI, Dose) resulted in 80.5% calibration *Y* variance explained by two factors. The simplest model based on inlet parameters QIN and TUI resulted in 79.7% calibration *Y* variance explained by one factor. With the addition of image analysis results (FTI), the calibration R^2 increased until 0.84. The best prediction model for TUO, about 87% calibration *Y* variance explained by two factors, included both inlet wastewater quality parameters (QIN and TUI), and some GLCM feature vectors (Variance, Prominence, Correlation and Contrast).

Even though the addition of FTI or GLCM variables to the models resulted in 4 and 7% increase of explained *Y* variance respectively, the main advantage of the image analysis supported prediction is a better estimation of the over ranged outlet turbidity values (more than 5 FNU). Figure 5 shows a comparison of two TUO estimation models. The first model (Figure 5(a)) is based only on wastewater inlet parameters and tend to underestimate TUO, which is acceptable for the TUO values below 5 FNU and close to it. However, there are two samples with considerably high TUO (over 10 FNU), which were underestimated by the model (predicted values lower than 5 FNU). The addition of the GLCM feature vectors (Variance, Prominence, Correlation and Contrast) as predictor variables increased the efficiency of estimation and resulted in the lower amount of underestimated effluent turbidity values.

Table 3. Results of partial least squares regressions for the effluent turbidity prediction					
Model, *X* variables	Number of factors	R^2 cal.	R^2c-val.	RMSEC	RMSECV
All inlet wastewater parameters QIN, TUI, PHI, CNI, Dose	2	0.805	0.793	1.919	2.013
All inlet WWP and FTI	2	0.829	0.81	1.795	1.885
QIN, TUI	1	0.797	0.789	1.96	1.998
QIN, TUI and FTI	1	0.837	0.828	1.755	1.795
QIN, TUI, Dose and FTI	2	0.841	0.812	1.732	1.885
QIN, TUI, Variance, Prominence, Correlation, Contrast	2	0.866	0.851	1.594	1.666

cogent ·· engineering

Figure 5. Comparison of two effluent turbidity prediction models: (a) TUO = f(QIN, TUI, CNI, PHI, Dose); (b) TUO = f(QIN, TUI, Variance, Prominence, Correlation, Contrast).
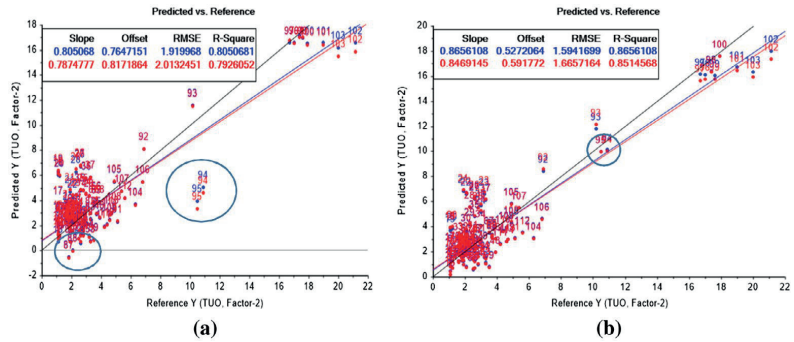
3.5. Difficulties and further research needs

Some weaknesses of the described image analysis installation are that the employed camera is quite expensive and difficult to be properly controlled. Also, it has the battery charge limit, since cannot be charged directly from the electrical plug. The system was not able to work continuously, and this is the main reasons the test data-set resulted in a quite low number of samples. In addition, the resulting images have a high resolution but require quite much space for the storage. Currently, all the calculations of GLCM features and FTIs are done in an external computer. However, we see the potential for the sensor prototype to be further developed. In the next stage, it should be a self-standing computational system with a cheaper camera, which can constantly work with the pre-set settings.

The further research studies are necessary to develop the fully automated floc sensor prototype. The concept of the system will be based on automated flocs image acquisition and its texture image analysis with further matching of the resulted data to different mathematical models, defining the optimal coagulant dose. The improvement of existing on-line dosage control system is a key focus of this research.

4. Conclusions

The images of flocs give a sharp indication of the changes of inlet wastewater parameters and/or coagulation conditions.

The images of flocs are unique for different wastewater qualities and coagulation conditions. GLCM textural features (quantified images of flocs) can distinguish and separate different wastewater coagulation conditions: normal, during the precipitation events, wastewater with the high inlet turbidity.

Floc texture index was introduced and calculated by summarising four GLCM feature vectors—Contrast, Entropy, Homogeneity and Variance. FTI can be used as an early indication parameter of the changes in wastewater qualities and coagulation conditions, which lead to the increase of effluent turbidity. However, further research is needed with the bigger calibration and validation data sets.

Effluent turbidity values can be predicted by few inlet wastewater parameters—flow and inlet turbidity, with $R^2 = 0.79$. The addition of processed flocs' images data increases the outlet turbidity prediction up till $R^2 = 0.87$.

The study shows a potential possibility of the floc sensor prototype to be developed. The images of flocs may be used online for troubleshooting and to improve the existing coagulant dosage control system.

cogent ·· engineering

List of abbreviations

ANN	Artificial neural networks
CNI	Inlet conductivity
DIA	Digital image analysis
DWTP	Drinking water treatment plant
FNU	Formazin nephelometric unit
FTI	Floc texture index
GLCM	Grey level co-occurance matrix
IDM	Inverse difference moment
NTMS	Nephelometric turbidimeter monitoring system
PC	Principal component
PCA	Principal component analysis
PDA	Photometric dispersion analyser
PHI	Inlet pH
PHO	pH after coagulant dosage
PLSR	Partial least squares regression
QIN	Inlet flow rate
RMSEC	Root mean square error of calibration
RMSECV	Root mean square error of cross-validation
SCADA	Supervisory control and data acquisition
SCD	Streaming current detector
Total P	Total Phosphorous
TUI	Inlet turbidity
TUO	Outlet turbidity
WWTP	Wastewater treatment plant

Author details
N. Sivchenko[1]
E-mails: nataliia.sivchenko@nmbu.no, nataliia.sivchenko@gmail.com
ORCID ID: http://orcid.org/0000-0002-1054-0794
H. Ratnaweera[1]
E-mail: harsha.ratnaweera@nmbu.no
ORCID ID: http://orcid.org/0000-0003-1456-2541
K. Kvaal[1]
E-mail: knut.kvaal@nmbu.no
ORCID ID: http://orcid.org/0000-0002-0579-7448
[1] Department of Mathematical Sciences and Technology, Norwegian University of Life Sciences, P.O. Box 5003-IMT, 1432 Aas, Norway.

