Modelling and Optimisation of Wastewater Coagulation Process

Modellering og optimalisering av koaguleringsprosess i avløpsvann

Philosophiae Doctor (PhD) Thesis

Subhash S. Rathnaweera

Dept. of Mathematical Sciences and Technology
Norwegian University of Life Sciences

Ås 2010

Thesis number 2010: 05
ISSN 1503-1667
ISBN 978-82-575-0917-

Summary

The efficiency and economics of chemical coagulation, one of the most robust wastewater treatment processes, is critically dependent on the optimal coagulant dosage which is strongly related to the influent quality. The process basics indicate that the optimal coagulant dosage is dependent on several parameters such as flow, particles, phosphates and pH which are not proportional to each other, thus cannot be represented by a single parameter alone. However, the usage of two or more parameters in wastewater coagulation control is yet to be seen in full scale applications as a common strategy. A survey among Norwegian wastewater treatment plants indicates that over 80% only used flow, or sometimes combined with pH, as the dosing control strategy. The recent developments in the on-line water quality sensors have brought about the practical possibilities to use them in treatment plants. Despite this, active usage to unveil the hidden potential of these instruments and control concepts is still scarce. The main challenge now lies in establishing robust and accurate mathematical models to describe the dosage vs influent & effluent parameters, as well as concepts to make the online measurements more accurate and valid. A multiple-parameter and multiple-model based coagulant dosage control (XCDC) concept has been developed, tested and elaborated in this thesis.

The basis for this concept was preliminarily evaluated by Lu (2003), where a single model based basic modelling results were presented. The present study elaborates an expanded investigation and results by selecting the best suitable multivariate calibration system for model development, a system to validate and manage online water quality monitoring data, a multiple-model system to manage non-validated values, full scale tests under various conditions and finally contributing to the restructuring of the software enabling universal implementation at wastewater treatment plants (WWTPs).

The studies were conducted at four WWTPs in both Norway and China. Full scale tests were conducted in NRA WWTP, Lillestrøm, Norway, HIAS WWTP, Hamar, Norway and Xiao Hong Men WWTP, Beijing, China. Furthermore, a pilot-scale study was conducted in Gaobeidian WWTP, Beijing, China.

Coagulation is a well defined process, which can be mathematically described (Ratnaweera *et al.* 1994). However, the construction of a conceptual model has been a challenge due to the complex nature of the process. Few successful attempts to construct relationships between coagulant dosage and the water quality parameters are reported. In the present study, three different multivariate analytical methods, MLR, PCR and PLSR, were evaluated to find the best suitable regression analytical method for the purpose. Considering the fact that most related studies report the use of these systems to describe the historical relationships, this thesis focuses on the validity of these models' predictability of future situations enabling usage in online process control.

A robust and accurate error recognising and validating system for online measurements is crucial when they are used in process control as they may create critical conditions. The

commonly available hardware and set-point based error detection methods are inadequate in real time process control. The experimental coagulant dosing control studied is evaluated with an efficient and accurate error recognising and validating system. The concept is based on a software based floating error detection system developed using multivariate calibration systems.

When the measurements are validated and errors are identified, a management system to minimise their impacts on the real time process control is necessary to secure the accuracy of the process. A robust multiple model based strategy was integrated to the multi parameter based experimental coagulant dosage control system. The concept is based on a set of models with a variable number of water quality parameters, enabling the activation of the best suitable dosage estimation equation at all times.

Though the experimental coagulant dosing control system required complicated programming structure, it was possible to integrate it in to a simple, commercially available Programmable Logical Controller (PLC). The PLC could then integrate in to the treatment plants' main supervisory control and data acquisition (SCADA) systems.

The system at NRA is successfully running with over 12% of coagulant saving with considerable reduction of the sludge production. Further studies suggest the ability of saving 16% or more compared with historical data. The HIAS system was successfully run until a shock-loading due to two very different influent types required usage of two model sets to manage variations. The experiments showed the possibility to save 5% to 15% compared with the traditional coagulant consumption. The treatment plants in Beijing, China showed savings up to 31% while maintaining the same effluent qualities. Further studies showed the possibility to reduce the coagulant demand 2.4 to 7.8 times by changing the present dosage strategies combined with use of better coagulants.

Sammendrag

Effektiviteten og økonomi av kjemiskfelling, en av de mest robuste avløpsrenseprosessene, er kritisk avhengig av en optimal koagulantdose som er sterkt knyttet til kvaliteten av innløpsvann. Basiskunnskap innen denne prosessen viser til at kogaulantdosen er avhengig av flere parametere som vannmengde, partikler, fosfater og pH – verdier som ikke kan representeres ved en enkelt parameter. Bruken av to eller flere parametere i doseringskontroll under fellingsprosessene finnes fortsatt ikkes i full skala applikasjoner som en felles strategi. En undersøkelse blant norske avløpsrenseanlegg viser at over 80% kun anvender vannmendge, evt kombinert med pH overstyring, som doseringskontrollstrategi. Den siste utviklingen i sanntids vannkvalitetsensorer åpner praktiske muligheter for å bruke dem i renseanlegg. Til tross for denne utviklingen er det meget lite aktiv bruk av disse instrumentene for å avsløre det skjulte potensialet av slike kontrollkonsepter i fellingsprosessen. Den største utfordringen ligger i å etablere robuste og nøyaktige matematiske modeller for å beskrive relasjoner mellom dosen og innløp og utløp, samt konsepter for å gjøre sanntidsmålinger mer nøyaktig og gyldig. En multiparameter og flermodell basert koagulant doseringskontroll (XCDC) er utviklet, testet og utarbeidet i denne avhandlingen.

Grunnlaget for dette konseptet ble først evaluert av Lu (2003) på en enkelt modell. Den nåværende studien presenterer en utvidet undersøkelse og resultater ved å velge de best egnede multivariate kalibreringssystem for modellbasert utvikling, et system for å kontrollere og administrere online vannkvalitet overvåkingsdata, flere modellsystem for å administrere ikke-validerte verdier, fullskala tester under ulike forhold og et avsluttende bidrag til restrukturering av programvaren som har universal implementering i ulike avløpsanlegg renseanlegg.

Studiene ble gjennomført på fire renseanlegg i både Norge og Kina. Fullskala tester ble gjennomført ved NRA i Lillestrøm, HIAS i Hamar og Xiao Hong Menn reneseanlegget i Beijing, Kina. I tillegg ble en pilotskala studie gjennomført i Gaobeidian renseanlegget i Beijing, Kina.

Koagulering er en veldefinert prosess som kan beskrives matematisk (Ratnaweera *et al.*1994). Bygging av en konseptuel modell har imidlertid vært en utfordring på grunn av prosessens komplekse natur. Få vellykkede forsøk på å matematisk beskrive forhold mellom koagulantdosering og vannkvalitetsparametere er rapportert. I denne studien ble tre ulike multivariate analysemetoder, MLR, PCR og PLSR, vurdert for å finne den best egnede regresjonsanalysemetoden. Tatt i betraktning at de fleste relaterte studier rapporterer bruken av disse systemene til å beskrive sammenheng i historiske data, fokuserer denne avhandlingen på gyldigheten av disse modellenes kapasitet til å forutse fremtidige situasjoner og om det åpner for bruk i online prosesskontroll.

Et robust og nøyaktig konsept for å oppdage feil i sanntidsmålinger og deres validering er avgjørende når de brukes i prosesskontroll da feilmålinger kan skape kritiske forhold. Den

vanlige maskinvarebaserte og set-point baserte feildeteksjonsmetoden er utilstrekkelig i sanntidsprosesskontroll. Det eksperimentelle koagulant doseringskontrollsystemet er evaluert med et konsept for effektiv og nøyaktig feilgjenkjennelse og validering. Konseptet er basert på et programvarebasert flytende feildeteksjonssystem utviklet ved hjelp av multivariate kalibreringssystem.

Når målingene er validert og feil er identifisert, er det nødvendig med et styringssystem for å minimere målefeilenes virkninger på sanntidsprosesskontrollfor å sikre nøyaktigheten av prosessen. En robust og fleremodell basert strategi ble integrert til et multiparameterbaserte eksperimentelle koagulant doseringskontrollsystem. Konseptet er basert på et sett av modeller med varierende antall vannkvalitetsparametere, slik at kun de best egnede doseringensestimatlikningene aktiveres til enhver tid.

Selv om det eksperimentelle koagulant doseringskontrollsystemet krever komplisert programmeringsstruktur var det mulig å integrere det på en enkel og kommersielt tilgjengelig programmerbar logisk styring (PLS). PLS kan deretter integreres i renseanleggets sentrale datastyringssystem (SCADA).

Systemet på NRA er vellykket og kjører fortsatt med over 12% av sparing av fellingsmidller med betydelig reduksjon av slamproduksjonen. Videre studier antyder muligheten til å spare på 16% eller mer sammenlignet med historiske data. Forsøkene ved HIAS var vellykket inntil en sjokkbelasting grunnet to svært ulike typer av innløp krevde bruk av et dobbelt modellsett for å håndterer estimering av den optimale doseringen. Forsøkene viste muligheten til å spare 5% til 15% sammenlignet med tradisjonelle fellingsmiddelforbruk. Renseanlegget i Beijing viste muligheten for å spare opp til 31% og samtidig opprettholde den samme utløpskvaliteten. Videre studier viste muligheten for å redusere fellingsmiddelforbruket med 2.4 til 7.8 ganger ved å endre det nåværende doseringspunktet samt bruk av bedre koaguleringsmidler.

This thesis is dedicated to my late father Somapala Rathnaweera,
whose ambition was to push me in to a world
which he could not see...

Acknowledgement

First, I would like to express my deep sense of gratitude to my scientific advisor Prof. Harsha Ratnaweera, for his invaluable guidance, inspirited encouragement and vast knowledge which has enabled me to develop an understanding of the subject and especially for his kindness and patience. He has helped me in numerous ways from the initiation to the completion of my study. This thesis would not have been possible without his support. I owe my deepest gratitude to Dr. Tor Håkonsen, my external supervisor of this study, for helping me to start studies in NRA, and continuously advising and encouraging me throughout my study. His knowledge and advice enriched my research. I offer my sincere gratitude to Prof. Oddvar Lindholm, the main supervisor of the study who supported me throughout my study with all the formalities, advice and guidance through these four years. I am honoured to be a student of his.

It is an honour for me to thank Mr. Ingar Tranum, the director and Mr. Eirik Rismyhr, the technical manager of NRA – Lillestrøm wastewater treatment plant, for accepting me as a fully integrated member of the NRA community. I am indebted to Mr. Stein Martin and all the colleagues who worked in the NRA wastewater treatment plant for constantly supporting me. They never denied me in any circumstances. Thank you all for the true friendly contribution. My special thanks are conveyed to Mr. Bernt Hellend, who has helped me since I began my studies in NRA. His wise practical knowledge helped me to establish all my practical skills with instruments. The great contribution of all the kind ladies especially Grete, Yelka and Ludmilla at the Noranalyse laboratory is deeply appreciated. They made available their support and knowledge in a number of ways during my study.

I would like to thank to Mr. Geir Hagen and Mr. Gjermund Sørensen at HIAS Iks, and their colleagues for helping me to make this study successful in HIAS wastewater treatment plant. Also I would like to show my gratitude to Prof. Hanchang Shi, Ms. Hou L, Dr. Pang and Dr. Qiu Yong for their valuable guidance in the studies held in Beijing China. The success of this study is greatly attributed to their support.

My heartiest gratitude is offered to Mr. Miroslav Hribljan and Mr. Dejan Josik from Indas Industry Assistance, for helping us to develop the XCDC software and also the valuable on-line support.

I am heartily thankful to the Norwegian University of Life Sciences (UMB) and the Department of Mathematical Sciences and Technology (IMT) for accepting me for the PhD studies and employing me throughout the study.

It is a pleasure to thank the Norwegian Institute for Water Research (NIVA), Doscon AS, NRA IKS and VA- support AS for funding my studies and enabling a dream come through.

I wish to express my sincere thanks to Mr. Dinindu Ratnaweera for helping me by reading and correcting English of this thesis.

I offer my profound thanks to my mother and family members for continuously encouraging me to pursue my studies. I wish to thank my loving wife Indika and daughter Samidi for their understanding, patience, help and permission to use an endless time on my studies, which I should have spent with them.

Lastly, I offer my heartiest regards and blessings to all those who supported me in any respect during my study.

List of Acronyms

AAO	Anaerobic – Anoxic - Oxic treatment
ANN	Artificial neural networks
CDC	Coagulant dosing control
CN	Conductivity
CNI	Influent conductivity
DAY	Day of week
DWTP	Drinking water treatment plants
GBD	Gaobeidian wastewater treatment plant, Beijing, China
HIAS	HIAS wastewater treatment plant, Hamar, Norway
MLR	Multiple linear regression
NRA	NRA wastewater treatment plant, Lillestrøm, Norway
OP	Orthophosphate
P	Phosphorus
PCR	Principal component regression
PH	pH
PHI	Influent pH
PHO	pH - after coagulation
PLC	Programmable logical controller
PLSR	Partial least squares regression
QI	Influent flow
R^2	Regression coefficient
RMSE	Root mean square error
SC	Streaming current
SCADA	Supervisory control and data acquisition
SCD	Streaming current detector
SCO	Streaming current - After coagulation
SS	Suspended solids
TEI	Influent temperature
TIM	Measured hour
TP	Total phosphorus
TU	Turbidity
TUI	Influent turbidity
TUO	Effluent turbidity
UV	Ultra violet
WW	Wastewater
WWTP	Wastewater treatment plant
XCDC	Experimental coagulant dosing control system
XHM	Xiao-Hong-Men wastewater treatment plant, Beijing

Contents

1 Introduction

1.1 Why coagulation is important

The chemical coagulation process has been popular in many countries due to its efficiency, flexibility and robustness against climatic and shock loads (Ratnaweera *et al*. 2002). Coagulation, flocculation and separation are the three major processes used in chemical water and wastewater treatment. Coagulation is induced by adding chemical coagulants to the water and, typically, letting the particles agglomerate in a flocculation basin (Kemira 2003). The flocculated particles are separated by sedimentation or filtration. Coagulation is the most important process of the three unit processes (Lu *et al*. 2003) and it can influence the other downstream processes of the system as well as the total outcome.

With stringent demands on removal of phosphorus (P) and suspended solids (SS) from wastewater (WW), most wastewater treatment plants (WWTP) are in need of improvement in respect of their current treatment processes. As influent quality and pollution loads become more challenging with time, the demand for effluent quality is becoming stricter. As a result, the biological P removal processes in many WWTP are experiencing difficulties with reaching goals on effluent P concentration in several parts of the world. Thus, chemical coagulation has become one of the best options for facing numerous challenges in WW treatment today.

1.2 Current challenges with the coagulation process

Though the capital costs of a chemical WW treatment plant are generally much lower compared to those of a biological P removal plant, the operational cost can be relatively high (Leentvaar *et al*. 1979). Hangouet *et al*. (2007) have reported that the chemical costs may represent up to 20% of the operating cost of an average treatment plant. This percentage varies from plant to plant with quality of influents, expected treatment quality and management of the WWTP. Using the proper coagulant and selecting the most appropriate coagulant dose are the two most important management factors in addition to providing optimal process conditions.

The coagulant introduced to water or WW is consumed in three different ways. Part of it reacts with the dissolved orthophosphate (OP) to precipitate, another portion is used for suspended particle (SS) removal by any of the coagulation mechanisms, such as double layer compression, charge neutralization, bridging or colloid entrapment (Zeta Meter Inc. 1993). The rest is removed as residuals in the treated effluent water. Thus the OP, SS and pH of WW are considered to be the main parameters influencing the coagulant demand (Ratnaweera 1991; Gillberg *et al*. 1996). In addition to these, most of the other influent quality parameters indirectly contribute to coagulation reaction. For example: temperature influences the coagulation mechanism in different ways. The solubility of inert matter and coagulants both increase with increasing temperature. Increased temperature increases the motion of particles, facilitating more collisions for rapid agglomeration of particles. On the other hand, Bache (1996) documented floc weakening at low temperatures, in certain mechanisms.

Flow, pH, turbidity (TU) or SS, and colour are the main parameters used for coagulant dosing control in drinking water treatment plants. In WWTPs, the phosphate is used as a parameter instead of colour.

Significant daily changes, seasonal changes and time changes of influent quality due to climatic factors are very common and well documented. Hansen (1996) has studied the day and night changes of WW at the plant gate and observed four times larger flow and BOD during the daytime compared to night. Holmquist (2004) has reported drastic variations of phosphate in influent WW during a single day. Fig. 1 illustrates the variation of flow, orthophosphate and turbidity during three consecutive days in NRA WWTP Lillestrøm (NRA), where a major part of the practical work of this study was conducted. It clearly shows the variation of parameters with time.

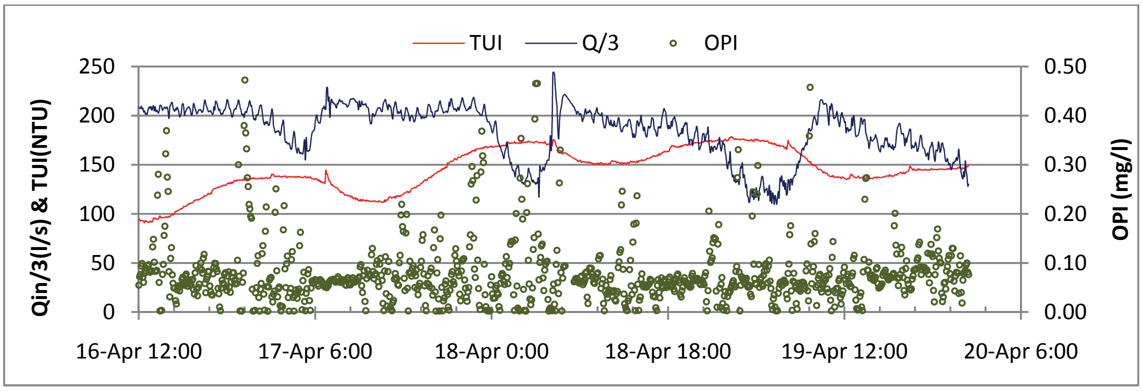


Fig.1. The WW quality variation after biological treatment of NRA WWTP during three consecutive days in 2008. QIN is the inflow in l/s, TUI is influent turbidity in NTU, OPI is influent Orthophosphate concentration in mg/l.

Fig. 1 confirms that none of the parameters are proportional to each other. Consequently, it is not possible to predict the optimal dosage based on one or two of these parameters, when the process is heavily dependent on all three parameters.

1.3 Dosing control – today's practice

It is reported that most chemical water and WWTP are adapted to use at least a flow proportional coagulant dosing control (CDC) system (Dentel 1991). Looking at Norwegian conditions, in a survey among major drinking water treatment plants (DWTP)s and WWTPs, Ratnaweera (2004) reported that 83% of WWTPs and 80% of DWTPs use only flow as a control parameter, sometimes in combination with a pH overrun function for CDC. Integration of quality parameters like TU or colour was found in less than 20% of the plants.

According to Ratnaweera (2004), some DWTPs with raw water sourced from lakes, where the raw water quality remains more or less unchanged over the year, may obtain satisfactory results with constant dosing, irrespective of the flow, particles, colour, etc. However, the same survey reported that 2% of Norwegian WWTPs use constant coagulant dosage irrespective of the flow and quality of wastewater. Obviously, such a CDC will be far from optimum. In constant dosing, although plant operators normally use laboratory Jar testing results to evaluate the appropriate doses, higher dosages are common to ensure adequate treatment efficiencies. This, most probably, will result in overdosing under normal conditions and under-dosing during shock loads.

The above mentioned survey shows that, although reliable, cheap and robust, on-line measuring equipment is available, their usage in CDC is not yet popular. There has been a significant positive development during the last few years in this regard, for example the potential use of phosphate measurement has been documented by several authors. Holmquist (2004) described a successful dosing control system with double point phosphorus measurement where the influent phosphate concentration varies rapidly within a few minutes. Devisscher (2002) evaluated a successful system to control the chemical dosage using OP measurements in activated sludge treatment.

Online TU, SS and colour measurements are more reliable for CDC and many more studies have been conducted regarding them. Zeghal *et al*. (1996) used outlet TU with feedback control in CDC in DWTP. However, it was also noted that, while it is possible to use feedback control in a DWTP where the water quality remains stable over days or weeks, it is impossible to use feedback control in a WWTP where the water quality could change within 15 minutes, combined with a retention time of several hours (Ratnaweera 2004). Hansen (1996) reported on the online measurement of SS to reduce the usage of coagulant and get better and more stable results at the WWTP. He suggested the possibility of controlling the dosage without paying too much attention to the load of suspended solids by having an online correlation between dosage and suspended solid concentration. Mels *et al*. (2002) designed a TU-related polymer dosing method to remove particles from activated sludge effluents. He further suggested that a TU-related polymer dosing system could even control the nitrogen concentration. Aguiar *et al*. (1996) reported that, whatever the origin of the raw water, the optimum dose of coagulant was 2.1 ± 0.2 mg Fe per mg of TOC. Although the TOC was not easily measurable online at the time of their study, the relation of TOC to total solids, which is measurable, was important for real-time CDC. Optical monitor is one of the interesting instruments that can be used for online CDC. Application of optical monitor in water and WW CDC has been reported by several authors. Huhang and Liu (1996) evaluated optical monitor successfully in a laboratory-scale process. Chou *et al*. (1994) used optical monitor to monitor the coagulation and flocculation process. They documented that optical monitoring can give an index for CDC. Eisenlauer *et al*. (1985 and 1987) documented the use of fibre optic sensors for CDC in water treatment.

Using the ultra violet (UV) absorbance of raw water is a well documented alternative to measuring the colour of water. Flower (2004) documented how the use of UV as input to a coagulant control system in water is relatively easy. Sergo *et al*. (2008) used the difference in UV absorption between influent and ultra filtered water in CDC.

The average dose was reported to be reduced from 100 ppm to 75 ppm, using a system based on continuous cationic species demand measurements in a WWTP in Vienna (Sailer, 2002), using a Streaming Current Detector (SCD). The SCD is an instrument which takes an on-line measurement related to the Zeta potential of colloids. The theory and function of SCD is well documented in the literature. (Dentel *et al.*1989; Dentel 1991 and 1995; Elicker *et al*. 1992; Walker *et al*. 1996). The Streaming Current (SC) has been used in many water treatment plants as a control measurement for the coagulation process. Dose control using SC in drinking water has been extensively researched (Mohomad and Dentel 1997; Adgar *et al*. 2005; Sueg-Young Oh *et al*. 2005). Several studies of the control of polymer dose in sludge treatment were also documented (Abu and Dental 1997; Briley 2002; Byun *et al*. 2007). However, the use of SCD as a feedback control in CDC in WWTP is still limited to laboratory and pilot scale tests.

Similar to the SC, to determine the negative charge concentration of an influent in the middle of the process, the raw water can be titrated with a cationic polyelectrolyte, or positively charged iron, or aluminium hydroxo complexes. Since the neutralization point of the negative charge concentration changes due to frequent variations in influent, the charge concentration must be repeatedly determined by titration. Laboratory scale colloidal charge titration for CDC was reported as early as 1967 (Kavamura 1967). An on-line charge titration unit is used for this purpose today. Researchers have studied and proven that the charge titration unit is a successful and available tool for CDC (Bernhardt and Schell 1996). Mattson *et al*. (1996) reported the successful use of a cationic demand measurement unit designed for treatment of storm water in the Rya WWTP, Goteborg, Sweden.

Several researchers have reported on the use of artificial neural networks (ANN) in water and WW technology for CDC. Successful evaluation of ANN in a pilot scale plant was reported by Baxter (2001). Wu and Lo (2008) reported on the use of ANNs and an adaptive network-based fuzzy inference system to control the coagulant dose in water. They concluded that the adaptive network-based fuzzy inference system is better than ANN systems in CDC. Han *et al*. (1997) integrated ANN and Fuzzy logic system to develop a CDC system based on TU, temperature, pH and alkalinity. The fuzzy model was used for normal influent WW and the ANN was used when the WW quality varied. Several other studies for CDC using ANN and fuzzy logic have been reported (Leeuwen *et al*. 1999; Valentin *et al*. 1999; Joo *et al*. 2000; Yu *et al*. 2000; Holger *et al*. 2003; Maier *et al*. 2004; Chen 2006) although successful full scale applications are yet to be reported.

Coagulation is a well defined process. For example, if three identical water samples are added to identical doses, and they coagulate, flocculate and sediment identically, the result will be three identical effluent samples. Such a process can be described mathematically. Construction of a conceptual model, however, may be a challenge due to the complex nature of the process. Based on this, Ratnaweera *et al*. (1994) proposed a concept for a multi-parameter based experimental coagulant dose control system (XCDC). Accordingly, if different influent conditions can be identified with corresponding doses and effluents, it should be possible to estimate the effluent results of new influents with respective dosages. In order to implement this concept in a practical way, the system needs to be modelled mathematically using as much as possible influent, effluent and dosing conditions.

Using a single model, a preliminary evaluation of this concept was made by Lu (2003) and the basic modelling results were presented. The present study elaborates on this work includes an expanded investigation with results obtained with the most suitable multivariate calibration system for model development, a system to identify and manage online water quality monitoring data and full scale tests under various conditions. Finally, the software was restructured to enable a practical implementation at WWTPs.

1.4 Role of online water quality monitoring instruments and challenges

While laboratory jar tests are generally accepted as a valuable tool in estimating the optimum dosages and process conditions like coagulation pH, it does not give a comprehensive understanding of full-scale conditions (Yu *et al*. 2000). Thus, on-line measurement-based dose detection systems are considered by several specialists to be the tool with highest potential.

During the last few decades, on-line water quality monitoring technology has been developed considerably (Jeppsson *et al*. 2002; Vanrolleghem 2003). Today, robust and cheap on-line quality sensors are easily available for most of the relevant parameters. Existing online monitors have been developed based on different measuring principles and different designs with respect to size, calibration and maintenance, measuring principle and reagents, sample requirement, response time and user friendliness (Henrik 1996). Measurements of parameters like flow, orthophosphates, pH, conductivity, SS, TU, Zeta potential, etc. are some examples which are widely used in industry today (Hansen 1996; Lu 2003), though only scarcely for CDC.

1.4.1 Instrument error detection

In real-time process control systems based on on-line measurements, the accuracy of the measurement is critical (Rieger *et al*. 2002). Therefore, a robust tracking system for instrumental errors is very important (Ratnaweera and Blom, 1995). Two different methods of error detection are available on many on-line sensors today. Hardware-based error detection is more common and performed in the sensor or instrument itself by simple logic circuits, installed by the manufacturer to identify measurable limits as well as rapid fluctuations of the measurements. This type of error detection cannot differentiate between unusual process conditions and instrument failures (Colby and Ekster 1997). Software-based error detection, the second category, is used in advance process control systems facilitated by SCADA or PLC network systems. This environment provides more space for analyzing measurements comparing historical values. However, the usage of advanced software based error detection systems is scarce in the water and WW sector. In this thesis, strategies based on the latter concept are elaborated both for detection and minimisation of impacts on dosage prediction. A non linear multivariate calibration method for validation of on-line instruments is discussed. Furthermore, important practical issues which may arise when using such methods on WW are discussed.

2 Experimental methods and procedures

2.1 Laboratory scale tests

The laboratory scale tests were conducted at the Nornalyse laboratory, which is a part of the (NRA), and at laboratories at other treatment plants.

Jar testing was used in screening analysis for selection of the best suitable coagulants for WW and to detect the range of suitable coagulant doses for the influent in each treatment plant.

KEMIRA Flocculator jar test apparatus was used for the studies presented in this thesis. The operation procedures were as follow: 1.0 l of sample was collected in a tall beaker and rapid mixing was started. As soon as the vortex was formed, a designated amount of coagulant was added. The procedure consisted of rapid mixing at 400 rpm for 1 min followed by slow mixing at 30 rpm for 10 min and completed with 30 min of sedimentation. Supernatants of the settled sample were taken for TU, pH, SS, COD, TP and OP measurements as needed.

In addition to the jar tests, several manual sampling campaigns were conducted during the study. The sampling campaigns were mainly used to determine the influent quality, evaluate the effluent quality when the dose was controlled by CDCs, and evaluate the on-line measurements in a study of the relationship between coagulant quality parameters.

Hourly grab sampling was conducted using ISCO automatic samplers. In the laboratory samples TU, pH, SS, COD, TP and OP were measured as needed. Time-scheduled grab sampling campaigns were conducted on the influent and effluent of the WWTP when required.

2.2 Full scale tests

Data collections and full-scale evaluation studies of XCDC were held in NRA WWTP, Lillestrøm, Norway (NRA), HIAS WWTP, Hamar, Norway (HIAS), Xiao-Hong-Men WWTP, Beijing (XHM), China and on a pilot scale in Gaobeidian WWTP, Beijing, China (GBD).

2.2.1 Full scale test facilities: NRA WWTP, Lillestrøm, Norway (NRA)

2.2.1.1 Introduction to the plant

NRA is built in a tunnel of a rock at Lillestrøm, Norway. This WWTP serves the population in three municipalities named Lørenskog, Rælingen and Skedsmo, and several small and large scaled industries in the area. Presently the plant is running with a capacity of about 110 000 p.e. with around 50,000 m^3/day.

The treatment scheme consists of four unit processes (Fig.2). The pre treatment process consists of grit chambers, sand traps and pre-sedimentation. The biological process is based on the floating biofilm reactor series. The effluent from the biological treatment is then treated chemically with coagulation followed by post sedimentation. Sludge from all unit processes is mixed together and treated in a sludge treatment process.

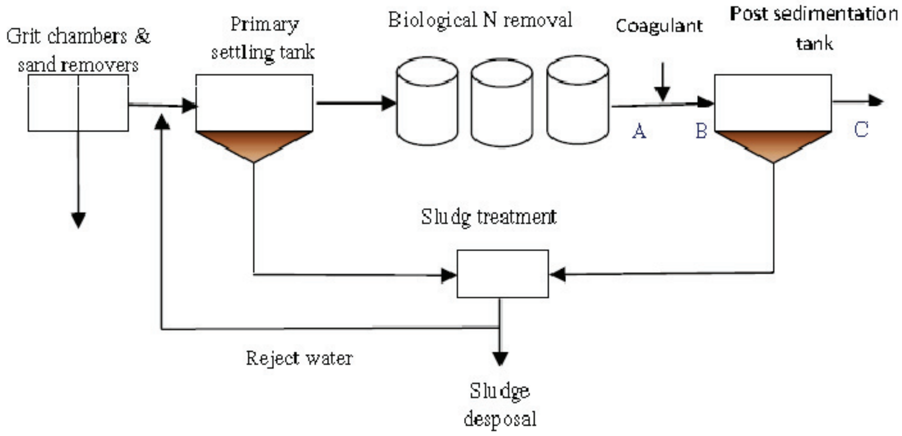


Fig. 2. Schematic diagram of NRA WWTP in Lillestrøm, Southern Norway (NRA). A, B and C *are sampling points for influent sampling, after dosing and effluent sampling respectively.*

2.2.1.2 Challenges related to coagulation

By using a commercial coagulant consisting of poly-aluminium chloride, NRA was managed with good effluent quality resulting in more than 98% of total phosphorus (TP) removal, 91% of COD and 80% of total nitrogen removal. Thus, overall effluent quality was not a challenge at this WWTP. With about 2.5 million NOK/year spending on coagulants alone, an economical optimisation and securing of even better effluent quality stability were still required.

The treatment process has undergone several improvements and modifications during the last few years. The main change was to include a biological treatment process and, at the same time, the influent for the coagulation process was changed from raw influent to the post biological effluent. Although the particles in the new influent to the coagulation process were drastically reduced, the coagulant demand remained more or less unchanged.

The NRA has a modern SCADA system including a real-time flow proportional dosing system with time related coefficients. Apart from the need to periodically adjust the time related coefficients and manually change the coagulant dose per m^3 of WW, it was assumed that there is potential to further optimise the dosing control for the reasons explained earlier. A multiple on-line parameter based real-time dose control system was designed and evaluated for this WWTP.

2.2.1.3 Online Water Quality Monitoring

Treated WW from the biological treatment flowed in two closed pipes and the coagulant chemical was injected into a pressurized part of each pipe in order to mix it well with the water. The treated water from two separate pipes collected into one vertical pipe and lifted up about 7 m to distribute to a horizontal canal. The canal distributes treated water to six sedimentation tanks. This structure of the plant did not allow to measure the quality of WW directly and therefore special sampling tanks were constructed for the measurement of influent turbidity (TUI), influent conductivity (CNI), influent pH (PHI) and temperature (TEI) before coagulant injection (A in Fig. 2), and for pH (PHO) and streaming current (SCO) after coagulant injection (B in Fig.2). The turbidity of effluent (TUO) was measured directly in the effluent ditch (C in Fig.2).

The sensors were connected to a ‘Hach Lange SC100’ integrated controller and the output signals from the SC100s were shared by the SCADA system of the WWTP and the XCDC Programmable Logical controller (PLC). The PLC for the XCDC was connected to the SCADA system and functioned as a ‘Modbus Slave’ to the main system. The dosing pump was controlled by the SCADA system using the information received from the XCDC PLC. All the measurements and dose predictions as a function of time were logged by the main SCADA system and these could be downloaded as 5 minute averages. Full scale application was started on 12th March 2009 and the XCDC is still in operation.

2.2.2 Full scale test facilities: HIAS WWTP, Hamar, Norway (HIAS)

2.2.2.1 Introduction to the plant

The Hedemarken Interkommunale Avløpssamband -HIAS (HIAS) is situated on the eastern bank of Norway’s largest inland water body, Lake Miøsa. The HIAS is owned by four municipalities, Hamar, Løten, Ringsaker and Stange. The WWTP serves about 50 000 people and the main food processing industries in the catchment area with a daily sewage flow of about 20 000m^3.

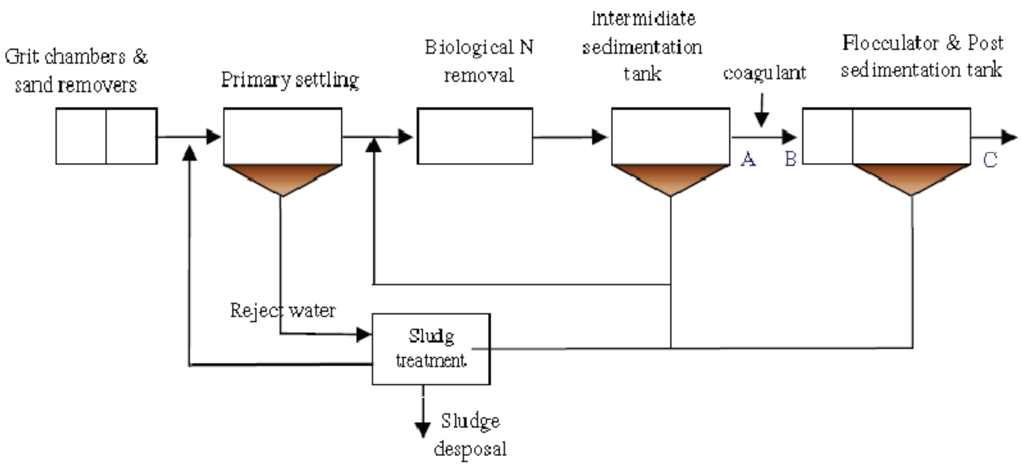


Fig. 3: Schematic diagram of HIAS WWTP in Hamar, South-Eastern Norway (HIAS). A, B and C are sampling points for influent sampling, after dosing and effluent sampling respectively.

The treatment process consists of primary mechanical treatment, grit and sand removers and pre sedimentation, biological treatment followed by a clarifier and chemical treatment process with post sedimentation. The treated water is released to Lake Mijøsa.

2.2.2.2 Challenges related to coagulation

The HIAS was managed with good effluent quality with more than 95% of total phosphorus (TP) removal, and 90% of COD removal. The plant is able to reach the demands of effluents using quite high coagulant dosages requiring frequent operator interventions. The coagulant, which is a commercial aluminium sulphate, cost 2.9 million NOK in 2008. The main challenge of the plant was to reduce coagulant use and also to reduce the requirement for human interventions. A multiple parameter-based experimental CDC (XCDC) was designed for use at this plant.

The next challenge identified in HIAS was the frequent shock loads that occurred in the influent of the chemical treatment. The influent for chemical treatment comes after a biological treatment followed by an intermediate sedimentation. The biological process has a designed maximum capacity of 270 l/s. In practice, when the influent exceeds 250 l/s, the surplus by-passes the biological process and mixes with the other portion of the biologically treated effluent just before the coagulation process. This has resulted in significant changes in the influent quality going into the coagulation step, resulting in shock loads. The quality of the bypassed water is largely different from the influent from the settling tank of the biological treatment. This situation could be generally treated as two different raw water types treated at the WWTP. The first XCDC models were unable to tackle this difference as described in this thesis elsewhere.

To overcome this challenge, we designed a system with two sets of algorithms. Here, one set of algorithms runs until the QI increases up to 250 l/s. Once the QI reaches the 250 l/s limit, the XCDC switches to the second algorithm, which was developed using data from by-passed water.

2.2.2.3 Online water quality monitoring

On-line TUI, CNI, PHI, OPI and TEI were measured in the influent ditch before the coagulant was introduced. PHO was measured in the ditch after coagulant dosing. OPO and TUO were measured in the effluents.

The sensors of all the instruments were connected to 3 ‘Dr Lange SC1000’ controller modules. The output signals from the SC1000 modules were shared among the SCADA system and the PLC of the XCDC. The PLC with XCDC software was installed and connected to the SCADA system as a ‘Modbus master’ unit and then it was able to read data, process data and write back data to the main frame. The dose pump was controlled by the main control system according to the dose predicted by the PLC. All the measurements and dose predictions with time were logged on the main SCADA system and could be downloaded as 10 minute averages. The instruments were well maintained by a skilled operator.

2.2.3 Full scale test facilities: Xiao-Hong-Men WWTP, Beijing, China (XHM)

2.2.3.1 Introduction to the plant

'Xiao-Hong-Men' is one of the 14 largest WWTPs in Beijing. The plant is located in a land area of 47 hectares in Chaoyang District, on the bank of the Liangshui River, which serves as the recipient for treated wastewater. The plant is designed to serve 2.42 million people and industries in a 223.5 km^2 area in Beijing. Present sewage inflow capacity to the plant is 600 000 m^3/d.

The treatment process consisted of four parallel lines. The process consisted of a physical pre-treatment with screening and two sedimentation tanks, biological AAO treatment process designed for both nitrogen and phosphorus removal, followed by post sedimentation tanks. Coagulant is added to the effluent of the AAO process and separation occurs in four sedimentation tanks.

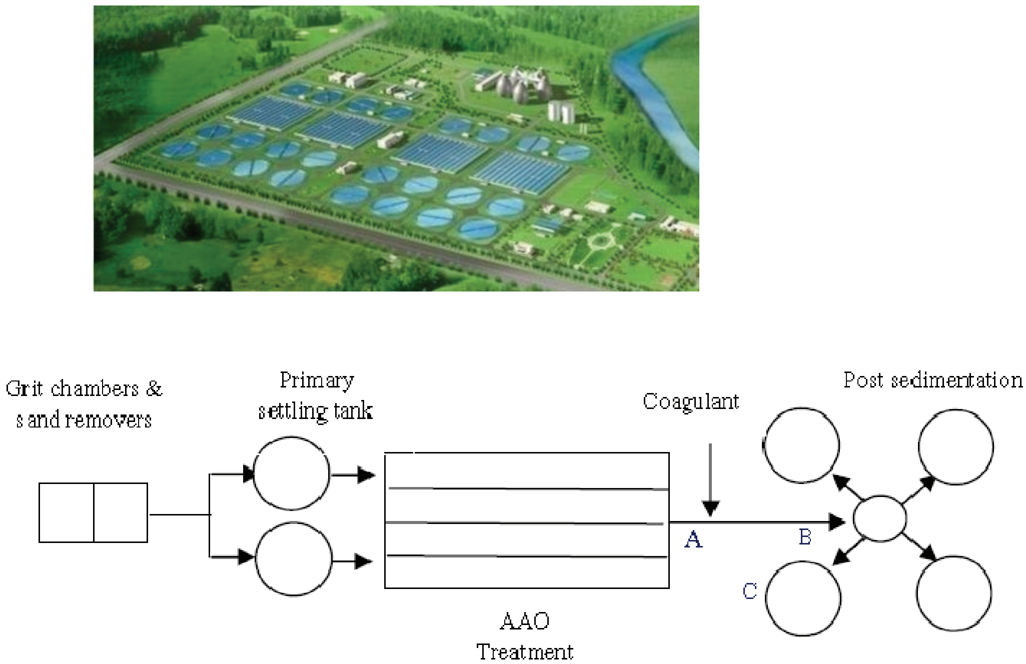


Fig. 4. Bird's eye view and schematic of XHM. Four treatment lines, each with two pre-sedimentation tanks, four AAO lines (each consisted of three lines for anoxic AAO processes) and four post-sedimentation tanks, are clearly visible.

2.2.3.2 Challenges related to coagulation

The plant used a commercial coagulant consisting of a mixture of 7.3% Al_2O_3 and 3.3% Fe_2O_3. Constant dose for four lines was delivered by four manually controllable pumps. The constant dose consumed large amounts of coagulant.

The AAO effluents, to which coagulant was added, were very high in SS, around 3500 mg/l, due to biomass. It is observed that a good part of this SS will be settled even without coagulants, due to the good flocculation and sedimentation properties of WW containing biomass. However, it is known that a portion of coagulant is consumed by the SS in water. Thus, it is possible to obtain an additional reduction of coagulant if they are added to the pre-settled wastewater.