References
Bache, D., & Gregory, R. (2007). *Flocs in water treatment*. London: IWA Publishing.
Baxter, C. W., Shariff, R., Stanley, S. J., Smith, D. W., Zhang, Q., & Saumer, E. D. (2002). Model-based advanced process control of coagulation. *Water Science and Technology, 45*, 9–17.
Bouyer, D., Coufort, C., Liné, A., & Do-Quang, Z. (2005). Experimental analysis of floc size distributions in a 1-L jar under different hydrodynamics and physicochemical conditions. *Journal of Colloid and Interface Science, 292*, 413–428. doi:10.1016/j.jcis.2005.06.011
Burgess, M., Curley, J., Wiseman, N., & Xiao, H. (2002). On-line optical determination of floc size. Part I: Principles and techniques. *Journal of Pulp and Paper, 28*, 63–65.
Chakraborti, R. K., Atkinson, J. F., & Van Benschoten, J. E. (2000). Characterization of alum floc by image analysis. *Environmental Science & Technology, 34*, 3969–3976. doi:10.1021/es990818o

Chang, Y., Liu, Q.-J., & Zhang, J.-S. (2005). Flocculation control study based on fractal theory. *Journal of Zhejiang University Science B, 6B*, 1038–1044. doi:10.1631/jzus.2005.B1038

Cheng, W. P., Chang, J. N., Chen, P. H., Yu, R. F., & Huang, Y. W. (2011). Turbidity fluctuation as a measure of floc size in a coagulation pilot study. *Desalination and Water Treatment, 30*, 98–104. doi:10.5004/dwt.2011.1878

Cheng, W. P., Kao, Y. P., & Yu, R. F. (2008). A novel method for on-line evaluation of floc size in coagulation process. *Water Research, 42*, 2691–2697. doi:10.1016/j.watres.2008.01.032

Chou, S., Lin, S., & Huang, C. (1998). Application of optical monitor to evaluate the coagulation of pulp wastewater. *Water Science and Technology*. doi:10.1016/S0273-1223(98)00343-6

Conners, R. W., Trivedi, M. M., & Harlow, C. A. (1984). Segmentation of a high-resolution urban scene using texture operators. *Computer Vision, Graphics, and Image Processing, 25*, 273–310. doi:10.1016/0734-189X(84)90197-X

Critchley, R., Smith, E., & Pettit, P. (1990). Automatic coagulation control at water-treatment plants in the North-West region of England. *Water and Environment Journal, 4*, 535–543. https://doi.org/10.1111/wej.1990.4.issue-6

Dentel, S. K., Thomas, A. V., & Kingery, K. M. (1989a). Evaluation of the streaming current detector-I. Use in jar tests. *Water Research, 23*, 413–421. doi:10.1016/0043-1354(89)90132-2

Dentel, S. K., Thomas, A. V., & Kingery, K. M. (1989b). Evaluation of the streaming current detector-II. Continuous flow tests. *Water Research, 23*, 423–430. doi:10.1016/0043-1354(89)90133-4

Haralick, R. M., Shanmugam, K., & Dinstein, I. (1973). Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics, 3*, 610–621. doi:10.1109/TSMC.1973.4309314

He, W., Nan, J., Li, H., & Li, S. (2012). Characteristic analysis on temporal evolution of floc size and structure in low-shear flow. *Water Research, 46*, 509–520. doi:10.1016/j.watres.2011.11.040

Jarvis, P., Jefferson, B., Gregory, J., & Parsons, S. A. (2005). A review of floc strength and breakage. *Water Research, 39*, 3121–3137. https://doi.org/10.1016/j.watres.2005.05.022

Juntunen, P., Liukkonen, M., Lehtola, M., & Hiltunen, Y. (2013). Dynamic soft sensors for detecting factors affecting turbidity in drinking water. *Journal of Hydroinformatics, 15*, 416–426. https://doi.org/10.2166/hydro.2012.052

Juntunen, P., Liukkonen, M., Lehtola, M., & Hiltunen, Y. (2014). Characterization of alum floc by image analysis in water treatment processes. *Cogent Engineering, 1*, 944767.

Li, T., Zhu, Z., Wang, D., Yao, C., & Tang, H. (2006). Characterization of floc size, strength and structure under various coagulation mechanisms. *Powder Technology, 168*, 104–110. doi:10.1016/j.powtec.2006.07.003

Liu, W. (2016). *Enhancement of coagulant dosing control in water and wastewater treatment process* (Doctoral dissertation). Norwegian University of Life Sciences, Ås.

Liu, W., & Ratnaweera, H. (2017). Feed-forward-based software sensor for outlet turbidity of coagulation process considering plug flow condition. *International Journal of Environmental Science and Technology, 14*, 1689–1696.

Liu, W., Ratnaweera, H., & Song, H. (2013, September 18–20). Better treatment efficiencies and process economies with real-time coagulant dosing control. In *11th IWA Conference on Instrumentation Control and Automation*. Narbonne.

Manamperuma, L., Ratnaweera, H. C., & Rathnaweera, S. S. (2013, September 18–20). Retrofitting coagulant dosing control using real-time water quality measurements to reduce coagulant consumption. In *11th IWA Conference on Instrumentation Control and Automation*. Narbonne.

Manamperuma, L., Wei, L., & Ratnaweera, H. (2017). Multi-parameter based coagulant dosing control. *Water Science and Technology, 75*, 2157–2162 http://doi.org/org/10.2166/wst.2017.058

Rasband, W. S. (1997/2016). *ImageJ, US National Institutes of Health, Bethesda, MD*. Retrieved from https://imagej.nih.gov/ij/

Ratnaweera, H., & Fettig, J. (2015). State of the art of online monitoring and control of the coagulation process. *Water, 7*, 6574–6597. doi:10.3390/w7116574

Sharp, E., Banks, J., Billica, J., Gertig, K., Henderson, R., Parsons, S. A., ... Jefferson, B. (2005). Application of zeta potential measurements for coagulation control: Pilot-plant experiences from UK and US waters with elevated organics. *Water Science and Technology: Water Supply, 5*, 49–56.