2.2.3.3 Online water quality monitoring

The XCDC trials were carried out in one of the treatment lines. Water from AAO tank effluents before mixing coagulants, after mixing coagulants and outlet of the post sedimentation tank were pumped into three collection tanks designed for measuring the quality parameters. The TUI, OPI, CNI, PHI and TEI were measured in the first tank with AAO effluents before the addition of coagulant. Sampling for OPI was done in a separate tank, thus it did not influence the other influent parameters. PHO was measured in the second tank immediately after coagulant addition and the TUO was measured in the sedimentation tank effluents in tank three. The dosing pump of line A was prepared for automatic control by using a frequency controller to control the pump rpm according to a 4-20mA current signal from the PLC.

The sensors of all the instruments were connected to a ‘Dr Lange SC1000’ controller. The output signal of the SC1000 was transferred to the PLC via analogue input cards. The predicted dose was set to transfer through an analogue output card as a 4-20mA current signal, which was used to control the frequency controller of the dosing pump.

10 minute averages of all the measurements and dose predictions were logged on the PLC and were easily downloaded as a text file. Full-scale testing was conducted during the period 17th June to 14th July 2009.

2.2.4 Full scale test facilities: Gaobeidian WWTP, Beijing, China (GBD)

2.2.4.1 Introduction to the plant

Gaobeidian WWTP is currently the largest sewage treatment plant with biological nutrient removal in China. The WWTP serves a catchment area of about 96 km^2, with design capacity of 1 000 000 m^3/d.

The treatment scheme at GBD is identical to that at XHM, with four treatment lines consisting of physical pre treatment, AAO biological treatment and post chemical coagulation process, followed by sedimentation tanks. As at XHM, the chemical treatment process came just after the AAO process. They used constant dosing with a commercial coagulant containing aluminium sulphate.

The XCDC studies were conducted in a pilot scale treatment plant in the GBD premises. The pilot plant was designed identical to the WWTP with two lines. Pre-sedimentation tank, AAO process and post sedimentation were the components. Initially coagulation was not practiced in the pilot. We designed the pilot plant to add coagulants after settling of the AAO effluent in one sedimentation tank. The coagulated particles were settled in another tank (Fig. 4). Influent to the pilot plant was raw WW from the WWTP and the flow was controlled to a constant 3 m^3 per hour. The capacity of each sedimentation tank was 34 m^3. The retention time was about 11 hrs.

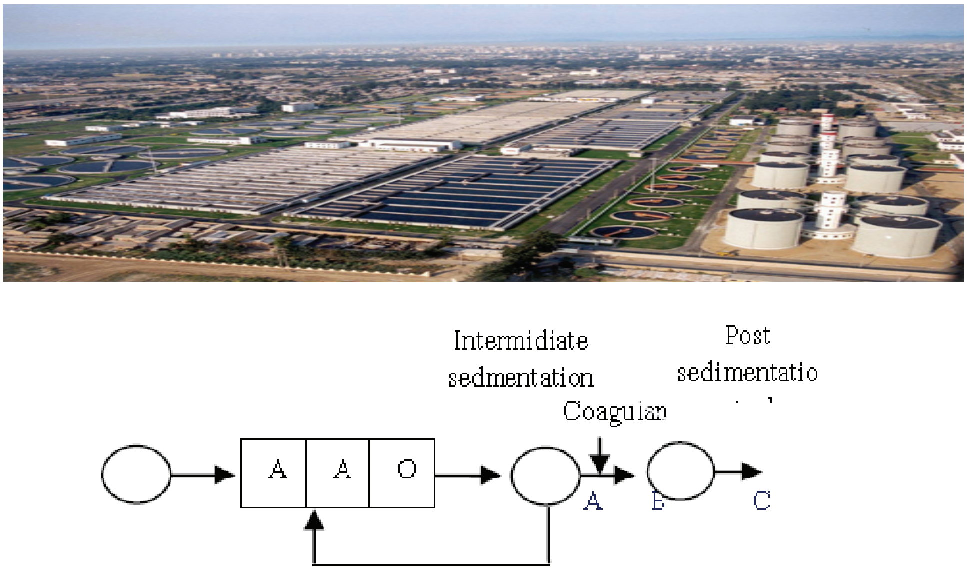


Fig. 5. Bird eye view and schematic of Gabeidian WWTP and schematic diagram of Gaobeidian pilot plant (GBD). A, B and C are sampling points for influent sampling, after dosing and effluent sampling respectively.

2.2.4.2 Challenges related to coagulation

The same challenges described under XHM also applied to GBD.

2.2.4.3 Online water quality monitoring

TUI, CNI, PHI, TEI and OPI were measured in the first post-sedimentation tank, before introduction of coagulant. PHI was measured in the middle of the second post-sedimentation tank where treated water was introduced to the tank. The effluent TUO was measured in effluents of the sedimentation tank. The sensors of all the instruments were connected to a Dr Lange SC1000 controller and, from the controller, the signal was transferred to the PLC controller via analogue input cards. The predicted dose was transferred from the analogue output card as a 4-20 mA current signal, which was used to control the dosing pump. Dosing was done by a digital peristaltic pump, the flow of which could be controlled with 4-20 mA analogue signals. 10 min averages of all the measurements and dose predictions were logged on the PLC and were easily downloaded as a text file.

2.3 Water quality analysis

2.3.1 Online analysis

2.3.1.1 Handling of error measurements

The on-line quality parameters explained in previous section and the dose prediction were logged as 5 or 10 minute averages. The logged data were easily downloaded from either the central SCADA system at the WWTP or from the PLC logs as facilitated by the WWTP. The logged data sets always contained visible errors for several reasons, such as when measurements were taken during maintenance and calibration, due to mal-functioning of sensors, poor maintenance of collection tanks and due to purposely generated error values for study purposes. As the first step, the data set was edited by eliminating error records. Many of these incidents, which were recorded in instrument maintenance records, were easy to identify and to remove from the data set. Some error values, detected by the XCDC software, could also be identified and removed using logged records. The other mal-functioned records were detected manually by comparing to maximum and minimum specified values. The edited data set was used for model calibration of the XCDC.

2.3.1.2 Hydraulic retention time for sample selection

In order to select the most suitable sample sets for calibration of XCDC algorithms, the effluent quality was used. In general, the effluent quality corresponding to a particular influent and dosage was identified with a time shift equal to the estimated retention time of the sedimentation tank.

In practice, sedimentation tanks perform neither under plug-flow conditions nor in complete mixed flow. The situation in a typical sedimentation tank is in between these two situations. The effluent consists of several portions from different batches of influent. In order to study the influent contribution at the end of the sedimentation tank, several tracer tests were conducted in post-sedimentation tanks at NRA.

During these tests, 50ml of Rhodamine B solution (concentration 150g/l) was introduced to the distribution channel in one of the four sedimentation tanks. The detector was placed in the outlet of the sedimentation tank and recorded the Rhodmine concentration in the outlet water against time. The reading was logged every 5 minutes and analysed. The water flow in the channel was also measured.

2.3.2 Experimental coagulant dosing control (XCDC) process

2.3.2.1 Model development and evaluation

Historical data samples with all the on-line parameters, together with respective coagulant doses and time and date, were collected for the preliminary model calibration. The TUO was adjusted as explained elsewhere, to explain the respective effluent quality against coagulant dose. Then the data set was edited by removing identifiable erroneous data, for example, the measurements during the calibration and maintenance work, etc. The effluent turbidity and relevant laboratory measurements were used efficiently to select the best-fit samples for model calibration.

Multivariate calibration is a strong tool for prediction of one or few variables using more than one other regression variable. There are several multiple variable regression calibration methods. First, the most suitable multiple regression calibration method for prediction of future coagulant dose was evaluated using online parameters of WW. Then the best calibration method selected was used to develop models for the XCDC system. The regression models were carefully developed using the statistical software UNSCRAMBLER version 9.8, which is a specialized software for multivariate analysis.

The coefficient of determination (R^2) is the percentage of the total variation in y-values that is explained by the regression equation. Calibration of R^2 and validation of R^2 were used to demonstrate how well the model explains calibration as well as validation data sets.

The root mean square error (RMSE) quantifies the difference between the real value and the estimated value from the model. RMSE for calibration set (RMSEC) and RMSE for Validation set (RMSEP) were used to study the prediction accuracy of the models (Esbensen 2000; Martens and Næs 1991)

The preliminary models were run and evaluated offline, i.e. without coupling to the dosing pump. At NRA, the 1st models were evaluated offline for 45 days, until the plant management was satisfied with the reliability of the dose prediction. After 45 days, the system started real-time dose prediction for the full scale plant. At HIAS, the models were evaluated offline for two weeks and in XHM, the 1st set of models was evaluated off-line for two days. After obtaining satisfactory estimates offline, a phase with active dosing started. An intensive effluent sampling campaign was conducted at the beginning of active dosing in each plant. The system was closely observed and necessary adjustments were made.

The active dosing system was carefully followed with observations and necessary dose adjustments. After two months in NRA and HIAS and four days in the Beijing plants, data were collected and a 2nd model set was calibrated and implemented in the system. The new sets of models were used to predict the dose with close observations.

In HIAS, the dose prediction was affected by sudden stress loads from the bypassed water in the plant. Therefore a new system was designed to handle the issue of bypassed water. The system is discussed under the chapter 3.6. in this thesis.

2.3.2.2 Dose tuning trials

In order to investigate the minimum possible dose of coagulants, a series of trials with modifications to dosage were conducted. As described elsewhere in this thesis, the estimated dosage from the current models still has room for improvement. The objective was to evaluate the level of tolerance to further dosage reduction.

In the procedure, the XCDC predicated dosages were further reduced by up to 40% in stepwise intervals of 10%. The effluent was sampled and analysed during the period.

2.3.2.3 XCDC system hardware and software

A Programmable Logical Controller (PLC) is a specialized microprocessor-based mini-computer developed from computer systems in the late 1960s. A PLC can carry out many types of control functions in industrial processes. The PLC works by looking at its inputs and depending upon their state, turning on/off its outputs according to the software entered by the user. Most PLCs run on the IEC 61131-3 standardized programming language. IEC 61131-3 is a vendor-independent standardized programming language (www.rtaautomation.com) established by the International Electrotechnical Commission (IEC), for industrial automation. The IEC61131-3 standard contains 5 different programming languages: Ladder diagrams (LD), Sequential Function Charts (SFC), Function Block Diagrams (FBD), Structured Text (ST), Instruction List (IL) (Mika Strömman 2002; Real Time Automation, Inc. (Web review))

We used BECKHOFF TwinCAT PLC, (www.beckhoff.com) which is based on the standard IEC61131-3 programming language and we used ST as the software development language.

3 Results and discussion

3.1 Hydraulic retention time and sample selection.

To detect the best relationship between influent, dosage and effluent quality, tracer tests were conducted.

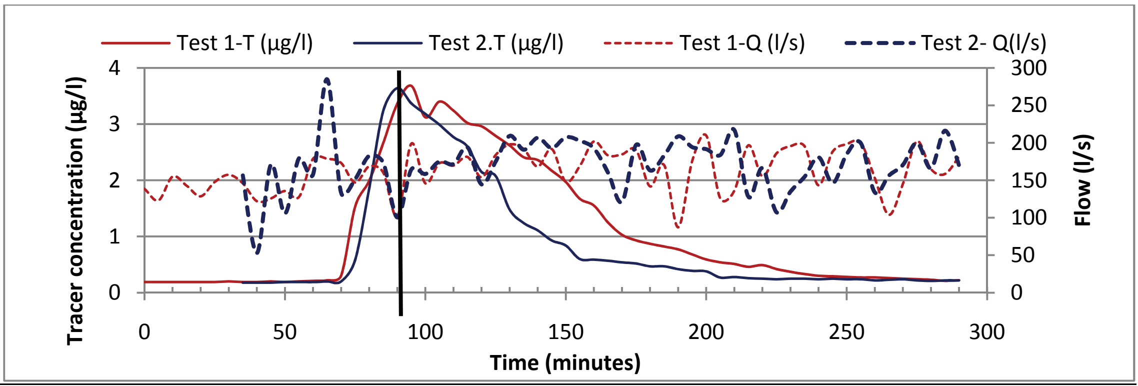


Fig. 6: Results of tracer tests, conducted in NRA WWTP to study flow in the sedimentation tank. In these tests, the tracer, Rhodamin B, was introduced to the inlet of the sedimentation tanks its concentration was detected in effluents as a function of time. Continuous lines represent the concentration of Rhodamin B in the effluents of the sedimentation tank. Dotted lines show the water flow during the test period.

Fig. 6 can be interpreted as follows. The sedimentation tank does not show plug flow behaviour. The Rodamine traces were visible after 50–70 min of dosing and continued during 150–200 min. Thus, it is assumed that a typical influent batch with a specific coagulant dosage will follow the same path. The maximum detected concentration of the tracer in all tests was reached within approximately 90 min which is about 70% of the theoretical hydraulic retention time of the tank.

The time to reach the maximum concentration from the first detection was about 30 min in both tests. This must be due to 'flow-through' areas (short circuiting) in the tank. The delayed removal showing as a long tail in the figure is due to possible dead zones, wall effects and blending effects of the sedimentation tank. The changing real time inflow also influenced the composition of effluents and does not allow us to model the composition of the effluents in a simple way.

With the experience of tracer tests, it is clear that the theoretical hydraulic retention times cannot be used in simplified modelling. However, it was found that a value of 70% of the theoretical retention time enabled efficient modelling

In order to use the effluent as a quality control parameter, the effluent quality parameters were shifted by the corrected retention times and considered as the representative effluent quality for the respective coagulant dose. The effluent quality changes were then compared to the influent quality changes to identify correlations with changes in parameters. Fine adjustments to the effluent time were done accordingly and then the shifted effluent was considered as the result of the corresponding coagulant dose.

3.1.1 Sample selection criteria

Table 1. Sample selection criteria at four different WWTP with different calibrations. TPO (lab) denotes the laboratory measurements of effluent TP. TUO and OPO were on-line measurements.

WWTP	1st calibration	2nd calibration	Remarks
NRA	TUO < 5NTU TPO(lab)< 0.5mg/L	TUO = 2 and 5NTU TPO(lab)< 0.5mg/L	In 2nd calibration, less than 2NTU was considered to be overdosing
HIAS	TUO < 8NTU OPO < 0.1mg/L	TUO < 2 and 7NTU OPO < 0.08mg/L PHO= 6.2 and 7	In 2nd calibration, the pH overrun fraction was used as a constant.
Xiao Hong Men	TUO < 10NTU TPO(lab) < 1mg/L	TUO less than 8NTU TPO (lab)< 1mg/L	
Gaobeidian	TPO < 1mg/L Some doses manipulated according to Jar test results.	TP(lab) < 1mg/L and manipulated doses Lab OPO was observed	TUO was not a good indicator. Only the TPO and dose were manipulated when necessary

The edited effluent quality was then successfully used for sample selection criteria. The sample selection criterions were varied in different WWTP and different calibration procedures. The table 1 below shows the criteria used in each treatment plant for different calibrations.

3.2 Instrumental errors (detection and management)

In the CDC system based on on-line WW parameters, the parameter quality is very important. In the XCDC, several different criteria were used to overcome possible error measurements contributing to dose prediction. The error detection criteria used were:

a) Wide practical range of parameter variation. Here we defined the maximum and minimal potential values for each parameter and the values were tagged as errors when they went out of this range. One of the drawbacks of this method was that the defined ranges are generally too wide and was not sensitive to seasonal changes and sudden changes of water quality. The other weakness was the difficulty in detecting mal-functions within the "valid" range.
b) When the measured value remained unchanged for less than x% for over 2 hours, they were tagged as error values due to a malfunctioning sensor.
c) Non logical measurement relationships: If PHO was larger than PHI; both measurements were considered to be error values. Here we considered only PHI and PHO values when they were not tagged as errors by other error detection methods.
d) A model based on novel error detection criteria is explained in the next chapter.

3.2.1 Model based novel method for error detection

The results presented here are from NRA. WW flow (QI), turbidity (TU), conductivity (CN), pH (PH), temperature (TE) day of week (DAY) and measured hour (TIM) were taken from online instruments. 3000 data samples collected during 12 days were used for the evaluation and demonstration of the method.

Each of TU (using DAY, TIM, Q, PH, CN, and TE), CN (using DAY, TIM, Q, PH, TU, and TE) and PH (using DAY, TIM, Q, TU, CN, and TE) were predicted from the rest of the parameters by the PLS regression method. The data predicting ability of the PLS regression models was evaluated using one long data set ("All", 3000 samples, 12 days), also with 3 random samples ("R1-R3") and 6 shorter data sets each with 500 consecutive samples("G1-G6"). Table 2 shows the regression statistics and Fig. 7 illustrates the prediction of TU compared with measurements.

Table 2. Statistics of regression models developed with all 3000 samples (ALL), three of the random samplings (R1-R3) and six data groups (G1-G6).

	All	R1	R2	R3	G1	G2	G3	G4	G5	G6
R^2 (error)	0.60	0.58	0.63	0.60	0.93	0.89	0.84	0.94	0.88	0.94
RMSE (deviation)	4.47	4.48	3.94	4.52	1.40	1.07	0.98	0.78	0.79	2.02

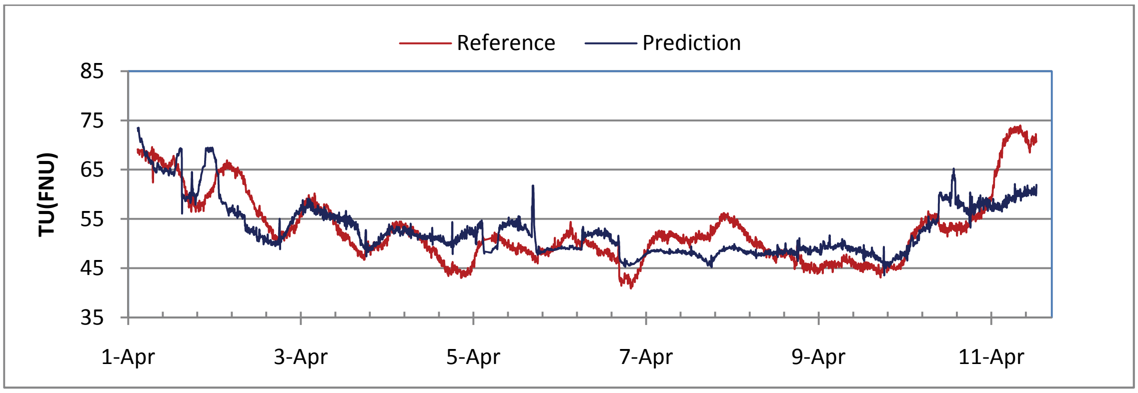


Fig. 7. Predictability of TU when modelling with all 3000 samples.

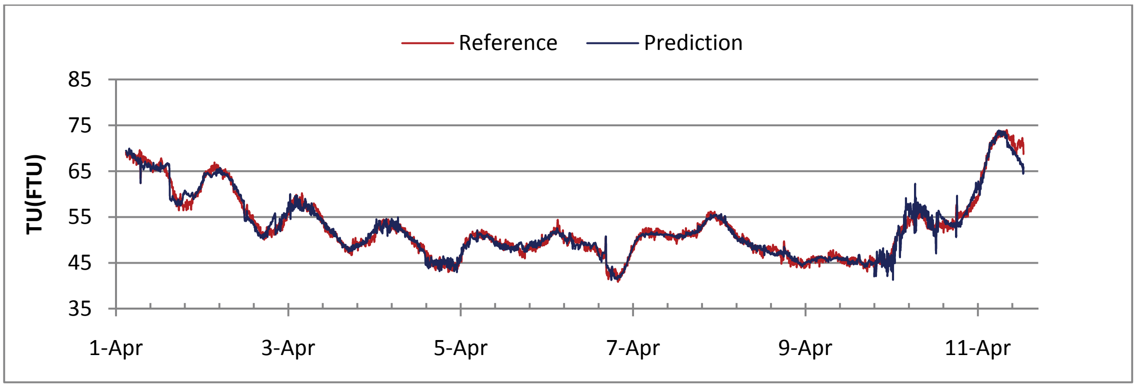


Fig. 8. Predictability of TU when modelling with 500 sample (two days) groups. The figure is a combined figure with all 6 predictions.

The predictability of TU significantly improves when 2-day data sets are used compared with longer sets. The reasons for this phenomena and possibilities for improving the conditions are further discussed in the appendix paper II in this thesis. The other two parameters (CN and PH) also show the same predictive trend.

The error limits were calculated as a percentage of allowable error for the each parameter. In this study, ±5%, ± 3% and ± 3% of the predicted values of TU, CN and PH, respectively, were selected as the error limits. Fig.9 illustrates how the alarm (or validation) ranges for measurements float with time, maximising the error detection accuracy. A, B and C are simulated error values which are well identified as errors beyond the validation levels.

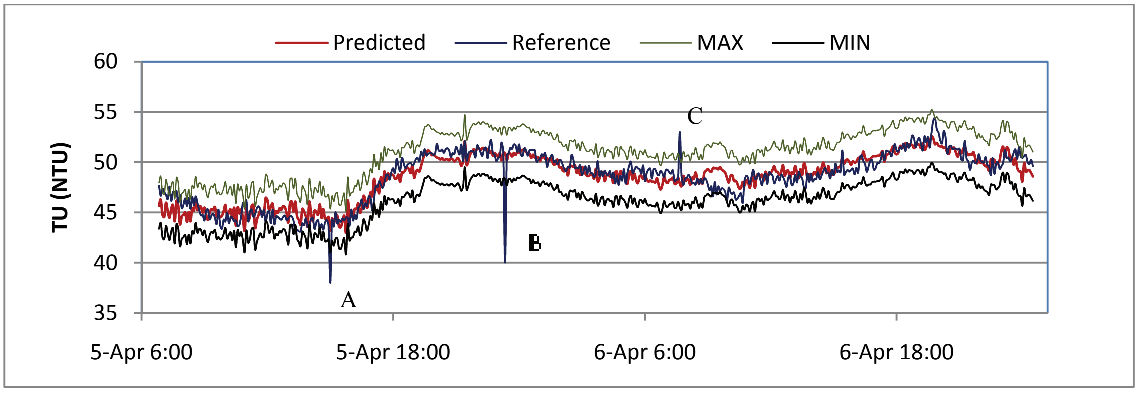


Fig. 9. Demonstration of error detection criteria for TU. A, B and C are the simulated error values. Predicted and Reference: predicted and measured turbidity values. Max and Min: error limits

This concept can be used practically to improve the validation of online instruments to improve process control (eg. where turbidity is a significant parameter in coagulant dosing control). Today, such instruments have wide static alarm levels, set by the manufacturer or user, which create errors in process control. As shown above, the best predictions can be obtained over smaller periods, which may require periodic calibration of the prediction equations. Using statistical software which functions both automatically and online, this issue can be solved efficiently.

3.3 Estimation of optimal dosage

3.3.1 Selection of calibration method

In the literature, various multiple variable calibration methods are used as suitable analytical methods for treating sample sets with specific analytical goals (Johnson and Wichern 1982). Among the different regression methods, multiple linear regression (MLR), principal component regression (PCR) and partial least squares regression (PLSR) are the three popular regression methods used for prediction of data. Henceforth, MLR, PCR and PLSR are compared analytically to select most suitable method for WW dose prediction.

NRA was taken as the case. Total inflow (QIN), turbidity (TUI), CNI, pH (PHI) and temperature (TEI) were measured before dosing with coagulant; while pH (PHO) and streaming current (SCO) were measured just after coagulant dosing and mixing. Also, the real-time coagulant dose (DOS), the date and time were recorded. 18 983 real-time data during three months from September to November 2008 were used.

The data set was divided into 6 groups with each containing 3 000 samples. Then each group of data was used to calibrate model algorithms by all three methods, for predicting DOSE using the other parameters, interaction effects and square effects. Each model that was developed was used to predict the complete sample set with 18 983samples.

In addition, the total data set and several representative samplings of the dataset were used to develop regressions using three different methods. The predictabilities of models developed by different regression methods as well as the same regression method were compared to each other.

Table 3: Correlation coefficient (R^2) and root mean square errors (RMSE) of both calibration (Cal) and validation (Val) samples in each algorithm, calibrated by MLR, PCR and PLSR analysis.

		(1)		(2)		(3)		(4)		(5)		(6)	
		1-3000		3001-6000		6001-9000		9001-12000		12001-15000		15001-18000	
		R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE
MLR	Cal	0.98	2.77	0.99	3.67	0.96	4.0	0.97	4.3	0.95	5.1	0.90	9.0
	Val	0.98	2.86	0.99	3.72	0.96	4.1	0.97	4.3	0.95	5.2	0.89	9.2
PLS	Cal	0.96	3.68	0.98	4.7	0.95	4.7	0.94	5.8	0.89	7.3	0.87	10.4
	Val	0.96	3.68	0.98	4.7	0.95	4.7	0.94	5.8	0.89	7.3	0.86	10.6
PCR	Cal	0.96	3.69	0.98	4.7	0.89	7.0	0.94	5.8	0.89	7.4	0.69	15.0
	Val	0.96	3.69	0.98	4.7	0.89	7.0	0.94	5.8	0.89	7.4	0.69	16.0

Table 3 shows that all MLR models showed the best model statistics. But Figs. 10, 11 and 12 show that PLSR and PCR always showed better prediction ability. When comparing PLS and PCR, PCR showed comparatively better predictability, but the PLS statistics were better.

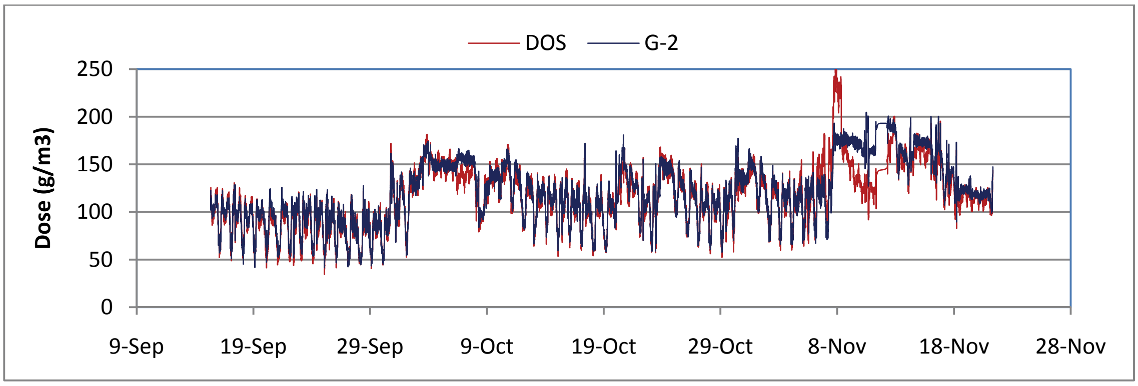


Fig. 10. The prediction of the 'complete dataset' by the PCR model for the 'group 2' data set. (DOS: true dose, G-2: data group with 3000 data)

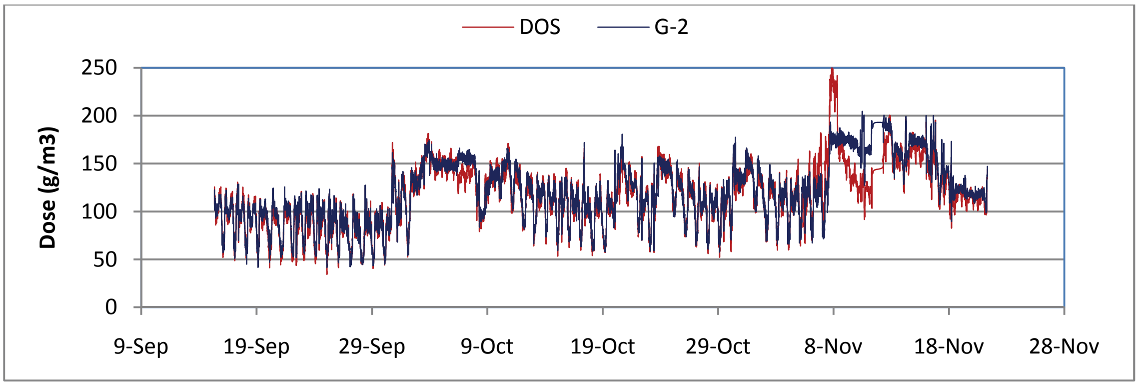


Fig. 11. The prediction of the 'complete dataset' by the PLSR model for the 'group 2' data set. (DOS: true dose, G-2: data group with 3000 data)

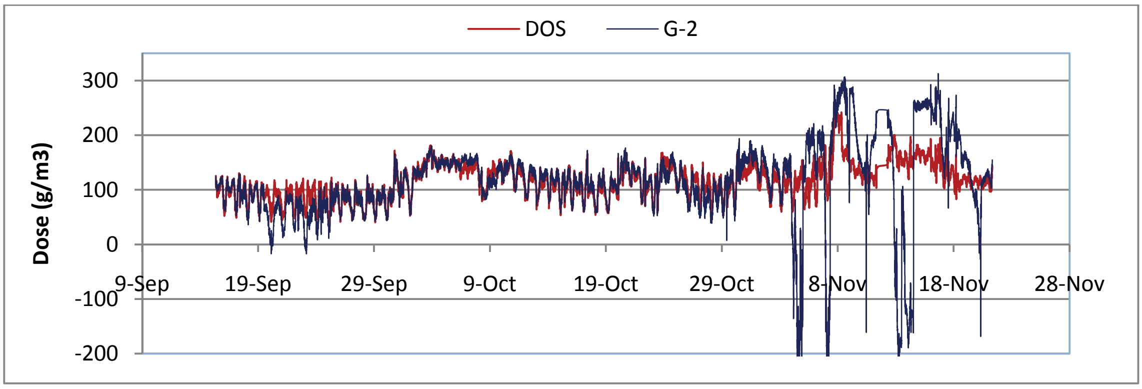


Fig. 12. The prediction of the 'complete dataset' by the MLR model for the 'group 2' data set. (DOS: true dose, G-2: data group with 3000data)

Figs. 10, 11 and 12 show a comparison of the actual DOSE value and predicted values of the complete data set according to the PLS, PLSR and MLR models developed with the group 2 data set. Group 2 was used in this demonstration because it was the best performing data set among all 6 groups. Both PCR and PLSR models show visually the same accuracies of prediction in data while MLR performs poorly.

MLR performs with better model statistics than the other two methods. But it shows poor prediction for out of range data used in the analysis. When we observe the results closely, we can see that the model making the best comparative prediction with its own data set (the data set used in its calibration) was the MLR model. This explains that the MLR calibration is suitable for explaining a data set, but not superior for predicting future values.

When compared to the other analyses which are not shown here, it was clear that the MLR performed poorly with negative dose prediction in most cases. PCR and PLSR did not predict any negative values and both showed acceptable predictability.

PCR and PLSR perform overall equally. But the PLSR model statistics were better, leading us to select PLSR as the best model calibration method for modelling WW coagulant dose.

3.4 Model calibration

The typical basic XCDC model is given below.

DOSE= *f*(TIM, DAY, TUI, CNI, PHI, TEI, OPI, PHO, SCO, interaction among variables, squares of the variables)

In model development, the interaction effects of TIM and DAY with other parameters were not included, because there was no clear relationship among them. The OPI was not available in NRA and SCO was only used in NRA. The other parameters were common to all studies.

In the calibration procedure, the 1st set of the model was run in the system for a specific duration and then the data from the 1st model run were used for the 2nd calibration. The model statistics in the second model set was improved in most studies except HIAS. The data used for the 1st calibration had dose prediction criteria in the WWTP. The dosing systems in WWTPs were either constant dosing or simple flow proportional dosing. The 2nd series of models was developed with the dose predicted by a XCDC model. Therefore it is clear that the 2nd series of models gave better performances. But it is interesting to see that the model performances of the 1st series of models were also were very good. This was demonstrated in dose predictions in NRA and Xiao Hong Men WWTPs. Fig.13 shows that the dose sensitivity has been improved in the second model.

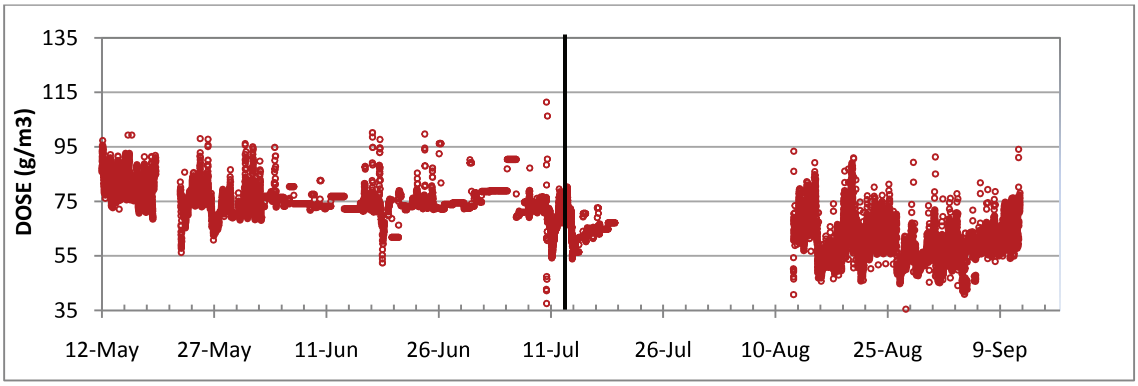


Fig.13. Active doses of the XCDC system. Up to the black vertical line, the preliminary models functioned and then secondary re-calibrated models started to function. (The empty space is lost data due to SCADA system renovation.)

The poor performance of the 2nd series of models in HIAS was due to the influence of shock loads to the WW during the running of the 1st model. This matter is discussed in chapter 3.6.1.

3.5 Dosing control in practice

3.5.1 CDCS results

Fig. 14 shows a comparison of time factor manipulated flow proportional dose in NRA with the CDCS dose prediction.

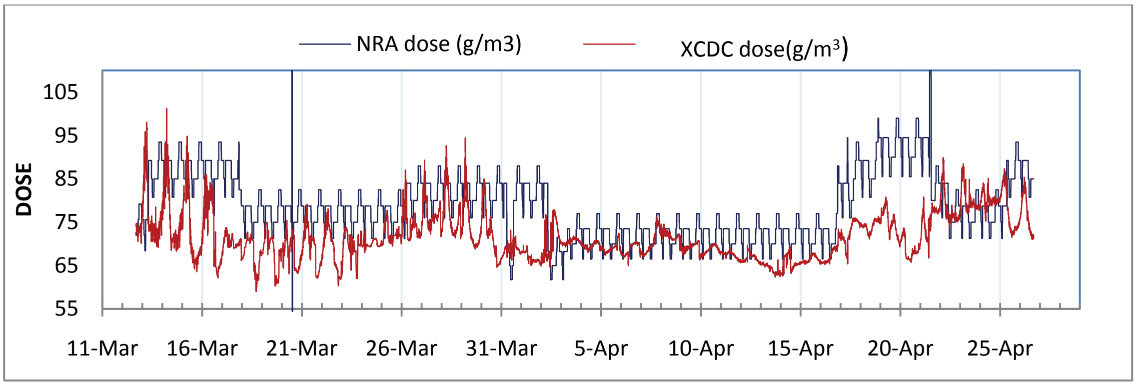


Fig. 14. The WWTP's traditional dosing method (NRA dose) compared to the new dose (XCDC dose) predictions, a comparison during the testing period without active dosing. The changes in NRA dose 'levels' are due to manual changes of plant operators. If not, the dosing should vary in the same base level.

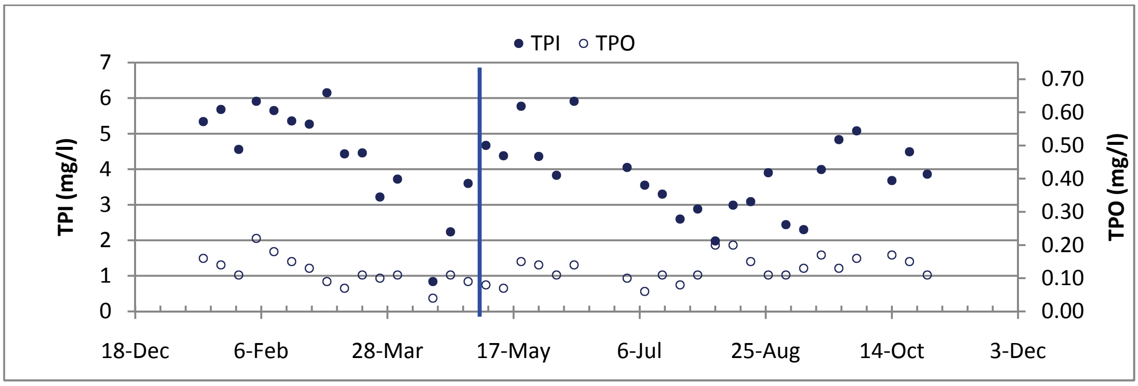


Fig. 15. Influent and effluent TP measurements before and after new dose control system. The blue vertical line divides the figure before and after XCDC implementation.

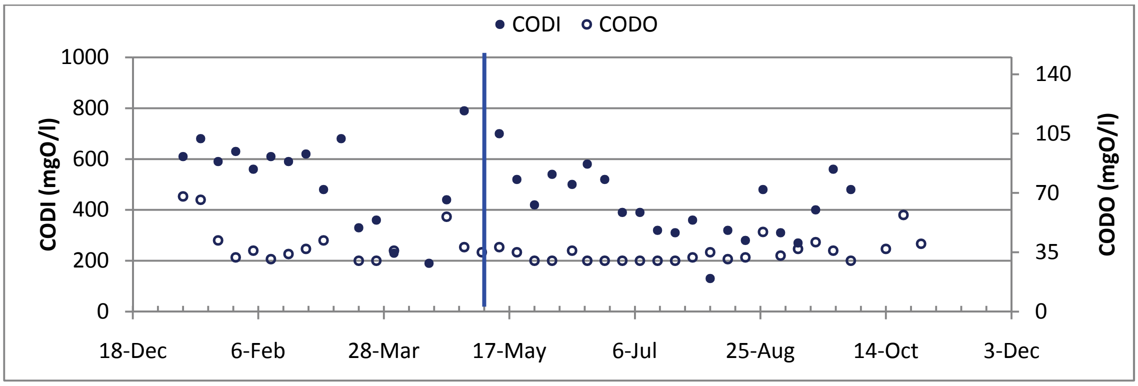


Fig. 16. Influent and effluent COD measurements before and after new dose control system. The blue vertical line divides the figure before and after XCDC implementation.

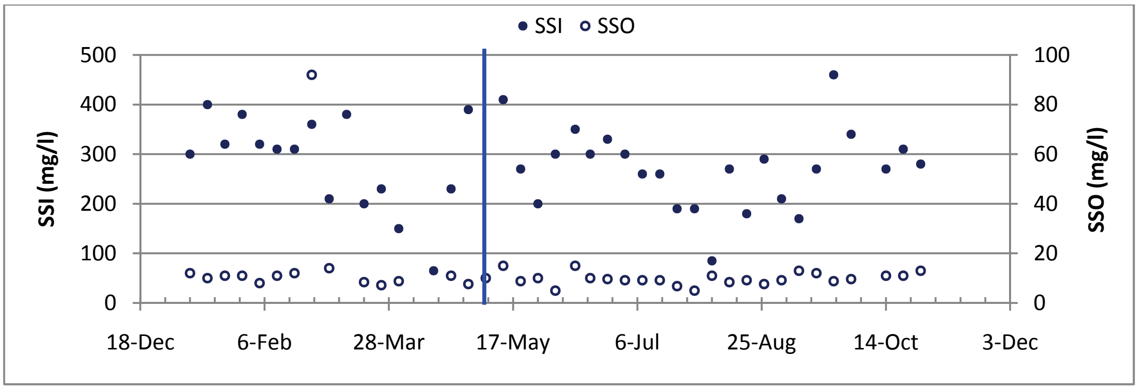


Fig. 17. Influent and effluent SS measurements before and after new dose control system. The blue vertical line divides the figure before and after XCDC implementation.

Figs. 15, 16 and 17 show the influent and effluent total phosphorus (TP), chemical oxygen demand (COD) and suspended solid (SS), respectively, in NRA WWTP. Daily composite samples were measured each week in the laboratory. The effluent there was well maintained by the XCDC.

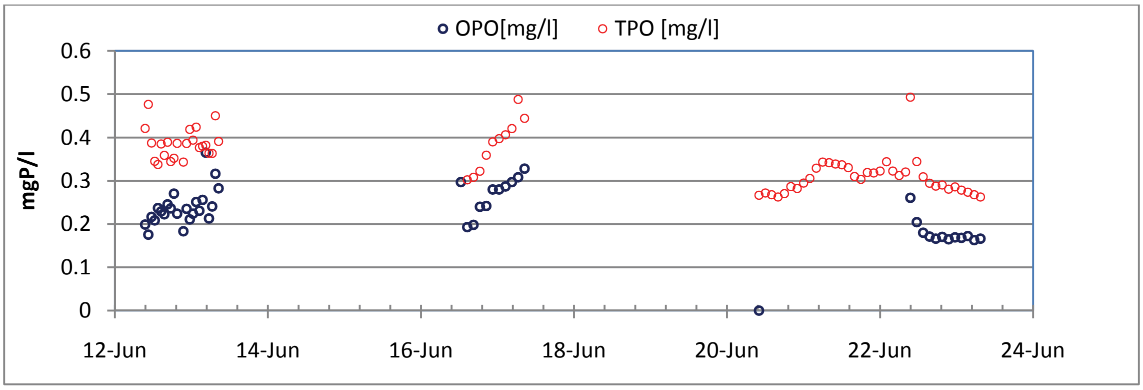


Fig. 18. Effluent OP and TP during the XCDC test at HIAS WWTP.