Sharp, E., Rigby, R., Hughes, R., Claronino, J., Maher, K., Patel, K., ... Jefferson, B. (2016, June 22–24). Using online zeta potential measurements for full-scale coagulation control in drinking water treatment. In *IWA Specialist Conference, Advances in particle science and separation: Meeting tomorrow's challenges*. Oslo.

Sivchenko, N., Kvaal, K., & Ratnaweera, H. (2014, June 15–18). Image analysis of flocs and mathematical modelling applied to coagulation-flocculation process. In *Poster, IWA Specialist Conference, Advances in particle science and separation: From mm to nm scale and beyond*. Sapporo.

Sivchenko, N., Kvaal, K., & Ratnaweera, H. (2016). Evaluation of image texture recognition techniques in application to wastewater coagulation. *Cogent Engineering, 3*(1). doi:10.1080/23311916.2016.1206679

VA-Support. (2012). *DOSCON - Experience at NRA. (03.07.2012)*.

Vahedi, A., & Gorczyca, B. (2011). Application of fractal dimensions to study the structure of flocs formed in lime softening process. *Water Research, 45*, 545–556. doi:10.1016/j.watres.2010.09.014

Valentin, N., & Denœux, T. (2001). A neural network-based software sensor for coagulation control in a water treatment plant. *Intelligent Data Analysis, 5*, 23–39.

Vlieghe, M., Coufort-Saudejaud, C., Frances, C., & Liné, A. (2014). In situ characterization of floc morphology by image analysis in a turbulent Taylor-Couette reactor. *American Institute of Chemical Engineers, 60*, 2389–2403. https://doi.org/DOI https://doi.org/10.1002/aic

Wu, C., Wang, L., Hu, B., & Ye, J. (2013). Influential factors of formation kinetics of flocs produced by water treatment coagulants. *Journal of Environmental Sciences, 25*, 1015–1022. doi:10.1016/S1001-0742(12)60150-8

Xu, Y., Chen, T., Cui, F., & Shi, W. (2016). Effect of reused alum-humic-flocs on coagulation performance and floc characteristics formed by aluminum salt coagulants in humic-acid water. *Chemical Engineering Journal, 287*, 225–232. doi:10.1016/j.cej.2015.11.017

Yavich, A. A., & Van De Wege, J. (2013). Chemical feed control using coagulation computer models and a streaming current detector. *Water Science & Technology, 67*, 2814. doi:10.2166/wst.2013.198

Yu, R., Chen, H., & Cheng, W. (2017). Applying online image analysis to simultaneously evaluate the removals of suspended solids and color from textile wastewater in chemical flocculated sedimentation. *Journal of Environmental Informatics, 29*(1).

Yu, R. F. (2014). On-line evaluating the SS removals for chemical coagulation using digital image analysis and artificial neural networks. *International Journal of Environmental Science and Technology, 11*, 1817–1826. doi:10.1007/s13762-014-0657-1

Zheng, C., Sun, D.-W., & Zheng, L. (2006). Recent applications of image texture for evaluation of food qualities—a review. *Trends in Food Science & Technology, 17*, 113–128. doi:10.1016/j.tifs.2005.11.006

cogent ᐧᐧengineering

Cogent Engineering **(ISSN: 2331-1916) is published by Cogent OA, part of Taylor & Francis Group.**

Publishing with Cogent OA ensures:

- Immediate, universal access to your article on publication
- High visibility and discoverability via the Cogent OA website as well as Taylor & Francis Online
- Download and citation statistics for your article
- Rapid online publication
- Input from, and dialog with, expert editors and editorial boards
- Retention of full copyright of your article
- Guaranteed legacy preservation of your article
- Discounts and waivers for authors in developing regions

Submit your manuscript to a Cogent OA journal at www.CogentOA.com

Paper IV

Sivchenko, N., Kvaal, K., & Ratnaweera, H. Floc sensor prototype tested in the municipal wastewater treatment plant. In press, *Cogent Engineering*.

Floc sensor prototype tested in the municipal wastewater treatment plant

N. Sivchenko*, K. Kvaal, H. Ratnaweera

Norwegian University of Life Sciences, PO Box 5003-IMT, 1432 Aas, Norway
* corresponding author: nataliia.sivchenko@nmbu.no

Abstract: A novel floc sensor prototype was tested in a Norwegian municipal wastewater treatment plant. The resulting images of flocs, captured using a specially designed software, were analysed by textural image analysis technique – grey level co-occurrence matrix (GLCM). The results of image analysis were merged with the coagulation process measurement data – inlet and outlet wastewater parameters. The data based only on GLCM textural features resulted in 96.6 % explained total variance by two principal components and distinguished two classes in the data – low and high outlet turbidity values. The predicted by PLSR coagulant dosages precisely followed the reference dosages, explained Y total variance by 3 factors equals 91.8 % for calibration and 77.9 % for validation.

Keywords: coagulation; wastewater treatment plant; image analysis; texture analysis; floc sensor; dosage prediction; Raspberry Pi

1 Introduction

Wastewater contains a significant amount of suspended and dissolved pollutants, which creates the need for its treatment before discharging into the environment. Coagulation is one of the methods in wastewater treatment to remove suspended solids, phosphates and other water contaminants. At first treatment plants, the quality control was done by manual water sampling after certain stages of treatment (Edzwald, 2010). The main problem with this approach was very long response time from sampling until obtaining the results making it impossible to adjust operating conditions to achieve optimal results. Nowadays for coagulation process, the flow-proportional dosing concept is usually used, while process optimisation often bases on the results from the jar tests (Ratnaweera & Fettig, 2015). Many modern WWTPs perform water quality control based on online measurements of water parameter (Vesilind, 2003). In this approach, the sampling and analysing of water parameters are typically automatized while process control is carried out by online monitors coupled with complicated optimal dosage control systems (Bourgeois, Burgess, & Stuetz, 2001; Ratnaweera, 1997, 2004; Ratnaweera & Fettig, 2015). However, the operational cost of such monitors and control systems are very high. In additions, some of this equipment, for example, orthophosphate and total phosphate monitors, have an entirely long response time, because of the nature of phosphorus chemical analysis process. The characteristics of coming to the WWTP wastewater are changing dramatically within short periods, so the control systems, which require a longer time for the response, are not applicable for dosage control.

Advanced dosing control systems based on multiple water quality parameters that could be measured online have confirmed to be successful (Ratnaweera, 1997; VA-Support, 2012). Such systems enable a reduction in operational costs by lowering coagulant consumption, reduce the sludge volumes and maintain the desired removal of particles and phosphates (Manamperuma, Ratnaweera, & Rathnaweera, 2013; Manamperuma, Wei, & Ratnaweera, 2017). The need of wastewater treatment processes optimisation is growing and requires the development of

accurate, robust, reliable and low-cost online dosing control systems (Ratnaweera & Fettig, 2015).

Raspberry Pi® is a single-board computer which was created last decade and became popular because of its low price. It was previously used in the research projects as a data logger of the optical sensor for continuous marine monitoring and water quality monitoring (Murphy, Heery, et al., 2015; Murphy, Sullivan, Heery, & Regan, 2015). Raspberry Pi was also tested as an alternative to Programmable Logic Controller (PLC) for the automation of a small water treatment plant (Lagu & Deshmukh, 2015).

During the last decades, image analysis techniques were often applied in water and wastewater treatment industry to determine such particles characteristics as floc size, fractal dimension, strength and breakage (Chakraborti, Atkinson, & Van Benschoten, 2000; Chakraborti, Gardner, Atkinson, & Van Benschoten, 2003; Jarvis, Jefferson, Gregory, & Parsons, 2005; Vlieghe, Coufort-Saudejaud, Frances, & Line, 2014). However, their applicability has been limited to laboratory scale tests due to complicated and inaccurate measurements in the field, hardware and software limitations. We have proposed and tested a new approach in image analysis of flocs – computing flocs characteristics on the images as textural features (Sivchenko, Kvaal, & Ratnaweera, 2014, 2016, 2017). Such method of image characterisation is comparatively easy since it analyses the whole image texture instead of calculating the shape characteristics of each particle in the image. Here we have tested the innovative extremely low-cost image acquisition system based on Raspberry Pi single-board computer and camera module. This paper presents the applicability of the concept for municipal wastewater treatment to be further developed into an actual sensor. The sensor is to be used for advanced dosage control to optimising the coagulation process.

2 Experimental setup and Methods

2.1 Wastewater treatment plant

Full-scale tests were conducted at the Skiphelle wastewater treatment plant (Drøbak, Norway) in August 2016. Skiphelle WWTP receives municipal wastewater from Drøbak city and the neighbourhood area. Average inlet flow is 4,600 m^3/day during the days without snowmelt and/or precipitations.

Skiphelle WWTP is a mechanical-chemical precipitation plant. The treatment process consists of the next stages: screens, two parallel pre-sedimentation basins, three sequenced coagulation chambers with the different velocity gradients, and two parallel sedimentation chambers. The plant also has the sludge dewatering and thickening system. The inlet and outlet water quality parameters are measured by online sensors and recorded (average values) with the 10 min interval. The data is available for observation in the plant's SCADA (Supervisory Control and Data Acquisition) system and through the DOSCON (DOSCON AS, Oslo, Norway) system. The retrieved water parameters included inlet wastewater flow (QIN), inlet pH (PHI), inlet turbidity (TUI), wastewater temperature (TMP), coagulant dosage (Dose), pH after coagulant dosage (PHO) and outlet turbidity (TUO). The plant operators perform daily sampling of inlet and outlet total Phosphorous (total P). The summarised data of water quality parameters for the tests period is given in Table 1. Coagulant used in the Skiphelle WWTP is polyaluminium chloride (ECOFLOCK 90, Feralco), 9 ± 0.3 % Al by weight, and density 1356 ± 25 kg/m^3.

Table 1. Inlet and outlet wastewater quality parameters of Skiphelle WWTP during the experiments period

Variables	Mean	Min	Max	Standard Deviation
Inlet flow, m^3/h	190.6	75.6	346.1	60.9
Inlet pH	7.4	7.3	7.9	0.1
Inlet turbidity, FNU	161.0	55.8	500.0	78.3
Ww temperature, ºC	14.4	12.9	15.1	0.5
pH after coagulant dosage	6.9	6.6	7.2	0.1
Outlet turbidity, FNU	4.5	1.6	12.7	4.0
Coagulant dose, ml/s	5.8	2.4	7.3	0.9
Coagulant dose, ml/m^3	115.6	31.6	194.3	22.9
Coagulant dose, mmol Al/l	0.53	0.14	0.88	0.1

2.2 Image acquisition and pre-processing

A special installation was designed to observe changes in flocs' structure *online*. The facility was set above the second flocculation chamber and consisted of the tube, peristaltic pump, acrylic cell for image acquisition, lamp as a light source for a camera, Raspberry Pi with a camera module and the screen. The same system but with the DSLR camera was presented in our previous research paper (Sivchenko et al., 2017). The imaging cell was disassembled for cleanings with water each 2-3 days.

Raspberry Pi 2 Model B V1.1 single-board computer was chosen for its low price (35$ range), compact size and programmability. It is an affordable and perspective tool to be developed into the actual sensor. A low-cost (15$) camera module (5-megapixel OV5647 sensor) with the changeable focal length was mounted to the Raspberry Pi (fig. 1a, b). A unique program was written in Python to control the camera and make changes in the camera settings. The interface of the application is shown in fig. 2. Using the program, we can set specific needed parameters, such as camera settings (fig. 2b), the time lapse between the images and start time. In time lapse window (fig. 2c) option "Interval" means the time interval in seconds between the captured series of images. It was decided that one image might be not enough to represent the flocs structures. Hence, option "Picture count" indicates a number of images in series (3 during this research) and option "Shotspeed" sets the time interval in seconds between the images in one set (5 seconds in this case). It is also possible to use the online camera mode and see the flocs movement through the cell on the monitor in real time. On demand, the video can be recorded choosing menu option "Record video" (fig. 2a).