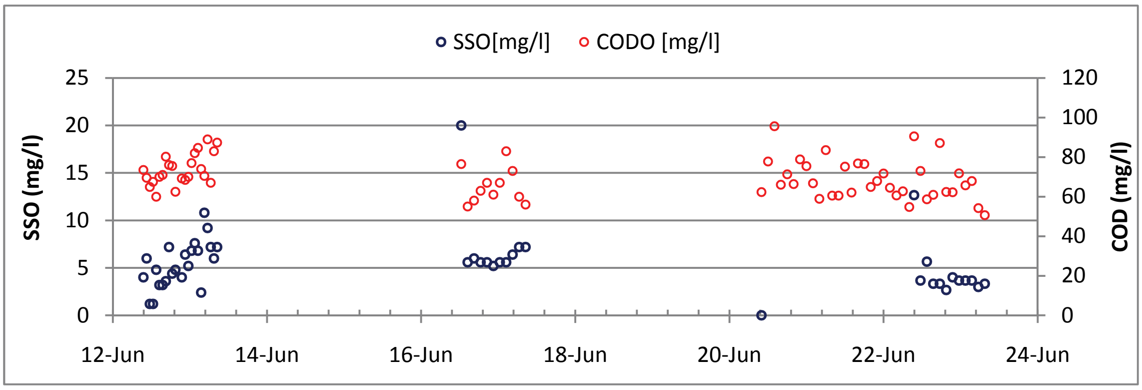


Fig. 19. Effluent SS and COD during the XCDC test at HIAS WWTP.

The figures above (Figs. 18 and 19) are the sampling analysis of the HIAS during the XCDC tests. It shows that the system was able to handle the dose correctly until the stress loads affected the prediction.

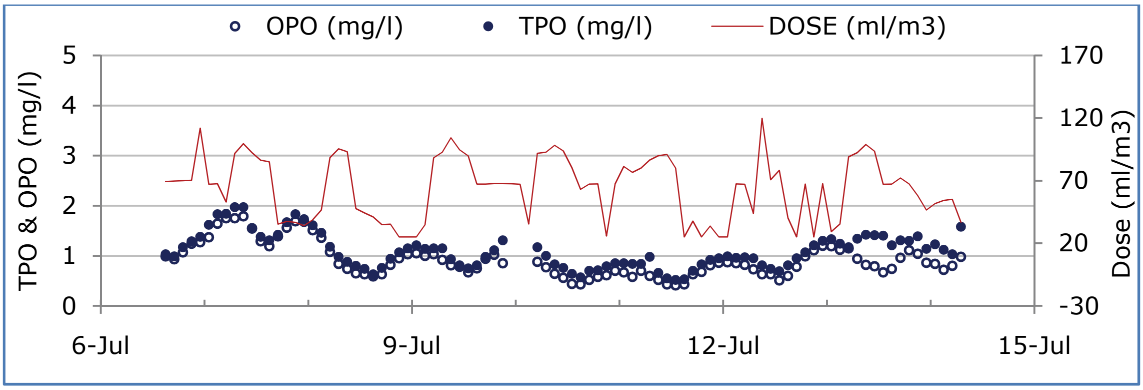


Fig. 20. TP and OP on the effluent of Xiao Hong Men plant.

Fig. 20 with data from Xiao Hong Men WWTP shows how the XCDC prediction influenced the effluent TP limit. In the periods with lower dose, the TP also showed an increasing trend. The most influential parameter there was QIN. The dose prediction was greatly influenced by the large fluctuations of flow.

The short study period at GBD gave some important experiences with the XCDC process. Fig. 21 shows a graph of a complete test. A is the period with constant dosing. B is the period run with the 1st series of calibrated models, C is period with the 2nd series of calibrated models and during D, the 2nd series calibrated models was altered according to specific criteria.

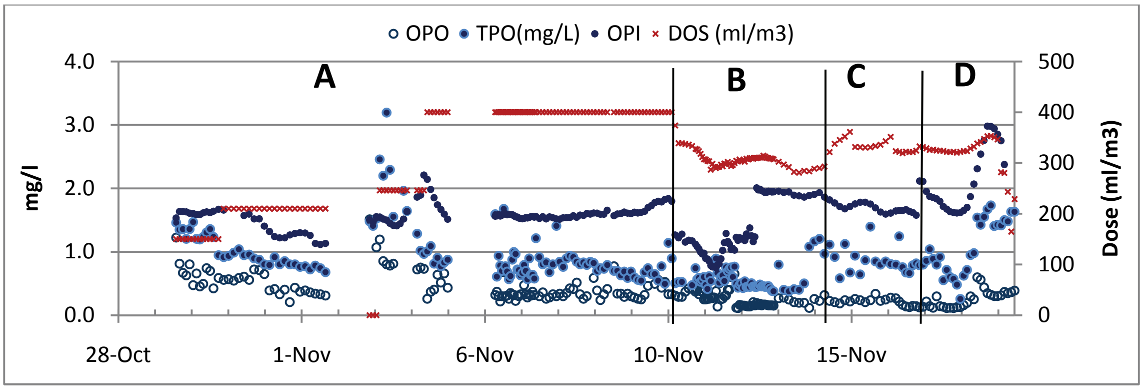


Fig. 21. OPO, TPO and Dose. During test period

The OPO and TPO were well controlled by the 1st model but the 2nd calibration did not give good results. In the second calibration, two important observations were made. As Fig. 22 shows, the model was too sensitive to parameter TIM. The influent waste was much more stable in quality. But TIM was varying from 0 to 23. This variation was too sensitive for the model developed with four days' data. To handle this situation, parameter TIM was removed from the model.

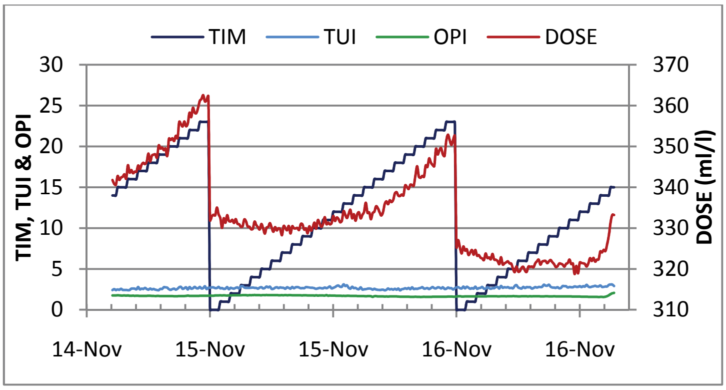


Fig. 22. The influence of TIM was too dominant. The variations of other parameters are not as large as TIM (0-23). Thus, TIM was removed from the model.

Fig. 23 shows that the sensitivity of dose prediction for increasing OPI. The OPI concentration was artificially changed during the study; the XCDCs may need to respond to such events which occur practically in WWTPs as shock loads. As can be seen in Fig. 21, the TP level was not successfully controlled by the coagulant while the OP level was controlled satisfactorily. This suggests that the coagulant used in the plant does not adequately settle the tiny suspended particles. This phenomenon was observed throughout the study period. The observation is discussed in Rathnawera *et al.* 2009.

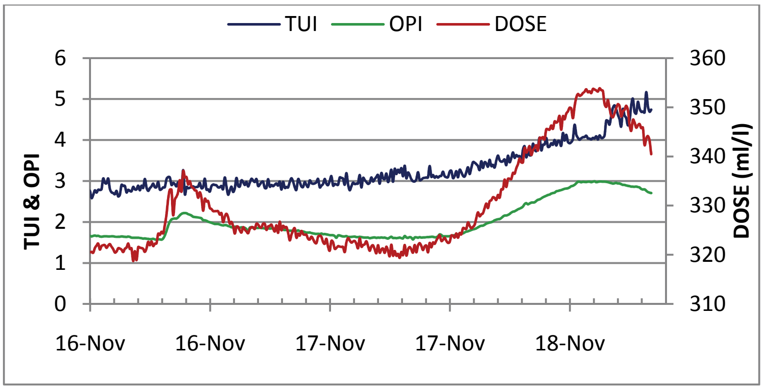


Fig. 23. Sensitivity of dose prediction to increasing OP concentration.

3.5.2 Cost savings from CDCS

The XCDC saves coagulants and reduces sludge production produced by excess aluminium hydrolysis. In addition to this, there are more savings such as labour cost for dose control, transportation of reagents and sludge, sludge treatment and environmental pollution. These are not quantified in this study in order to simplify the calculations.

Estimated annual savings in NRA was approximately 160 t of commercial coagulant, which is equal to 12% of the average traditional coagulant consumption. Thus, the annual savings by XCDC is approximately 250 000 NOK. Furthermore, XCDC improvement studies have proven the ability to reduce the dose by up to 15.6% of the traditional dose.

In terms of sludge production, the content of dry Al_2O_3 in 1 t of coagulant is 0.171 t. The dry Al_2O_3 reduction per year is 27.36 t. At a dryness of 30%, this reduces to 144.8 t of surplus sludge production.

The HIAS was running satisfactorily for about 2 months time until the problems explained earlier were experienced. The system was designed to solve the problem successfully, but yet to evaluate due to some failures in some instruments in the plant. The estimated savings during XCDC operation was about 5 to15% of the annual consumption. Thus, a saving between 140,000 to 418,950 NOK per year is anticipated at HIAS. The XCDC was satisfactorily controlled only during a few days at XHM. When using data during that period, a 25 to 31% saving was calculated.

3.5.3 XCDC improvements

To check the possibility of further improving dose, process sensitivity investigations were conducted at NRA.

The plant was run with reduced predicted doses in order to measure the dose-response curve and the lowest dose that can be achieved in a full scale plant. The actual dose was reduced by up to 40% of the predicated dose, in stepwise intervals of 10%. The effluent quality during this test period is presented in Fig. 24. As expected, a clear increase in effluent quality was found at reduced dose. With a dose of 70% of prediction, two sudden increases of effluent turbidity were experienced. One of these was due to a mal-function of the sludge pump on the sedimentation tanks and once the pump was functioning the TUO was compensated. This had no relationship with coagulation. The second was due to a false dose response to a sudden shock load. The situation was controlled by increasing the dose to 80%. The testing was continued up to 40% of the dose prediction. The sudden peak at 70% indicates that the percentage reductions of a dose prediction will not respond correctly to the changes of parameters.

These results indicate that it has potential in reducing the actual dosage at least by 20–30% of the predicted doses. It is necessary to recalibrate the model with longer data series to cover extreme conditions.

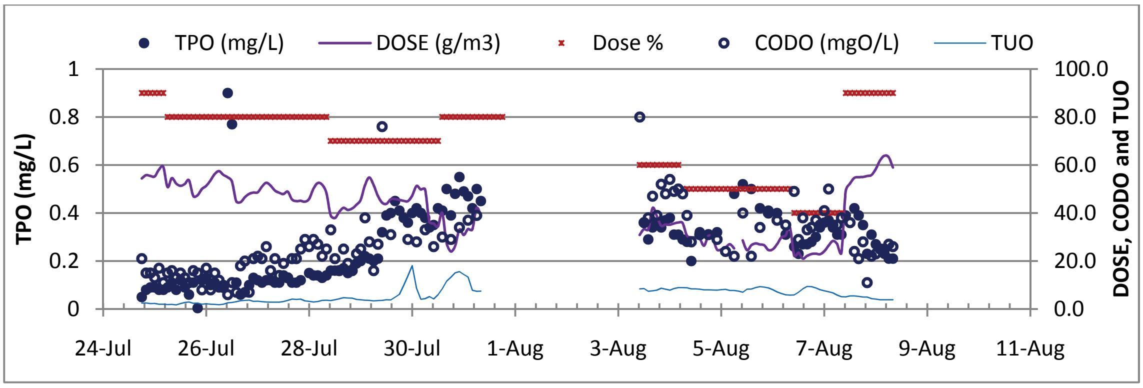


Fig. 24. Effluent TU changes with dose reductions. The first sudden peak in TU was due to a mal-function of the sludge pump and the second peak was due to poor dose prediction in response to sudden changes of parameters.

3.6 Dosing control: challenges and solutions

3.6.1 Different quality influents in HIAS

In HIAS, WW to the chemical treatment comes from biological treatment via the intermediate sedimentation tank. Since it has passed through the sedimentation procedure the water is thin with TUO of about 10NTU. The biological treatment has a maximum designed capacity of 270 l/s. In some periods of the year, the QIN exceeds the maximum capacity of the biological treatment and then the extra water is bypassed directly to the chemical treatment process. These events largely change the water quality in influents to the chemical treatment. Fig. 25 demonstrates the changes in water quality with bypassed water.

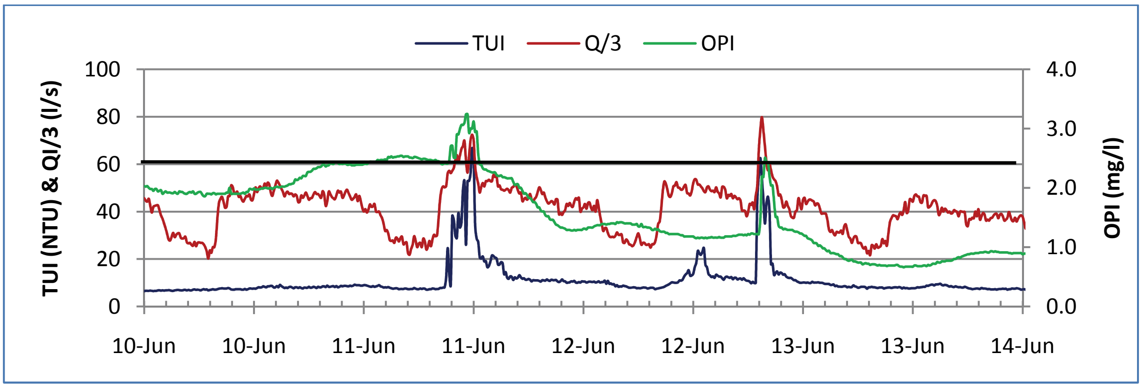


Fig. 25. Shock load at HIAS. The black horizontal line is the overflow limit in biological treatment.

The statistics of the 1st calibration model was extremely good with 0.98 R^2 and 1.40 RMSE in the model with all parameters. The system ran successfully for about 60 days and, during the rainy season, the trial system started overdosing with poor dose predictions. The system was re-calibrated with new data. The new calibration had comparatively poor statistics (0.81R^2 and 8.7 RMSE) even after extensive removal of outliers during modelling. These models did not satisfactorily overtake the shock loads. Later it was identified that, during the first modelling period, the influent was quite homogeneous as the process has no bypass. The bypass had created a new type of influent, which needed to be treated with a different model.

As a solution, the datasets from shock loads were isolated and modelled as a separate set of data. The XCDC software was modified to run two sets of models. The second set of models was set to overtake the dose predictions when the QIN exceeded 250 l/s. Here, we used 250 l/s because the plant used to bypass water from when the QIN is 240–250 l/s. The figure below shows the dose prediction from the two model system.

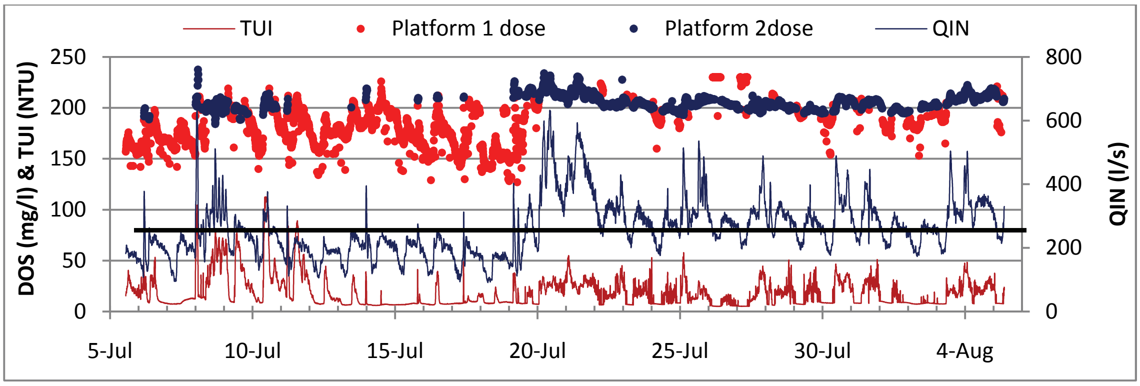


Fig. 26. The two platform model system in function. Platform dose 1 is the dose prediction from the model used for normal inflow. Platform 2 dose is the dose prediction simulated using the model developed using only samples with QIN over 250 l/s. The XCDC software has been designed to shift the platform from 1 to 2, when QIN exceeds 250l/s.

3.6.2 Dosing points in Chinese WWTPs

Coagulants were added to the AAO tank effluents prior to sedimentation in both WWTPs in China. The AAO effluents had more than 3000 mg/l SS due to biomass. It was also observed that, even without coagulants, the SS of this water was easily settled after a few hours to TUO < 10NTU in both cases. In GBD, the TU was below 5 NTU after 5 hours retention time without coagulants.

The literature describes that a part of the coagulant used in WW is consumed by SS in the water, and actually the portions for SS and phosphate removals “compete” with each other. When the SS portion in the water is large, the consumption is larger. In the Chinese WWTPs, the high SS which can be settled without chemical coagulation, competes for available coagulants using a portion of valuable coagulants. It was found in this study that the AAO effluents consume double the coagulant needed for settled water. Figs. 27 and 28 show the laboratory Jar test results comparing AAO effluents and settled effluents in two plants.

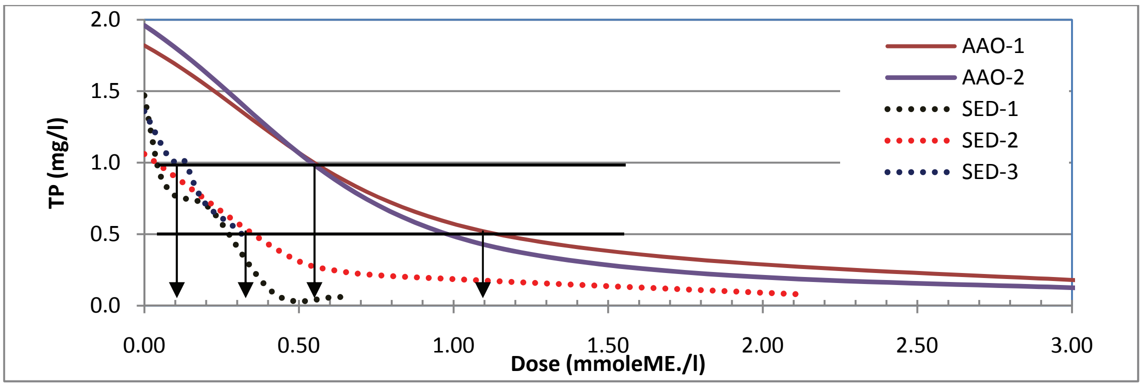


Fig. 27. Comparison of AAO effluent and settled water. Xiao Hong Men WWTP

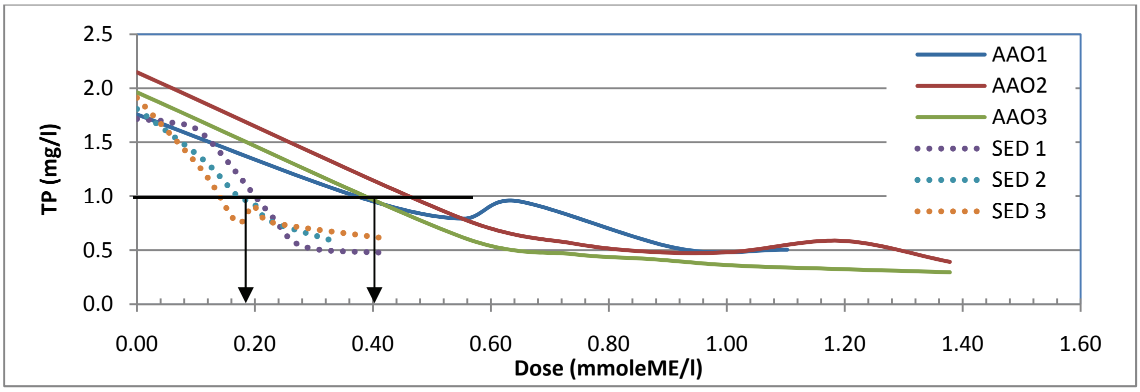


Fig. 28. Comparison of AAO effluent and settled water. Gaobeidian WWTP

The two figures illustrate that the coagulant demand for removing TP below 1 mg/l from AAO was twice that for settled water for GBD and about seven times for XHM WW.

3.7 Role of specific online monitors

3.7.1 Influencing parameters

QI, TUI, CNI, PHI, TEI, OPI, PHO and SCO were the on-line measuring parameters used in the study. The system at NRA WWTP did not have OPI, but SCO was used as a feedback parameter instead. The SCO was not used by any of the other treatment plants.

Using more parameters in the model improves the quality of the model. The figure below shows how the model R^2 and RMSE values varied as more variables were added to the model calibration in NRA.

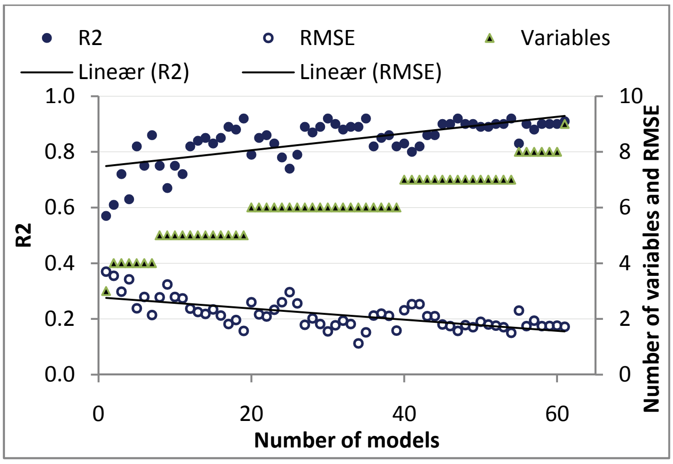


Fig. 29. Change of R^2 and RMSE values with changing number of variables in equations. The two lines indicate the trend in the statistics.

The figure shows that the R^2 values gradually increase with increasing number of variables in the mode. Also, the RMSE trends reduce with increasing number of variables. A complete table of this information is available in the appendix.

3.7.2 Streaming current as a feedback control parameter

As Abu-Orf and Dental (1997) mentioned, their newly installed SCD unit took a few days to stabilize. The SCO measurement used in this study also became much more stable a few days after installation, demonstrating the correct selection of place and installation of the SCD had no errors.

The following example was taken from a coagulant dose changing study. In the study, the coagulant dose was reduced continuously by 10% at a time. Fig. 30 below demonstrates the relationship between TUO and SCO in two different TUO 'peak' situations. The first 'peak' in TUO was not due to a failure in treatment but was due to a malfunction of the sludge pump. But the second 'peak' was definitely due to inadequate coagulant dosing due to a sudden change of flow and influent quality. It is interesting to observe that the SCD has not responded to the first 'peak' while responding the second 'peak'. In addition, the SCO and TUO continuously respond in opposite directions.

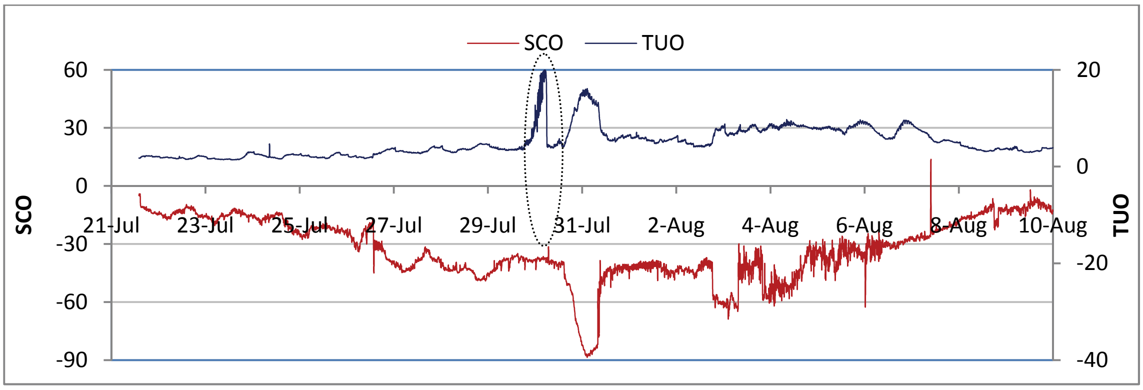


Fig. 30. Relationship between SCO and TUO. The peak marked with a dotted circle in TUO occurred as a result of a mistake by the plant operators. Apart from this peak, there is a clear inverse relationship between TUO and SCO that is revealed (the TUO here has been shifted by 90minutes).

Fig. 31 shows the behaviour of SCO and PHO in response to dose changes. It shows a clear relationship between dose, SCO and PHO. In the test, the dose was increased on 31st July. The PHO has been visibly reduced with increasing dose and PHO increases with reducing dose. The relationship between SCO and dose is also well demonstrated at the time when the SCO goes very low and suddenly rose again with increasing dose. The above two figures demonstrate the appropriateness of SCD measurement as feedback information.

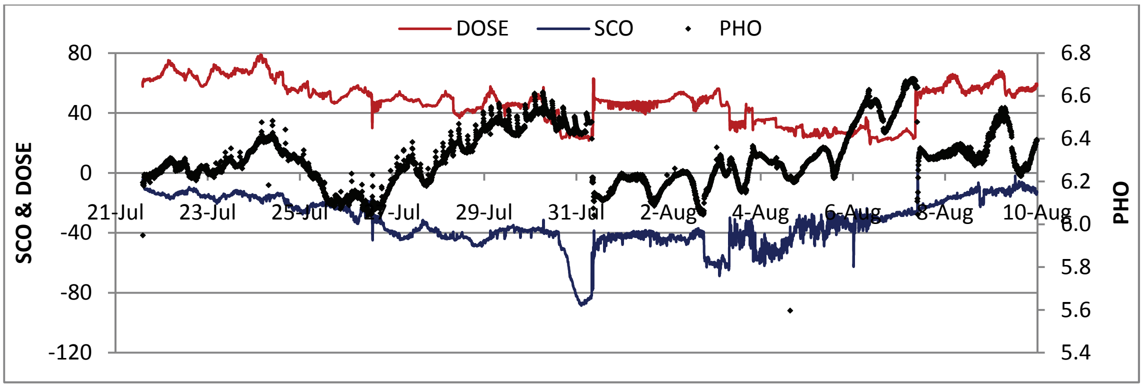


Fig.31. Relationship between dose, SCO and PHO.

This demonstrates that the reduction of dose was increasing pH and at the same time affecting the charge neutralization of colloids. As a result, the poor charge neutralization process resulted in poor effluent quality. This also indicates that the SCO is a better indicator of the treatment process compared to TUO. TUO as feedback information has two major drawbacks. The time delay due to retention time is one drawback. The other is that it can be influenced by other factors not related to process control. Using SCO as an indicator for WWTP operators for process monitoring would be more efficient.

Table 4. Evaluation of the influence of SCD on CDCS models with different combinations of variables. QI, DAY and TIM were used in all the equations and other parameters were changed to make different combinations. (Y: Contained in the model, - : Not contained in the model)

No of variables	DAY	TIM	QI	CNI	PHI	TEO	PHO	SCO	TUI	R^2	RMSE
4	Y	Y	Y	-	-	-	-	Y	-	0.72	2.98
5	Y	Y	Y	-	-	-	-	Y	Y	0.75	2.78
5	Y	Y	Y	-	-	-	Y	Y	-	0.75	2.79
6	Y	Y	Y	-	Y	-	Y	Y	-	0.79	2.56
6	Y	Y	Y	Y	-	-	-	Y	Y	0.89	1.79
7	Y	Y	Y	Y	-	Y	-	Y	Y	0.90	1.74
7	Y	Y	Y	Y	-	Y	Y	Y	-	0.90	1.78
8	Y	Y	Y	Y	Y	Y	-	Y	Y	0.90	1.74
8	Y	Y	Y	Y	Y	Y	Y	Y	-	0.90	1.76
9	Y	Y	Y	Y	Y	Y	Y	Y	Y	0.91	1.72

Table 4 shows some of the possible combinations of models including SCO. The table shows that the R^2 increases as more parameters are added to the equation. However, the R^2 and RMSE values of the model with only SCO were 0.72 and 2.98, which represents considerably good model performance. It changes to 0.91 and 1.72 when all the variables are available. This shows that using SCO as a feedback parameter will improve the security of the system.

The major drawback of SCD is the current price. In addition to this, it requires the frequent attention of operators to prevent it from clogging and to clean the sensor, compared with other instruments. Although Adgar *et al*. (2005) mention strong disturbances by pH, showing a negative correlation, this was not experienced during the tests presented here.

3.8 Process sensitivity to instrumental errors and handling of them

A strong criterion for parameter error detection, which is explained in the above chapter, was integrated into the XCDC. In order to avoid or minimize the negative effect of error dose prediction due to false parameter variables, we designed and evaluated a multiple model-based system. Theoretically, when there are 7 variables, 127 different combinations of equations can be built. Out of them, 1 equation includes all the parameters, which function when all the instruments function well. The next 7 equations excluding one parameter out of seven can be considered to be the situation when one on-line parameter is detected as malfunctioning. The 15 different equations excluding two parameters are equivalent to a situation with 2 parameters caught false. The rest of the 104 equations are obtained when more than 3 parameters are excluded at once (Fig. 32).

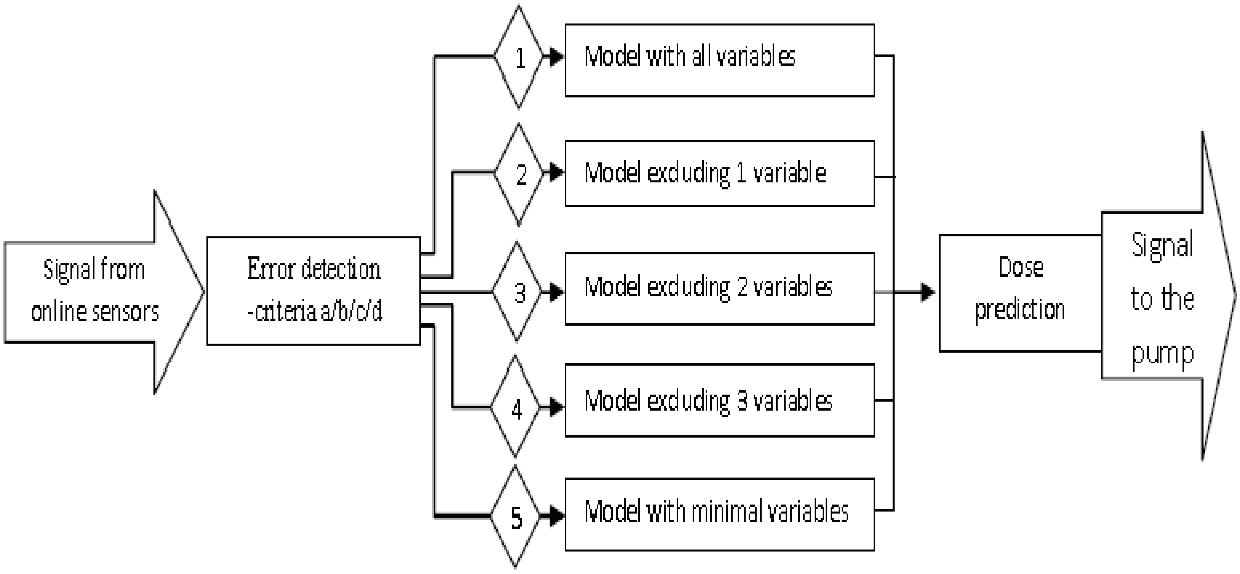


Fig. 32. Structure of error detection function of the multiple model-based XCDC system. Here: 1: All signals functioning, 2: Single parameter malfunctioning, 3: Any combination of two parameters malfunctioning, 4: Any combination of three parameters malfunctioning, 5: Any combination of malfunctioning parameters which is not covered by the first 4 steps.

In the present study, parameter QI was considered to be always correct, because the plant has several flow metres which are reliable and were not reported to have frequent errors. This made the situation simple and reduced the possible equations from 127 to 63. One equation was developed with all the parameters and then 6 more equations were developed, eliminating one parameter at a time. It was considered that the possibility of two instruments malfunctioning simultaneously was just as rare, since the instruments were maintained well during the study period. Thus, only one possible situation of two instruments simultaneously malfunctioning as a control was used. Failures in three or more instruments were considered very rare assuming good maintenance. However, one combination with 3 malfunctioning instruments was tested in the 2nd set of models. A flow proportional prediction equation was used as a basic model, to overtake situations which are not covered by the combinations of equations used in the system. In NRA WWTP, a basic model was developed with parameters QI, TIM and DAY in the 1st models. In second calibration, the model developed with TIM and DAY were not good enough (R^2 = 0.57, RMSE = 3.7). Therefore a simple linear QI proportional equation was selected. Accordingly, in preliminary models 9 different models were used and, in the second calibration models, 10 different models.

In HIAS, 9 models were used in both sets of calibration. In the studies in China, assuming the instruments were well maintained, two basic models, one with all parameters and with minimal parameters, were used. The following figures show some incidents of model shifting in NRA and HIAS.

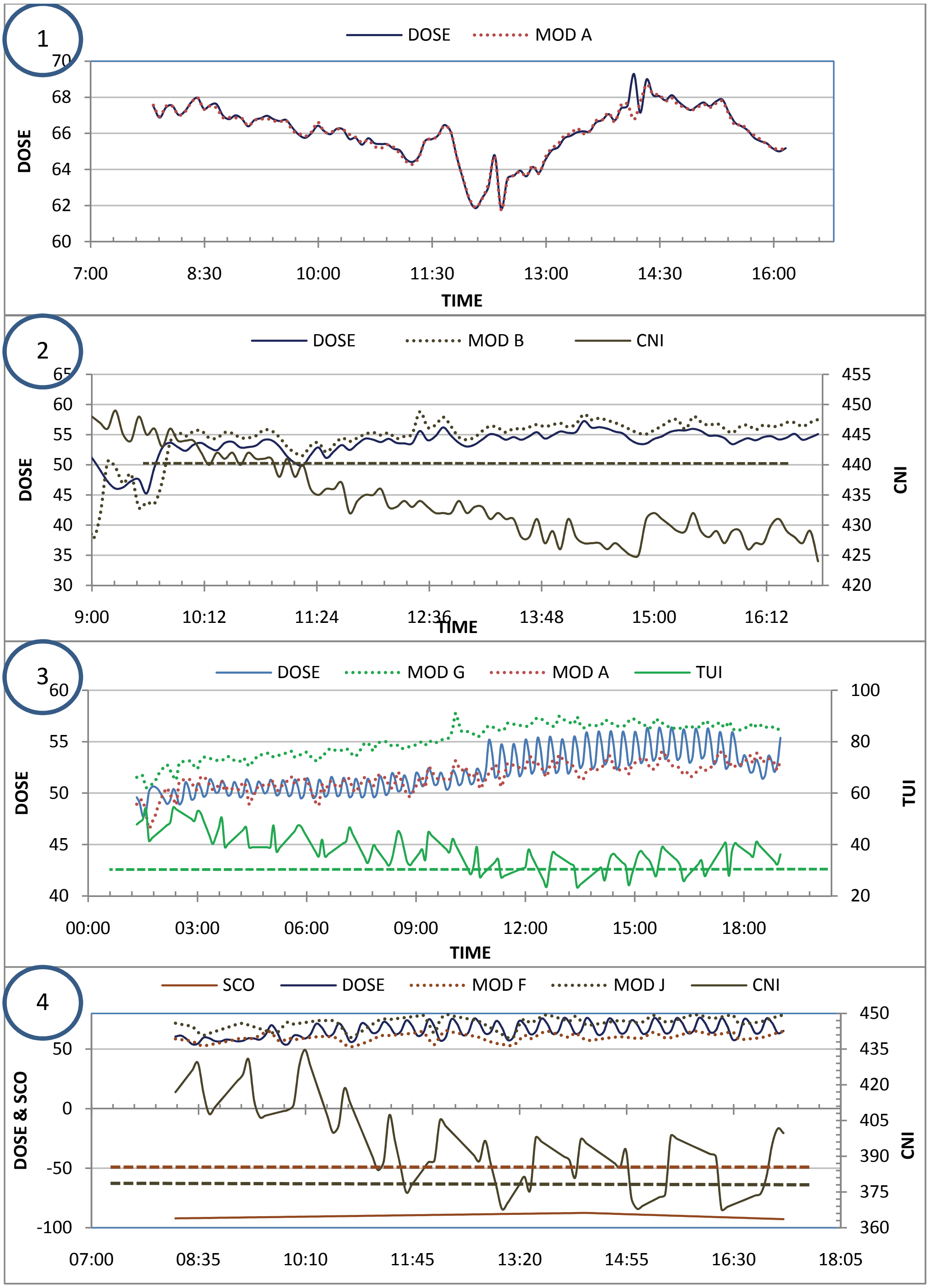


Fig. 33. Model shifting with error detection in parameters. The figures consist of real dose prediction value (DOSE), responsible parameters (abbreviations for parameters) and the simulated doses relevant to the erroneous values (MOD x, x = A-F). The horizontal lines are the error detection limits. Related error detection limit and parameter are in same colour.1: With all instruments working, 2: CNI detected error, 3: Parameter TUI is in the margin and identified as errors from time to time. Model is, subsequently, shifting from full model (A) to

TUI error model (G), 4: CSO is permanently out of the limits. CNI is critical. Model moving among error SCO (F) and Minimal model (J) with CNI detects as error or not.

The results showed that error handling by model shifting was very successful in both WWTPs.

4 Conclusions

- The evaluated experimental Coagulant Dosing Control System (XCDC) was successfully operated with well controlled optimal dose prediction in NRA WWTP.
- Among most popular multiple variable analytical methods, PLSR and PCR regression methods performed superior to MLR in the prediction of new variables from the 'pool' used for calibration and validation. PLSR was finally selected as the best method, due to its best overall performance.
- The robustness of the CDC system depends strongly on the accuracy of online measurements. The hardware-based simple error detection available in some sensors is not sufficient for this purpose. Software-based simple logical error detection methods were successfully and easily integrated into the XCDC. Furthermore, strong model-based floating error detection criteria have also been investigated in the present study
- A system based on multiple models was used to eliminate malfunctioning parameters from dose prediction. The multiple models system functioned successfully when detecting a malfunctioning parameter.
- At HIAS, a significantly different influent quality occurred due to the bypassed water. These sudden changes of influent during a short period of time were not sufficiently overshot by a single set of XCDC algorithms. The influents were treated as if they were from two locations (or two WWTPs), thus two different sets of equations were proposed to ensure the robustness of the XCDC system.
- Having more water quality parameters improved the quality of the model. It is found that when one or more parameters were removed from the model, the statistical quality of the model got worse. The significance of each variable varies according to the importance of the parameter in the system. For example, the CNI was the most significant parameter in XCDC at the NRA WWTP, while TUI was most significant at the HIAS WWTP.
- The streaming current (SC) was found to function as unique and efficient feedback information for the XCDC in NRA.
- Priorities of different WWTPs in different countries are different. Although the majority of WWTPs in the world use coagulation to remove both P and SS, the majority of Chinese

WWTPs are mainly concerned with P removal. They add coagulants to the effluents of traditional biological treatment with AAO treatment prior to sedimentation. The effluents have a very high SS which otherwise is able to settle well by itself, probably due to bio-flocculation or self coagulation. According to current practice with simultaneous coagulation, it is believed that part of the coagulant is unnecessarily wasted.

- The coagulant consumption in AAO effluents (simultaneous coagulation) was generally 200% or more compared to the demand of the settled water. To address the P removal demands in Chinese regulations, efficiently and cost effectively, changing the dosing point from AAO tank effluent to after settling was suggested, although it may require some retrofitting of constrictions.

- The studies showed that the XCDC was able to reduce 12% to 14% of the annual coagulant consumption in a Norwegian WWTP, while retaining effluent quality. Even higher savings are possible but require closer monitoring. Savings up to 30% by XCDC was shown in Chinese plants.

5 Recommendations for further studies

- Using effluent quality for the selection of sample sets can be further optimised. Modelling of sedimentation tank effluent hydraulics should be the focus in the next developments.

- A dual model-based system was designed for HIAS WWTP. The purpose of the system is to handle water of two different qualities in the same treatment plant. This could be applicable to other WWTPs with bypass or periodic addition of reject water and septic waste. The system needs to be evaluated at full scale.

- Feedback parameters for CDC are very important. PHO is the commonly used feedback parameter. The streaming current was evaluated and found to be a good feedback control parameter for CDC. Other cheap feedback control parameters need to be evaluated.

6 References

Abu-Orf, MM & Dentel, SK 1998, ‘Automatic control of polymer dose using the streaming current detector’, *Water Environment Research*, vol.70, no. 5, pp. 1005-1018.

Adgar, A, Cox, CS & Jones, CA 2005, ‘Enhancement of coagulation control using the Streaming Current’. *Bioprocess Biosyst Eng*, vol. 27 pp. 349-357.

Aguiar, A, Lefebvre, E, Rahni, M & Legube, B 1996, ‘Relationship between raw water TOC and the optimum coagulant dose’, *Env. Tech*, vol.17, no. 4, pp. 381-389.

Bache, DH, Ransool, E, Johnson, C & McGilligan, JF 1996, ‘Temperature influences and structure in the sweep floc domain’. *In:Chemical water and wastewater treatment VIII*, HH Hahn, E Hoffmann, H Ødegaard (eds), IWA Publishing, London, pp. 31-40.

Baxter, CW, Zhang, Q, Stanly, SJ, Shariff, R, Tupas, RRT & Stark, HL 2001, ‘Drinking water quality and treatment, The use of artificial neural networks’, Canadian journal of civil engineering, vol. 28, pp. 26-35.

Bernhardt, H & Schell, H 1996, ‘Experience in coagulant control by use of a charge titration unit’*, Aqua - Journal of Water Supply*, *Research and Technology*, vol. 45, no.1, pp. 19-27.

Briley, DS, Knappe, DRU 2002, ‘Optimizing ferric sulphate coagulation of algae with streaming current measurement’*, J. AWWA*, vol. 94, no. 2, pp. 80-90.