Camera settings during the tests period were next (fig. 2b): ISO – 400, colour effect – grey scale photo (u,v=128,128 – settings for Raspberry Pi camera), 1/336 shutter speed (2974 microseconds).

3 images were captured every 10 minutes with the interval 5 seconds between the images (sets of images further in the text). The size of the image-capturing zone in the cuvette was 3.2 cm × 9.6 cm. In order to obtain flocs with the proper depth of field, the black metal stripe was placed in the centre of the cuvette, which also became a background for the flocs. The choice of the background colour was based on the fact that the wastewater flocs are greyish

coloured. Thus, using a contrasting background, it is easier to perform the further image analysis. A floodlight LED (anslut® 427-624, 600 lumens) was used as a light source.
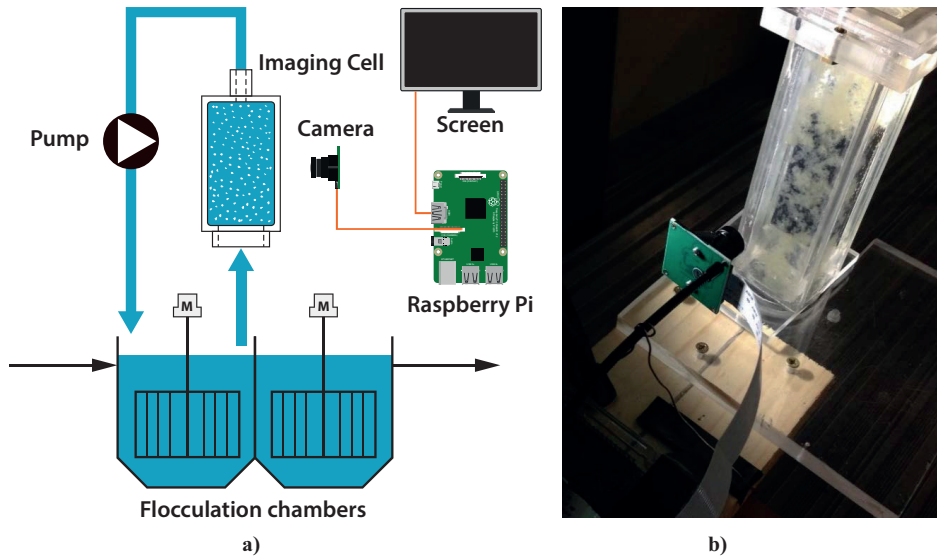

Figure 1 a) Schematic representation of the installation; b) a photograph of the part of installation – Raspberry Pi camera and imaging cell with flocs.

The obtained images have a resolution of 5 megapixels each. They were processed in the open source image analysis software *ImageJ v.1.49* (Rasband, 1996/2016) that based on plugins and macros. For each image 1360 × 1360 pixels (3 cm × 3 cm) area was cropped by manual investigation of the area.
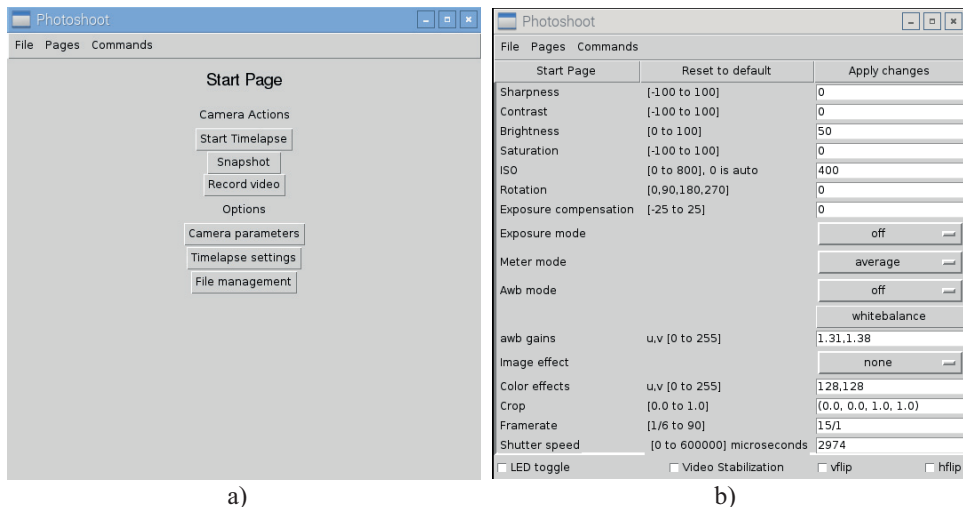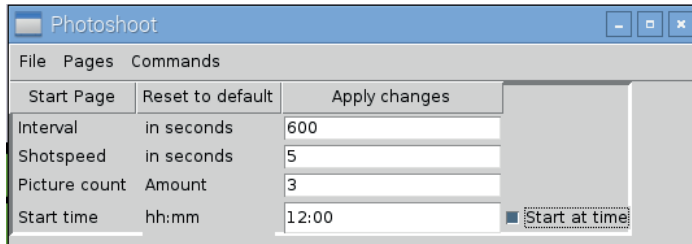
c)

Figure 2 Print screens of the developed program for the Raspberry Pi camera control: a) main page with the menu; b) camera parameters settings; c) time laps settings.

2.3 Image analysis by Grey level co-occurrence matrix (GLCM)

GLCM is one of the methods to analyse texture in the image. Previously it was successfully tested on the laboratory scale basis with model wastewater (Sivchenko et al., 2016) and in full scale with municipal wastewater using an expensive DLSR camera (Sivchenko et al., 2017). The GLCM measurements can be further implemented and computed by Raspberry Pi board itself.

ImageJ plugin "GLCM Texture Too" v. 0.009 was used to obtain the GLCM feature vectors. The resulting output was given as a vector of the next 4 parameters per each image: Contrast, Entropy, Homogeneity and Variance. Hence, the data matrix was obtained with the size 588×4. The detailed description, explanation, and equations for above GLCM texture features can be found in the literature (Conners, Trivedi, & Harlow, 1984; Haralick, Shanmugam, & Dinstein, 1973; Zheng, Sun, & Zheng, 2006). GLCM textural features for two images in a set were averaged and used for calibration, while the textural information from the third image in a set was used for validation during principal component analysis (PCA).