Byun, S, Jae-hyun, K, Myung-hak, K, Ki-young, P & Seockheon, L 2007, ‘Automatic control of polymer dosage using streaming potential for waterworks sludge conditioning’, *Separation and Purification Technology*, vol. 57, pp. 230–236.

Chen, CL, & Hou, PL 2006, ‘Fuzzy model identification and control system design for coagulation chemical dosing of potable water’, *Water Science & Technology, Water Supply*, vol.6, no. 3, pp. 97–104.

Chou, S, Lin, S, & Huang, C 1998, ‘Application of optical monitor to evaluated the coagulation of pulp water’, *Water Research*, vol. 37, no. 12, pp. 111-119.

Colby, S & Ekster, A 1997, ‘Control-Loop Fault Detection’, *Water Environ, Technol,* vol. 9, pp. 11- 43.

Dentel, SK 1991a, ‘Characterizing coagulation processes with the streaming-current detector’. *Water Supply*, vol. 9, no. supp, pp. 71-75.

Dentel, SK 1991b, Coagulant control in water treatment, *Critical reviews in environmental control*, vol. 21, no.1, pp. 41-135.

Dentel, S K 1995, ‘Use of the streaming current detector in coagulation monitoring and control’, *J. Water Supply Res*, *And Tech.-AQUA,* vol. 44, no. 2, pp. 70-90.

Dentel, SK & Kingery, KM 1989, 'Theoretical principal of streaming current detection', *Water science and technology*, vol. 21, pp. 443-453.

Devisscher, M, Bogaert H, Bixio, D, Van de Velde, J & Thoeye, C 2002, 'Feasibility of automatic chemicals dosage control a full-scale evaluation', *Water Science and Technology*, vol. 45, no.4–5, pp. 445–452.

Eisenlauer, J & Horn, D 1985, 'Fiber-optic sensor for dose control in flowing suspension', *Colloids and Surfaces*, vol. 14, pp. 121-134.

Eisenlauer, J & Horn D 1987, 'Fiber-optic on-line flocculent dose control in water treatment operations', *Colloids and Surfaces*, vol. 25, no. 3-4, pp. 111-129.

Elicker, ML, Resta, JJ, Hunt, JW & Dentel, SK 1992, 'Fundamental considerations in Use of steaming, curent detector in CDC'. *5th international Gothenburg symp*, Chem, Treat, Nice, France.

Esbensen, KH 2000, *Multivariate data analysis in practice*, CAMO as, Oslo, Norway.

Flower, PRA 2004, 'Optimization of potable water quality and treatment chemical usage through specially developed dosing control systems in the city of Cape Town', *In: Chemical water and wastewater treatment VIII*, HH Hahn, E Hoffmann, H Ødegaard (eds), IWA Publishing, London, pp. 15-28.

Gillberg, L, Nilsson, D & Åkesson, M 1996, 'The influence of pH when precipitating orthophosphate with aluminum and iron salts' *In: Chemical water and wastewater treatment VIII*, HH Hahn, E Hoffmann, H Ødegaard (eds), IWA Publishing, London, pp. 95-105.

Han, TH, Nahm, ES, Woo, KB, Kim, CJ & Ryu JW 1997, 'Optimization of coagulant dosing process in water purification systems', Proceedings of SICA annual conference, Tottori, Japan, pp.1105-1109.

Hangouet, JP, Pujol, R, Bourgogne, P, Ropert, D & Lansalot G 2007, 'Optimizing Chemical Dosage in Primary Settling Tanks', *Chemical Water and Wastewater Treatment IX*, Proceedings of the12th Gothenburg Symposium 2007, Hermann H. Hahn Erhard Hoffmann Hallvard Ødegaard (eds), IWA Publishing, London, pp. 59-67.

Hansen, B 1996, 'Dosing control of coagulants based on on-line monitoring of suspended solids in sewage treatment plants', *Chemical water and wastewater treatment IV*, HH Hahn, E Hoffmann, H Ødegaard (eds), IWA Publishing, London, pp. 137-145.

Henrik AT & Kisby, K 1996, 'N and P online meters. Requirements, maintenance and stability', *Wat. Sci. and Tec*, vol. 33, no. 1. pp. 147-157.

Holger, RM, Morgan, N & Christopher WKC 2003, 'Use of ANN for predicting optimal alum doses and treated water quality parameters', *Environmental Modelling & Software*, vol. 19, no. 5, pp. 485-494.

Homquist, J 2004, 'PIX dosing controlled by double point measurement of phosphorous', *Chemical water and wastewater treatment VIII,* HH Hahn, E Hoffmann, H Ødegaard (eds), IWA Publishing, London, pp. 29-35.

Huhang, C & Liu, CB 1996, 'Automatic control for chemical dosing in laboratory scale coagulation process by using an optical monitor' *Water Research*, vol. 30, no. 08, pp. 1924-1929.

Jeppsson, U, Alex, J, Pons, MN, Spanjers, H & Vanrolleghem, PA 2002, 'Status and future trends of ICA in wastewater treatment- A European perspective', *Wat. Sci. Tech*, vol. 45, no.4–5, pp. 485–494.

Johnson, RA & Wichern, DW 1982*, Applied multivariate statistical analysis*. Fifth edition, Prince hall, USR, NJ 07458.

Joo, DS, Choi, DJ & Park, H 2000, 'Determination of optimal coagulant dosing rate using an artificial neural network', *Water SRT – Aqua*, vol. 49, pp. 49-55.

Kawamura, S 1967, 'Coagulant dosage control by colloid titration technique', *Journal of the American water works association*, vol. 59, no. 8, pp. 1003-1013.

Kemira, 2003, *About water treatment.* Agneta Lindquist(edt), Kemira Kemwater, Helsingborg.

Leentvaar, J, Yawema TSJ & Roersma, RE 1979, 'Optimisation of coagulant dose in coagulation, Flocculation of sewage', *Water Research*, vol. 13, pp. 229-236.

Leeuwen van, J, Chow, CWK, Bursill, D & Drikas, M 1999, 'Empirical mathematical models and artificial neural networks for the determination of alum doses for treatment of southern Australian surface waters', *J. Water Supply Res. And Tech.-AQUA*, vol. 48, no. 3, pp. 115-123.

Lu, L, 2003, 'Model based control and simulation of wastewater coagulation', PhD thesis, Agriculture University of Norway.

Lu, L, Ratnaweera, H & Lindholm, O 2003, 'Coagulant dosage control in chemical wastewater treatment plants- a review of modeling approaches', *Vatten*, vol.59, pp. 227-235.

Lu, L, Ratnaweera, H, Lindholm, O & Lileng, K 2002*,* 'Model-Based Real Time Control of Coagulant Dosing'. *Proceedings International IWA Conference on Automation in Water Quality Monitoring*, N. Flesischmann, Lengergraber, G & Haberl, R (eds), University of Agricultural Sci., Vienna, pp. 421-424.

Maier, H R, Morgan, N, Chow, CWK 2004, 'Use of artificial neural networks for predicting optimal alum doses and treated water quality parameters', *Environment Modelling & Software*, vol. 19, no. 5, pp. 485-494.

Martens, H & Næs, T 1991, *Multivariate calibration*. John Wiley & Sons Ltd, New York.

Mattsson, A, Sörensen J, Bengtsson J 2004, 'Strom water treatment with cationic demand dosing control', In: *Chemical water and wastewater treatment VIIl*, HH Hahn, E Hoffmann, H Ødegaard, (eds), IWA Publishing, London, pp. 69-76.

Mels, AR, van Nieuwenhuijzen AF & Klapwijk, A 2002, 'Turbidity related dosing of organic polymers to control the denitrification potential of flocculated municipal wastewater', *Chemical water and wastewater treatment VII*, HH Hahn, E Hoffmann, H Ødegaard (eds), IWA Publishing, London, pp. 71-79.

Mika, S 2002, 'Implementation of the SA/SD-method in PLC-programming', *AS-116.140 Postgraduate Seminar of Information and Computer Systems in Automation*, Tekniska högskolans studieportal Noppa.

Mohomad, M &. Dentel SK 1997, 'Polymer dose assessment using the Streaming', *Water environment research*, vol. 69, no. 6, pp. 1075-1085.

Rathnaweera,S, Ratnaweera,H & Lindholm O 2009, 'Multi-parameter based dosing control as an efficient tool for improved phosphate removal by coagulation- experiences from Beijing', Paper presented at International Forum on Environment Simulation and Pollution Control at Beijing, 13-14 Nov. 2009.

Ratnaweera, H 1991, 'Influence of the degree of coagulant preploymerization on wastewater coagulation mechanisms', Doctoral thesis, Norwegian Institute of Technology, Dissertation publishing, University Microfilms International, Michigan, USA.

Ratnaweera, H 2004, 'Coagulant dosing control- a review', *Chemical water and wastewater treatment VIII*, HH Hahn, E Hoffmann, H Ødegaard (eds), IWA Publishing, London, pp. 3-13.

Ratnaweera, H, Blom, H & Aasgaard, G 1994, *Flexible Coagulant Dosing Control System based on Real-Time Wastewater Quality Monitoring,* Chemical water and wastewater treatment III, Hahn, H.H. and Klute, R (eds), Springer-Verlag, Berlin, 105-116.

Ratnaweera, H, Blom, H 1995, 'Optimization of coagulant dosing control using real-time models selective to instrumental errors', *Water Supply*, vol. 13, no. 3-4, pp. 285-289

Ratnaweera, H, Lu, L & Lindholm, O 2002, 'Simulation Program for Wastewater Coagulation', *Chemical Water and Wastewater Treatment VII*, H.H.Hahn, E.Hoffmann and H. Oedegaard (eds), Gothenburg, Sweden, IWA publishing, London, pp. 253-260.

Real Time Automation, Inc., Control IEC 61131-3 the Fast Guide to IEC 61131-3 Open Control Standard & Software. http://www.rtaautomation.com/iec61131-3/

Rieger, L, Alex, J, Winkler, S, Boehler, M, Thomann, M & Siegrist, H 2002, 'Progress in sensor technology- progress in progress control? Part I: Sensor property investigation classification'. In: *Proceedings International IWA Conference on Automation* in Water Quality Monitoring. N. Fleischmann, G. Langergraber and R. Haberl (eds), University of Agricultural Sciences, Vienna. pp. 65-72.

Sailer, E 2002, 'Fully automatic dose control of coagulants by using particle chage', *Chemical water and wastewater treatment VII,* HH Hahn, E Hoffmann, H Ødegaard, IWA Publishing, London, vol. 7, pp. 81-90.

Sergio, G, Salinas, R, Maria, DK, Aleid, D, Hilde, P & Jan, CS 2008, 'Optimization of PACl dose to reduce RO cleaning in an MS', *Dealination*, vol. 220, pp. 239-251.

Sueg-Young, O, Doo-Gyoon, B, Jae-Moon, H & Hyun-Sung, S 2005, 'Automatic Control on Dosing Coagulant as to Stream Current', *ICCAS2005, KINTEX*, Gyeonggi-Do, Korea.

Valentin, N, Denoeux, T, Fotoohi, F 1999, 'Modeling of coagulant dosage in a water treatment plant', *'99 Fifth International Conference on Engineering Applications of Neural Networks,* Warsaw, Poland, pp.13-15.

Vanrolleghem, PA &. Lee, DS 2003, 'On-line monitoring equipment for wastewater treatment processes, State of the art', *Water Science and Technology*, IWA Publishing, vol. 47, no. 2, pp. 1.

Walker, CA, Kirby, JT, Dentel, SK 1996, 'The streaming current detector, A quantitative model', *J. Colloid and interface science*, vol. 182, no. 1, pp. 71-81.

Wu, Guan-De & Lo, Shang-Lien 2008, 'Predicting real-time coagulant dosage in water treatment by artificial neural networks and adaptive network-based fuzzy inference system', *Engineering Applications of Artificial Intelligence,* vol. 21, pp. 1189– 1195.

Backhoff, www. backhoff.com- Backhoff PLC tutorial. (*download.visualcomponents.net/.../Tutorials/Beckhoff/PLCAdd-on%20Beckhoff.pdf)*

Yu, R F, Kang, S F, Liaw, SL & Chen, MC 2000, 'Application of artificial neural network to control the coagulant dosing in water treatment plant', *Water Science & Technology*, vol. 42, no.3-4, pp. 403–408.

Zeghal, S, Philippe, J, Sauvegrin, P & VignolesCh 1996, 'Chemical addition control for phosphorus removal in primary sedimentation Tanks', *In:Chemical water and wastewater treatment VIII*, ed HH Hahn, E Hoffmann, H Ødegaard, IWA Publishing, London, pp.147-15.

Zeta meter inc, 1993, *Everything about Coagulation*. Zeta-meter inc. Staunton, Virginia.

7 Apendix - Publications

7.1 Multiple Model-Based Coagulant Dosage Control System

Subhash Rathnaweera, Harsha Ratnaweera, Tor Håkonsen and Oddvar Lindholm
Paper submitted to the International Journal of Environmental Technology and Management (IJETM)

7.2 Modelling coagulant dosage in wastewater treatment plants, using MLR, PCR and PLSR statistical analysis.

Subhash Rathnaweera, Harsha Ratnaweera, Tor Håkonsen, Oddvar Lindholm, Ellen Sandberg., Paper submitted to the International Journal of Environmental Science and Technology

7.3 Improving process control by advanced error detection using floating validation ranges of online measurements

Subhash Rathnaweera, Harsha Ratnaweera, Oddvar Lindholm, Tor Håkonsen
Paper is accepted to the 7th Leading Edge Conference on Water and Wastewater Technologies (2-4 June 2010).

7.4 Multi-parameter based real-time coagulant dose control system for wastewater treatment

Subhash Rathnaweera
Paper submitted to the International Journal of Environmental Modelling & Software.

7.5 Multi-parameter based dosing control as an efficient tool for improved phosphate removal by coagulation- experiences from Beijing

Subhash Rathnaweera, Harsha Ratnaweera, Oddvar Lindholm
Paper presented at International Forum on Environment Simulation and Pollution Control at Beijing, 13-14 Nov. 2009.

Paper I

Multiple Model-Based Coagulant Dosage Control System.

Subhash Rathnaweera, Harsha Ratnaweera, Tor Håkonsen, Oddvar Lindholm

The paper submitted to the International Journal of Environmental Technology and Management (IJETM)

Multiple Model-Based Coagulant Dosage Control System

Subhash Rathnaweera*†, Harsha Ratnaweera**, Tor Håkonsen*** Oddvar Lindholm*

*University of Life Sciences, Aas, Norway
**Norwegian Institute for Water Research (NIVA), Oslo, Norway
*** VA-Support AS, www.va-support.no
† Corresponding author (subhash.rathnaweera@umb.no)

Abstract

A novel real-time water quality parameter-based coagulant dosing control system is evaluated in full scale treatment plants. Erroneous measurements from real time monitoring equipment may create critical conditions for the coagulant dosing control system. The usage of multiple models which excludes erroneous parameters is evaluated. The need for several model sets even in one treatment plant is discussed to improve dosage predictions when the influent is subject to variable treatment methods and foreign inputs like periodic septic discharge. The robustness and accuracy of the evaluated automated dosing control system can be improved significantly by following these approaches.

Introduction

Optimizing coagulant dosage in water and wastewater treatment has become more and more critical in order to maintain effluent quality with minimal effect on the environment at minimal cost. The chemical costs of an average wastewater treatment plant (WWTP) represent up to 20% of the operating cost of a treatment plant (Hangouet *et al.* 2007)

It is well documented that optimal dosage has a strong effect on the influent water quality parameters such as turbidity, colour, pH, phosphate, temperature, etc. Also, the wastewater quality changes significantly during a season, a week or a day and time of the day (Sagberg *et al*. 1990; Buttler *et al.* 1995). This indicates that coagulant dose would be varying with time according to changes in quality.

With the development of cheap and convenient online sensors for water quality measurements, their usage in real-time coagulant dosing control is becoming popular today. Use of flow meters, sometimes together with one or two of these quality parameters, are probably the most common applications in water and wastewater treatment facilities around the world. An investigation into Norwegian water and wastewater treatment plants (WWTP) has reported that 80% of drinking water treatment plants and 83% of wastewater treatment plants control their coagulant dosage either by flow proportional dosage or with pH overrun functions (Ratnaweera 2004).

Using more influent parameters would be more efficient to determine the optimal coagulant dose instead of using one or two of them. A novel coagulant dose control system using multiple parameter measurements (XCDC) has been developed and preliminarily evaluated (Ratnaweera *et al*. 2002; Lu 2003; Lu *et al*. 2003; Rathnaweera *et al.* 2010a).

Although novel online measuring instrument sensors are less complicated and more reliable (Olsson *et al*. 1998), poor performance due to technical or installation failures are still common and to be expected in systems. The XCDC system is based on measurements from online measuring sensors. Thus, malfunctioning sensors will critically influence the dose prediction. The XCDC system developed was enabled with a fail-safe error detection system (Rathnaweera *et al.* 2010b). The system was designed for real-time validation of online measurements and which eliminates the error parameters from the dose prediction process in order to ensure the precision of the dose prediction. This paper presents the error parameter elimination criteria of the XCDC system, that can be utilised in water and wastewater processes.

It is well known that the compositions of wastewater in WWTPs differ from one another (Henze 1997). This shows that a model developed for one type of water is not applicable to another type without a proper calibration. Even in one WWTP, one may experience two or more wastewater types requiring individual calibration. This paper presents a successful modelling concept which solves the calibration challenge at one WWTP arising from its variable water quality.

Methods and materials:

The studies were held in four WWTPs. NRA, Lillestrøm, Norway, is a WWTP with daily inflow of about 50 000 m^3. The treatment process consisted of mechanical pre-treatment, biological treatment with floating bio-film reactors and chemical coagulation process. Instrumentation of online sensors directly in the treatment process was not possible in the plant, so this was carried out using a bypass line in a sampling chamber. Online influent turbidity (TUI), influent conductivity (CNI), influent pH (PHI) and temperature (TEI) were measured in a specially constructed influent collecting tank placed before chemical dosing. After dosing, pH (PHO) and Streaming Current (SCO) were measured in another collecting tank after dosing with coagulant.

HIAS is a WWTP in Hamar, Norway. The plant treats about 20 000 m^3 per day, and consists of mechanical pre-treatment, biological treatment followed by intermediate sedimentation and a chemical coagulation process. The online measurements were done directly in the ditches and tanks of the system. TUI, CNI, PHI, orthophosphate (OPI) and TEI were measured in the influent ditch before coagulant was introduced. PHO was measured in the ditch after dosing with coagulant.

Xiao Hong Men (XHM) is one of the 14 large WWPs in Beijing, China. The daily sewage absorption of the plant is around 600 000 m^3. Mechanical pre-treatment, biological treatment followed by sedimentation tanks is the main structure of the plant. Chemical coagulants are added to the biological treatment effluents and coagulated particles are settled in post sedimentation tanks. Online sensors were placed in collecting tanks to which wastewater was pumped from the treatment plant. The TUI, OPI, CNI, PHI, OPI and TEI were measured in the first tank with AAO effluents before addition of coagulant. PHO was measured in the wastewater pumped after mixing of coagulant.

Gaobeidian WWTP is currently one of the largest sewage treatment plants in China. The WWTP has a design capacity of 1 000 000 m^3/day. XCDC testing was carried out in a pilot scale plant with capacity of 72 m^3/day. The model-based system was only partially studied in this trial.

The dosing control tests were carried out based on the XCDC concept (Rathnaweera *et al.* 2010a). The XCDC software was designed with several online measurement validation functions for error detection.

- Set points for maximum and minimum values for each parameter were defined and when the measurement went out of the range, it was taken as an error value (Fig. 1*a*). One of the drawbacks of this method was that the defined range had to be large and was not sensitive to seasonal changes and sudden changes of water quality. This was not able to detect possible malfunctioning of instruments occur within the defined range.
- When a measurement was unchanged (repeating within a defined smaller range) for more than two hours, the measurement was detected as an error (Fig. 1b). The software re-checks the values continuously and once the value starts to change continuously for 10 minutes, the signal is automatically considered to be valid.
- Non logical measurement relationships: If PHO measurement is larger than the PHI measurement continuously for more than one minute, both measurements were considered to be error values (Fig. 1c). Once the PHO remains lower for one minute, the system identifies both PHI and PHO as functioning. This rule is only activated when the PHI and PHO values are detected as valid by other validation concepts.
- Model-based novel error detection criteria, which is detailed elsewhere (Rathnaweera *et al*. 2010b).

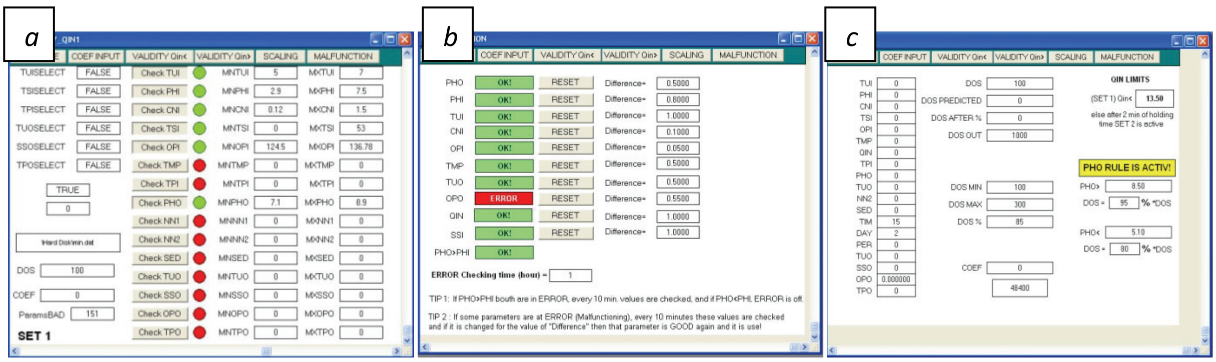


Fig.1: Measurement validation interface in the XCDC software. a) Interface for potential maximum and minimum values. b) Interface for identification of repeating value. The user must specify the accepted minimum value for each variable. c) Interface for checking the difference between PHI and PHO.

In order to avoid or minimize the negative effect of erroneous dose predictions due to false online measurements, a multiple model-based protection system was designed and evaluated. Theoretically, when there are 7 variables, 127 different combinations of algorithms can be made. Out of them, 1 algorithm includes all the parameters, which is the situation with all the instruments functioning well. The next 7 algorithms exclude one parameter out of seven. This can be considered as the situation when 1 on-line parameter is detected as mal functioning. The 15 different algorithms excluding 2 parameters are equivalent to the situation when 2 parameters are caught false. The rest of the 104 algorithms correspond to excluding more than 3 parameters at once.

The software was designed to select the appropriate prediction algorithm from the pool of algorithms to eliminate parameters detected as error values. (Fig. 2)

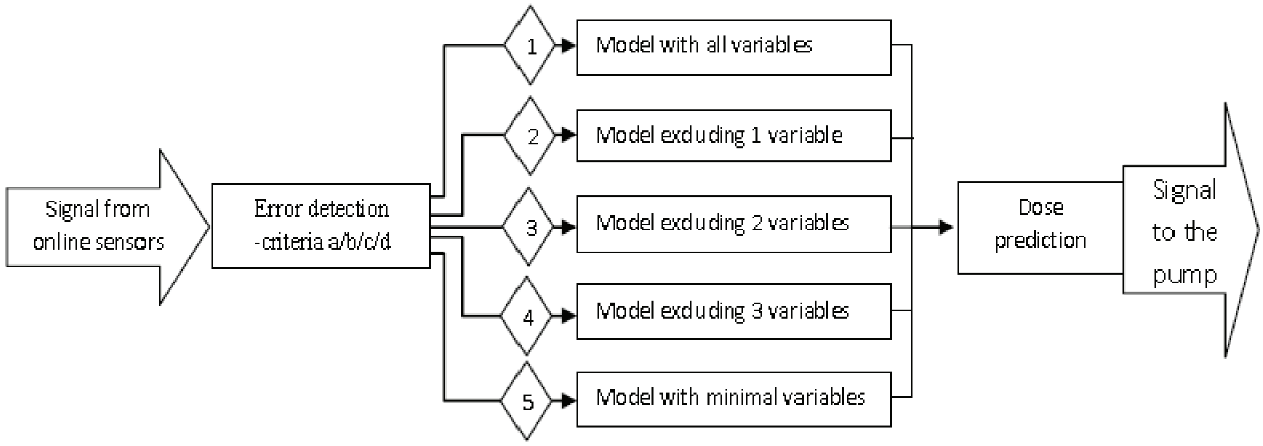


Fig. 2: Structure of measurement validation function in multiple model-based XCDC system. Here; 1: All signals functioning, 2: One parameter malfunctioning, 3: Any combination of two parameters malfunctioning, 4: Available combinations of three parameters mal functioning, 5: Any combination of malfunctioning parameters not covered by the first 4 steps.

Result and discussion

Reduction of possible algorithms

Using all possible 127 algorithms in the system is labour and time consuming as well as practically unimportant. We have considered simplification of this stage based on experience from the WWTPs. With general maintenance practices, we can expect failures of one parameter or less frequent failures of two parameters simultaneously. Failures in three or more parameters out of seven parameters would be very rare when the system is maintained with normal diligence.

In all the systems of the present study, the parameter influent flow (QI) was always assumed correct, because the plants had several flow meters which are reliable and did not experience frequent errors. This reduced the possible algorithm combinations to 63, with 6 susceptible variables.

A basic algorithm with all parameters was used in all four studies. 6 more algorithms were developed by eliminating one parameter at a time in NRA and HIAS WWTPs. In XHM WWTP, we used few combinations with one parameter failure in the 1st calibration. We assumed that situations with 2 and 3 simultaneous parameter failures are rare, when instruments are well maintained during the study period. We selected one possible scenario for two instrument failures in NRA and HIAS WWTPs and one for failure of three instruments in NRA and XHM WWTP. A simple flow proportional prediction algorithm using QI, TIM and DAY, were used as a basic model in NRA, HIAS and XHM WWTPs and a model proportional to TIM and OPI concentration was used in the Gaobeidiang pilot scale study. We calibrated two sets of model algorithms in each plant. In the second calibration at NRA, the model developed with TIM and DAY was not good enough (R^2 0.57, RMSE 3.7). Therefore we used a simple linear QI proportional algorithm.

Accordingly, for the 1st models, we used 9 different algorithms and 10 different algorithms in the second calibration models. In HIAS WWTP, 9 models were used in both sets of calibrations. In the XHM WWTP, we used 7 models in the 1st calibration and 2 basic models for the 2nd calibration. Assuming the instruments were well maintained during the study period of the Gaobeidian pilot plant, two basic models, one with all parameters and another with a minimum number of parameters, were used. Table 1 below shows the parameters and calibration statistics of the models.

Table 1: Models with various parameters with respective R^2 and RMSE values for NRA and HIAS. 7 models (A-J) are presented and parameters included in the models are indicated with a "Y" NRA had SCO but not OPI. HIAS had OPI but not SCO.

										OPI/	NRA		HIAS	
		DAY	TIM	QI	PHI	CNI	TUI	TEI	PHO	SCO	R2	RMSE	R2	RMSE
1st set of models	1A	Y	Y	Y	Y	Y	Y	Y	Y	Y	0.84	11.6	0.98	1.51
	1B	Y	Y	Y	Y	Y	Y	Y	Y	-	0.89	9.30	0.98	1.35
	1C	Y	Y	Y	Y	Y	Y	Y	-	Y	0.89	9.50	0.93	3.17
	1D	Y	Y	Y	Y	Y	Y	-	Y	Y	0.88	9.90	0.97	2.22
	1E	Y	Y	Y	Y	Y	-	Y	Y	Y	0.88	9.90	0.98	1.55
	1F	Y	Y	Y	Y	-	Y	Y	Y	Y	0.89	9.50	0.96	2.32
	1G	Y	Y	Y	-	Y	Y	Y	Y	Y	0.89	9.50	0.90	1.40
	1H	Y	Y	Y	Y	-	Y	-	Y	Y	0.86	10.50		
	1I	Y	Y	Y	-	-	-	-	-	-	-	-	0.86	4.49
	1J	-	-	Y	-	Y	Y	Y	-	Y			0.95	2.57
2nd set of models	2A	Y	Y	Y	Y	Y	Y	Y	Y	Y	0.91	1.72	0.86	8.70
	2B	Y	Y	Y	Y	Y	Y	Y	Y	-	0.90	1.75	0.85	8.90
	2C	Y	Y	Y	Y	Y	Y	Y	-	Y	0.90	1.74	0.78	10.70
	2D	Y	Y	Y	Y	Y	Y	-	Y	Y	0.88	1.94		
	2E	Y	Y	Y	Y	Y	-	Y	Y	Y	0.90	1.76	0.82	9.80
	2F	Y	Y	Y	Y	-	Y	Y	Y	Y	0.83	2.30		
	2G	Y	Y	Y	-	Y	Y	Y	Y	Y	0.90	1.74	0.86	8.70
	2H	Y	Y	Y	Y	Y	Y	Y	-	-	0.90	1.76	0.80	10.20
	2I	Y	Y	Y	-	-	-	Y	Y	Y	0.83	2.33		
	2J	-	-	Y	-	-	-	-	-	-	-	-	0.68	13.10
	2K	Y	Y	Y	-	Y	Y	Y	-	Y			0.85	8.90
	2K	Y	Y	-	Y	Y	-	Y	Y	Y			0.66	13.60

Model statistics indicate that the second set of models, except HIAS, were clearly improved with better R^2 values and smaller RMSE values compared to the first set of models.

The data used for the calibration of the first set of models were the data with existing flow proportional dose in the WWTP, when the XCDC was working offline. Thus the dose here had no relationship with water quality parameters. Data with dose predictions from the first models were used for the 2nd model calibrations. Logically, the 2nd set of models should be better models due to the larger and better data sets. For all WWTPs apart from HIAS, this was the case. The 2nd set of models at HIAS was poorer than the 1st set. This phenomenon will be discussed later in this paper.

The algorithm with all the parameters is prominent among others (table 1). The algorithms excluding more than one parameters show poor statistics compared to the algorithm with all parameters. The same phenomenon was observed in Table 2 and visualized in Fig. 3. Table 2 is consistent with the model statistics of 60 different combinations of variables, developed from the 2nd data set of NRA. The trend lines of Fig. 3 show that the R^2 is clearly increasing while RMSE is decreasing with the number of parameters included in the model. This can be simply explained as a statistical phenomenon when increasing the number of X variables, the R^2 value will be increased, irrespectively of the precision of prediction. But a simultaneous reduction of the RMSE indicates an improvement in the predictability of models with increased number of variables. In addition to the statistical explanation, it is well documented that the coagulation process is strongly related with water quality parameters (Ratnawera *et al*. 1994; Lu *et al.* 2002). The results show that including more parameters has improved the accuracy of the system.

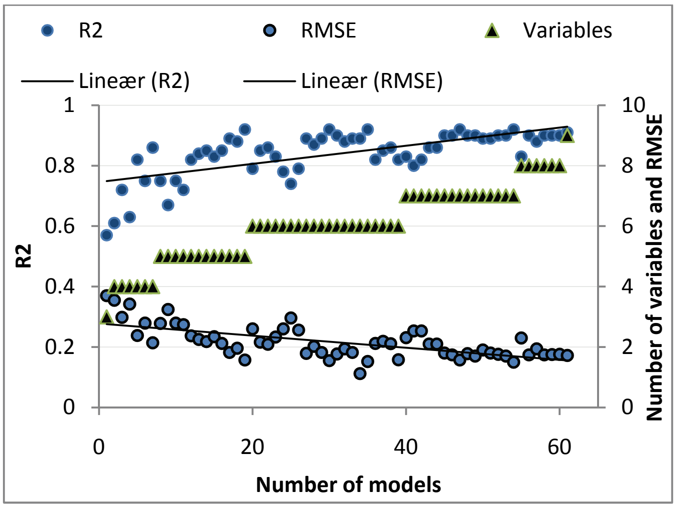


Fig.3. Change of R^2 and RMSE values with changing the number of variables in the algorithms. The two lines indicate the trends in the statistics.

Table 2. R^2 and RMSE values of 60 different models with different combinations of parameters

	No of variables	DAY	TIM	QI	CNI	PHI	TEO	PHO	SCO	TUI	R^2	RMSE
0	3	Y	Y	Y	-	-	-	-	-	-	0,57	3,70
1	4	Y	Y	Y	-	-	-	-	-	Y	0,61	3,55
2	4	Y	Y	Y	-	-	-	-	Y	-	0,72	2,98
3	4	Y	Y	Y	-	-	-	Y	-	-	0,63	3,42
4	4	Y	Y	Y	-	-	Y	-	-	-	0,82	2,38
5	4	Y	Y	Y	-	Y	-	-	-	-	0,75	2,79
6	4	Y	Y	Y	Y	-	-	-	-	-	0,86	2,14
7	5	Y	Y	Y	-	-	-	-	Y	Y	0,75	2,78
8	5	Y	Y	Y	-	-	-	Y	-	Y	0,67	3,24
9	5	Y	Y	Y	-	-	-	Y	Y	-	0,75	2,79
10	5	Y	Y	Y	-	Y	-	-	-	Y	0,72	2,74
11	5	Y	Y	Y	-	Y	-	-	Y	-	0,82	2,37
12	5	Y	Y	Y	-	Y	Y	-	-	-	0,84	2,25
13	5	Y	Y	Y	-	-	Y	-	-	Y	0,85	2,18
14	5	Y	Y	Y	-	-	Y	-	Y	-	0,83	2,34
15	5	Y	Y	Y	Y	-	-	-	-	Y	0,85	2,12
16	5	Y	Y	Y	Y	-	-	-	Y	-	0,89	1,82
17	5	Y	Y	Y	Y	Y	-	-	-	-	0,88	1,96
18	5	Y	Y	Y	Y	-	Y	-	-	-	0,92	1,57
19	6	Y	Y	Y	-	-	-	Y	Y	Y	0,79	2,60
20	6	Y	Y	Y	-	-	Y	-	Y	Y	0,85	2,16
21	6	Y	Y	Y	-	-	Y	Y	-	Y	0,86	2,09
22	6	Y	Y	Y	-	-	Y	Y	Y	-	0,83	2,33
23	6	Y	Y	Y	-	Y	-	-	Y	Y	0,78	2,60
24	6	Y	Y	Y	-	Y	-	Y	-	Y	0,74	2,96
25	6	Y	Y	Y	-	Y	-	Y	Y	-	0,79	2,56
26	6	Y	Y	Y	Y	-	-	-	Y	Y	0,89	1,79
27	6	Y	Y	Y	Y	-	-	Y	-	Y	0,87	2,02
28	6	Y	Y	Y	Y	-	-	Y	Y	-	0,89	1,82
29	6	Y	Y	Y	Y	-	Y	-	-	Y	0,92	1,55
30	6	Y	Y	Y	Y	-	Y	-	Y	-	0,9	1,77
31	6	Y	Y	Y	Y	Y	-	-	-	Y	0,88	1,94
32	6	Y	Y	Y	Y	Y	-	-	Y	-	0,89	1,82
33	6	Y	Y	Y	Y	Y	-	Y	-	-	0,89	1,12
34	6	Y	Y	Y	Y	Y	Y	-	-	-	0,92	1,52
35	6	Y	Y	Y	-	Y	Y	-	-	Y	0,82	2,12
36	6	Y	Y	Y	-	Y	Y	-	Y	-	0,85	2,19
37	6	Y	Y	Y	-	Y	Y	Y	-	-	0,86	2,11
38	6	Y	Y	Y	Y	-	Y	Y	-	-	0,82	1,58
39	7	Y	Y	Y	-	-	Y	Y	Y	Y	0,83	2,31
40	7	Y	Y	Y	-	Y	-	Y	Y	Y	0,80	2,53
41	7	Y	Y	Y	-	Y	Y	-	Y	Y	0,82	2,53
42	7	Y	Y	Y	-	Y	Y	Y	-	Y	0,86	2,10
43	7	Y	Y	Y	-	Y	Y	Y	Y	-	0,86	2,10
44	7	Y	Y	Y	Y	-	-	Y	Y	Y	0,90	1,80
45	7	Y	Y	Y	Y	-	Y	-	Y	Y	0,90	1,74
46	7	Y	Y	Y	Y	-	Y	Y	-	Y	0,92	1,57
47	7	Y	Y	Y	Y	-	Y	Y	Y	-	0,90	1,78
48	7	Y	Y	Y	Y	Y	-	-	Y	Y	0,90	1,70
49	7	Y	Y	Y	Y	Y	-	Y	-	Y	0,89	1,90
50	7	Y	Y	Y	Y	Y	-	Y	Y	-	0,89	1,80
51	7	Y	Y	Y	Y	Y	Y	-	-	Y	0,90	1,76
52	7	Y	Y	Y	Y	Y	Y	-	Y	-	0,90	1,70
53	7	Y	Y	Y	Y	Y	Y	Y	-	-	0,92	1,50
54	8	Y	Y	Y	-	Y	Y	Y	Y	Y	0,83	2,30
55	8	Y	Y	Y	Y	-	Y	Y	Y	Y	0,90	1,74
56	8	Y	Y	Y	Y	Y	-	Y	Y	Y	0,88	1,94
57	8	Y	Y	Y	Y	Y	Y	-	Y	Y	0,90	1,74
58	8	Y	Y	Y	Y	Y	Y	Y	-	Y	0,90	1,75
59	8	Y	Y	Y	Y	Y	Y	Y	Y	-	0,90	1,76
60	9	Y	Y	Y	Y	Y	Y	Y	Y	Y	0,91	1,72

Model shifting

Instrumental failures in the well maintained systems of the study were rare. In order to evaluate model shifting and its effect, selected measurements were programmed to be erroneous. This was easily done by changing the measurement validation limits in the system to detect them as errors by the software and temporally disconnecting the sensors.

Fig. 4 is an elucidation of model changes during the period of 12th May to 24th may 2009, with the 1st set of algorithms in NRA. It clearly shows that the predicted dosage varies with the model. However, a significant effluent quality reduction did not result from the model changes, although it may have caused higher dosages.

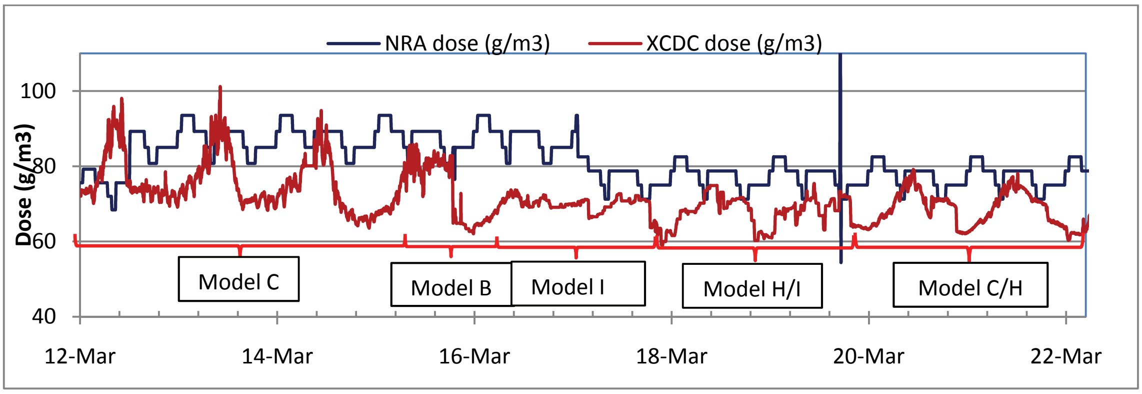


Fig. 4: Dose prediction changes with model shifting. The NRA dose was the dose with the traditional dosing system in the plant. XCDC dose was the dose prediction according to the new system. The algorithms in the function were changed to respond to detection of errors in the parameters.

Fig. 5. presents examples of predicted and measured dosages at NRA, HIAS and XHM, as a result of model shifts due to parameter validation. The real-time dosages (DOSE) are presented with the estimated dosages using various models (MOD x, where x is the model number) for comparison. Fig. 5.1 shows the results with a model containing all parameters, while the rest shows the dosages as a combination of two or more models during parameter validation. Each figure includes one or two parameters with their respective validation criteria (same coloured vertical line), which triggers the selection of a different model excluding or including the given parameter. The automatic model shifting with validation of parameters functioned well in practice enabling an accurate and robust system.

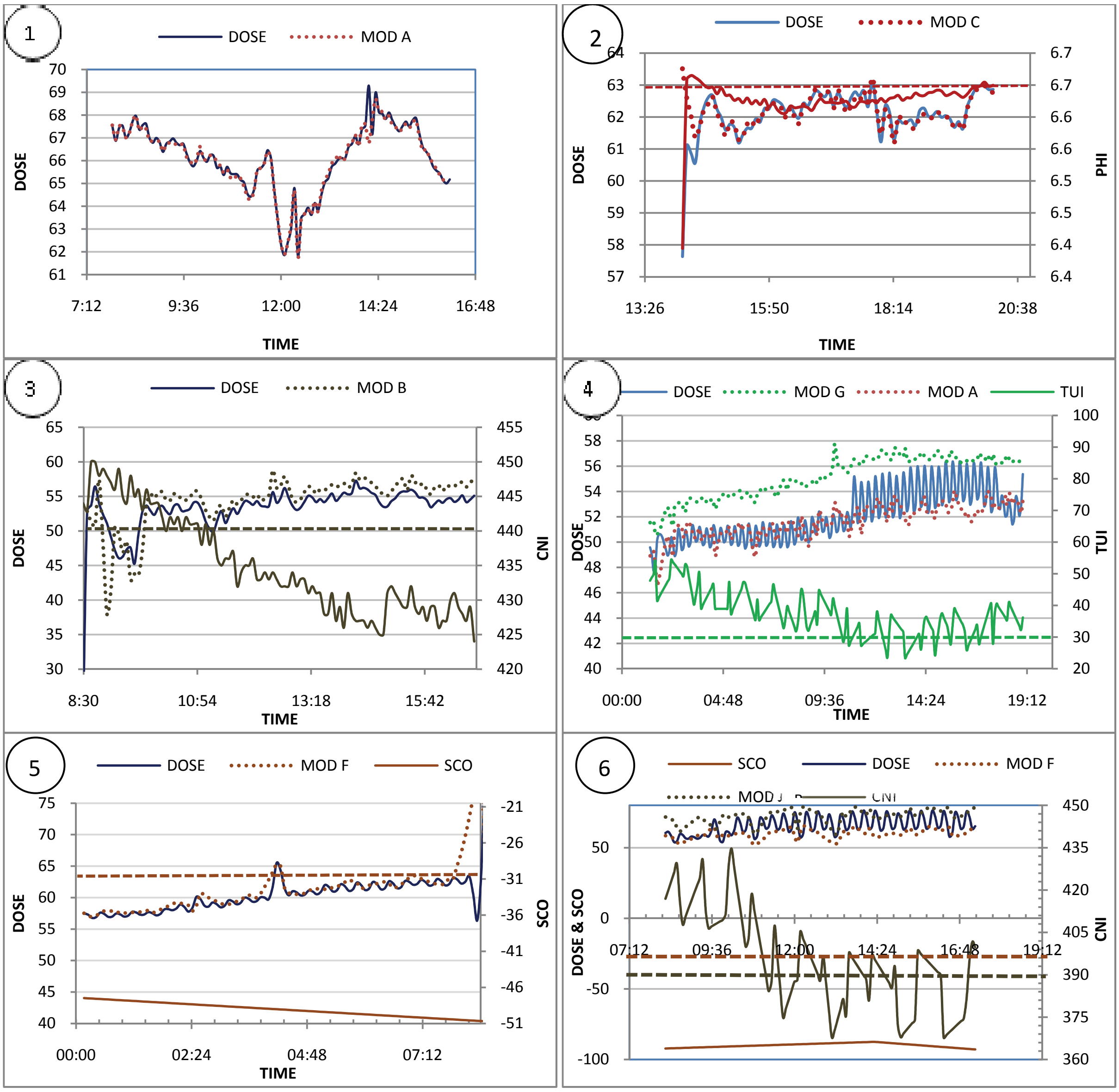


Fig. 5. Dose prediction by different models as result of validation of parameters. Each figure consists of predicted dosage in the system (DOSE), estimated using one or more models depending on the validity of parameters. The vertical lines indicate the validation criteria for the parameter given in the right Y-axis. Figs. 5-1, 5-2, 5-3 and 5-5 used only one model as the relevant parameters were valid during the relevant period. Figs. 5-4 and 5-6 had invalid parameters at times thus shifting between 2 models.

Two platform based control system at HIAS

In the sample set for the 1st calibration, the influent was quite homogeneous as the process had no serious bypassing events and thus the model statistics were extremely good with 0.98 R^2 and 1.40 RMSE. The system ran successfully for about 45 days until the rainy season began and the system started to respond with overdosing with poor dose prediction.

During the recalibration, we learnt that most of the samples related to the rainy days were rejected as outliers by the statistical software. Thus, outlier selection was done extremely carefully during the re-calibration. The new calibration had comparatively poor statistics (0.86 of R^2 and 8.7 of RMSE) even with extensive removal of outliers. Further, these models did not satisfactorily describe the situation during rainy days.

A further analysis of the situation showed a difference in influent quality depending on flow. Wastewater of HIAS WWTP passes a biological treatment followed by an intermediate sedimentation process before the chemical treatment. In rainy seasons and during snow melting events, the QIN to the WWTP exceeds the maximum capacity of the biological treatment. The maximum designed capacity of the biological treatment unit is 270 l/s, thus at the limit of 240 to 250 l/s, the excess water bypasses the biological stage and goes directly to the chemical stage. These events considerably influenced the water quality of influent creating a significantly different influent quality to the chemical treatment stage.

The bypass produced two water types for the chemical treatment process. The appropriate doses for two water types could not be predicted with one algorithm. A two model platform-based system has been introduced to address the two water types. Fig. 6 demonstrates the changes in water quality with bypassed water.

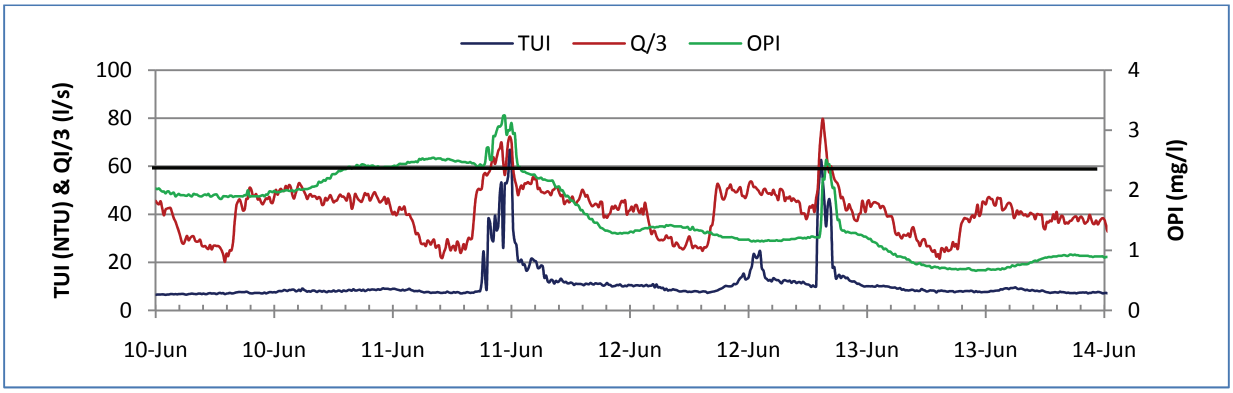


Fig.6: Shock loads at HIAS. The black horizontal line is the overflowing limit during biological treatment.

The statistical software UNSCRAMBLER 9.8 was used for multivariate calibration in these studies.

The data showed that the TUI and OPI were changed in small quantities during the period flow without bypassing the biological treatment. The TUI and OPI were increased drastically with increasing QIN during the bypassing events. It is obvious that the demand of coagulants must be increased with increasing TUI and OPI. The model developed to predict the dose for quiet water with smaller variability of parameters was not able to correctly address the larger variables of the water. Thus the model responded with enormously higher dose predictions which were not acceptable.

In the re-calibration process, the less frequent untypical samples in the data set were identified as outliers. On the other hand, the rapid variability of data did not show a good response to dose, since the dose was not correctly adjusted to face the events. These were the two reasons for poor performance of the re-calibration. Due to these factors, the re-calibrated models did not perform satisfactorily.

As a solution, the datasets from rainy days were isolated and modelled as a separate set of data. This model was better than the second calibration models, with an R^2 was 0.92 and RMSE was 4.3. Furthermore, almost all the data were used in calibration except five extreme outliers.

The XCDC software was thus modified to run two sets of models in platforms. In the first platform, we used the models calibrated without rainy day data. The second platform consisted of models calibrated using rainy day data. Fig. 7 shows the software interfaces designed for two model sets used in HIAS.

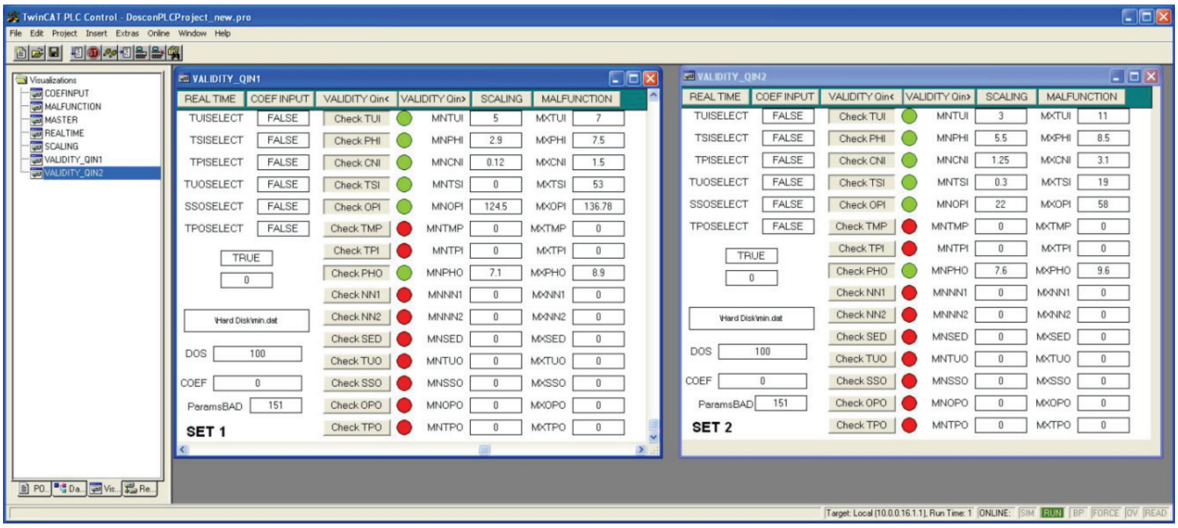


Fig. 7: Software interface for the two platform system.

The second set of models was set to overtake the dose predictions when the QIN exceeded 250 l/s. Fig. 8 below shows the dose prediction from the two model system.

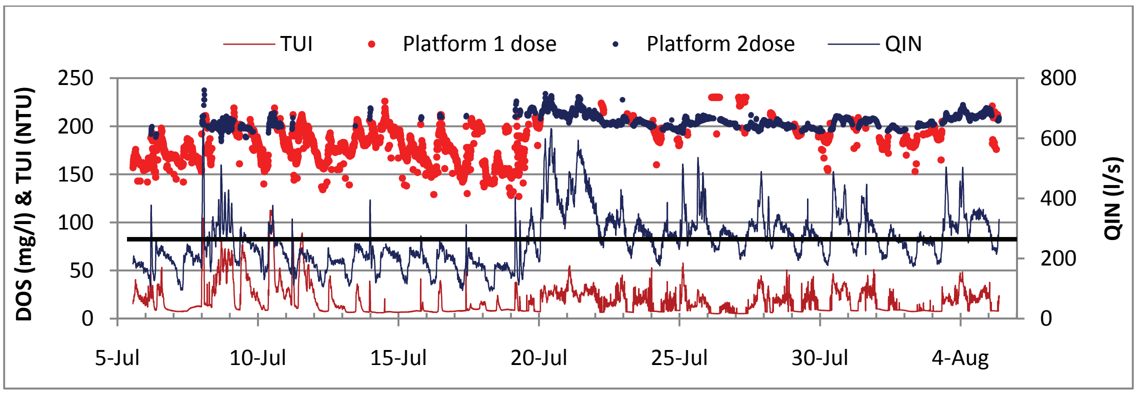


Fig.8: Two platform model system in operation. Platform 1 dose is the dose predicted by the model used for normal inflow. Platform 2 dose is the dose prediction simulated using the model developed exclusively with samples with QIN over 250 l/s. The XCDC software has been designed to shift the platform from 1 to 2, when QIN is over 250 l/s.

Conclusions

- When more parameters are used in the model, the model statistics and accuracy of the prediction improves.
- When online measurements are validated, a model system including or excluding various parameters can be utilised to efficiently predict the coagulant dosage.
- Some of the models can be excluded from model combinations by considering the low probability of selected instrumental errors to simplify the dosing concept
- Internal treatment differences (bypasses) and internal shock loads (septic) may require the influent to be treated as two or more types requiring several model sets.
- Using a multiple model-based strategy for the XCDC system makes the system more precise, reliable and robust.

References

Hangouet, JP, Pujol, R, Bourgogne, P, Ropert, D & Lansalot, G 2007, 'Optimizing Chemical Dosage in Primary Settling Tanks'. *Chemical Water and Wastewater Treatment IX, Proceedings of the12th Gothenburg Symposium* , Hermann H. Hahn Erhard Hoffmann Hallvard Ødegaard (eds), IWA Publishing, London, pp. 59-67.

Lu, L 2003, 'Model based control and simulation of wastewater coagulation', PhD thesis, Agriculture University of Norway.

Lu, L, Ratnaweera, H &Lindholm, O 2003, 'Coagulant dosage control in chemical wastewater treatment plants- a review of modelling approaches', *Vatten, vol.* 59, pp. 227-235.

Lu, L, Ratnaweera, H, Lindholm, O & Lileng, K 2002, 'Model-Based Real Time Control of Coagulant Dosing'. *Proceedings International IWA Conference on Automation in Water Quality Monitoring*, N. Flesischmann, Lengergraber, G. and Haberl, R. (eds), U. of Agricultural Sci., *Vienna*, pp. 421-424.

Olsson, G, Aspegren, H & Marinus, KN 1998, 'Operation and control of wastewater treatment - a Scandinavian perspective over 20 years', *Water Science and Technology*, IWA Publishing, vol. 37, no. 12, pp. 1-13.

Rathnaweera, S, Ratnaweera, H, Håkonsen, T & Lindholm, O 2010b. 'Error detection technique for online water quality measurement for increasing accuracy in alarms and control functions'. Accepted manuscript for the 7th Leading Edge Conference on Water and Wastewater Technologies (2-4 June 2010). (submitted for publication)

Rathnaweera, S, Ratnaweera, H, Lindholm, O & Håkonsen, T 2010a, 'Multi-parameter based real-time coagulant dose control system for wastewater treatment', (submitted manuscript).

Ratnaweera, H 2004, 'Coagulant dosing control- a review', *Chemical water and wastewater treatment VIII*, HH Hahn, E Hoffmann, H Ødegaard (eds), IWA Publishing, London, pp. 3-13.

Ratnaweera, H, Blom, H & Aasgaard, G 1994, 'Flexible Coagulant Dosing Control System based on Real-Time Wastewater Quality Monitoring', *Chemical water and wastewater treatment III*, Hahn, H.H. & Klute, R(eds), Springer-Verlag, Berlin, pp. 105-116.

Ratnaweera, H, Lu, L & Lindholm, O 2002, 'Simulation Program for Wastewater Coagulation', *Chemical Water and Wastewater Treatment VII*, HH. Hahn, E. Hoffmann & H. Oedegaard, Gothenburg (eds), Sweden, IWA publishing, London, pp. 253-260.

Sagberg, P, Sæter, R & Bagggerud, BA 1990, 'Increasing the surface load at a direct precipitation plant, VEAS, Norway'. *In: Chemical water and wastewater treatment*, H. H. Hahn and R. Klute (eds.). Springer, Berlin Heidelberg New York, pp. 271-282.

Paper II

Modelling coagulant dosage in wastewater treatment plants, using MLR, PCR and PLSR statistical analysis.

Subhash Rathnaweera, Harsha Ratnaweera, Tor Håkonsen, Oddvar Lindholm, Ellen Sandberg
Paper submitted to the International Journal of Environmental Science and Technology

Modelling coagulant dosage in wastewater treatment plants, using MLR, PCR and PLSR statistical analysis.

Subhash Rathnaweera*†, Harsha Ratnaweera**, Tor Håkonsen***, Oddvar Lindholm*, Ellen Sandberg*

*University of Life Sciences, Aas, Norway
**Norwegian Institute for Water Research (NIVA), Oslo, Norway
*** VA-Support AS, www.va-support.no
† Corresponding author (subhash.rathnaweera@umb.no)

Abstract

Multivariate analysis is a powerful tool for predicting one or more variables using more than one predictor variable. Multiple linear regression (MLR), Principal component regression (PCR) and Partial least squares regression (PLSR) are the three common regression analytical methods that can be used for data prediction purposes. These methods are evaluated for the chemical coagulation process in wastewater to elucidate the relationship between coagulant dose and water quality parameters and to study the comparative predictive powers of the models. Although MLR showed the best model statistics compared to the other two methods, the predictive power for data outside its calibration range was poor. This indicates that MLR is appropriate for data description purposes but not for long-term predictions in wastewater processes. PCR and PLSR showed almost similar performance with respect to model statistics and data prediction. The statistics of PLSR was slightly better compared to PCR. Thus PLSR is suggested as the best performing regression method for modelling wastewater with the perspective of long-term prediction of dose.

Introduction

Coagulation is a well defined process. For example, if three identical water samples have coagulant added at identical dosages, and they are left to coagulate, flocculate and sediment identically, the result will be three identical effluent samples. Such a process can be described mathematically. Construction of a conceptual model, however, may be a challenge due to the complex nature of the process (Ratnaweera *et al.* 1994). Few attempts to construct relationships between coagulant dosage and water quality parameters are reported. Lu (2003) attempted to construct a relationship using partial least squares regression which is a multivariate calibration method, while attempts using logical concepts like fuzzy logic or artificial neural network are also reported.

Multivariate analyses are used for a number of different purposes. Esbensen (2000) divides the objectives of multivariate analysis into three main groups. The first is 'Data description', which is the exploration of data to find relationships between multiple variables. The next group is 'discrimination and classification', which is separation of groups of data according to the relationships between variables. The third group is 'regression and prediction' which is the approach that relates two sets of variables to each other and uses one set to predict the relationship of the other.

Multivariate calibration is a powerful tool for predicting one or several variables using more than one other regression variable. There are several ways to perform multiple variable data analysis. The method that is used for analysis should be selected according to the goal to reach. Thus, developing a clear goal for multiple variable analysis is very important.

Multiple linear regression (MLR), principal component regression (PCR) and partial least squares regression (PLSR) are the modelling methods most used among several methods of multivariate calibration. Due to the simplicity and ease of calculation, MLR has been the most popular method for many decades. The drawback of MLR is that, when there are a large number of variables, for example the number of variables is larger than the number of observations, or if the variables are strongly inter-correlated (co-linear), it will fail as a predictor (Robert *et al.* 1999). Multiple co-linearity is a most common situation when the number of variables is greater than four or five.

Recently, with the development of instrumentation, computing power and chemometrics, so-called full regression methods, like PCR and PLSR have become popular. As Risvik (2007) explains, PCR consists in simplifying the data set by reducing multi-dimensional data sets to lower dimensions for analysis. It can be used for exploratory data analysis and to make predictive models.

PLS was developed in the 1960s by Herman Wold (Tobias 1997). PLS is a regression method combining both MLR and PCR. PLS regression is based on projections onto latent structures derived from principal components analysis and it has been proven to be superior to MLR whenever the explanatory variables are not independent of each other, or when multi co-linearity is present.

Most authors advocate that PLS regression is a better predictive tool than the other two considered methods. Dane *et al*. (2001) documented that the best results for their study were obtained by PLS regression. Fülöp and Hancsók, (2008) found that PLS had better prediction efficiency than PCR in their comparison of calibration models based on near infrared spectroscopy data. Clementi *et al*. (1997) documented that PLS performs better than MLR for protein structure prediction. Garcia-Olmo *et al.* (1998) documented some advantages and disadvantages of PLS and MLR and concluded that MLR was good enough for quality control of the Iberian pig industry, because they found that MLR and PLS had the same predictive power. Lipp (1996) compared MLR, PCR and PLSR in the quantitative determination of foreign oils in butter fats and concluded that MLR is a suitable method for this purpose.

Discussion of relevant statistical theories is beyond the goal of this paper. The theory of MLR, PCR and PLS can be found in the literature (Borga *et al.* 1992; Martens and Næs 1991; Martens *et al.* 2001; Wold *et al*. 2001; Johnson & Whichern 2002; Esbensen, 2000; etc.).

Several studies have been conducted in different fields of science comparing the predictive powers of MLR, PCR and PLSR. Still, comparison studies in the field of water and wastewater are scarce. It was also difficult to find a comparison of statistical regression methods in order to predict future data over a longer time period, which can be used as a simulation model for prediction over a long time duration.

In this paper, we compare three different multivariate analytical methods, MLR, PCR and PLSR in order to find the most suitable regression method for the simulation of real-time coagulant dosage control models for wastewater treatment. Considering the fact that it is difficult to find studies for predicting values out of the calibration data set, we especially focus on the predictive power of each method of regression.

Method

Data Preparation

The NRA waste water treatment plant (WWTP) at Lillestrøm is one of the largest WWTPs in Norway. The WWTP is a recipient of both sewage and industrial wastewater from 3 neighbouring municipalities with a capacity of 50,000 m^3/day. The treatment process consisted of physical pre-treatment, biological treatment with bio-film reactors and chemical coagulation. The data were collected using online sensors before and just after the addition of coagulant.

The total inflow (QI), turbidity (TUI), conductivity (CNI), pH (PHI) and temperature (TEI) were measured before dosing the coagulant; pH (PHO) and streaming current (SCO) were measured just after coagulant dosing and mixing. The measurements, including real-time coagulant dosage (DOSAGE), date and time were logged at 5 min intervals, by the supervisory control and data acquisition (SCADA) system of NRA. 20,000 samples from September to November 2008 were collected for the study.

All online measuring instruments were strictly calibrated and maintained according to instrument guidelines, during the period of study.

The data set was edited to remove known erroneous data, for example during instrument cleaning and calibration periods. 1,087 data samples (5% of the total data set) were discarded from the 20,000 data in the editing process and the remaining 18,983 data samples were used for further analysis.

Regression procedure

Regression models belonging to three methods were developed using the statistical software UNSCRAMBLER version 9.8, which is specialized software for multivariate analysis.

Before the calibration of models, variables were standardized so that all variables contributed with equal weight to the model. This was done by first centring (subtracting the mean value from the variable) and then weighting by dividing the standard deviation.

Cross validation, which helps ensure the fitness of models and will also demonstrate the predictive ability of the model, was used in all analytical methods (Martens *et al.* 2001). Eliminating extreme outliers from the data set is essential for better prediction model (William *et al.* 1998). 'U score versus T score' plots as well as 'residual y variance versus leverage' plots were used to remove minimal outliers from PCR and PLSR. In MLR, a 'residual y variance versus leverage' plot was used for the same purpose.

Having too many X variables in the model may over-fit the model. Over-fitted models will show larger R^2 values, but will have low predictive ability (Clementi *et al*. 2002). Jack-knife validation, which is defaulted in the UNSCRAMBLER cross-validation run with uncertainty test (CAMO 1999), was used to identify and remove insignificant X variables from the PCR and PLSR models (Martens and Martens 2000; Martens *et at*. 2001). At the same time, the p-values in the ANOVA table were used to identify insignificant variables in MLR. The 'prediction versus measure' plot indicates the fit of the model. Best fit models should give a straight line through the origin with a slope of one. In this analysis we used these plots to evaluate the fitness of equations (Martens and Næs 1991).

The coefficient of determination (R^2) is the percentage of the total variation in the y-values that is explained by the regression equation. We used calibration R^2 and validation R^2 to demonstrate how good the model explains calibration as well as validation data sets. The root mean squares error (RMSE) quantifies the difference between the real value and the estimated value in the model. RMSE for the calibration set (RMSEC) and RMSE for the validation set (RMSEP) were compared to show the relative difference between the prediction and calibration sets (Esbensen 2000; Martens and Næs 1991).

Test 1. Batch prediction evaluation

The total 18983 samples were divided into six groups, named G-1 to G-6, having 3,000 data samples in each, as follows: 1–3000, 3001–6000, 6001–9000, 9001–12000, 12001–15000 and 15001–18000 respectively.

DAY (Day of the week beginning from Monday was categorized from 1 to 7), TIM (Each hour of a day was categorized from 1 to 24), QIN, TUI, CNI, PHI, TEI, PHO and SCO, with their interaction and square terms, were used as explanatory variables for the DOS as response variable.

DOS = *f*(TIM / DAY / QIN / TUI / CNI / PHI / TEI / PHO / SCO / Cross effects of variables and Square effects of variables)

Each of the abovementioned groups with 3000 data was used to calibrate the statistical models by MLR, PCR and PLSR analysis. All the above mentioned procedures were followed in model calibration.

Each model developed in this way was used to simulate the dosage of the complete dataset (18,983 samples) and the predicted results were represented graphically in order to compare the prediction and deviation from actual data.

Both model statistics and figures were used to compare the predictability of three statistical modelling methods and the most suitable method was finally selected as a calibration model for online prediction of coagulant dosage in wastewater treatment systems.

Test 2. Use of complete data set

The complete dataset (18,983 samples) was used in the same way for calibration of algorithms according to each regression method. Predictions were plotted against the actual dosage values, to evaluate the predictive strength.

Results and discussion

Description of data

Descriptive statistics of the variables in the complete dataset are illustrated in Table 1. The table shows that the dataset is characterised by large variability among variables. Fig. 1 graphically shows the variability of samples during the period of sampling.

Table 1: Descriptive statistics of the data set. (SD: Standard deviation)

	Mean	SD	Minimum	Maximum	Skewness
Day	4.0	2.0	1.0	7.0	0.0
TIM	11.5	6.9	0.0	23.0	0.0
QIN	550.9	151.1	172.9	964.5	0.1
CNI	436.0	105.0	254.0	710.0	0.8
TEI	15.4	3.1	8.7	21.4	-0.7
TUI	134.8	85.0	51.6	499.5	2.9
PHI	6.7	0.4	5.6	8.6	-0.5
PHO	6.3	0.4	4.5	7.3	-0.4
SCI	-29.5	34.7	-320.8	50.4	-1.5
DOS	116.0	31.6	34.3	252.2	0.3

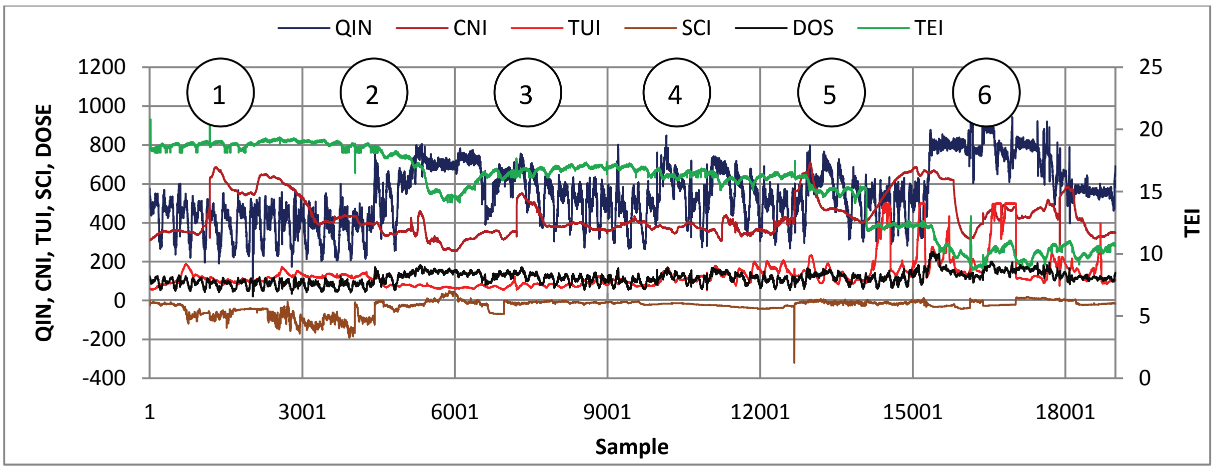


Fig. 1: Variability of variables among grouping. 1 to 6 in circles indicate the six defined data groups. The figure visualizes the variability of the parameters in different groups.

Table 2: A comparison of mean and standard deviation (SD) of different data groups.

	G-1		G-2		G-3		G-4		G-5		G-6	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Day	3.5	1.9	4.5	2.0	3.5	1.9	4.5	2.0	3.5	1.9	4.5	2.1
TIM	11.6	6.8	11.4	7.0	11.5	6.8	11.6	7.0	11.3	6.9	11.7	6.9
QIN	411.0	84.6	508.3	163.3	578.3	108.7	536.5	118.3	521.6	101.6	746.8	111.2
CNI	503.1	124.6	377.2	59.1	381.0	64.7	376.2	24.4	496.0	100.9	495.2	105.4
TEI	18.8	0.2	17.7	1.5	16.5	0.7	16.4	0.4	14.6	1.6	10.3	1.0
TUI	111.3	26.9	99.6	26.5	78.1	13.4	113.7	29.8	167.3	87.7	236.9	134.1
PHI	6.7	0.4	6.7	0.3	6.7	0.3	6.7	0.4	6.7	0.4	6.8	0.3
PHO	6.3	0.4	6.4	0.3	6.4	0.3	6.3	0.4	6.3	0.4	6.2	0.3
SCI	-59.1	32.3	-56.8	55.9	-15.2	19.2	-21.4	9.4	-16.9	14.0	-12.1	18.1
DOS	87.8	19.0	108.5	35.2	120.7	21.4	113.5	24.5	112.6	22.6	153.5	28.8

The variation of variables in different groups can be seen in Table 2. Especially the mean values and SD values of QIN, CNI, TUI and SCI vary widely between different groups. Thus DOS (the Y variable) is shown as a variation in different groups. Variables QIN, CNI, TUI and SCO show large variation between different groups compared to the other parameters. Comparing the relevant values for each parameter in Tables 1 and 2, it is visible that the characteristics of each group vary from the other. Fig. 1 visualizes the behaviour of variables.

Table 3: Correlation between each pair of variables and also between each variable and dosage.

TIM	-0.01								
QIN	0.09	0.23							
CNI	0.11	0.01	-0.12						
TEI	-0.02	-0.01	-0.64	-0.09					
TUI	-0.05	0.02	0.20	0.34	-0.46				
PHI	-0.03	-0.01	0.03	0.03	-0.10	0.11			
PHO	-0.04	0.01	-0.10	0.02	0.14	-0.06	0.69		
SCI	-0.03	0.03	0.51	-0.16	-0.51	0.01	0.09	-0.05	
DOS	0.14	0.27	0.92	0.00	-0.56	0.23	0.04	-0.08	0.50
	DAY	**TIM**	**QIN**	**CNI**	**TEI**	**TUI**	**PHI**	**PHO**	**SCI**

According to Table 3, PHO vs. PHI (0.69), TEI vs. QIN (-0.64), SCI vs. Qin (0.51), SCI vs. TEI (0.51) and TUI vs. TEI (-0.46) show correlation between variables. This shows a possibility of interference with MLR from a 'multi co-linearity' problem.

The correlations between dosage and each parameter were not that strong, except flow which was due to the flow proportional dosing system in the treatment plant. This poor correlation between X variables and the Y variable suggested the use of cross effects and square effects of X variables in a multiple variable analysis process.

Comparison of regression methods

Table 4 shows the R^2 and RMSE of each model developed by MLR, PCR and PLSR. According to the table, all three methods produced good regression statistics.

Table 4: Correlation coefficient (R^2) and root mean square errors (RMSE) of both calibration (Cal) and validation (Val) samples in each algorithm, calibrated by MLR, PCR and PLSR analysis.

		(G-1)		(G-2)		(G-3)		(G-4)		(G-5)		(G-6)	
		R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE
MLR	Cal	0.98	2.77	0.99	3.67	0.96	4.0	0.97	4.3	0.95	5.1	0.90	9.0
	Val	0.98	2.86	0.99	3.72	0.96	4.1	0.97	4.3	0.95	5.2	0.89	9.2
PLS	Cal	0.96	3.68	0.98	4.7	0.95	4.7	0.94	5.8	0.89	7.3	0.87	10.4
	Val	0.96	3.68	0.98	4.7	0.95	4.7	0.94	5.8	0.89	7.3	0.86	10.6
PCR	Cal	0.96	3.69	0.98	4.7	0.89	7.0	0.94	5.8	0.89	7.4	0.69	15.0
	Val	0.96	3.69	0.98	4.7	0.89	7.0	0.94	5.8	0.89	7.4	0.69	16.0

The MLR models have comparatively the best performance with largest R^2 and smallest RMSE over the other considered analytical methods. This implies the MLR models perform best for the data set used for calibration and validation.

Compared to MLR, PC and PLS regressions always produce slightly lower statistics. Comparing PCR and PLSR in regressions for groups (1), (2), (4) and (5), they show the same R^2 and RMSE values. Among the groups, group 6 shows the lowest R^2 value and the largest RMSE values amongst all analytical methods. This was due to the larger variation of variables during the period.

Accordingly, MLR would be the best method for data description purposes in wastewater. The other two methods also perform well for this purpose.

Each model was then used to predict the dosages of the complete data set. In the figures below, (Figs. 2, 3 and 4) two selected groups of different calibration methods are compared. The G-2 data set represents reasonably well the X variables in the complete data set (Fig. 1 and Table 3). In addition, the G-2 models have the best statistics. The predictive power of G-2 is greater than the predictive power of the other groups. G-2 can be selected as the best data set for predicting the total data. In the figures, G-5 gave an average performance and is compared to G-2.

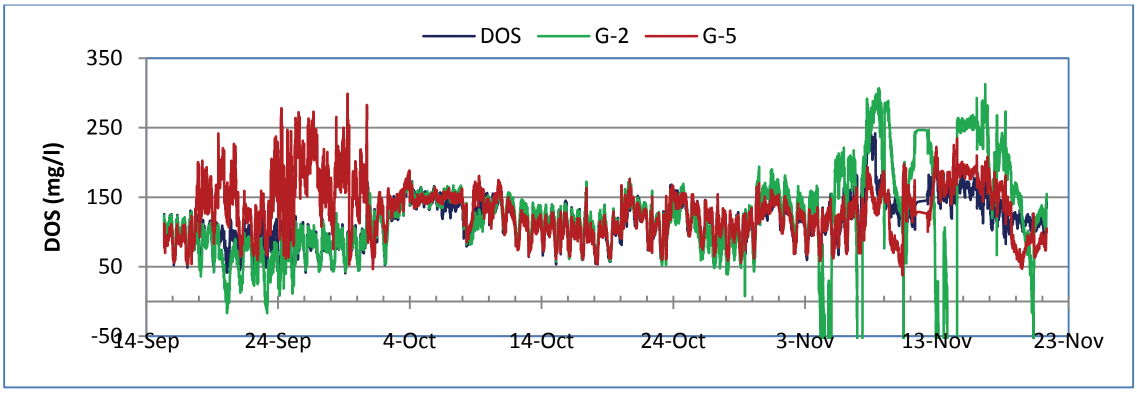


Fig. 2: Predictions of MLR models for G-2 and G-5.

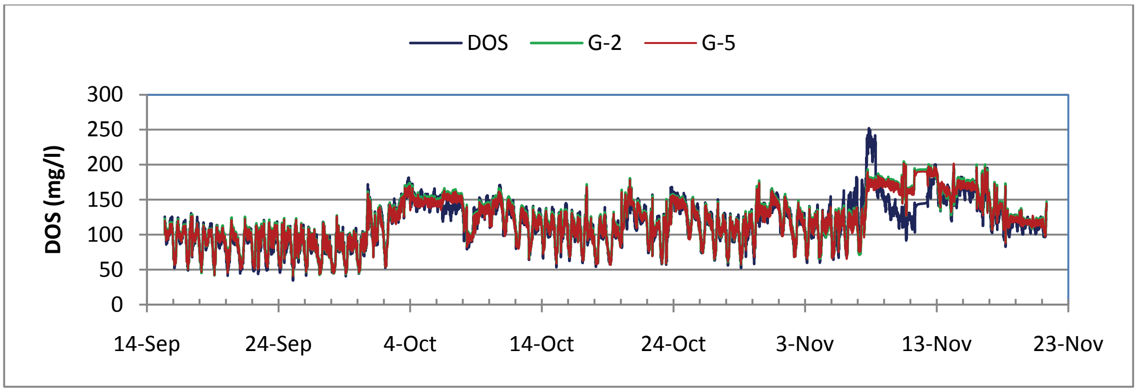


Fig. 3: Predictions of PCR models for G-2 and G-5.

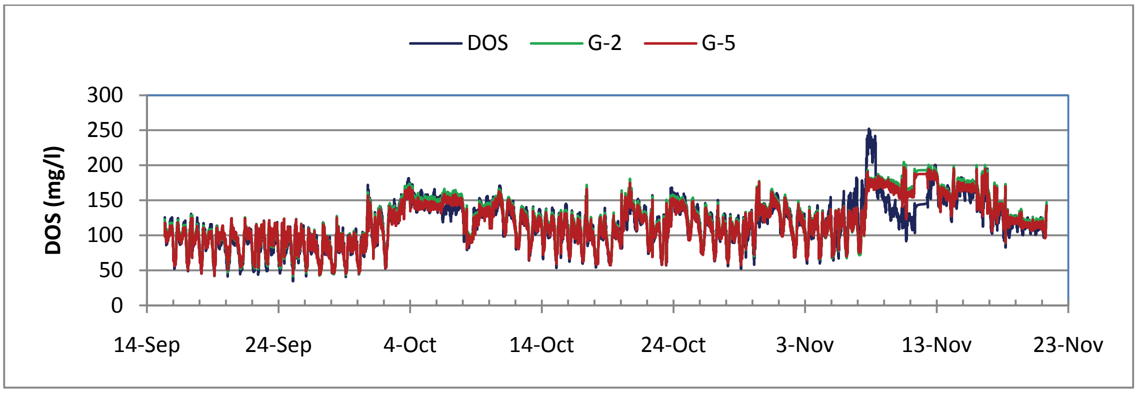


Fig 4: Predictions of PLSR models for G-2 and G-5.

Fig. 2 indicates clearly that although the model statistics were better in MLR, the long-term predictive power is poor. It indicates that the multiple linear regression method is good at explaining a given data set but not suitable for use in the long-term prediction of coagulant dosage in WWTP. Over-fitting with too many predictor variables is one reason for this (Clementi *et al.* 2002*)*. Another reason is the co-linearity problems among X variables. This causes the model to have poor predictive power (Martens and Næs 1991; Johnson and Whichern 2002).

When comparing PLSR and PCA predictions, in most cases the predictability of both models was about the same. In groups C and F, PCR had comparatively better prediction performance. But the statistics of PLSR perform better than the PCR models in C and F. G-6 had the poorest prediction and model statistics. This can be explained by the behaviour of variables in the 3000 samples in this group. Out of the X variables in G-6, TUI is not stable and varies rapidly. Also Qin was larger than all other data sets. The conductivity is also changing within the group. Thus the Y variable does not change rapidly to compensate for the changes in the Xs and this results in the poor performance of the model. This indicates the importance of careful selection of representative samples for statistical calibration.

Use of complete data set

Figs. 5, 6 and 7 indicate the predictive power of the models when the whole data set is used for model calibration. Table 5 shows the statistics of the three models. All three models exhibited very good model statistics and better predictive power for its own data set. PCR has the lowest R^2 and largest RMSE values compare to the others. The best statistics were, as usual, in MLR and it seems the best predictor model for its own data (Fig. 5). But the above experience indicates the poor predictability of MLR for new long-term data.

PLSR does not deliver the best model statistics compared to MLR. But the predictive power of the model with the large data set can be considered acceptable. Fig. 7 shows the performance of PLSR regression.

Table 5. Statistics of calibration models using complete data set.

		R2	RMSE
PLS	**CAL**	0.92	8.9
	VAL	0.92	8.9
PCR	**CAL**	0.87	11.5
	VAL	0.87	11.5
MLR	**CAL**	0.93	8.5
	VAL	0.93	8.5

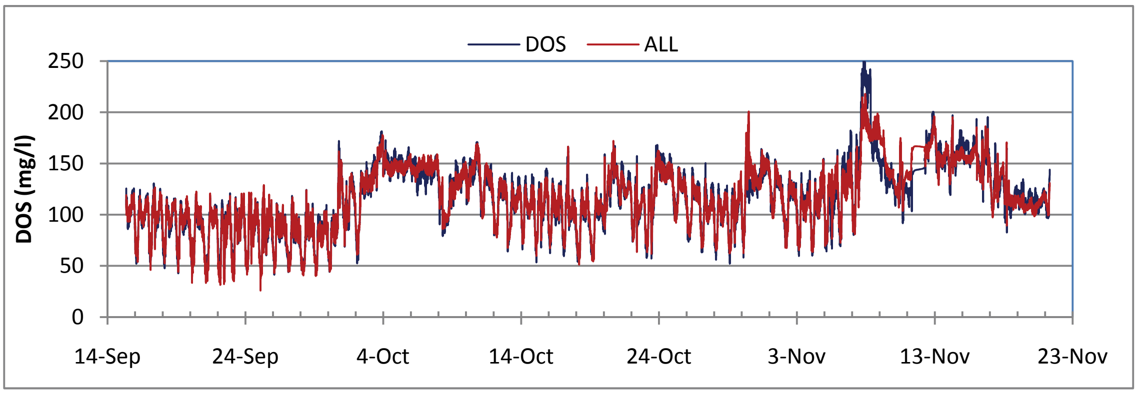


Fig. 5: Measured Y variables (DOS) and predicted Y variable by MLR calibration model using complete data set.

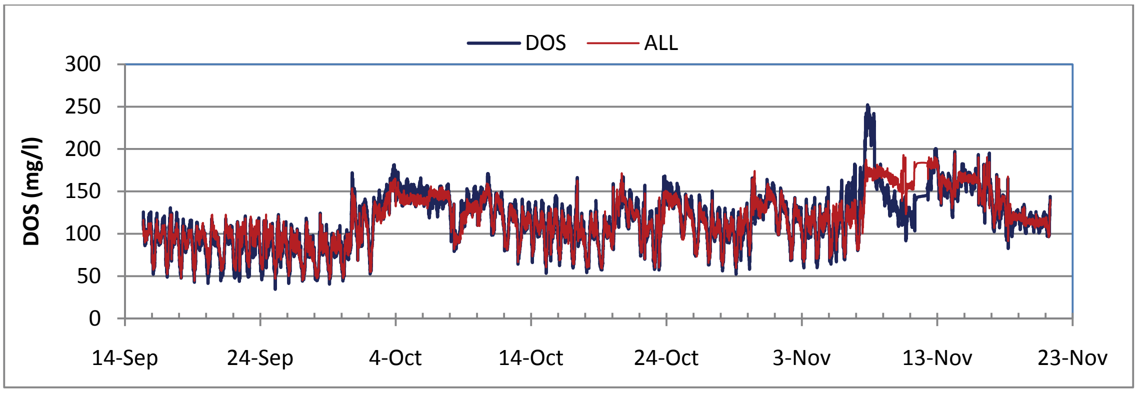


Fig. 6: Measured Y variables (DOS) and predicted Y variable by PCR calibration model using complete data set.