Calibration data set included inlet and outlet measurements of the coagulation process, coagulant dosage and corresponding 4 GLCM textural features of images of flocs (average values from 2 images for each 10 min process data). Validation data set included inlet and outlet measurements of the coagulation process, coagulant dosage and corresponding 4 GLCM textural features of images of flocs (retrieved from the third image captured for the 10 min process data).

2.4 Conjugation of two data sets

Three images for each 10 min were chosen to be representative. The measured GLCM feature vectors were averaged for each 2 images and matched with the retrieved water quality parameters. According to the tracer tests conducted in Skiphelle WWTP, the outlet turbidity values were 45 min shifted to meet the response lag between the coagulant injection point and outlet from the sedimentation tank. After the removal of missing values and outliers, the resulting data set included 196 samples. In a real-world situation, the removal of outliers should be treated carefully due to a warning of unwanted conditions.

2.5 Multivariate statistical analysis and modelling

The resulted data matrix was processed in statistical software *The Unscrambler® X 10.3* (CAMO Software AS, Norway) and in MATLAB using PLS toolbox. Principal component

analysis (PCA) was performed to find the relationships between water quality parameters and images of flocs – GLCM feature vectors. PCA is a statistical data analysis technique to reduce the dimensionality of the data set, overview and describe the interrelationships among variables and to find so-called hidden structures in the data. Partial least squares regression (PLSR) was performed to predict coagulant dosage based on the combination of water quality parameters and GLCM texture features. PLSR is a statistical regression method to model the response variable using a large number of predictor variables while those variables may highly correlate.

3 Results and Discussion

Previously we have tested three different texture image analysis systems applied to coagulation process in the laboratory scale (Sivchenko et al., 2016). It was found that the textural information retrieved from the images of flocs is related to the coagulant dosages. Thus, the quantified information from the images of flocs can be used directly for the dosage prediction. Afterwards, the special installation for image acquisition was placed in the Skiphelle WWTP. The captured by DSLR camera images of flocs were used to predict the outlet turbidity values after sedimentation (Sivchenko et al., 2017). It was also shown that such image analysis technique could work as an early warning system of coagulation failure.

This work aimed to continue the development of a floc sensor prototype using cheap hardware available in the market. Since the difference in images resolution for Nikon D600 DSLR camera and Raspberry Pi camera module is significant, 24 MP and 5 MP, respectively, it was essential to check whether the last can produce flocs images of enough quality to use them for the coagulant dosage prediction.

The significant advantage of a Raspberry Pi computer is that it is entirely programmable and easy to control the camera module. It is also possible to develop a sophisticated self-standing system, which includes image acquisition, storage and processing – all done in Raspberry Pi. Besides, the single-board computer has a Wi-Fi module, so the captured/processed information can be sent right away to the server, database cloud or remote PC.

During this research, the images of flocs were captured with the specified frequency and stored on the Raspberry Pi memory card, afterwards collected manually and processed on the remote PC. Here we present the results of analysed images of flocs by GLCM and corresponding them measurement data from the coagulation process.

Figure 3a shows the results of PCA for 4 GLCM feature vectors of the flocs images. With only two principal components (PCs) the total explained variance equals 96.6 % (PC1=63.3 %) for calibration and 90.4 % (PC1=52.2 %) for validation. PC1 is mainly explained by Contrast, Entropy and Homogeneity (loadings values 0.55, 0.54 and -0.51, respectively), while GLCM textural parameter Variance mainly contributes to PC2 (loading value 0.69).

The results of PCA performed for all inlet (QIN, TUI, PHI) and outlet (PHO, TUO) process measurements, 4 textural features of the flocs images and coagulant dosage in ml/s, are shown in figure 3b. Data class, which corresponds to the high outlet turbidities (marked in red), is better separated by two PCs. However, the data set in this case included variable TUO. Even though the scores are visually separated better, the total explained variance of the data set is

much lower. Total explained variance for calibration: PC1=38.4 %, PC2=55.9 %, PC3=70.4 %, PC4=79.7 %; for validation: PC1=18.2 %, PC2=29.9 %, PC3=30.4 %, PC4=60.8 %.
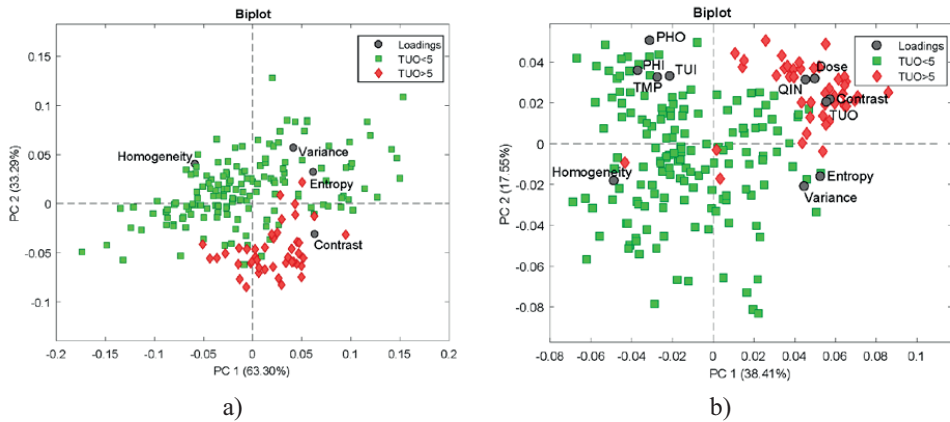

a)   b)

Figure 3 Scores plot of PCA, PC1 vs PC2 for: a) data with 4 GLCM textural features of the images of flocs; b) data containing wastewater quality parameters, GLCM textural features and the coagulant dosage

For the prediction of coagulant dosages the data matrix was divided into calibration and test data sets, 60 % and 40 % of the data, respectfully. The data was divided based on the outlet turbidity values. The X matrix for PLSR included inlet wastewater parameters – QIN, TUI, PHI, TMP; after the dosage measurement PHO; an hour of the day; and GLCM textural features – Contrast, Entropy, Homogeneity and Variance. The response Y was coagulant dosage in ml/s. The PLSR model was calibrated on the data values which correspond to the outlet turbidity measurements between 1.9-5 FNU (desired range of effluent turbidity for the Skiphelle WWTP). The test data set included measurements related to outlet turbidity values (TUO) higher than 5 FNU.

Figure 4 shows the results of PLSR – coagulant dosage prediction. The continuous red line is the reference dose (dosages used in the WWTP). Black squares represent the dosage prediction of calibration data with corresponding TUO in a range 1.9-5 FNU. Black dots are the dosage predictions which correspond to TUO less than 1.9 FNU. Black diamonds are the dosage predictions which correspond to TUO above 5 FNU. Minimum desired value of effluent turbidity – 1.9 FNU and maximum value 5 FNU are marked by dashed green lines, TUO min and TUO max, respectfully.

Overall, the predicted coagulant dosages precisely follow the reference dosages. Prediction R^2 equals 0.92 for calibration and 0.78 for validation with three factors, root mean square error (RMSE) for calibration is 0.182 and 0.297 for validation. The area with high TUO represents a rain event, dosage was manually adjusted by plant operators and tend to be under-estimated. The predicted by PLSR model dosages (black diamonds, fig. 4) suggest having higher coagulant use for the wet-weather period.
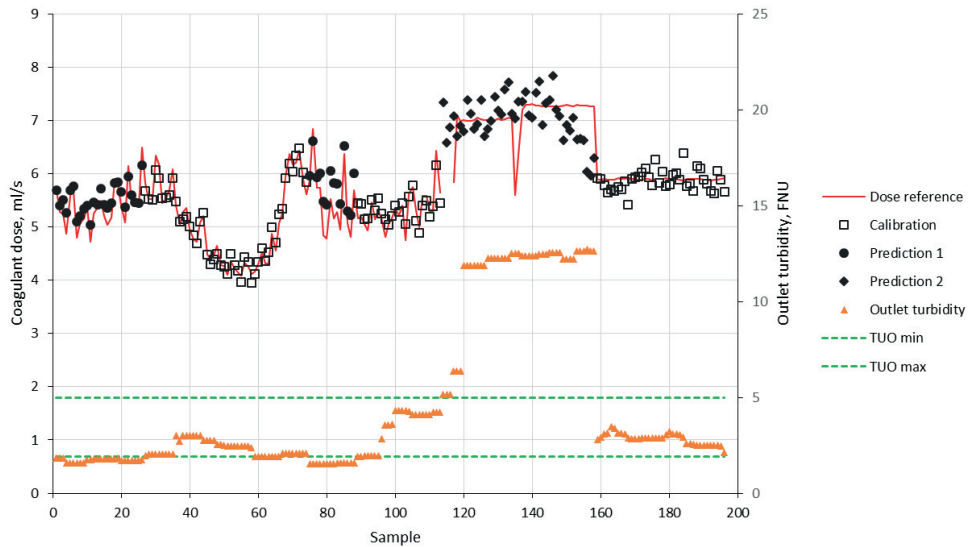
Figure 4 Comparison of the predicted coagulant dosages by PLSR and reference dosages with corresponding them effluent turbidity measurements.

Figure 5 represents the sample flocs images of three classes in the data set – coagulation process conditions when resulting outlet turbidity was lower than 1.9 FNU, range of desired in the plant effluent turbidity 1.9-5 FNU and conditions lead to high TUO values (higher than 5 FNU). The actual values of TUO for the presented images are written in brackets. For each class, the average GLCM textural features are presented. The predicted coagulant dosages for the particular presented images of flocs are also shown.

Investigating the average values of four GLCM parameters, we can conclude that the images of flocs with higher Homogeneity (visually these flocs also seem to have a higher density) result in good treatment efficiency – low outlet turbidity measurements. Images of flocs with the higher Contrast, Entropy and Variance values point on the coagulation failure and require higher dosages of the coagulant. Visually such flocs are more separated from each other, and probably it results in worse sedimentation abilities of the flocs. However, the sedimentation rates of the different floc structures were not studied in this research.
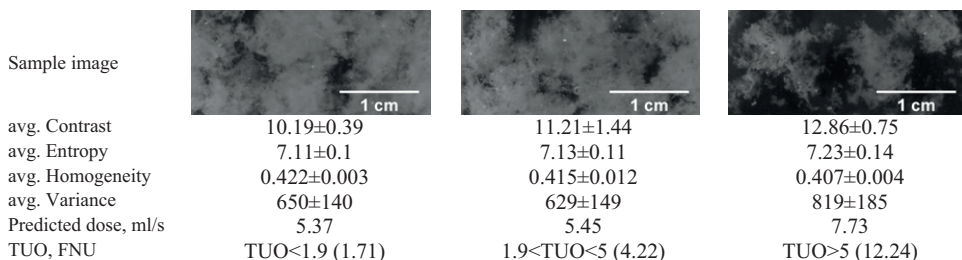
Sample image			
avg. Contrast	10.19±0.39	11.21±1.44	12.86±0.75
avg. Entropy	7.11±0.1	7.13±0.11	7.23±0.14
avg. Homogeneity	0.422±0.003	0.415±0.012	0.407±0.004
avg. Variance	650±140	629±149	819±185
Predicted dose, ml/s	5.37	5.45	7.73
TUO, FNU	TUO<1.9 (1.71)	1.9<TUO<5 (4.22)	TUO>5 (12.24)

Figure 5 Comparison of the sample images corresponding to different outlet turbidity measurements and the ranges of their GLCM textural features.

Further detailed investigations of flocs and their sedimentation abilities are needed to find the correlations between flocs images and sedimentation rates.

Regarding the further floc sensor prototype development, the Raspberry Pi and camera module are to be sealed into the waterproof stainless steel case and put directly into the coagulation chamber for online image analysis. The software in Raspberry Pi should be extended to include automated GLCM texture analysis algorithm and rewriting function of the stored images. The appropriate light source is also to be found and mounted next to the camera lens.

Conclusions

The tested innovative image acquisition system – floc sensor prototype, proved to produce the images of flocs enough quality to be further used for troubleshooting, coagulation process faults detection and more efficient coagulant dosage prediction.

Images of flocs quantified by textural image analysis technique – GLCM, were able to distinguish coagulation conditions, which lead to insufficient wastewater treatment with high outlet values of turbidity.

Coagulant dosages were predicted by wastewater quality parameters and 4 GLCM textural features with R^2=0.92 for calibration and R^2=0.78 for validation (three factors).