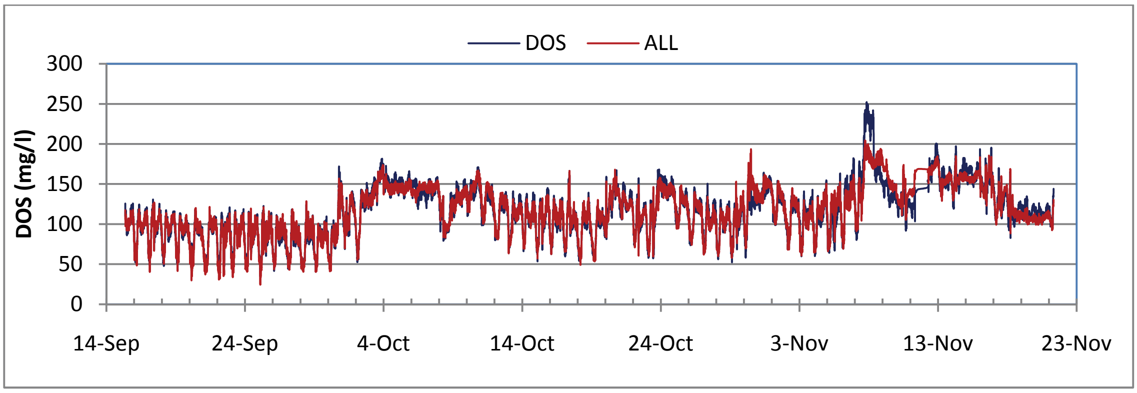


Fig. 7: Measured Y variables (DOS) and predicted Y variable by PLSR calibration model using complete data set.

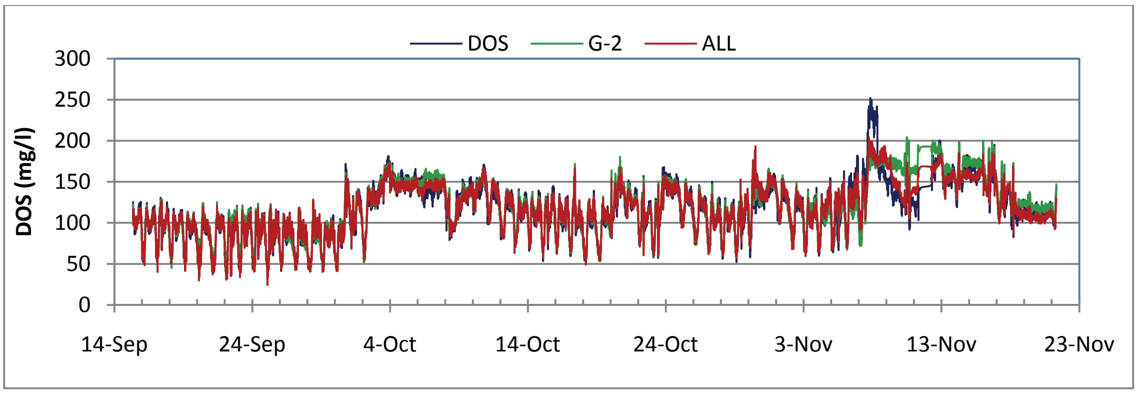


Fig. 8: Comparison of PLSR model calibrated using data of group B with PLSR model developed using total data set. The line 'ALL' is the prediction by the complete data set.

Fig. 8 compares the predictive power of the PLSR calibration model for group 2, which was the best performing model among the 6 groups and the PLSR calibration model for the complete data set. It shows that the predictive power of both regressions were good. But close observations shows better performance with the complete data set. Thus, PLSR calibration model is more precise when the data set is expanded. Since the expanded data set includes large variability of X variables, the sensitivity of the model for varying new X variables is comparatively higher than a model fixed using a small data set with small variability of X variables.

Conclusions

- Multiple linear regression (MLR), principal component regression (PCR) and partial least squares regression (PLSR) are shown to be capable of describing the dosage (DOS) using the other parameters in all the groups. This shows that the DOS has either a direct or indirect relationship with all the water quality parameters that were considered.
- Among the three multiple variable calibration systems, PLSR and PCR performed more or less equally in the prediction of dosage from wastewater quality parameters. But the statistical performance was better in PLS regression. This indicates that the overall performance of PLS was better for the given purpose.
- MLR models have better model statistics, but do not make good predictions for future data that are varying. MLR is better at explaining a dataset than at predicting future values in wastewater.
- The predictive power of a model is highly dependent on the behaviour of the X variables in the data set used for calibration. Therefore it is very important to select more representative data series to calibrate models for chemical dosage control in WWTP.
- Using a large dataset covering all possible behaviours of wastewater quality changes during a longer time is important for robust dosage control system development.
- In this study, we suggest PLSR as the most appropriate and safe technique for developing a robust system for coagulant dosage control.
- We have also discussed a procedure for better calibration of models using statistical software.

References

CAMO 1999, *Manual addendum for the Unscrambler*. CAMO ASA, Oslo, Norway.

Clementi, M, Sergio Clementi1, Gabriele Cruciani, Manuel Pastor, Andrew M. Davis & Darren R. Flower 1997, ‘Robust multivariate statistics and the prediction of protein secondary structure content’, Short communication, *Protein Engineering*, vol. 10, no. 7, pp. 747–749.

Dane, AD, Gerard J. Rea, Anthony D. Walmsley & Haswell, S J 2001, ‘The determination of moisture in tobacco by guided microwave spectroscopy and multivariate calibration’, *Analytica Chimica Acta*, vol. 429, pp. 185–194.

Esbensen, KH 2000, *Multivariate data analysis in practice,* CAMO as, Oslo, Norway.

Fülöp, A & Hancsók, J 2008, ‘Comparison of calibration models based on near infrared spectroscopy data for the determination of plant oil properties’, *Department of Hydrocarbon and Coal Processing, Institute of Chemical and Process Engineering,* University of Pannonia, Veszprém, P.O.Box 158, H-8201, Hungary. (http://www.aidic.it/icheap9/webpapers/425Fulop.pdf)

Garcia-Olmo J, Garrido-Varo, A & De Pedro E 1998, 'Advantages and disadvantages of multiple leaner regression and partial least squire regression equations for prediction of fatty acids', University of Cordoba, Spain, 253-258. (http://scholar.google.com/scholar?cluster=15147206392524298946&hl=en)

Johnson, RA & Wichern, DW 2002, '*Applied multivariate statistical analysis*, Fifth edition, Prince hall, USR, NJ 07458.

Lippm, M 1996, 'Comparison of PLS, PCR and MLR for the quantitative determination of foreign oils and fats in butter fats of several European countries by their triglyceride composition', *Z Lebensm Unters Forsch*, vol. 202, pp.193-198.

Lu, L 2003, 'Model based control and simulation of wastewater coagulation', PhD thesis, Agriculture University of Norway, pp. 126.

Martens, H, Martin Høy, Frank Westad, Ditte Folkenberg & Magni Martens 2001, 'Analysis of designed experiments by stabilized PLS Regression and jack-knifing', *Chemometrics and Intelligent Laboratory Systems* vol. 58, pp. 151–170.

Martens H, Martens M 2000, 'Modified jack-knife estimation of parameter uncertainty in bi-linear modeling PLSR', *Food Quality and Preference,* vol. 11, no. 1–2, pp. 5–16.

Martens, H. & Næs, T 1991, *Multivariate calibration*, John Wiley & Sons Ltd, New York.

Ratnaweera, H, Blom, H & Aasgaard, G 1994, 'Flexible Coagulant Dosing Control System based on Real-Time Wastewater Quality Monitoring', *Chemical water and wastewater treatment* III, Hahn, H.H. and Klute, R (eds), Springer-Verlag, Berlin, pp. 105-116

Risvik Henning 2007, 'Principal Component Analysis' (PCA) & NIPALS algorithm. (folk.uio.no/henninri/pca_module/pca_nipals.pdf)

Robert, NR & Randall, DT 1999, 'Multivariate Methods for Process Knowledge Discovery. The Power to Know Your Process, Paper 252-26, SAS Institute Inc., Cary, NC: SAS Institute. (*www2.sas.com/proceedings/sugi26/p252-26.pdf)*

Tobias, Randall D 1997, 'An Introduction to Partial Least Squares Regression. SAS Institute Inc', Cary, NC: SAS Institute. (ftp.sas.com/techsup/download/technote/ts509.pdf)

William JE & Stephen L. Morgan 1998, 'Outlier Detection in Multivariate Analytical Chemical Data', *Anal, Chem*, vol.70, pp. 2372-2379.

Wold, S, Michael S & Lennart E 2001, 'PLS-regression: a basic tool of chemometrics', *Chemo* (eds) *metrics and Intelligent Laboratory Systems*, vol. 58, pp. 109–130.

Paper III

Improving process control by advanced error detection using floating validation ranges of online measurements

Subhash Rathnaweera, Harsha Ratnaweera, Oddvar Lindholm, Tor Håkonsen

Paper is accepted to the 7th Leading Edge Conference on Water and Wastewater Technologies
(2-4 June 2010)

Improving process control by advanced error detection using floating validation ranges of online measurements

Subhash Rathnaweera*†, Harsha Ratnaweera**, Oddvar Lindholm*, Tor Håkonsen***

*University of Life Sciences, Aas, Norway
**Norwegian Institute for Water Research (NIVA), Oslo, Norway'
***VA-Support AS, www.va-support.no
† Corresponding author (subhash.rathnaweera@umb.no)

Abstract

A robust and accurate tracking system for instrumental errors in online water quality measurements is important as these errors may create critical conditions when used in process control. The most common hardware-based error detection which is performed within the sensor itself may be of limited value. Software-based error detection is a valuable supplement to this. A multivariate statistical method to identify instrumental errors is proposed using full-scale wastewater plant data from Norway and China. Where universal relationships were less accurate, periodic relationships over selected time intervals were found to increase the accuracy of the validation models. Using novel statistical software suitable for online calibrations, an efficient and accurate error detection system can be established for water and wastewater treatment plants.

Introduction

Plant automation and process real-time control (RTC) are common in the water and wastewater treatment industry today. Online measurements of water quality play a critical role in RTC systems. With the development of technology, robust and reliable online sensors for most of the quality and quantity parameters that can be used in RTC are now available on the market.

In RTC systems based on online measurements, the accuracy of the measurement is essential (Rieger *et al*. 2002). Instrumental errors may create critical conditions where they are used in process control. In online sensors for water monitoring, false measurements are common for several reasons. For example, instrument malfunctions, poor calibration of a pH electrode, poor cleaning of the measuring windows of TU sensors, dead electrodes, construction failures such as particles gathering around sensors and collecting tank failures are all very common reasons. In addition to those mentioned, communication failures between the controller and the sensor as well as the controller and the RTC system (programmable logical controller (PLC) or supervisory control and data acquisition (SCADA) system) could produce false values as measurements. Therefore a robust and accurate real-time validation system for detecting error measurements is very important.

The error detection systems for online measurements can be categorised on two levels. These are hardware-based error detection and software-based error detection.

Hardware-based error detection is more common and is performed in the sensor itself. Mostly this is controlled by simple logic electrical circuits, installed by the manufacturer, to identify measurable limits as well as rapid fluctuations in the measurements (Hargesheimer *et al*. 2002). This type of error detection can be categorised as a low level error detector, because it cannot differentiate between unusual process conditions and instrument failures (Colby and Ekster 1997). Hardware level error detection is used in most water quality online sensors but it will not detect errors in enough detail as discussed above.

Software-based error detection is used in advanced process control systems facilitated by SCADA or PLC network systems. This environment provides more space for analysing the measurements and comparing to historical values. Rule-based detection, model-based detection and classification and recognition-based detection are the three different approaches for software-based error detection.

Rule-based detection systems are based on 'if-then' reasoning. This method is suitable for most simple measurements. This method is not appropriate for more complicated situations which cannot be simply solved by an 'if–then' scenario. Classification and recognition-based detection is mainly based on Kohonen classifiers and Artificial Neural Network (ANN) technology (Lorense *et al*. 1997). Here, the system classifies the data into an n-dimensional input space (where n equals the number of samples) based on the closeness of the data points. If the new measured value cannot be classified in the database, the system displays an alarm to the operator about the detection of an error (Hargesheimer *et al.* 2002). Model-based error detection is based on dynamic models of monitored systems and processes. The sensors or online analysers are compared to the process models on a real-time basis. Instrument failure is detected by the distance of the readings from the prediction (Venkatasubramanian *et al.* 2003). According to Isermann and Balle (1997), model-based methods are the most frequently applied methods for fault detection in most online systems in industrial applications.

Although SCADA or PLC network systems are common in most water and wastewater treatment plants, software-based error detection for online measurements are not very common. In this paper, we discuss a novel method of model-based validation of on-line measurements in water and wastewater.

Multivariate statistical calibration is a powerful tool for handling more than one X variable controlling one or more Y variables (Martens and Neas 1989; Johnson and Wichern 1982). Partial least squares regression (PLSR) is a reliable multivariate analytical method for prediction of data. PLSR is a regression method combining both MLR and PC regression proposed by Herman Wold in 1960 (Tobias 1997). Use of this analytical method was not common until the development of powerful computers for calculation. Today very specialized statistical software is available in the market for this purpose.

Relationship of one quality parameter to one or more measurements of water or wastewater has been documented. Blom (1996) showed a good correlation of total phosphate with several measurements like flow (Q), orthophosphate (OP), pH (PH), suspended solids and iron. Lu (2003) was able to predict OP by Q, PH, TU conductivity (CN) and temperature (TE) with 0.96 R^2. He also found relationships with up to 0.92 R^2 between alkalinity and Q, TU, CN, and temperature. Good correlations of 0.97 R^2 was found between total phosphate (TP) and OP as well as TP and TU in the effluent of biological treatment processes (Raphael 2009). He further reported several strong correlations amongst two quality parameters in wastewater. Very good correlations between online TU and laboratory TP and TU with COD and SS were reported by Zeghal *et al*. (1996). Much more documentation about the strong relationships among water quality parameters is available in the literature (Mels *et al*. 2002; Hansin 1996). In this study we used these relationships for validation of measurements.

Development of statistical software capable of online PLSR analysis has opened a new door for novel developments of PLS calibration-based sciences (www.Camo.com). Industrial usage of online regression calibration systems is still not very common. The integration of online PLS calibration software with ‘Near infrared spectroscopy’ in food quality control (www.foss-nirsystems.com) is an example. In this paper, we suggest an error detection concept integrated with online calibration methods.

Methods

The study was conducted at the NRA wastewater treatment plant (WWTP) in Lilleastrøm, Norway. The WWTP consists of a mechanical pre-treatment and a biological treatment unit followed by chemical coagulation treatment. Real-time Q, TU, CN, PH and TE were measured using online sensors, installed between the biological treatment unit and chemical treatment unit.

Five minute averages of online measurements during eleven days, from 1st to 12th of April 2009, were collected. The data were treated by removing known erroneous data such as measurements during instrument maintenance and calibration. Three thousand samples were selected for the study.

Based on the concept presented in the introduction, the data set was used to predict the TU, CN and PH from the other parameters. Prediction equations are given below.

TU~f(Q, CN, PH, TE, interactions of variables, squires of variables)

CN~f(Q, TU, PH, TE, interactions of variables, squires of variables)

PH~f(Q, CN, TU, TE, interactions of variables, squires of variables)

During PLS calibration, data standardisation (which makes all variables contribute with equal strength to the model) and cross validation (which improves the fitness and predictive power of the model) were used (Martens *et at*. 2001). ‘U versus T score’ plots as well as ‘residual y variance versus leverage’ plots were used to detect and remove outliers. To avoid over-fitting of the model, (Clementi *et al.* 2002) insignificant variables were removed according to a Jack-knife validation routine available in the software (CAMO 1999).

A plot of the prediction versus measured values indicates the fitness of the model. Best fit models should give a straight line through the origin with a slope of one. In this analysis, we used these plots to evaluate the fitness of the equations. The coefficient of determination (R^2) is the percentage of total variation in the y-values that is explained by the regression equation. The root mean squares error (RMSE) quantifies the difference between the real value and the value estimated by the model (Esbensen 2000; Martens and Næs 1989). We used R^2 and RMSE in order to demonstrate the quality of the calibrated equations. The statistical software UNSCRAMBLER 9.8 was used for the statistical calibration analysis.

The complete sample set, and 6 representative sets of the total data formed by picking every 3rd, 5th, 9th, 15th and 30th sample starting from the first sample, and every 30th sample starting from the 7th sample. Finally, 6 groups with five hundred consecutive samples (about 2 day's data), named G1 to G6, were used to calibrate the regression equations and predict TU, CN and PH as described above. The model fit and the precision of predictions made by the equations were studied. Based on the precision of prediction and model strength, the most suitable sampling method for the prediction of data was selected. The selected model algorithms were used for further development of the measurement validation method.

In order to find the error detection limits, 3%, 5% and 10% of the predicted values were compared.

To demonstrate the validation method, TU was used as an example. In order to evaluate the error detection ability during sudden changes of parameters, 3 error values (named A, B and C) were simulated in the data series. The error values were simulated by the authors based on their experience.

Method validation was conducted in a pilot scaled WWTP at the Gaobeidian WWTP in Beijing, China. In this study, we used an additional parameter Ortho-phosphate (OP), in addition to the above mentioned parameters. The data collection was held from 28th October to 7th November in Beijing. The data were logged at 10 minute intervals. The samples were tested to prove the method described above using the same criteria.

Results and discussion

a). Suitable sample size for better prediction,

The variable statistics of the data set is given in the Table 1. According to Table 1, it is possible to note that the variables vary widely within the data set. Varying variables during the study was considered as an advantage for the study, since it will demonstrate the ability to track true errors as opposed to natural variations.

Table 1: Variable statistics of data set.

	CN	PH	Q	TE	TU
Min	376.00	6.24	636.05	5.67	40.83
Max	723.00	6.62	928.84	8.99	73.97
Mean	496.37	6.46	759.96	7.15	52.84
Sdev	75.88	0.08	23.27	0.82	7.10
Skwness	0.73	0.27	-0.20	0.26	1.09

Table 2 shows model statistics of prediction equations developed using the complete set of data. According to the R^2 values, the equation for TU represents only 60% of the data. CN and pH equations represent 77% and 59% of the data, respectively. The RMSE value of pH is very small and TU is also at a low level. But the RMSE of CN is large and will affect the precision of the prediction made by the equation.

Table 2: Statistics of prediction algorithms developed with all 3000 samples.

	TU	PH	CNI
R^2	0.60	0.59	0.77
RMSE	4.47	0.05	36.20

Fig.1 displays the prediction of TU using the developed algorithm. The figure shows that the precision of prediction by the model equation is not accurate for even its own dataset.

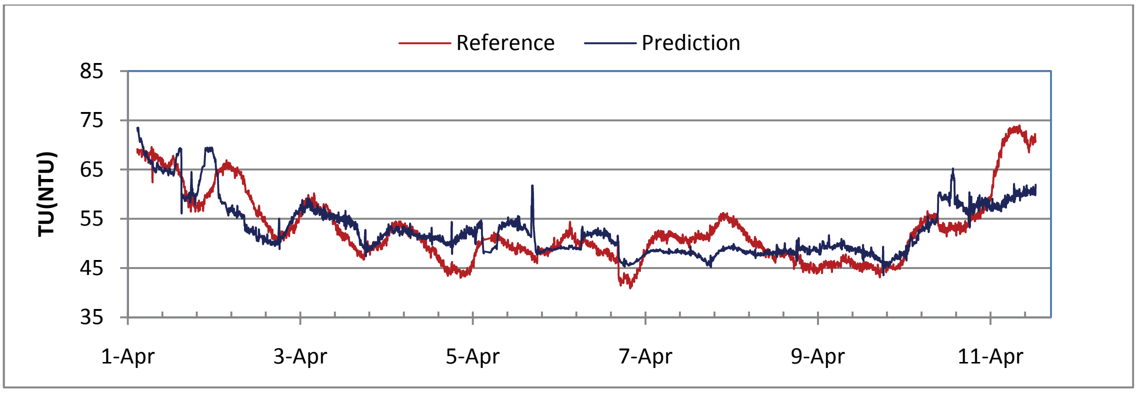


Fig 1: Prediction of TU using all 3000 data for algorithm calibration. R^2 = 0.60 and RMSE = 4.47

Next, we tried different numbers of random samples to develop PLS algorithms that give a good representative equation for the dataset. Table 3 shows the R^2 and RMSE values of the various equations. According to the table, the equations have not been significantly improved by reducing the sample size by sampling the complete dataset.

Table 3: Statistics of algorithms for TU produced by random sampling of the total dataset. The sampling was done every 3rd, 5th, 7th, 9th, 15th and 30th sample starting from sample 1 and also every 30th sample starting from sample 7.

n	3000	1001 (every 3rd)	601 (every 5th)	429 (every 7th)	334 (every 9th)	251 (every 15th)	101 (every 30th)	101 (every 30th from 7)
R^2	0 .60	0 .58	0 .63	0 .60	0 .57	0 .61	0 .52	0 .59
RMSE	4 .47	4 .48	3 .94	4 .52	4 .40	4 .06	3 .30	3 .10

The R^2 values and RMSE values for different equations developed using the sequential sets of data are given in Table 4. The table shows that the model equations for all 6 groups have been improved with R^2 values above 84 and RMSE values less than 2.02. The figure for the prediction of TU with different groups of equations is given below (Fig.2). It illustrates that the accuracy of prediction is better compared to the equation based on the complete data set (Fig.1).

Table 4: Statistics of TU prediction algorithm developed with 500 sample groups.

	G1	G2	G3	G4	G5	G6
R2	0.93	0.89	0.84	0.94	0.88	0.94
RMSE	1.40	1.07	0.98	0.78	0.79	2.02

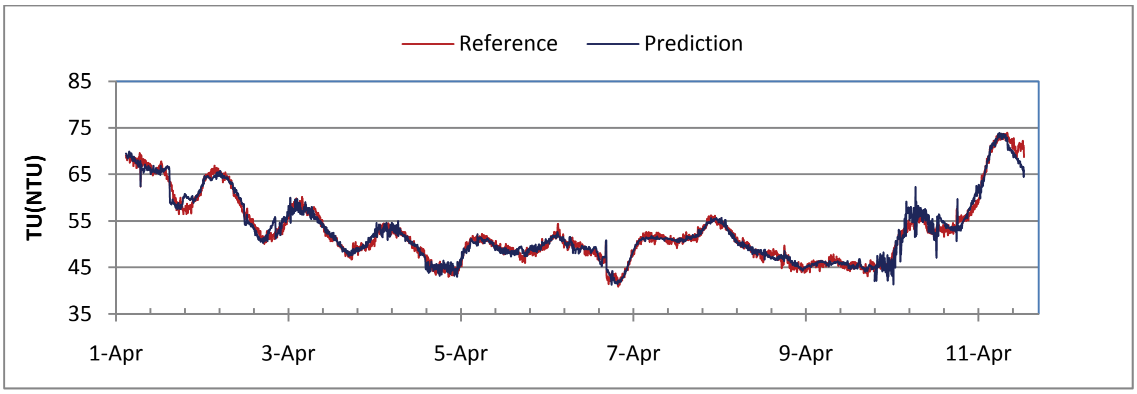


Fig. 2: Using all 500 samples of each dataset for algorithm calibration.

Observations with CN and PH were the same. Thus, based on the above experience, we decided on a grouping of 500 consecutive data samples for the prediction of the parameters as being more reliable for the three parameters of NRA wastewater under consideration.

The number of samples in a group must be decided by the practitioner according to the variables, water quality changes, expected accuracy of the subjected instrument and experience.

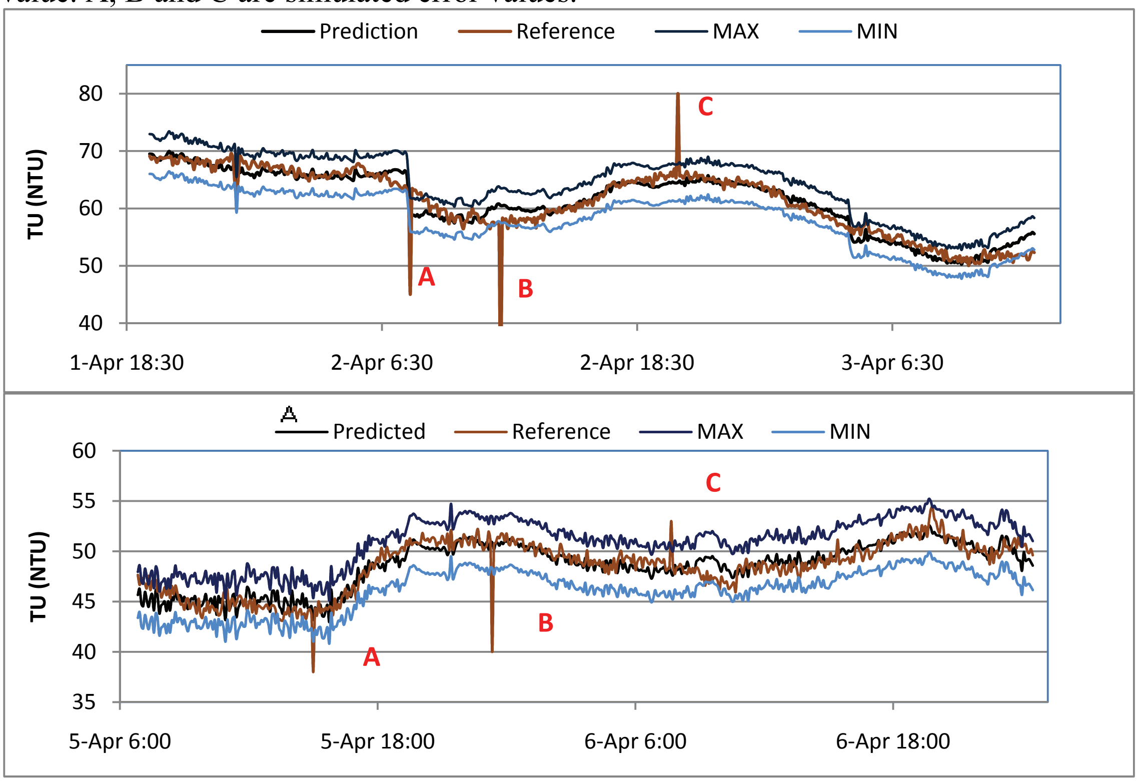
b). Predictability and error detection criteria

The error detection limits need to be reasonable $\pm$ values for the predicted parameter. Using a percentage of the prediction was considered to be the best limit detection. Table 5 shows min, max and mean values of the sample set, and error limits at 10%, 5% and 3% of calculations. An error limit of $\pm$10% showed too wide an error interval which does not precisely detect errors. 3% to 5% were shown to be better error limits, which were within 2 to 4 times of the RMSE values. Therefore, 5% of the average TU, which was on average 2 times the RMSE value, 3% of the average CN, on average about 3.5 times the RMSE value, and 3% of average pH which represented 4 times the RMSE value were used. The error limit values must be selected by the practitioner according to the variables, water quality, expected accuracy of the subjected instrument and experience.

Table 5: Variable statistics and error detection limits. We selected the 5% error range for TU, 3% error range for CN and 3% error range for PH, for the error detection limits.

	TU			CN			PH		
	max	min	mean	max	min	mean	max	min	mean
Value	74.0	40.8	52.8	723.0	334.0	496.2	6.6	6.3	6.5
10 %	7.4	4.1	5.3	72.3	33.4	49.6	0.7	0.6	0.6
5 %	3.7	2.0	2.6	36.2	16.7	24.8	0.3	0.3	0.3
3 %	2.2	1.2	1.6	21.7	10.0	14.9	0.2	0.2	0.2

Fig. 3 shows the prediction of samples, error detection limits and simulated error values of each group. Any value which does not belong within the error limits is detected as an error value. A, B and C are simulated error values.



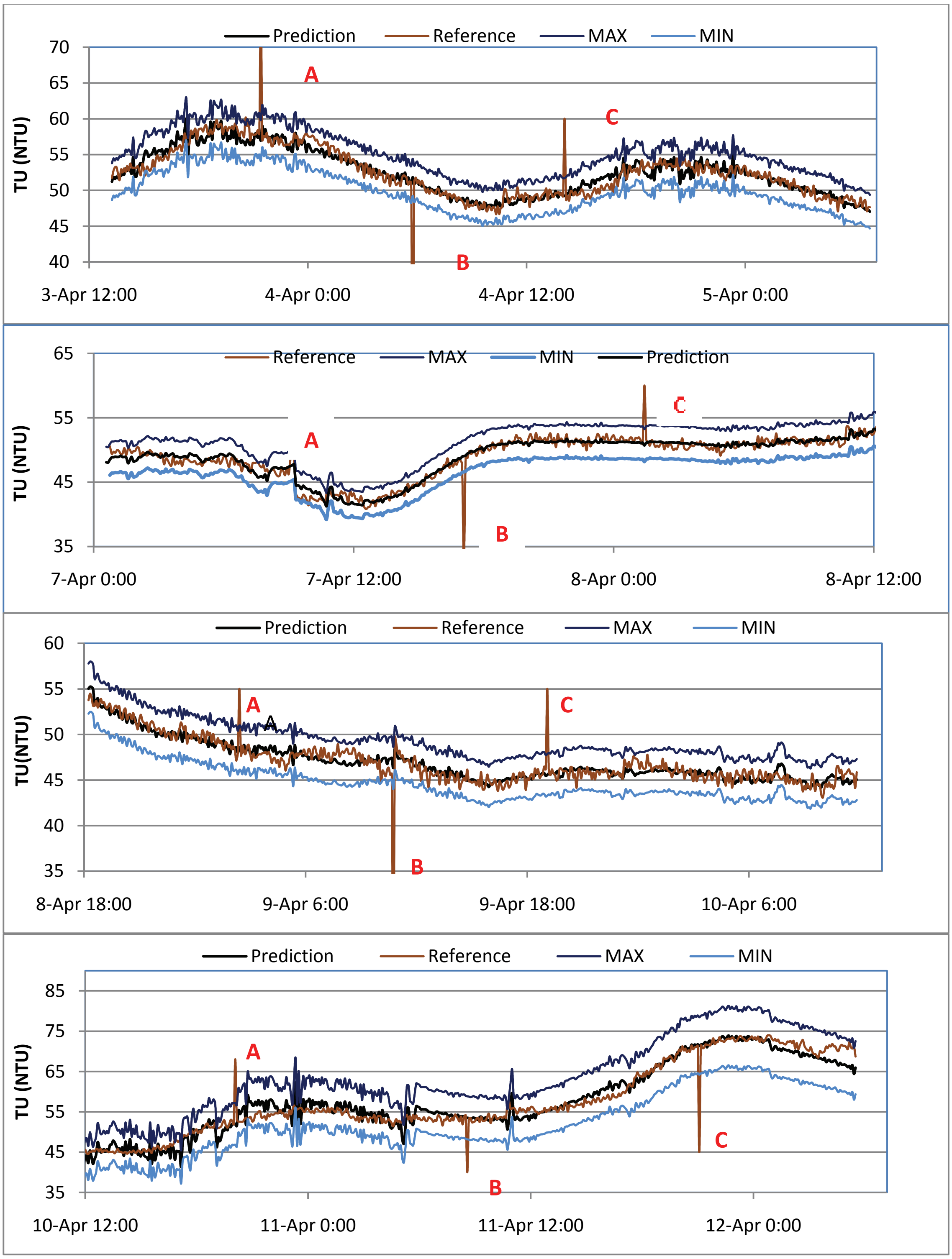


Fig. 3: Continuation of error detection ability of TU, using Q, pH, CN and TE. The error detection limit is $\pm$ 5% of the prediction denoted by MAX and MIN.

As Fig. 3 shows, measurements within 5% of the correct value are accepted and the rest are detected as error values. Fig. 3 shows the detectability for TU as an example. The situation for CN and PH also followed the same trend.

C. Error detection for pH and CN

Figs 4 and 5 show the prediction for the first group of samples and the detected errors for PH and CN. A, B and C are error values used by the authors in order to assess the error detection ability of the systems. For CN and PH, we used 3% of the prediction as the error limit. The figures show that error detection based on this method is acceptable for both CN and PH measurements.

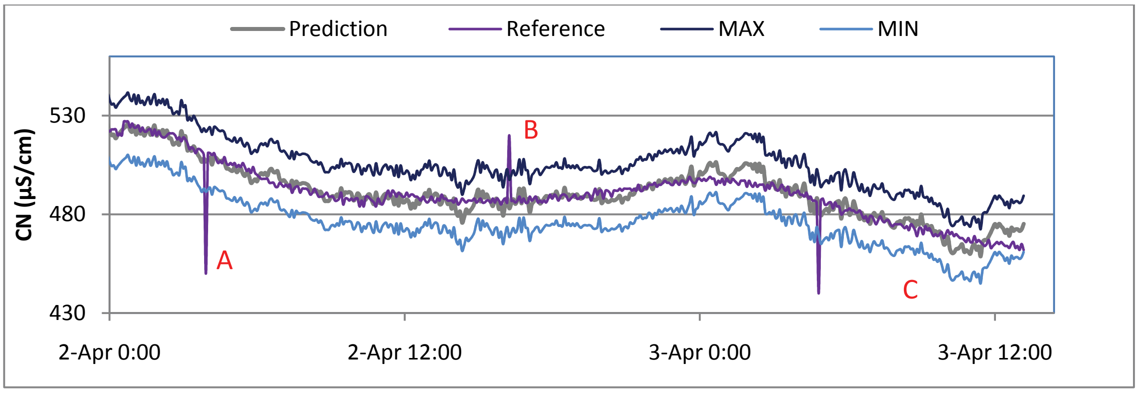


Fig. 4: Error detection on CN (+/- 5 % error intervals) R^2 = RMSE=

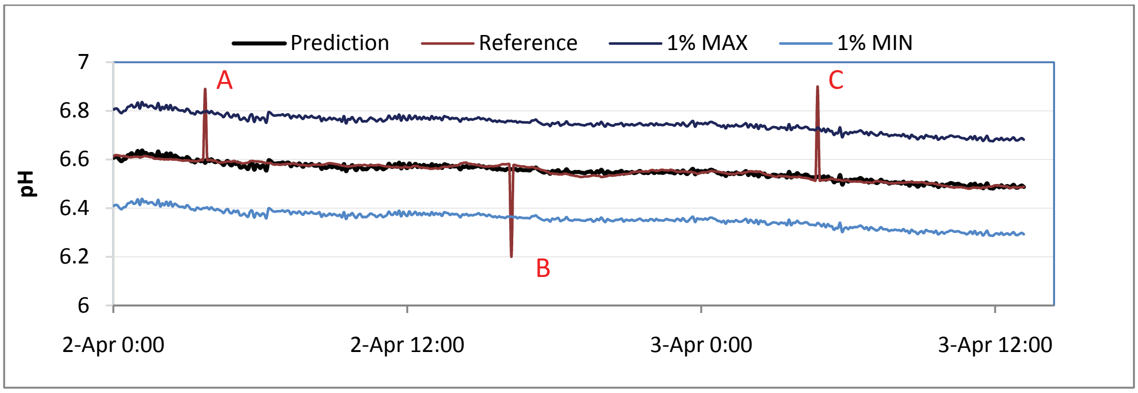


Fig. 5: Error detection on PH (+/- 3% error intervals)

D. Method validation

The validation of the method proposed above at Gaobeidian WWTP in Beijing is presented below. The data were collected after post sedimentation of the AAO treatment process in the pilot scale treatment plant. As in Table 6, the regression coefficients for models of OP and PH, using 500 samples, were good. Fig. 6 demonstrates the error detection ability of each of the abovementioned parameters. Here again, A, B and C are three simulated values to detect the tracking of sudden changes of variables. The results show the method is acceptable for the studied parameters of the WWTP.

Table 6: Model statistics of OP, PH and TU predictions in the Gaobeidian WWTP, Beijing.

	OP	PH
R^2	0.89	0.92
RMSE	0.04	0.06

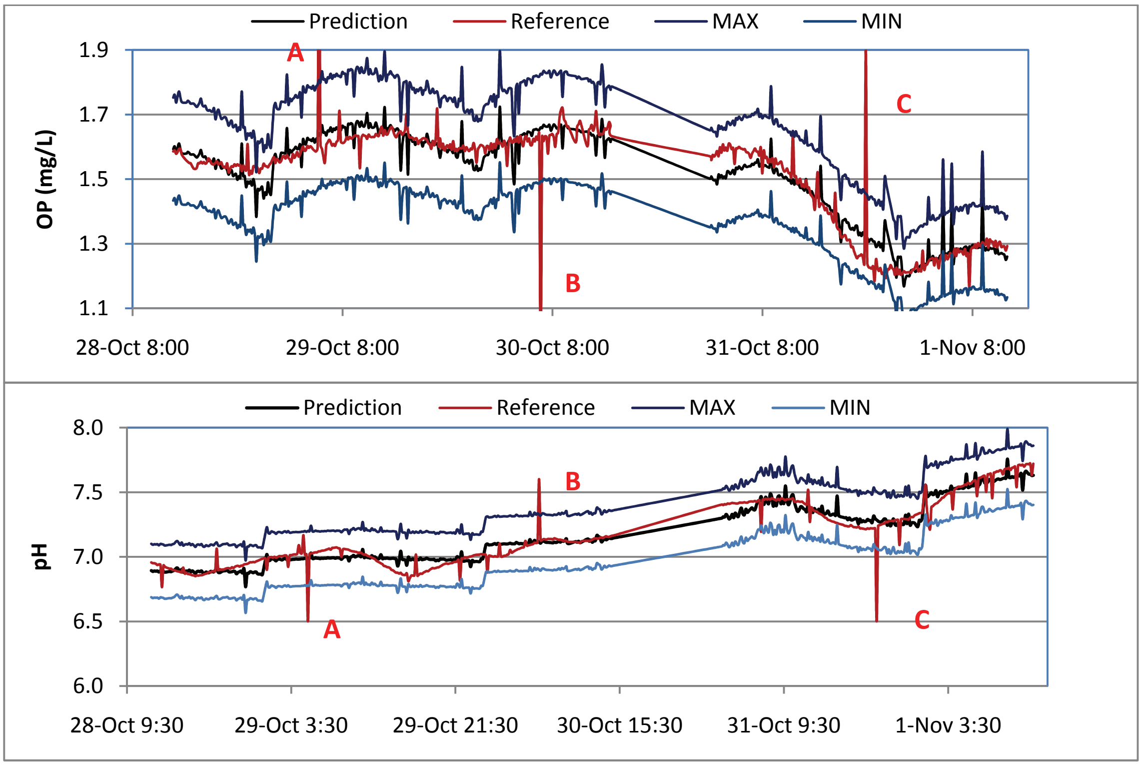


Fig. 6: OP and PH predictions, error limits and error detection are demonstrated in the figures. The error limits for both were 3%.

d). Practical use of the method

According to the error detection concept for online instruments, a high precision in the prediction of the subjected variable is anticipated. On the other hand, we use several X variables which induce no feedback response and probably have no direct relationship with the Y variable. Therefore the prediction algorithms will not have much tolerance for large variations of the Y variable for long periods of time. The algorithms calibrated using large datasets corresponding to long time durations with larger variations of variables were not good at predicting the Y variable accurately. Therefore, developing one universal model for error detection throughout the year is, at least for the wastewater under consideration, not possible.

Relatively small sample numbers, with limited variation of variables, were better at predicting the Y variable more precisely. This made us use a smaller sample number for error detection criteria.

According to the proposed error detection method, it is suggested that the algorithm calibration be done continuously using 'real-time functional' statistical software. Each of the 500 'correct' sample sets may be calibrated individually and the algorithm will be used to predict the next sample. The next sample measurement will be evaluated by error limits detected by the predicted value. Using 'correct' values for sample calibration is very important to prevent wrong calibrations. Online calibration of algorithms is one of the challenges of this method, but it is now possible using novel statistical software designed for online calibration.

Conclusions

- Online water quality can be accurately described by other parameters of the same sample. The accuracy of predicting water quality parameters from other parameters varies with sample size. Thus, the use of a universal algorithm for one parameter is not possible for the wastewater used for testing. Instead, it is possible to use relatively small datasets for more precise prediction with a narrow error range.

- Using the predictability, a novel method for error detection of online parameters used for RTC is successfully demonstrated.

- The number of samples and the error ranges for error detection need to be decided by the practitioner according to parameter variables, water quality and the expected accuracy of the subjected instrument.

- Online calibration of equations is suggested. Statistical software for online calibrations is available on the market.

- Full-scale studies are recommended to further verify the proposed concept.

References

Asok, R & Rogelio, L 1991, 'An introduction to sensor signal validation in redundant measurement systems', IEEE Control Systems Magazine. Vol. 11, no. 2, pp. 44-49.

Blom, HA 1996, 'Indirect measurement of key water quality parameters in sewage treatment plants', *Journal of Chemo metrics*, vol. 10, pp. 697-706.

Camo 1999, *Manual addendum for the Unscramble*, CAMO ASA, Oslo, Norway.

Clementi, M, Sergio Clementi1, Gabriele Cruciani, Manuel Pastor, Andrew M. Davis and Darren R.Flower 1997, 'Robust multivariate statistics and the prediction of protein secondary structure content', Short communication, *Protein Engineering*, vol. 10, no. 7, pp. 747–749.

Colby, S & Ekster, A 1997, ‘Control-Loop Fault Detection’. *Water Environ, Technol*, vol. 9, pp. 11. 43.

Hargesheimer, EE, Osvaldo, C & Jarka, P 2002, ‘Online monitoring for drinking water utilities’, *CRS PROAQUA*, pp. 320-323.