It was found that images of flocs with high values of homogeneity parameter are related to low outlet turbidity values. While images of flocs with higher contrast, entropy and variance values are associated with low treatment efficiency (high effluent turbidity).

The study shows a potential possibility of the floc sensor prototype to be developed into an actual sensor, based on textural image analysis of flocs. The sensor is to be used to improve an existing coagulant dosage control system.

List of abbreviations

FNU	Formazin nephelometric unit
GLCM	Grey level co-occurrence matrix
DSLR	Digital single-lens reflex (camera)
LED	Light-emitting diode
MP	Megapixels
PC(n)	Principal component
PCA	Principal component analysis
PHI	Inlet pH
PHO	pH after coagulant dosage
PLC	Programmable Logic Controller
PLSR	Partial least squares regression
QIN	Inlet flow rate
RMSE	Root mean square error
SCADA	Supervisory control and data acquisition
TUI	Inlet turbidity
TUO	Outlet turbidity
Ww	Wastewater
WWTP	Wastewater treatment plant

Acknowledgment

References

Bourgeois, W., Burgess, J. E., & Stuetz, R. M. (2001). On-line monitoring of wastewater quality: a review. *Journal of Chemical Technology & Biotechnology*, *76*(4), 337–348. http://doi.org/10.1002/jctb.393

Chakraborti, R. K., Atkinson, J. F., & Van Benschoten, J. E. (2000). Characterization of Alum Floc by Image Analysis. *Environmental Science & Technology*, *34*(18), 3969–3976. http://doi.org/10.1021/es990818o

Chakraborti, R. K., Gardner, K. H., Atkinson, J. F., & Van Benschoten, J. E. (2003). Changes in fractal dimension during aggregation. *Water Research*, *37*(4), 873–883.

Conners, R. W., Trivedi, M. M., & Harlow, C. A. (1984). Segmentation of a high-resolution urban scene using texture operators. *Computer Vision, Graphics, and Image Processing*, *25*(3), 273–310. http://doi.org/10.1016/0734-189X(84)90197-X

Edzwald, J. K. (2010). *Water quality and treatment : a handbook on drinking water*. AWWA. McGraw-Hill, New York.

Haralick, R. M., Shanmugam, K., & Dinstein, I. (1973). Textural Features for Image Classification. *IEEE Transactions on Systems, Man, and Cybernetics*, *3*(6), 610–621. http://doi.org/10.1109/TSMC.1973.4309314

Jarvis, P., Jefferson, B., Gregory, J., & Parsons, S. A. (2005). A review of floc strength and breackage. *Water Research*, *39*(14), 3121–3137.

Lagu, S. S., & Deshmukh, S. B. (2015). Raspberry Pi for automation of water treatment plant. *2015 International Conference on Computing Communication Control and Automation*, 532–536. http://doi.org/DOI 10.1109/ICCUBEA.2015.109

Manamperuma, L., Ratnaweera, H. C., & Rathnaweera, S. S. (2013). Retrofitting coagulant dosing control using real-time water quality measurements to reduce coagulant consumption. In *11th IWA Conference on Instrumentation Control and Automation*. Narbonne, France, 18–20 September 2013.

Manamperuma, L., Wei, L., & Ratnaweera, H. (2017). Multi-parameter based coagulant dosing control. *Water Science and Technology*, *75*(9), 2157–2162. http://doi.org/org/10.2166/wst.2017.058

Murphy, K., Heery, B., Sullivan, T., Zhang, D., Paludetti, L., Lau, K. T., … Regan, F. (2015). A low-cost autonomous optical sensor for water quality monitoring. *Talanta*, *132*, 520–527. http://doi.org/10.1016/j.talanta.2014.09.045

Murphy, K., Sullivan, T., Heery, B., & Regan, F. (2015). Data analysis from a low-cost optical sensor for continuous marine monitoring. *Sensors and Actuators, B: Chemical*, *214*, 211–217. http://doi.org/10.1016/j.snb.2015.02.023

Rasband, W. S. (1997/2016). *ImageJ, US National Institutes of Health, Bethesda, MD.* Retrieved from http://imagej.nih.gov/ij/

Ratnaweera, H. (1997). *Chemical wastewater treatment: a concept for optimal dosing of coagulants*.

Ratnaweera, H. (2004). Coagulant dosing control - a review. *Chemical Water and Wastewater Treatment, IWA Publishing, London*, *7*, 3–13.

Ratnaweera, H., & Fettig, J. (2015). State of the Art of Online Monitoring and Control of the Coagulation Process. *Water*, *7*(11), 6574–6597. http://doi.org/10.3390/w7116574

Sivchenko, N., Kvaal, K., & Ratnaweera, H. (2014). Image analysis of flocs and mathematical modelling applied to coagulation-flocculation process. In *Poster, IWA Specialist Conference, Advances in particle science and separation: from mm to nm scale and beyond.* Sapporo, Japan, 15-18 June 2014.

Sivchenko, N., Kvaal, K., & Ratnaweera, H. (2016). Evaluation of image texture recognition techniques in application to wastewater coagulation. *Cogent Engineering*, *3*(1). http://doi.org/10.1080/23311916.2016.1206679

Sivchenko, N., Ratnaweera, H., & Kvaal, K. (2017). Approbation of the texture analysis imaging technique in the wastewater treatment plant. *Cogent Engineering*, *4*(1373416). http://doi.org/10.1080/23311916.2017.1373416

VA-Support. (2012). *DOSCON - Experience at NRA. (03.07.2012)*.

Vesilind, P. (Ed.). (2003). *Wastewater treatment plant design*. IWA publishing.

Vlieghe, M., Coufort-Saudejaud, C., Frances, C., & Line, A. (2014). In situ characterization of floc morphology by image analysis in a turbulent Taylor-couette reactor. *American Institute of Chemical Engineers*, *60*(7), 2389–2403. http://doi.org/DOI 10.1002/aic

Zheng, C., Sun, D.-W., & Zheng, L. (2006). Recent applications of image texture for evaluation of food qualities—a review. *Trends in Food Science & Technology*, *17*(3), 113–128. http://doi.org/10.1016/j.tifs.2005.11.006

Norwegian University
of Life Sciences