Esbensen, KH 2000, *Multivariate data analysis in practice*, CAMO as, Oslo, Norway.

Hansen, B 1996, ‘Dosing control of coagulants based on on-line monitoring of suspended solids in sewage treatment plants,, In: Chemical water and wastewater treatments VIII, HH Hahn, E Hoffmann, H Ødegaard eds (eds), IWA publishing, London, pp 137-145.

Harald Martens, Martin Høy, Frank Westad, Ditte Folkenberg & Magni Martens 2001, ‘Analysis of designed experiments by stabilized PLS Regression and jack-knifing’, *Chemometrics and Intelligent Laboratory Systems*, vol. 58, pp. 151–170.

Camo: http://www.camo.com/rt/Products/Unscrambler/olup.html

Foss-nirsystems: http://www.foss-nirsystems.com/software.html

Isermann, R &. Ballé, P 1997, ‘Trends in the application of model based fault detection and diagnosis of technical processes’, *Control Eng. Practice,* vol. 5, no. 5, pp. 709-719.

Johnson, RA & Wichern, DW 1982, *Applied multivariate statistical analysis*, Fifth edition, Prince hall, USA, NJ 07458.

Lorenz, A, M. Blum, H. Ermert, T. Senge 1997, ‘Comparison of Different Neuro-Fuzzy Classification Systems for the Detection of Prostate Cancer in Ultrasonic Images. IEEE Ultrasonics Symposium Proceedings, 1997 - ieeexplore.ieee.org.

Lu, L 2003, ‘Model based control and simulation of wastewater coagulation’, PhD thesis, Agriculture University of Norway, pp. 126.

Martens, H & Næs T 1989, *Multivariate calibration*, John Wiley & Sons Ltd, New York.

Mels, AR, van Nieuwenhuijzen, AF & Klapwijk, A 2002, *Turbidity related dosing of organic polymers*. In: Chemical water and wastewater treatments VIII, HH Hahn, E Hoffmann, H Ødegaard (eds), IWA publishing, London, pp 71-80.

Raphael, M 2009, ‘Indirect Measurement of Water QualityParameters in Wastewater Treatment Plants’, Bachelor Thesis, University of Applied Sciences, Ostwestfalen-Lippe, Germany.

Rieger, L, Alex, J, Winkler, S, Boehler, M & Siegist, H 2002, ‘Progress in sensor tecknology – progress in process control? Part 1: Sensor property investigation classification’. In: N.Fleischmann, G. Langergraber & R. Haberl (Eds.) University of Agriculture sciences, Vienna, pp. 65-72.

Practical solutions:
http://www.varianinc.com/image/vimage/docs/products/dissolution/shared/practicalsolutions/20070802_Practical_Solutions_Vol. 2.

Tobias, Randall, D 1997, 'An Introduction to Partial Least Squares Regression. SAS Institute Inc', Cary, NC: SAS Institute.

Venkatasubramanian Venkat, Raghunathan Rengaswamy, Kewen Yin & Surya N. Kavuri 2003, 'A review of process fault detection and diagnosis Part I: Quantitative model-based methods', *Computers and Chemical Engineering,* vol. 27, pp. 293-311.

Zeghal, SP,Sauvegrain, PJ & Vignoles, C 1996*, Chemical addition control for phosphorus removal in primary sedimentation tank,. In: Chemical water and wastewater treatments VIII*, HH Hahn, E Hoffmann, H Ødegaard eds (eds), IWA publishing, London, pp 147-158.

Paper IV

Multi-parameter based real-time coagulant dose control system for wastewater treatment

Subhash Rathnaweera

Paper submitted to the International Journal of Environmental Modelling & Software

Multi-parameter based real-time coagulant dose control system for wastewater treatment

Subhash Rathnaweera*†

*University of Life Sciences, Aas, Norway

† Corresponding author (subhash.rathnaweera@umb.no)

Abstract

Coagulant demand is strongly related to the quality parameters of the treated water or wastewater. Affordable and reliable online sensors for most quality parameters are now readily available, yet rarely used in process control. A robust coagulant dose control system based on online measurements of water quality parameters is presented. Popular multivariate analytical method, partial least square regression was used to build the relationship between the coagulant dose and wastewater quality parameters. The system was tested on four wastewater treatment plants in Norway and China. Coagulant savings up to 15% in Norwegian plants and up to 31% in Chinese plants were observed. This paper presents the method, function and experiences of the full-scale implementation of the system in different wastewater treatment plants.

Introduction

The chemical coagulation process has become popular in many countries due to its efficiency, flexibility and its robustness for climatic and shock loads (Ratnaweera *et al.* 2002). The coagulation process is induced by adding chemical coagulants to the water and, typically, letting the particles be agglomerated in a flocculation tank. The flocculated particles are separated by sedimentation, flotation or filtration.

Capital costs of a chemical wastewater treatment plant (WWTP) are generally much lower compared to the biological P removal plant. Nevertheless, the operational cost of chemical treatment plants could be relatively high (Leentvaar *et al.* 1979). Hangouet *et al.* (2007) reported that the chemical costs may represent up to 20% of the operating cost of an average treatment plant. This percentage varies from plant to plant with the quality of influent, required treatment quality and the management of the WWTP. Choosing the optimal coagulant and coagulant dosage are the two most important management factors in addition to providing the optimal process conditions.

Ortho phosphates (OP), suspended solids (SS) or turbidity (TU) and the pH of wastewater are considered as the main parameters influencing the coagulant demand (Ratnaweera 1991, Gillberg *et al.* 1996). In addition to those, most of the other influent quality parameters indirectly contribute to the coagulation reaction. Significant daily changes, seasonal changes and time changes of influent quality due to climatic factors are very common and well documented. Hansen (1996) has studied the day and night changes of wastewater at the plant gate and observed four times larger flow and BOD during the day time compare to night. Holmquist (2004) has reported drastic variations of phosphate in influent wastewater during one day. Figure 1 illustrates the variation of Flow, OP and TU during three consecutive days in NRA WWTP Lillestrøm (NRA), where a major part of the practical work of this study was conducted. Figure 1 confirms that none of the parameters are proportional to each other. Subsequently, it is not possible to predict the optimal dosage based on one or two of these parameters, when the process is heavily dependent on all three parameters.

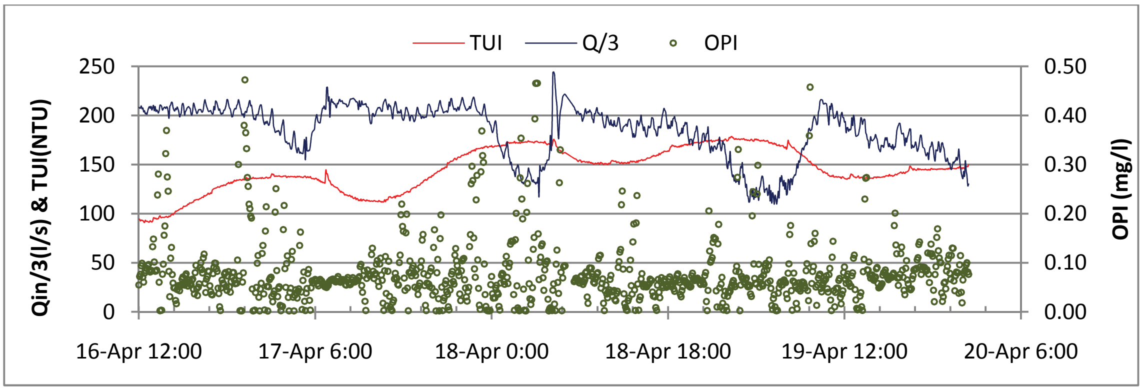


Fig.1.The WW quality variation after the biological treatment of NRA WWTP during three consecutive days in year 2008. QIN is the inflow in l/s, TUI is influent turbidity in NTU, OPI is influent Orthophosphate concentration in mg/l.

It is reported that most of the chemical water and wastewater treatment plants are adopted to use at least a flow proportional coagulant dosing control (CDC) system (Dentel 1991). Looking in to the Norwegian conditions, in a survey among major drinking water treatment plants (DWTP)s and WWTPs, Ratnaweera (2004) reported that 83% of the WWTPs and 80% of the DWTPs use flow as the only control parameter, sometimes in combination with pH over-run function for CDC. Integration of quality parameters like TU or colour was found in less than 20% of the plants.

According to Ratnaweera (2004), some of the DWTPs with raw water source from lakes, where the raw water quality remains more or less unchanged over the year, may obtain satisfactory results with constant dosing, irrespective of the flow, particles or colour etc. However, the same survey reported that 2% of Norwegian WWTPs use constant coagulant dosage irrespective to the flow and quality of the wastewater. Obviously, such a CDC will be far from the optimum. In constant dosing, though plant operators normally use laboratory Jar testing results for evaluating the appropriate dosages, higher dosages are common to ensure adequate treatment efficiencies. This, most probably, will result in over dosing under normal conditions and under-dosing during shock loads.

During the last few decades, online water quality monitoring technology has been developed considerably. (Jeppsson *et al*. 2002, Vanrolleghem 2003). Today, robust and affordable online quality sensors are readily available for most of the relevant parameters. Measurements of parameters like flow, OP, pH, conductivity, SS, TU, Zeta potential, etc. are some examples which are widely used in the industry today (Hansen 1996, Lu 2003). The above mentioned survey shows that though reliable, cheap and robust online measuring equipments are available and used in process control, usage of most of them in CDC is not yet widespread. There has been a significant positive development during the last few years in this regard (Holmquist, 2004, Zeghal *et al.* 1996, Hansen 1996, Baxter 2001, Mels *et al*. 2002, Devisscher 2002, Flower 2004, Adgar *et al.* 2005, Sueg-Young Oh *et al.* 2005, Wu and Lo 2008, Sergo *et al.* 2008.)

Coagulation is a well defined process. E.g., if three identical water samples are added identical dosages, and coagulate, flocculate and sediment them identically, the result will be three identical effluent samples. Such a process can be mathematically described. Construction of a conceptual model, however, may be a challenge due to the complex nature of the process. Based on this, Ratnaweera *et al.* (1994) proposed a concept for multi-parameter based experimental coagulant dosage control system (XCDC). Accordingly, variable influent conditions with corresponding dosages should enable estimating the effluent results. In order to practically implement this concept, the system needs to be mathematically modelled using as much as possible datasets with influent, effluent and dosing conditions.

In this paper, strategies related to the latter concept are elaborated, both on detection and minimization of negative impacts on dosage prediction. A non-linear multivariate calibration method for validation of online instruments is discussed.

Method and materials

Wastewater treatment plants

The studies were conducted in four municipal WWTPs, two were in Norway and two were in China.

NRA is a WWTP, built inside of a rock in Lillestrøm, Norway. The wastewater treatment plant serves the population and several small and large scale industries in three municipalities Lørenskog, Rælingen and Skedsmo. Presently the plant is running with a capacity of about 110 000 p.e. and 50 000 m^3/day.

The treatment scheme of NRA consists of mechanical pre-treatment with screening, sand removing and pre-sedimentation, floating bio-film bio-reactor for nitrogen removal process followed by a chemical coagulation process. Coagulants are injected to the opposite direction of the sewage flow to a pressurized pipe in order to develop flash mixing conditions. The treated water is then distributed to six sedimentation tanks. The traditional coagulant dosage was a real-time flow proportional dosing system with time related coefficients. Additionally, the amount of coagulant per cubic meter of water was altered by operators according to both influent and effluent quality. The effluent of the NRA WWTP generally reached the effluent quality demands. In the year 2007 the plant maintained the effluent quality with phosphorus removal above 95%, Nitrogen removal 74% and removal of organic matter was 95% (RA-2, 2007).

The HIAS WWTP is situated on the eastern bank of the Norway's largest inland water body, Lake Mjøsa. The WWTP serves about 50 000 people and industries, mainly food processing industries, in three municipalities, with daily sewage flow about 20 000m^3. The treatment process consists of primary mechanical treatment, biological treatment followed by a clarifier and chemical treatment process with post sedimentation. The chemical treatment, where XCDC was studied, consisted of coagulant injection, polymer addition and then eight treatment lines with flocculation chamber and sedimentation tank.

Xiao-Hong-Men (XHM) is one of the 14 largest WWTPs in Beijing. The plant is located in a land area of 47 hectares in Chaoyang District, at the bank of Liangshui River, which serves as the recipient for treated wastewater. The plant is designed to serve 2.4 million people and industries in 223.5km^2 area in Beijing. Present sewage inflow capacity to the plant is 600 000m^3 in a day. The treatment process consisted of four parallel lines with physical pre-treatment, biological Anoxic, Anaerobic and Aerobic (AAO) treatment process followed by post sedimentation tanks. Chemical coagulation has been started recently by adding coagulant to effluents of the AAO process and separation is carried out in four sedimentation tanks.

Gaobeidian (GBD) WWTP is currently the largest sewage treatment plant in China. The WWTP serves a catchment area of about 96 km^2, with design capacity of 1 000 000m^3/d. The treatment process of GBD WWTP was identical to the XHM WWTP. The XCDC studies were conducted in a pilot scale treatment plant in the GBD WWTP premises. The pilot plant was designed with same treatment process of the WWTP with 72m^3 inflow per day. Initially the coagulation was not practiced in the pilot. We designed the pilot plant to add coagulants after settling of the AAO effluent in one sedimentation tank. The coagulated particles were settled in another tank. Influent to the pilot plant was raw wastewater from the WWTP and the capacity of each sedimentation tank was 34m^3.

Data collection for CDC

In all the treatment plants influent flow (QI), turbidity (TUI), conductivity (CNI), pH (PHI) and Temperature (TEI) were measured before the coagulant injection. The influent orthophosphate (OPI) was measured in HIAS, XHM and GBD WWTPs. The pH after injecting coagulants (PHO) was a common measurement in all the trials while the streaming current (SCO) was measured only in NRA. At the far end of sedimentation tanks the effluent turbidity (TUO) was measured. In HIAS, the effluent orthophosphate concentration (OPO) was also measured together with TUO.

The two Norwegian WWTPs had modern Supervisory Control and Data Acquisition (SCADA) systems including a real-time flow proportional dosing system. All the XCDC sensors were connected the SCADA systems of the plants. The XCDC system was designed to run in a programmable logical controller (PLC) with a special software. The PLC coordinated with SCADA system to get information, processed and send the dosage prediction signal back to the SCADA, which controlled the dosing pump. Data were logged in the main SCADA system and were able to download as time averages. In the two Chinese plants, instruments were directly connected to the PLC and the dosage pump was controlled by the PLC output. The data was logged in the PLC as 10mins averages.

In order to validate the accuracy of the online measurements, several campaigns of lab analysis were conducted. Daily composite samples of influent and effluents of two Norwegian plants were also measured each week in the laboratory.

All the online instruments were calibrated and maintained properly as the directions of the manuals. Laboratory analyses were conducted according to Norwegian standards.

Data selection and model calibration,

The difference of the water quality between influent and effluent can be considered as the result of the treatment process of the WWTP, which consisting with coagulation, flocculation and sedimentation process. The efficiency of chemical process is strongly influenced by the coagulant dosage and subsequently reflects on the effluent quality.

The mathematical relationship between the dosage and effluent quality can not easily be modelled due to the influences of the fluctuating water flow, the retention time and behaviour of the sedimentation tanks. The tracer tests conducted in sedimentation tanks of NRA WWTP showed that 0.70 of the hydraulic retention time gave a reasonable relationship between influent quality and the effluent quality. We used this relationship to select the samples for calibration.

The effluent measurements were adjusted by shifting 0.70 times of the hydraulic retention time of each plant, to explain the respective coagulant dosage. The sample selection criteria were varied in each selection due to available effluent measurements, demanding effluent quality, treatment needs and duration of the study.

Multivariate calibration

Partial Least Square Regression (PLSR) is one of the most reliable regression methods to predict one or more dependent variables using more than one explanatory variable. Authors have studied the efficiency of PLSR for future dosage prediction of wastewater, compared to multiple leaner regression and principle component regression. The study proved the suitability of PLSR over the other two methods for dosage prediction. A detailed explication about mentioned assay is available elsewhere.

In present study we systematically used the PLSR method to calibrate models to predict coagulant dosage using the online measurements of wastewater. The calibrated algorithm consisted of all the variables, cross effects of the variables and square effects of the variables.

Coefficient of determination (R^2) is the percentage of the total variation in the y-values that is explained by the regression equation. Calibration R^2 and validation R^2 were used to demonstrate how good the model explains calibration as well as validation data sets. Root mean square error (RMSE) quantifies the difference between the real value and the estimated value by the model. RMSE for calibration set and RMSE for Validation set were used to study the prediction accuracy of the models (Esbensen 2000, Martens and Næs 1991). Statistical software Unscrambler 9.8 was used for all calibration analysis in the study.

Result and discussion

Data selection

As explained above, effluent measurements were adjusted by shifting 0.70 times of the hydraulic retention time of each plant, to explain the respective coagulant dosage. The sample selection criteria were varied in each selection due to available effluent measurements, demanding effluent quality, treatment needs and duration of the study. TUO was the common online quality parameter used in all the studies. In HIAS, we used online OPO as quality parameter, while laboratory measurements of TPO and OPO were used in the other treatment plants. We identified a need of using PHO for sample selection in HIAS since the influent WW experienced higher coagulation pH conditions in the plant. TUO was failed as a quality parameter in GBD pilot test because the different between TUI and TUO were not significant. Thus we had to depend on the laboratory OPO and TPO measurements for the sample selection. Shorter duration of studies in GBD and XHM forced to use manipulated dosages in sample sets. These manipulations were carefully made in order to save limited calibration samples. The 2nd calibration of HIAS was not successfully functioned due to frequent sock loads. We had to calibrate 3rd models set using the QIN as a quality control parameter. (All these special situations will be discussed separately in this paper.) Table 1 presents the criteria used in different trials.

Table 1. The sample selection criteria of four different WWTP in different calibrations. TPO(lab) denotes the laboratory measurements of effluent TP. TUO and OPO were online measurements

WWTP	1st calibration	2nd calibration	Remarks
NRA	TUO < 5NTU TPO(lab)< 0.5mg/l	TUO = 2 and 5NTU TPO(lab)< 0.5mg/l	In 2nd calibration, less than 2NTU was considered over dosing
HIAS	TUO < 8NTU OPO < 0.1mg/l	TUO < 2 and 7NTU OPO < 0.08mg/l PHO= 6.2 and 7	In 2nd cal. pH overrun faction was used stick. Calibrated 3rd model set using QIN>250l/s
Xiao Hong Men	TUO < 10NTU TPO(lab) < 1mg/l	TUO less than 8NTU TPO (lab)< 1mg/l	Dosage was manipulated when necessary
GBD	TPO < 1mg/l Manipulated some dosages according to Jar test results.	TP(lab) < 1mg/l and manipulated dosages Lab OPO was observed	TUO was not a good indicator. Only the TPO and dosage was manipulated when necessary

Model calibration and function

The typical formula of model algorithm is given below.

DOSAGE= f(TIM, DAY, TUI, CNI, PHI, TEI, OPI, PHO, SCO, interaction among variables, squares of the variables)

The interaction effects of TIM and DAY with the other parameters were excluded from the models, because it did not showed a clear relationship. The OPI was not available in NRA while SCO was only used in NRA. The other parameters were common for all the studies. Two model calibrations were conducted in each test to find a robust system with optimal dosage calibrations. The 1st calibration was done using the historical data with WWTP's traditional dosage system. The 2nd model set was calibrated using the data collected with XCDC dosage control.

The calibrated models were run in the plant without active control of the dosage for 45 days in NRA and 14 days in HIAS in order to observe the real-time behaviour of predictions. Fig. 1 shows the observation in NRA during the observed period. The figure shows that the XCDC predicted dosages were satisfactorily followed the pattern of the dosage prediction in the plant with more sensitivity. Active dosing was started once the dosage prediction was proved successful.

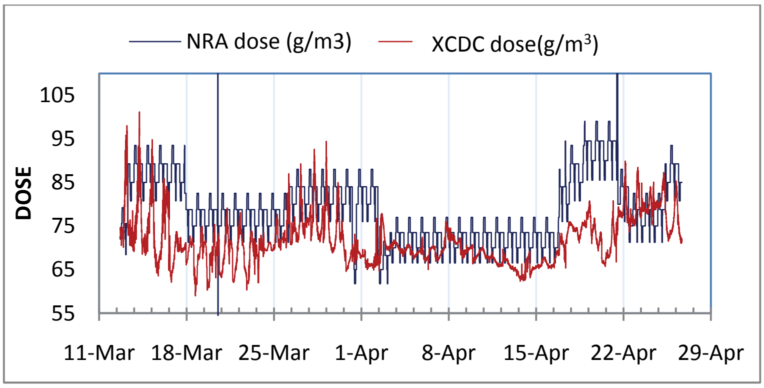


Fig. 1: The WWTP's traditional dosing method (NRA dosage) compared to the new dosage (XCDC dosage) predictions, a comparison during the testing period without active dosing. Changes in the NRA dosage 'levels' are due to manual changes by the plant operators. If not, the dosing should vary in a same base level.

The duration of 1st model running was varied in each plant. The 1st models run for 2 months in NRA and HIAS. It was only 4 days in XHM and GBD due to the short testing period. The 2nd models were calibrated using the data with 1st XCDC run. Fig. 2 shows that the 2nd models were much more sensitive and produced lower average dosage compared to the 1st models. This high sensitivity was identified in all the tests except HIAS WWTP.

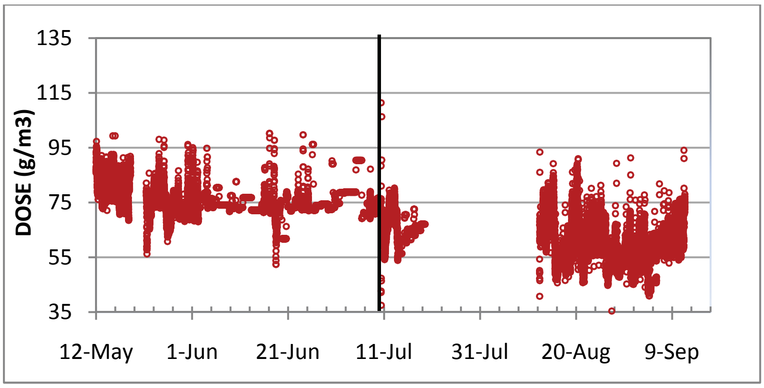


Fig.2: The active dosages of XCDC system in NRA. The preliminary models were well functioning up to the time noted with the back vertical line, and then secondary re-calibrated models were functioning. (The empty space is lost data due to SCADA system renovation.)

The 2nd model of GBD appeared too sensitive to the parameter TIM, which was hour of the day (Fig. 3). The influent quality in the pilot test was not rapidly changed while the TIM was varied from 0 to 23. This variability of the TIM was influencing the dosage prediction with sensitive 2nd model. We removed the variable from the algorithm in order to overcome the situation.

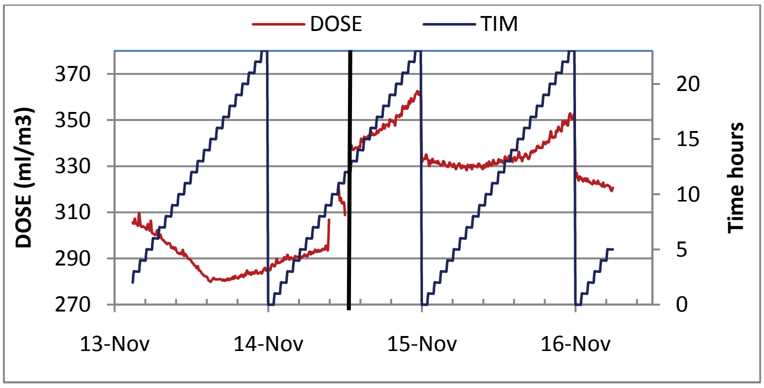


Fig. 3: The 2nd calibration was too sensitive to the TIM. 1st model (first half in before black line) did not show any correlation to TIM while the 2nd model (after the line) it started to follow the trend.

Too many outliers were observed in the 2nd calibration of HIAS system. Furthermore, the model statistics were not improved by the 2nd calibration and models did not function properly either. Finally we identified that the session with 1st models had been subjected to few poor dosage prediction events due to shock loads producing 'outlier data' for the calibration. In addition to that, the period in which the 2nd models were run was subjected to frequent shock loads. (Fig. 4. Left)

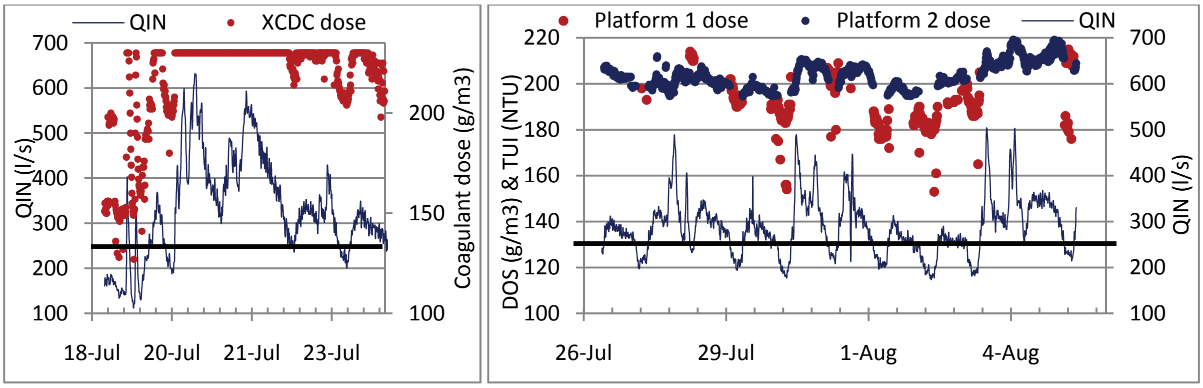


Fig. 4: Left: *the shock load during the 2nd model functioned in XCDC system. The dosage has been stopped at the maximum limit.* Right: *The two platform model system in function. Platform dosage 1 is the dosage prediction by the model used for normal inflow. Platform 2 dosage is the dosage prediction simulated using the model developed using only the samples with QIN over 250l/s. The XCDC software is been designed to shift the platform from 1 to 2, when QIN is over 250l/s.*

Effluent of Biological process with subsequent sedimentation was the normal influent for the chemical treatment in HIAS. This water was quite dilute with less variability. The biological process has a designed maximum capacity of 270l/s. In practice whenever the QIN exceeds 250l/s, the surplus water bypassed the biological treatment and directly entered the chemical process. This bypassed water strongly affected the quality of influent producing "shock loads" to the chemical treatment. These events frequently occurred during snow melting and raining. In order to solve this problem, we categorized the events as two water types in the same WWTP. Hence, we calibrated a separate model for bypassing events by extracting samples with QIN over 250l/s. XCDC system software was designed to run two models in two platforms and shift the platform from 1 to 2, when QIN is over 250l/s. the Fig. 4 (Right) demonstrates the dosage prediction according to the QIN.

XCDC improvements

To see the possibility of further improving the dosage, process sensitivity investigations were conducted in NRA.

The plant was run with reduced predicted dosages in order to see the dosage-response relationship and the lowest dosage that can be achieved in a full scale plant. The actual dosage was reduced up to 40% of the predicated dosage, stepwise by 10% intervals. The effluent quality during this testing period is presented in fig 5. As expected, a clear increase of the effluent quality was found with reduced dosage. With the dosage of 70% of the prediction, two sudden increases of effluent turbidity were experienced. One out of them was due to a malfunction in the sludge pumping of sedimentation tanks, and once the pump is in function the TUO was compensated. This had no relationship with the coagulation. The second was due to false dosage response of a sudden shock load. The situation was controlled by increasing the dosage to 80%. The testing was continued up to 40% of the dosage prediction.

The sudden peak at 70% indicates that the percentage reductions of a dosage prediction will not correctly response the changes of parameters.

These results indicate that it has a potential to reduce the actual dosage at least by 20% - 30% of the predicted dosages. It is necessary to recalibrate the model with longer data series to cover extreme conditions.

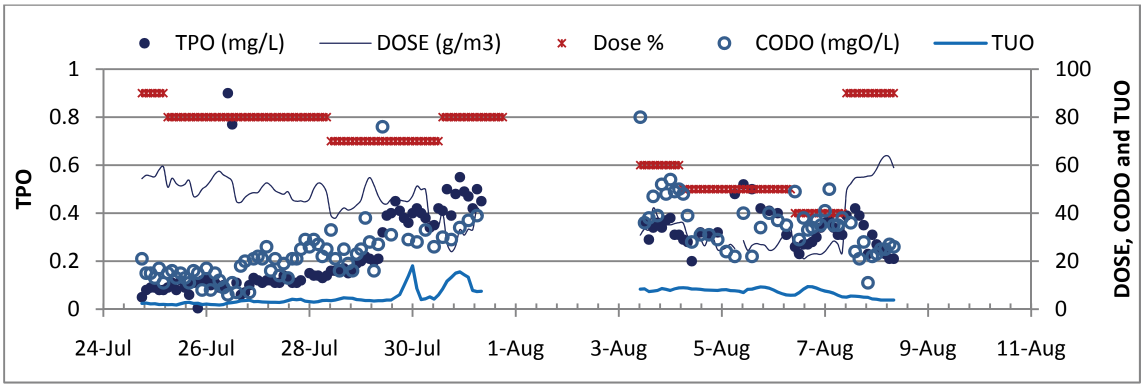


Fig.5. Effluent TU changes with the dosage reductions. The first sudden peak of the TU was due to malfunctioning of sludge pump and the second peak was due to poor dosage prediction against the sudden changes of parameters.

System performance

XCDC system in NRA was started on12th May 2009 and successfully runs as the permanent dosage control system of the WWTP. Figure 5 shows that the coagulant consumption of NRA WWTP was reduced with XCDC system, compared to the traditional, time adjusted flow proportional dosage system.

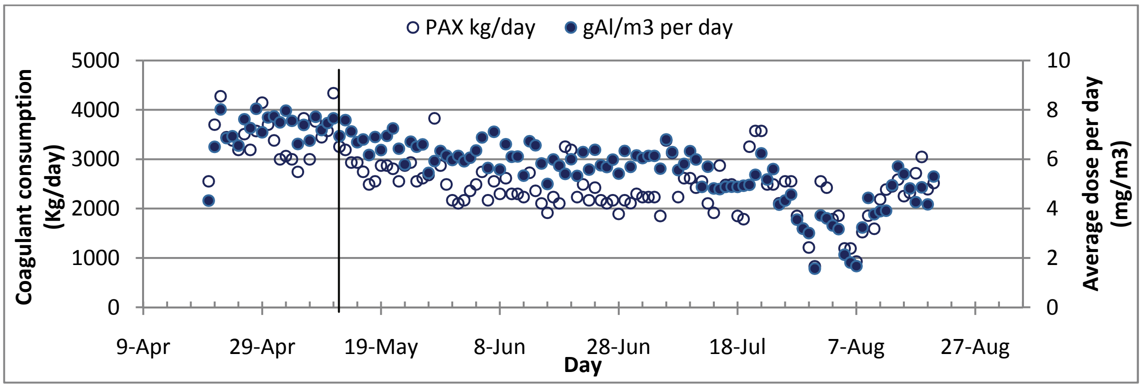


Fig. 4: Reduction of coagulant consumption using a multi-parameter model (after 11th May, black vertical line) compared with the flow proportional model (before 11th May). The solid circles show the consumption in kg-PAX/day and the non-filed circles shows the dosing in g-PAX/m3.

Figure 5 compares the effluent quality and removal of TP, SS and COD, before and after XCDC system function in the NRA WWTP. It is clear that the treatment quality was well maintained by XCDC with lower coagulant consumption. The system was successfully run for two months until frequent shock loads disturbed the system in HIAS. Figure 6 shows the results of three sampling campaigns conducted during the period. Though the dosage prediction was influenced by few shock loads during the period, effluent quality was maintained by securing the dosage by maximum limit. Figures indicate that the effluent of HIAS was well maintained during the study period.

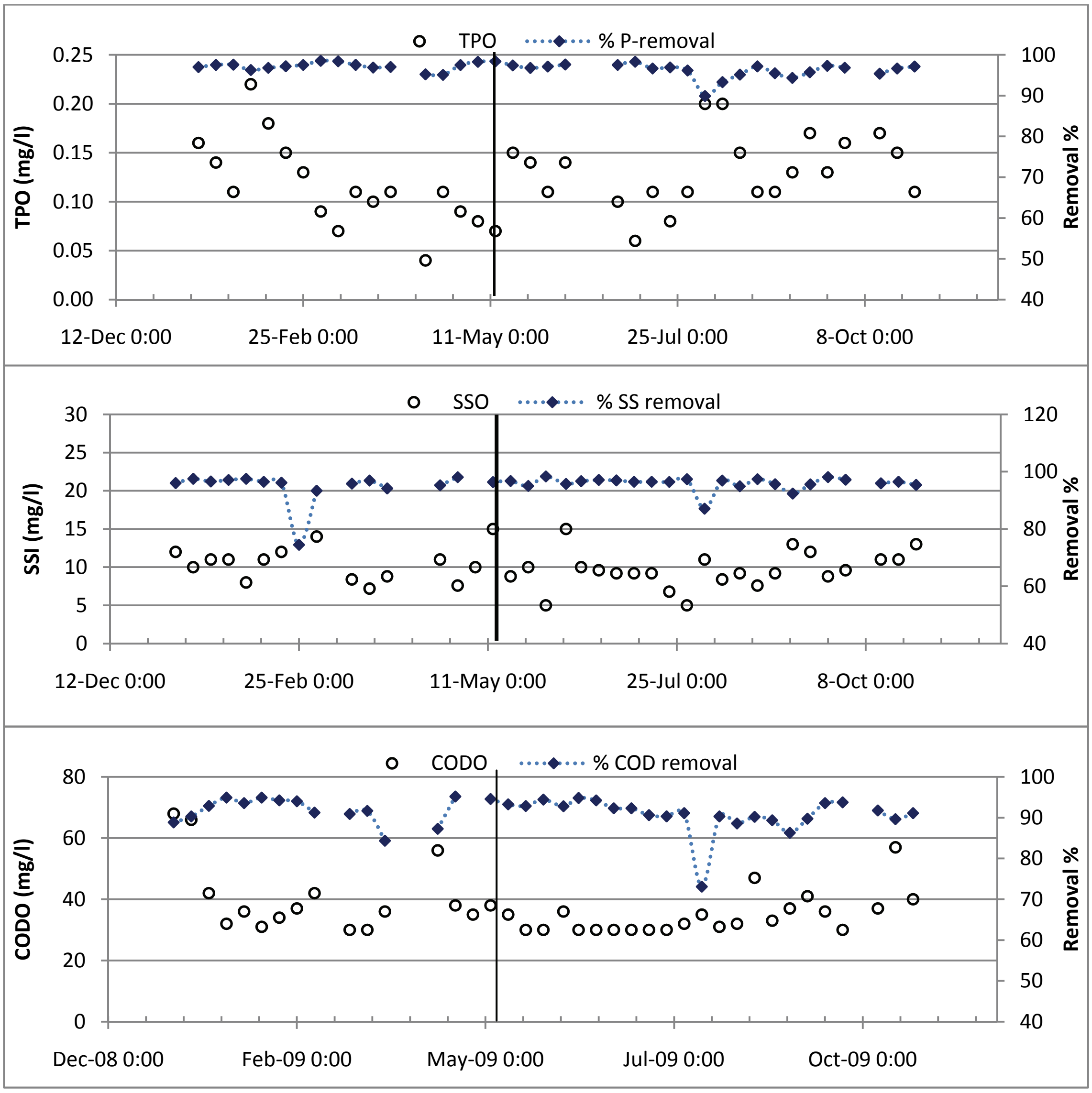


Fig. 5: Effluent quality and removal efficiency of the NRA WWTP with and without the XCDC system. the black vertical line indicates the starting point of XCDC system.

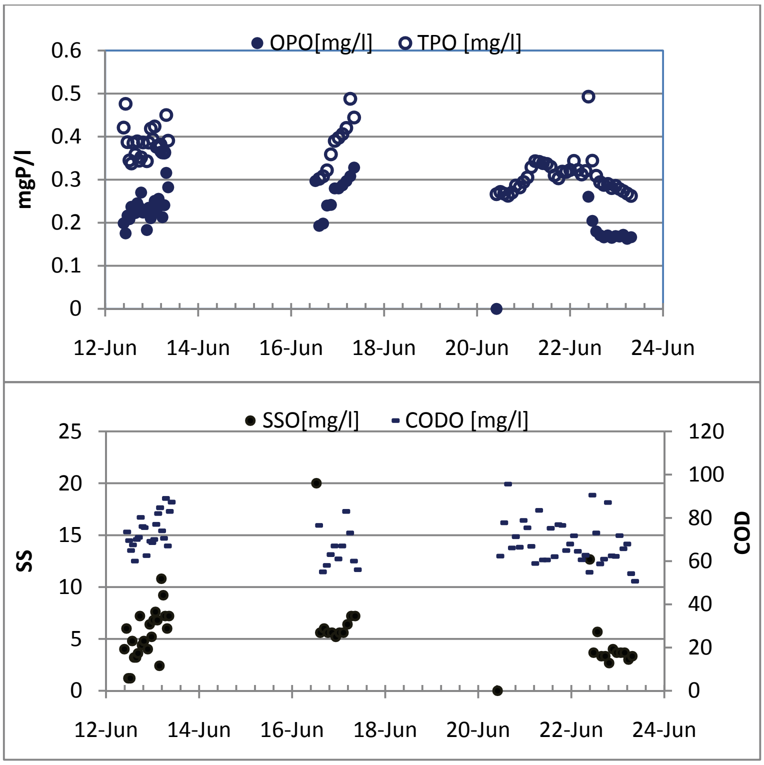


Fig. 6. Effluent quality during three sampling campaigns in HIAS. Upper: OP and TP of effluents and the lower effluent SS and COD.

The studies in the XHM and GBD were short. The most influential parameter for the XHM study was QIN. The dosage prediction was majorly influenced by the large fluctuations of the flow. The fig. 6 from XHM WWTP shows that the XCDC prediction influenced on the effluent TP limit. The system was only able to control the TPO level below 2 mg/l. In GBD pilot test, we changed the OPI concentration artificially, using $NaH_2PO_4.2H_2O$, to see the response of the XCDC models on sudden increasing of OPI. The figure indicates the sensitivity of the model. The system was able to maintain the TPO concentration well below 1mg/m^3 (Fig:8)

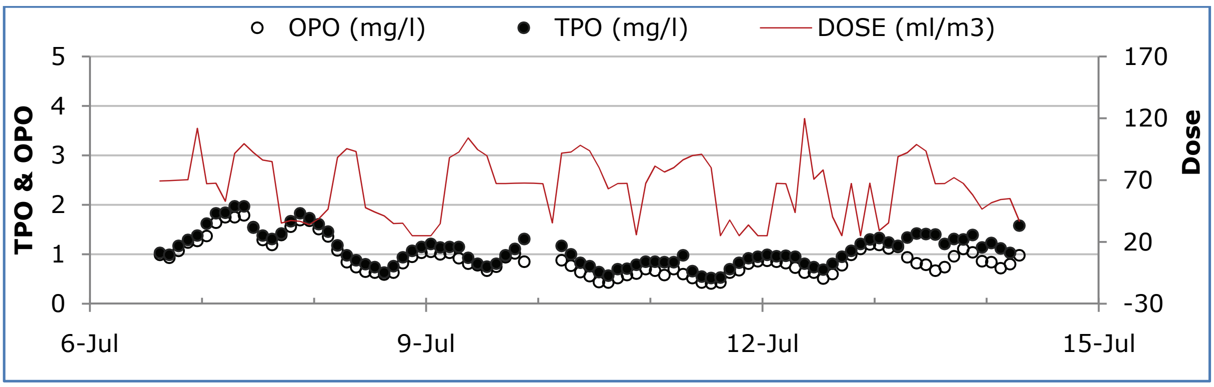


Fig. 7. TP and OP on the effluent of XHM plant during the study.

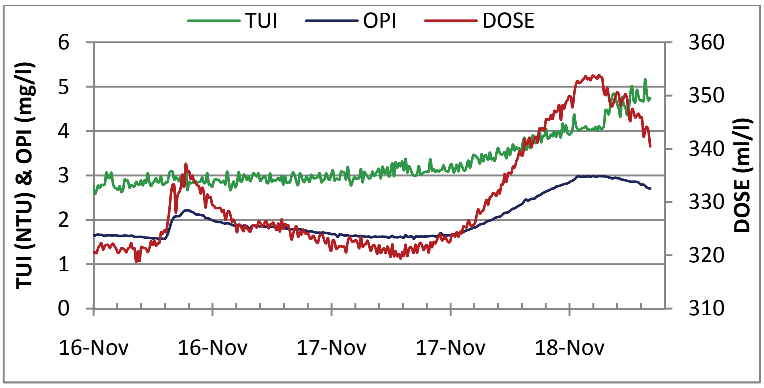


Fig. 8: The sensitivity of model to the increasing OPI in GBD.

Savings

The XCDC saves coagulants and reduces sludge production produced by excess Aluminum hydrolysis. In addition to that there are more savings such as labour cost for dosage controlling, transportation of reagents and sludge, sludge treatment and environment pollution. These are not quantified in this study to simplify the calculations.

Estimated annual savings in NRA was approximately 160t of the commercial coagulant, which equals to 12% of the average traditional coagulant consumption. Thus, the annual savings by XCDC is approximately 250 000 NOK. Furthermore, the XCDC improvement studies have proven the ability to reduce the dosage up to 15.6% of the traditional dosage.

In terms of sludge production, the content of dry Al_2O_3 in 1t of coagulants is 0.171t. The dry Al_2O_3 reductions per year is 27.36t. At a dryness of 30%, this will reduce 46.5t of surplus sludge production.

The HIAS was satisfactorily running for about 2 months time until the earlier explained problems were experienced. The system is designed to solve the problem successfully, but yet to evaluate due to some failures in some instruments in the plant. The estimated savings during the XCDC running was about 5%-15% of the annual consumption. Thus a saving between 140 000 to 400 000 NOK per year is anticipated at HIAS. The XCDC was satisfactorily controlled only during a few days at XHM. When use the data during that period, a 25 to 31% of savings were calculated.

Conclusions

The optimal coagulant dosage shows a strong correlation to the wastewater quality parameters. Thus, constant dosing in Chinese WWTP or simple flow proportional dosing control systems in Norwegian WWTPs did not reach the optimal dosage for the changing influents. The evaluated model based coagulant dosage control system was a robust dosage control solution which was successfully practiced in the studied WWTPs.

Re-calibration of models improved the model qualities while increasing the sensitivity of models estimating a precise coagulant dosage. This indicates that further tuning of the models is possible with more data during longer run of the system.

Influent to the chemical treatment in HIAS WWTP experienced frequent shock loads which created two different water types. These shock loads were not successfully handled with only one model. This problem was successfully treated with a dual platform system with two sets of equations.

The XCDC system had the potential to reduce the coagulant consumption up to 15% in studied Norwegian plants and up to 30% in the plants studied in China, with maintaining stable effluent quality. The process sensitivity investigations indicated the competence to reduce the dosages up to 30% more.

References

Adgar, A, Cox, CS & Jones, CA 2005, 'Enhancement of coagulation control using the SCD', *Bioprocess Biosyst Eng*, vol. 27, pp. 349-357.

Baxter, CW, Zhang, Q, Stanly, SJ, Shariff, R. Tupas, RRT & Stark, HL 2001, 'Drinking water quality and treatment: The use of artificial neural networks', *Canadian journal of civil engineering*, vol. 28, pp. 26-35.

Dentel, SK 1991a, 'Characterizing coagulation processes with the streaming-current detector'. *Water Supply*, vol. 9, no. supp, pp. 71-75.

Devisscher, M, Bogaert, H, Bixio, D, Van de Velde, J & Thoeye C 2002, 'Feasibility of automatic chemicals dosage control a full-scale evaluation', *Water Science and Technology*, vol. 45, no. 4–5, pp. 445–452 © IWA Publishing 2002.

Esbensen, KH 2000, Multivariate data analysis in practice. CAMO as, Oslo, Norway.

Flower, PRA 2004, 'Optimisation of potable water quality and treatment chemical usage through specially developed dosing control systems in the city of Cape Town', *In:Chemical water and wastewater treatment VIII,* HH. Hahn, E Hoffmann, H Ødegaard, IWA Publishing, London, vol. 8, pp. 15-28.

Gillberg, L, Nilsson, D & Åkesson, M 1996, The influence of pH when precipitating orthophosphate with aluminum and iron salts.

Hangouet, J P, Pujpl, R, Bourgogne, P, Ropert, D & Lansalot G 2007, '*Optimising chemical dosage in primary settling tanks', In; Chemical water and wastewater treatment IX*, HH. Hahn, H Ødegaard, IWA publishing, London, pp. 59-67.

Hansen, B 1996, 'Dosing control of coagulants based on on-line monitoring of suspended solids in sewage treatment plants', *Chemical water and wastewater treatment IV*, H. H. Hahn, E. Hoffmann, H. Ødegaard, IWA Publishing, London, vol. 4, pp. 137-145.

Homquist, J, 2004, ‘PIX dosing controlled by double point measurement of phosphorous’, *Chemical water and wastewater treatment VIIl*, IWA Publishing, London, Chemical water and wastewater treatment pp. 29-35.

Jeppsson, U, Alex, J, Pons, M N, Spanjers, H & Vanrolleghem, PA 2002, ‘Status and future trends of ICA in wastewater treatment – A European perspective’, *Wat. Sci. Tech.*, vol. 45, no. 4–5, pp. 485–494.

Leentvaar, J, Yawema, TS J & Roersma, RE 1979, ‘Optimisation of coagulant dosage in coagulation’, Flocculation of sewage, *Water Research*, vol. 13, pp. 229-236.

Lu, L 2003, ‘Model based control and simulation of wastewater coagulation’, PhD thesis, Agriculture University of Norway, pp. 126.

Martens, H & Næs, T 1991*, Multivariate calibration,* John Wiley & Sons Ltd, New York.

Mels, AR, van Nieuwenhuijzen, AF & Klapwijk, A 2002, ‘Turbidity related dosing of organic polymers to control the denitrification potential of flocculated municipal wastewater’, *Chemical water and wastewater treatment VII*, H. H. Hahn, E. Hoffmann, H. Ødegaard (eds) IWA Publishing, London, vol. 7, pp. 71-79.

RA-2, 2008, ‘Årsrapport- 2007’, *AS Sentralrensanlegget* RA-2, Standveien, 2010 Strømmen.

Ratnaweera, H 2004, ‘ Coagulant dosing control- a revie’, *Chemical water and wastewater treatment VIII*, HH Hahn, E Hoffmann, H Ødegaard (eds), IWA Publishing, London, pp. 3-13.

Ratnaweera, H 1991, ‘Influence of the degree of coagulant preploymerization on wastewater coagulation mechanisms’*,* Doctoral thesis, Norwegian Institute of Technology, Dissertation publishing, University Microfilms International, Michigan, USA. 166.

Ratnaweera, H, Blom, H & Aasgaard, G 1994, ‘Flexible Coagulant Dosing Control System based on Real-Time Wastewater Quality Monitoring’, *Chemical water and wastewater treatment III*, Hahn, HH & Klute, R, Springer-Verlag, Berlin, 105-116.

Ratnaweera, H, Lu, L & Lindholm, O 2002, Simulation program for wastewater coagulation. In: Chemical water and wastewater treatments VII, HH Hahn, E Hoffmann, H Ødegaard (eds), IWA publishing, London, pp 253-260

Sergio, G, Salinas Rodriguez, Maria D. Kennedy, Aleid Diepenveen, Hilde Prummel & Jan C. ‘Schippers 2008, ‘Optimization of PACl dosage to reduce RO cleaning in an MS’, *Desalination* 220, pp 239-251.

Sueg-Young Oh, Doo-Gyoon Byun, Jae-Moon Hwang & Hyun-Sung Song 2005, ‘Automatic Control on Dosing Coagulant as to Stream Current’. ICCAS2005, KINTEX, Gyeonggi-Do, Korea.

Vanrolleghem, PA & Lee, DS 2003, 'On-line monitoring equipment for wastewater treatment processes: state of the art', *Water Science and Technology*, IWA Publishing, vol. 47, no.2, pp.1-

Wu, Guan-De and Lo & Shang-Lien 2008, 'Predicting real-time coagulant dosage in water treatment by artificial neural networks and adaptive network-based fuzzy inference system', *Engineering Applications of Artificial Intelligence,* vol. 21, pp. 1189– 1195.

Zeghal,S, Philippe,J, Sauvegrin,P & Vignoles,C 1996, 'Chemical addition control for phosphorus removal in primary sedimentation Tanks', *Chem warter & ww treatment IV*, vol. 4, pp. 147-15.

Paper V

Multi-parameter based dosing control as an efficient tool for improved phosphate removal by coagulation- experiences from Beijing

Subhash Rathnaweera, Harsha Ratnaweera, Oddvar Lindholm

Paper presented at International Forum on Environment Simulation and Pollution Control at Beijing, 13-14 Nov. 2009.

Multi-parameter based dosing control as an efficient tool for improved phosphate removal by coagulation- experiences from Beijing

Subhash Rathnaweera*†, Harsha Ratnaweera**, Oddvar Lindholm*

*University of Life Sciences, Aas, Norway

**Norwegian Institute for Water Research (NIVA), Oslo, Norway

† Corresponding author (subhash.rathnaweera@umb.no)

Abstract.

As the influent quality and the pollution loads to wastewater treatment plants become challenging with time, the requirement for effluent quality is also becoming more stringent. As a result, the biological P removal processes in many wastewater treatment plants are experiencing difficulties to reach the goals of effluent P concentration in several parts of the world. Most of the Chinese wastewater treatment plants, which were originally designed for biological P removal by AAO biological treatment process, have started chemical coagulation. A novel system of coagulant dose control was studied in two wastewater treatment plants. The novel system showed an ability of saving up to 31% of the present coagulant consumption maintaining the present effluent quality with present coagulation system of the plant.

Most of the retrofitted plants simply add the coagulants to the AAO effluents containing huge amount of biomass, which unnecessarily consume the coagulants. Changing the dosing point from AAO effluents to a polishing stage after sedimentation would be helpful in reducing the coagulant demand by 2.4 to 7.8 times. This paper discusses the experiences of coagulant dose control and improvements of P removal in Chinese wastewater treatment plants.

Introduction

Sudden algal blooms and eutrification in water bodies are the results of water pollution due to excess nutrients concentration. The average molar ratio of nitrogen, phosphorus (P) and carbon in algal protoplasm is approximately 15:1:105. The algal growth will be limited if any of those nutrients are limited in this ratio. Thus, very small amounts of P can cause substantial algal growth. Thus controlling P sources to water bodies would be much more efficient compare to nitrogen (Jiang and Graham 1998).

Municipal wastewater (WW) is generally the main source of nutrients to the recipient water bodies.

WW treatment is designed to remove nutrient from WW before it released to the recipients. Most of the wastewater treatment plants (WWTP) are originally designed with Induced biological nitrogen and P removal. By implementing stringent regulations for environment conservation in the country the demand for effluent quality is becoming strict. As a result, the biological P removal processes in many WWTP are experiencing difficulties to reach the goals of effluent P concentration. The chemical coagulation is probably the most cost-efficient methods to retrofit existing biological WWTPs to achieve these goals, compared with upgrading the biological phosphate removal. Thus, many WWTPs in China choose to introduce coagulants to the biological stage (simultaneous precipitation), as it requires minimal constructional changes while it secures a significant improvement of P-removal.

Most WWTPs use flow-proportional dosing as the dosing control system. However, it is well documented that influent concentration of particles, phosphates and pH vary with time and are key parameters influencing the coagulation process (Ratnaweera 1991). Also the wastewater quality is changing significantly during a season, a week or a day (Sagberg *et al.* 1990). Thus, flow proportional dosing system will not be an optimal system.

With the development of cheap and convenient online sensors for water quality measurements, real-time dosing control using quality parameters is being largely practiced today. Use of one or two of these quality parameters along with the water flow is becoming popular among water and wastewater treatment facilities throughout the world (Ratnaweera 2004). Though, using one or two of the water quality parameters obviously improves the dosing strategy with limited savings of coagulants compared to the constant dosing; it will not perfectly reach the optimal coagulant dose, due to the other changing variables. Novel coagulant dose control system (XCDC) has been developed based on online monitoring of most relevant wastewater parameters in the influent, and using them in a statistical manipulation to establish a prediction algorithm for the optimal dosing requirement (Ratnaweera *et al*. 1994, Lu *et a*l. 2003).

Many Chinese WWTPs choose to introduce coagulants to the biological stage (simultaneous precipitation), as it requires minimal constructional changes while it secures a significant improvement of P-removal. However, in order to be cost efficient, the retrofitting must meet two requirements: (a) it is necessary to coagulate the wastewater which are well separated from the biomass, which will otherwise require excessive amounts of coagulants due to extremely high MLSS (b) an efficient coagulant dosing control system which estimates the optimal dosing based on real-time monitoring of key parameters influencing the coagulant consumption.

This paper discusses the experiences by full-scale and pilot scale implementation of an experimental coagulant dosing control system in WWTPs in Beijing, China and pros and cons about the present coagulation systems, experienced during the studies.

Method and Materials

The studies were conducted in two Chinese wastewater treatment plants (WWTP). Xiao-Hong-Men WWTP (XHM) is one of the largest WWTPs in Beijing. The plant is located in Chaoyang District, at the bank of Liangshui River, which serves as the recipient for treated wastewater. The plant is designed to serve 2.42million people and industries in 223.5km^2 area in Beijing. Present sewage inflow capacity to the plant is 600 000m^3 in a day. Treatment process consisted of four parallel lines with physical pre-treatment, biological Anoxic, Anaerobic and Aerobic (AAO) treatment process followed by post sedimentation tanks. Chemical coagulation has been started recently by adding coagulant to effluents of AAO process and separation is occurred in four sedimentation tanks. The XCDC trials were carried out in one of the treatment lines. Water from AAO tank effluents before mixing coagulants, after mixing coagulants and outlet of the post sedimentation tank were pumped in to three collection tanks designed for measuring the quality parameters. The TUI, OPI, CNI, PHI and TEI were measured in the first tank with AAO effluents before coagulant addition. Sampling for OPI was done in a separate tank, thus it did not influence the other influent parameters. PHO was measured in the second tank with immediate after coagulant addition and the TUO was measured in the sedimentation tank effluents in the tank three. The dosing pump of line A was prepared for automatic control by using a frequency controller to control the pump rpm according to the 4-20mA current signal from the PLC.

Gaobeidian WWTP (GBD) is currently the largest sewage treatment plant in China. The WWTP serves a catchment area of about 96 km^2, with design capacity of 1 000 000m^3/d. The treatment process of GBD was identical to the XHM. The XCDC studies were conducted in a pilot scale treatment plant in the GBD WWTP premises. The pilot plant was designed with same treatment process of the WWTP with 72m^3 inflow per day. Influent to the pilot plant was raw wastewater from the WWTP. We designed the pilot plant to add coagulants after settling of the AAO effluent in one sedimentation tank. The coagulated particles were settled in another tank. TUI, CNI, PHI, TEI and OPI were measured in the first post-sedimentation tank, before coagulant introduction. PHI was measured in the middle of the second post-sedimentation tank where treated water was introduced to the tank. The effluent TUO was measured in effluents of the sedimentation tank. Dosing was done by a digital peristaltic pump, of which the flow was able to be controlled by 4-20mA analogue signals.

In both systems, sensors of all the instruments were connected to Dr Lange SC1000 controller, and from the controller the signal was transferred to the PLC controller via analogue input cards. The predicted dose was transferred from the analogue output card as 4-20mA current signal, which was used to control the dosing pump. 10 min averages of all the measurements and dose predictions were logged in the PLC and were easily downloaded as text file.

The laboratory scale tests were conducted in the treatment plants. KEMIRA Flocculator jar test apparatus was used for jar testing to detect the range of suitable coagulant doses for the influent in each treatment plant. In addition to the jar tests, hourly grab sampling was conducted using ISCO automatic samplers and several manual samplings were conducted during the study. In the laboratory samples TU, pH, SS, COD, TP and OP were measured as needed.

Data collection was started with constant dosing. Preliminary data was collected during several days with close observations of the effluents sample measurements. Manual changes to the dosage were done as necessary. The preliminary data was collected and edited to avoid known error measurements like the measurements during maintenance and calibration of data.

Sample selection for modelling was made according to the effluent quality. In XHM, we used the TUO as a quality control parameter while TUO was not a sufficient parameter in GDB. In both plants hourly effluent total phosphate (TPO) were measured and we used it for sample selection. Dosage was slightly changed in some samples to save limited data for calibration. The table 1 contains the sample selection criteria used in two plants in both calibrations.

Table 1. The sample selection criteria of two WWTPs in different calibrations. TPO(lab) denotes the laboratory measurements of effluent TP. TUO and OPO were on-line measurements.

WWTP	1^{st} calibration	2^{nd} calibration	Remarks
XHM	TUO < 10NTU TPO(lab) < 1mg/l	TUO less than 8NTU TPO (lab)< 1mg/l	Dosage was manipulated when necessary
GBD	TPO < 1mg/l Manipulated some dosages according to Jar test results.	TP(lab) < 1mg/l and manipulated dosages Lab OPO was observed	TUO was not a good indicator. Only the TPO and dosage was manipulated when necessary

The selected sample set was used to develop the relationship between the coagulant dosage and the online measurements using PLS regression. The 1^{st} models were run in system for about four days. The data during the 1^{st} models were undergone through the sample editing and selection criteria and used to re-calibrate the models.

Multivariate statistic calibration is a strong tool to handle more than one X variables against one or more Y variables (Martens and Neas 1989, Johnson and Wichern 1982). Partial least square regression (PLS) is one of the reliable multivariate analytical methods for prediction of data. PLS regression analysis was used for developing the dosage prediction models using online measured parameters. The statistical software Unscrambler 9.8 was used for analysis.

Result and Dissuasion

Multiple parameter based dosage control

The typical formula of model algorithm is given below.

> DOSAGE= *f*(TIM, DAY, TUI, CNI, PHI, TEI, OPI, PHO, interaction among variables, squares of the variables)

Two models were developed in each calibration. Model A was included all the online measurements. The Model B was developed with minimal number of parameters to overtake the dosage system if any of the parameters malfunctioned. In XHM we used QI for the model B and in GBD, used OPI as the basic parameter for the Model B. The model statistics of the two systems are given in table 2. The model B was developed only with Q in XHM and OPI in GBD, they were not able to statistically develop. Thus simple mathematical relationships were developed.

Table 2. Model statistics of two calibrations in two WWTPs.

	1st calibration models				2nd calibration models			
	Model A		Model B		Model A		Model B	
	R2	RMSE	R2	RMSE	R2	RMSE	R2	RMSE
XHM	0.93	0.02	-	-	0.99	5.6	0.99	0.2
GBD	0.75	0.2	-	-	0.86	4.8	-	-

In XHM, we were restricted for a limited dosage level by the WWTP, and the XCDC was unable to reduce TPO below 0.5mg/l. We were able to control the dosage below 2mg/l and most of the times were controlled below 1mg/l.

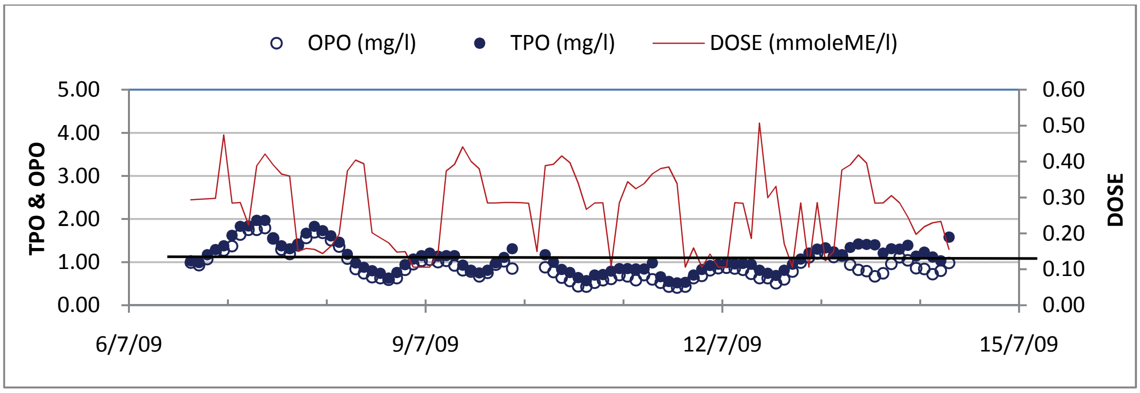


Fig 1. With data from Xiao Hong Men WWTP shows that the XCDC prediction influenced on the effluent TP limit. In the periods with lower dosage, the TP also show increasing trend. The most influential parameter there was QIN. The dosage prediction was majorly influenced by the large fluctuations of the flow.

The short study period at GBD gave some important experiences to the XCDC process. Fig. 2 shows a picture of a complete test. A was the period with constant dosing. B is the period run with 1st calibrated models, C is period with 2nd calibrated model and during D, and the 2nd calibrated models were altered with specific criteria.

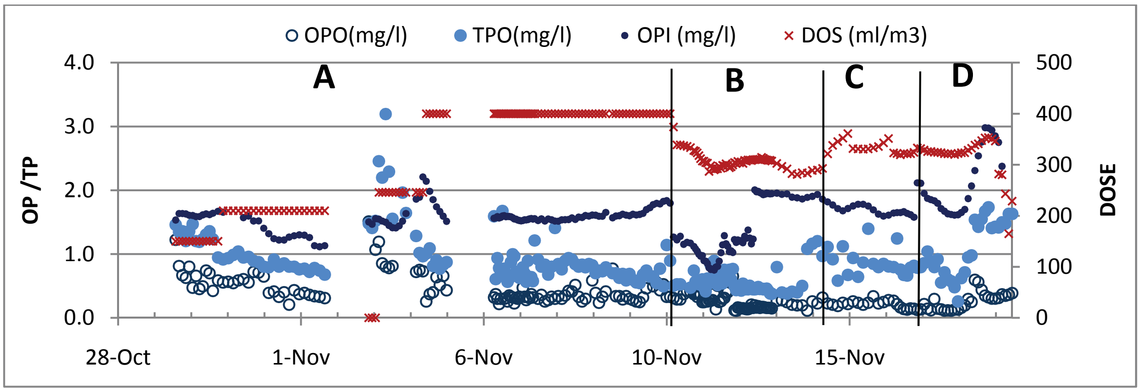


Fig. 2: OPO, TPO and Dosage. During test period

The OPO and TPO were well controlled by the 1st model but the 2nd calibration did not give good results in the sense of TP removal. The figure shows that the OPO was being well controlled under 0.5mg/l, during the 2nd calibration equations, though TPO was failed.

In the second calibration, two important observations observed. As figure 3 shows, the model was too sensitive to the parameter TIM. The influent quality the treatment facility was not largely varying. But the TIM was varying from 0 to 23. This variation was too influential to the model developed using four day's data. To handle the situation, parameter TIM was removed from the model.

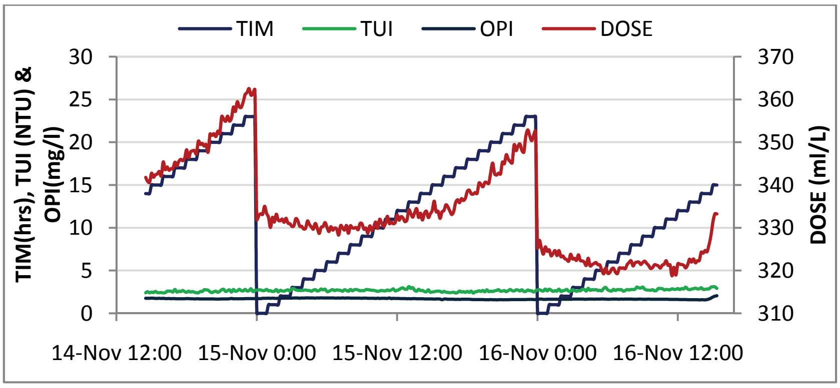


Fig. 3: The influence of TIM was too large. The variations of other parameters are not as large as TIM (0-23). I eliminated TIM from the model.

The figure 4 shows that the dosage prediction system was a good sensitivity tool for the OPI changes.

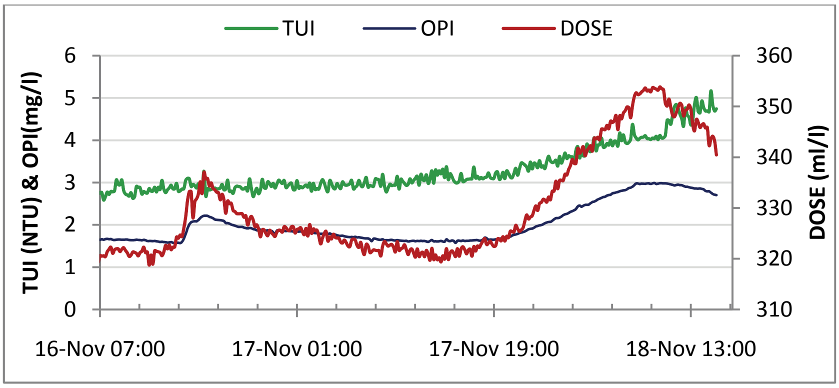


Fig. 4. The sensitivity to the OPI

Is it possible to reach effluent TP demands for Class 1A WWTP?

The traditional biological treatment process has difficulties to reach the effluent P demand in most Chinese Class 1A WWTP (0.5mgP/l or below). Thus, most of the WWTPs have started chemical coagulation as a solution. Although XHM and Gaobeidian WWTPs have started coagulant dosing, they were not able to reach class 1A effluent demand. Today both plants are struggling to maintain Class 1B effluent quality with below 1mgP/l. The figure x below shows the effluent quality of four treatment lines of XHM WWTP. The plant tested different lines for coagulation and all four lines were stared to dosage with chemicals from August. The figure shows that the plant still was not able to reduce the TPO below 0.5mg/l level.

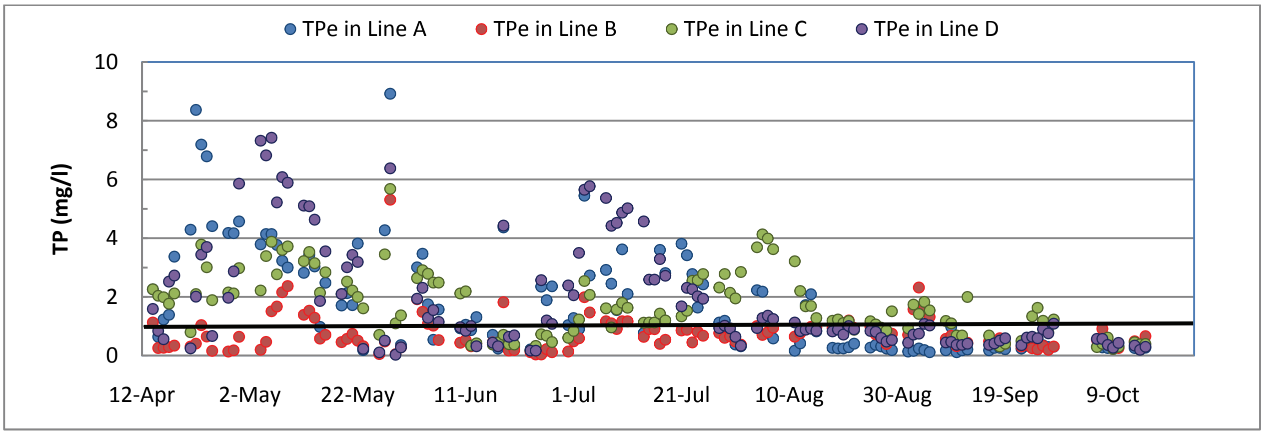


Fig. 5: Effluent TP level measurements in four treatment lines during the year 2008. The coagulation in all four treatment lines was commenced from month August.

The Fig. 6 compares the OPO levels in two different dosage levels, which were 250l/hr and 800l/hr. The plant was only able to maintain the effluent P level at 1mg/l with larger coagulant dosages.

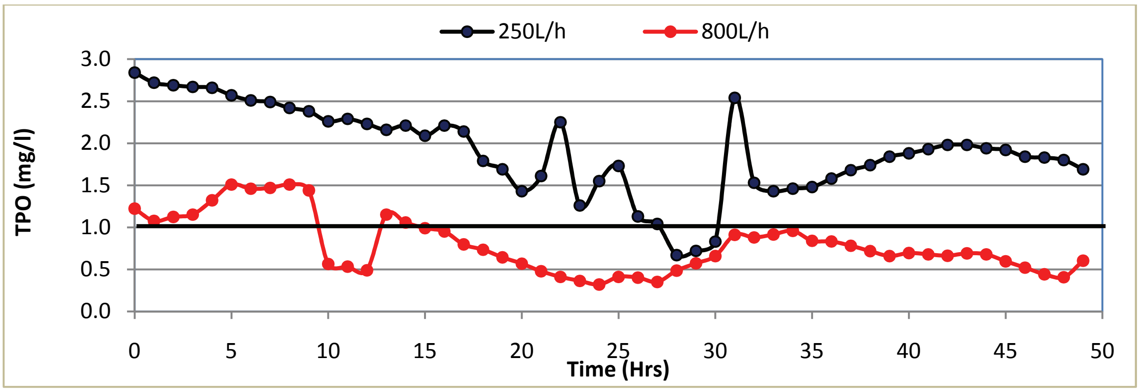


Fig. 6: Effluent OP level control by two different dosage levels. Dosage was controlled constant with 250l/hr and 800l/hr.

Coagulants were added to the AAO tank effluents prior to sedimentation in both WWTPs. The AAO effluents were with more than 3000mg/l SS due to the biomass. It was also observed that the SS of this water easily settled during few hours, even without coagulants to TUO < 10NTU in both cases. In GBD, the TU was below 5 NTU after 5hour of retention time without coagulants.

Literature describes that a part of the coagulant use in wastewater is used by SS in the water, and that the portions for SS and phosphate removals actually "compete" with each other. When the SS portion in the water is large, the consumption is larger. In the Chinese WWTPs, the high SS which is able to be settled without chemical coagulation is competing for the available coagulants using a portion of valuable coagulants.

AAO effluent vs. settled water

The figs. 7 and 8 are laboratory jar test results to compare the coagulation efficiency of the AAO effluents with huge SS and the settled WW after the secondary sedimentation. Both Gaobeidian and XHM wastewaters showed a significant improvement in treatment efficiencies with settled water.

Two different coagulant types were used in two plants. In XHM WWTP, both water types were able to reach 0.5mgP/l level in the laboratory. Furthermore, the average coagulant demand to reach 0.5mgP/l in AAO effluents was 1.09mmoleME./l while the same was 0.30mmoleME./l for settled WW. This shows a potential reduction of 2.75 times metal coagulant by dosing settled WW compare to AAO effluents, where the WWTP presently dosage. In order to reach 1mgP/l level, the AAO effluents consumed 7.8 times of coagulants compare to the settled WW.(Fig. 7)

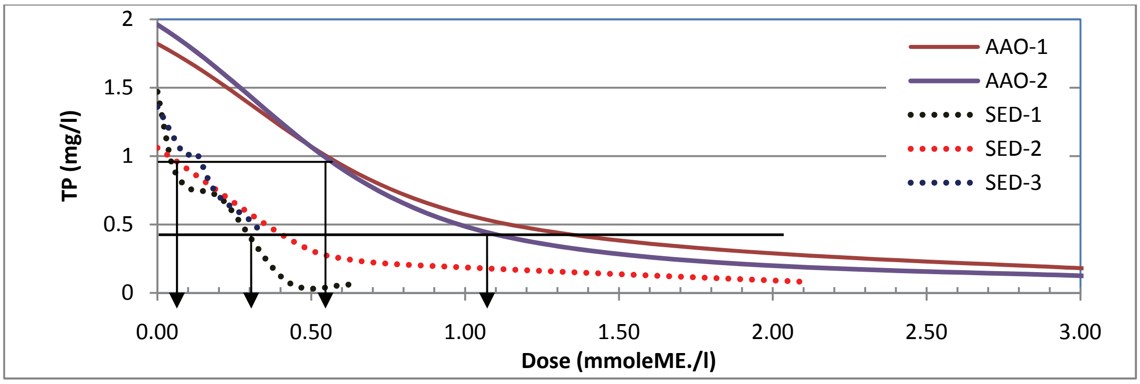


Fig. 7: Comparison of AAO effluents and settled GBD WW by means of coagulant consumption to reach 0.5mgP/l level.

Fig. 8 shows the same comparison done in GBD WWTP. It was hardly reach 0.5mgP/l level using the coagulant used in GBD WWTP. Thus we compared the coagulant demand to reach 0.1mgP/l. The coagulant demand of AAO effluents was 0.44mmols ME/l while 0.18mmoleME/l for settled WW. There was 2.4 times deference between two points.

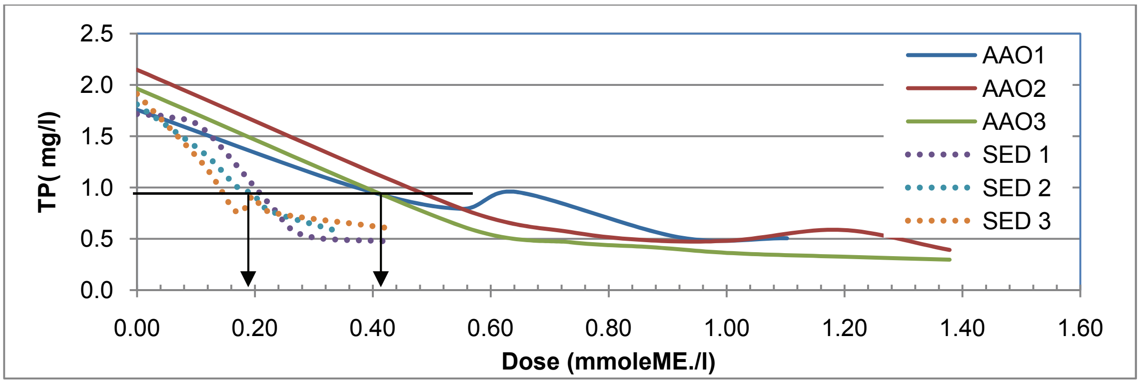


Fig. 8: Comparison of AAO effluents and settled XHM WW by means of coagulant consumption to reach 0.5mgP/l level.

Three important observations were made in the above study.

- Dosing settled water instead of AAO effluents will save 2.4 to 7.8 times of coagulants
- The actual coagulant demand to control TP in the lab was much lower compared to the practical dosage demand in the plant.
- The coagulant used in GBD was not suitable to reach Class A1 effluent quality

Changing the dosing point to dosage settled WW may require somewhat complicated constrictions and investments in transport systems between the sedimentation tanks in the final stage. In many cases, it is believed that the existing sedimentation tanks of a biological WWTP can be reorganised in such a way that about ¼ of them could be allocated to a post coagulation step, provided the surface loads are acceptable. Here, it may be necessary to use appropriate flocculants to enhance the surface loads, if that will be limiting case.

Coagulant selection

Selecting the best suitable coagulant for specific WW is one of the most critical factors for successful TP removal. We observed poor performance of coagulants in GDB WWTP. As Fig.8 shows, the coagulant hardly reached the 0.5mgP/l effluent limit, even in the Jar tests. In the XCDC testing, we observed that it successfully reduced the level OP, while TP level was not controlled successfully (Fig. 2). The same phenomenon was observed in Jar testing too (Fig 9).

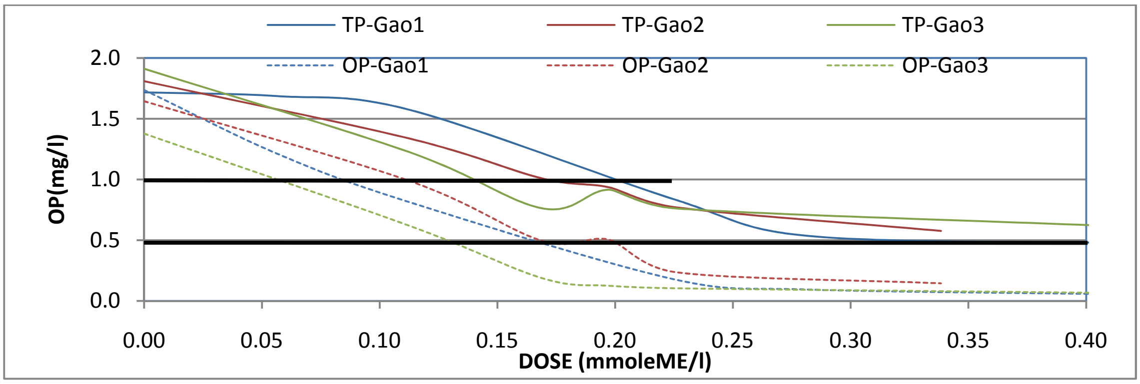


Fig. 9: OP and TP removal efficiency of the coagulant used in GBD WWTP. TP was not reached class 1A effluents.

In order to see if the coagulant is a compelling reason for the poor TP removal in GBD, we tested four commertial coagulant prodicts in laboratory. Table 3 contains the details of the coagulants used.

Table 3. Coagulants used for the performance comparison test

Product	Code name	Active ingredient	% Al	% Fe	Sp. gravity
Coag. used in DBD	Gao	Aluminium	6.1	-	1.24
Coag. used in XHM	Xia	Aluminium + Iron	7.2	3.3	1.30
PAX-18	P18	Aluminium	9	-	1.37
PAX-xl60	xl60	Aluminium	9	-	1.31
PAX-xl36	Xl36	Aluminium	6,9	-	1.29

When comparing Fig. 9 and Fig. 10, it is clear that performance of all four coagulants were better for TP removal compared to the coagulant used in GDB WWTP. All four coagulants were able to reduce the effluent TP below 0.5mg/l. Furthermore, it was observed that PAX 18 and PAX xl60 performed best among other coagulants. The coagulant demand to reach TP level of 0.5mg/l was considerably lower with PAX 18 and PAX xl60, compared to the other two coagulants.

This study shows that in order to reach the new requirements of P- removal, Chinese WWTP should consider using alternative coagulants for their WW.

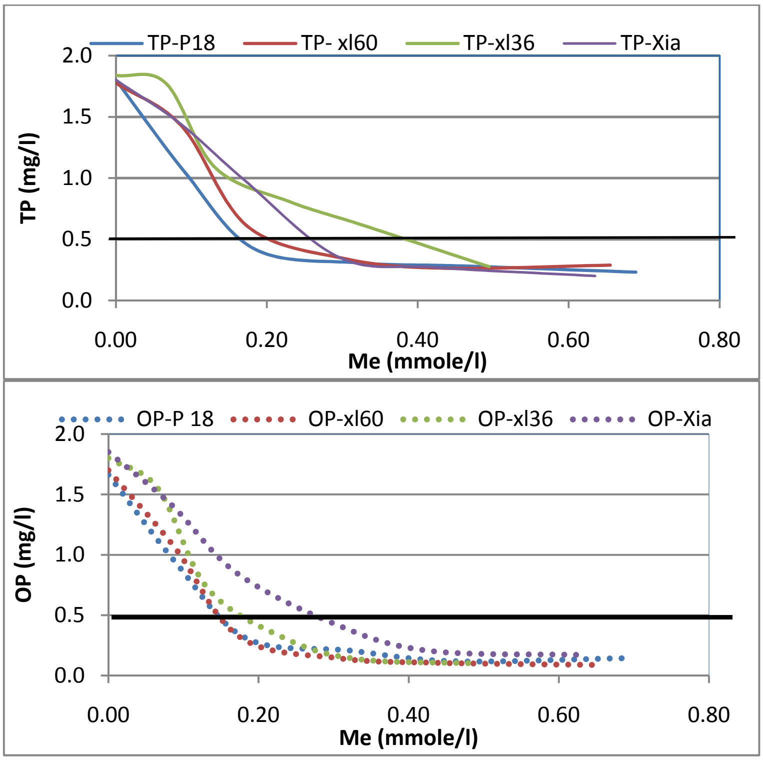


Fig. 10: Treated TP removal (upper) and OP removal (lower) with the increasing coagulant dosage.

Potential savings

The XCDC saves coagulants and reduces sludge production produced by excess metals. In addition to that, there are some more savings such as labour cost for dosage controlling, transportation of reagents and sludge, sludge treatment and environment pollution. These are not quantified in this study to simplify the calculations.

As the experience in XHM WWTP, we experienced a saving of 25 to 31% coagulants with XCDC system compared to the traditional dosing at AAO effluents. Furthermore, it was showed above, it is possible to reduce further 2.4 to 7.8 times of the coagulant consumption by changing the dosing point from AAO effluents to settled water. Thirdly, selecting the best suitable coagulant will help to reduce the dosage further.

Conclusions

A multiple parameter based coagulant dosage control system was successfully able to control the effluent TP limit below 1mgP/l with 25% to 30% coagulant savings, in studies conducted in Beijing.

Reaching Chinese standards for Class 1A WWTP was not possible during the studies. The studies showed selecting a more appropriate coagulant type for WW will help to reach the strict effluent quality demands.

Changing the dosing point from AAO effluents to a polishing stage after sedimentation would be helpful to reduce the coagulant demand by 2.4 to 7.8 times. Changing the dosing point may require somewhat complicated constrictions and investments in transport systems between the sedimentation tanks in the final stage. In many cases, it is believed that the existing sedimentation tanks of a biological WWTP can be re-organised in such a way that about ¼ of them could be allocated to a post coagulation step, provided the surface loads are acceptable. Here, it may be necessary to use appropriate flocculants to enhance the surface loads, if that will be limiting case.

The observed mechanism of biomass sedimentation will be interesting to investigate further.

References

Jiang, JQ & Graham, NJD 1998, 'Pre-polymerized inorganic coagulants and phosphorus removal by coagulation - A review', *Water Sa*, Vol. 24, no. 3, pp. 237-244.

Johnson, RA & Wichern, DW 1982, *Applied multivariate statistical analysis,* Fifth edition, Prince hall, USR, NJ 07458.

Lu, L, Ratnaweera, H & Lindholm , O 2000, 'Coagulant dosage control in chemical wastewater treatment plants- a review of modelling approaches', *Vatten*, vol. 59, pp. 227-235.

Lu, L, Ratnaweera, H, Lindholm O & Lileng K 2002, 'Model-Based Real Time Control of Coagulant Dosing', *Proceedings International IWA Conference on Automation in Water Quality Monitorin*g, N. Flesischmann, Lengergraber, G. and Haberl, R. (eds), U. of Agricultural Sci., Vienna, pp. 421-424.

Martens, H & Næs, T 1989, *Multivariate calibration,* John Wiley & Sons Ltd, New York.

Ratnaweera, H 1991, 'Influence of the degree of coagulant pre-polymerisation on wastewater coagulation mechanisms', Doctoral thesis, Norwegian Institute of Technology, Dissertation publishing, University Microfilms International, Michigan, USA, 166.

Ratnaweera, H 2004, Coagulant dosing control- a review', *Chemical water and wastewater treatment VIII*, HH Hahn, E Hoffmann, H Ødegaard (eds), IWA Publishing, London, vol. 8, pp. 3-13.

Ratnaweera, H, Blom, H & Aasgaard, G 1994, 'Flexible Coagulant Dosing Control System based on Real-Time Wastewater Quality Monitoring', *Chemical water and wastewater treatment III*, Hahn, HH & Klute, R, Springer-Verlag, Berlin, pp. 105-116.