# Full title: Model-based measurement error detection of a coagulant dosage control system

# Short title: A novel method of measurement error detection

**W. Liu[1]*, H. Ratnaweera[2] and K. Kvaal[2]**

[1]Inner Mongolia University of Finance and Economics, School of Resources and Environmental Economics, Hohhot, China
[2] Norwegian University of Life Sciences, PO Box 5003-IMT, 1432 Aas, Norway.
*The corresponding author, E-Mail: weilliu_2015@126.com, Tel.: +86 158 4711 6172.

# Full title: Model-based measurement error detection of a coagulant dosage control system

# Short title: A novel method of measurement error detection

**Abstract:** Online instruments are widely used in wastewater treatment plants and drinking water treatment plants for process monitor and control. Although maintenances of online sensors are important routine works, potential measurement errors of online sensors are challenging not only monitoring of coagulation process but also the coagulant dosage control system, what this paper is focusing on. In order to estimate and detect the potential measurement error, this paper proposes a concept of model-based measurement error detection. Relying on the model of the outlet software senor, the difference between simulation and measurement of outlet turbidity can be used as indicator of inlet measurement error. Based on the concept, this paper enables to quantify the measurement errors and build up a novel detection method. In addition, the paper compares the proposed detection method with a traditional method-the normal variation range. The results show that the proposed method has a better efficiency to detect the measurement error.

**Keywords:** Error detection; model; coagulation, online sensors, normal distribution

---

## 1. Introduction

In order to monitor, simulate and control wastewater treatment and drinking water treatment, online sensors are widely used. Relying on the online sensors, performance of each treatment section can be monitored, and corresponding operations can be carried out in time. Hence, measurement accuracy of online sensors is very important to process monitor and control. Although online sensors are regularly maintained, the measurement errors occur frequently (Thomann et al. 2002). Consequently, the measurement errors lead to inaccurate or even wrong operations, which could result in poor treatment efficiency and high operation costs. Accuracy of online measurements is still a weak point in the control chain (Winkler et al. 2004; Rieger et al. 2005; Rieger and Vanrolleghem 2008). Therefore, the detection methods of online measurement errors are necessary to study.

### *1.1 Background*

In wastewater treatment plants (WWTPs) and drinking water treatment plants (DWTPs), measurement errors of online sensors generate from several reasons. Firstly, particles, grease and crystallized coating tend to stick on sensor surfaces, which hinder the sensors to touch with wastewater. Secondly, aging issue causes unstable working status and measurement deviation from true value. Thirdly, due to the communication interruption between online sensors and the plant control system, the measurement values cannot be updated. Fourthly, human mistakes such as inaccurate calibrations and wrong settings also result in the measurement errors. In practice, since inlet water quality continually changes with time, the measurement errors in the variation curve show as different styles such as peak, constant, and drift.

Detection methods of the measurement errors are comprehensively studied, which are divided into two categories in general. Model-free approach focuses on single parameter measurement instead of considering the correlation among multiple parameters. Because of simple implementation, the model free approach such as normal distribution, discordant test and Rosner test are widely used (Edward and Charles 2014). Model-based approach, as another category, is to detect whether the correlation among many parameters are either statistically correct or in agreement with chemical or physical properties of the system (Robinson et al. 2005; Lo et al. 2016). Regarding the model-based approach, the empirical models are becoming popular for measurement error detection and several methods prove to be useful such as Artificial Neural Network (ANN), Partial Least Square Regression (PLSR) and Principal Component Analysis (PCA) (MacGregor et al. 1994; Misra et al. 2000; Venkat et al. 2003).

The coagulation process, what this paper is focusing on, is a multivariate non-linear system (Rathnaweera 2010; Maier et al. 2010). The outlet qualities of coagulation process, taking wastewater treatment as example, highly depend on coagulant dosage and inlet parameters including pH, turbidity, phosphate, temperature and so on. In order to achieve expected outlet qualities (controlled variables), coagulant dosage as a key manipulated variable should be close to the optimum value to deal with rapidly changeable inlet qualities (disturbing variables) (Liu and Ratnaweera 2016). In other side, erroneous online measurements of inlet qualities leading to inaccurate the dosage prediction cause unexpected outlet qualities. Hence, outlet qualities highly are dependent on working status of these inlet sensors when the coagulant dosage control system is running. In practice, although routine maintenances (clean and calibration) are regularly carried out by plant workers, measurement errors of online sensors usually occur and are challenging the reliability of the dosage control system. Thus, the error detection of inlet measurements is very necessary for a multi-parameter dosage control system.

Modelling of coagulant dosage control has been widely studied (Maier et al. 2004; Ratnaweera and Fettig 2015). Since there has been no a conceptual model so far due to the complexity of coagulation process, relationships between inlet parameters, dosage and outlet parameters are generally expressed by empirical models (Ratnaweera and Fettig 2015). Hence, in this paper, the model based-error detection of inlet measurements depends on the relationship between inlet parameters, dosage and outlet parameters, which is calibrated by methods of empirical model

There are many methods of empirical model, such adaptive network-based fuzzy inference system (ANFIS), ANN, PLSR, PCR, multiple linear regression (MLR) and so on. Rathnaweera (2010) concluded that PLSR has better model statistics than MLR and PCR when calibrating coagulant dosing model with same datasets. Wu and Lo (2008) proved that ANN has better performance than ANFIS for coagulant dosing control. Huang et al. (2009) highlights that performing heuristic resoning is a limitation of the ANN model while it is difficult to design and adjust automatically. Moreover, a multi-parameter dosage control system, using PLSR as the method of model calibration, was tested and applied in many WWTPs and DWTPs achieving acceptable results (Rathnaweera 2010; Liu et al. 2013). Therefore, this paper selects PLSR as modeling tool for the model-based error detection. Equation 1 presents the PLSR model structure of the multi-parameter dosage control system. By receiving online signals including wastewater flow (QIN), inlet pH (PHI), inlet turbidity (TUI), pH after coagulation (PHO), inlet conductivity (CNI), temperature (TMP) and outlet turbidity (TUO), the control system calculates real-time coagulant dosage. Hence, basing on the modelling method and the model structure, this paper aims to develop a model-based measurement error detection for the coagulant dosage control system.

$$Dosage\ prediction = f(QIN,\ TUI,\ PHI,\ PHO,\ CNI,\ TMP,\ TUO,\ interaction\ among\ variables,\ variables\ squares)$$

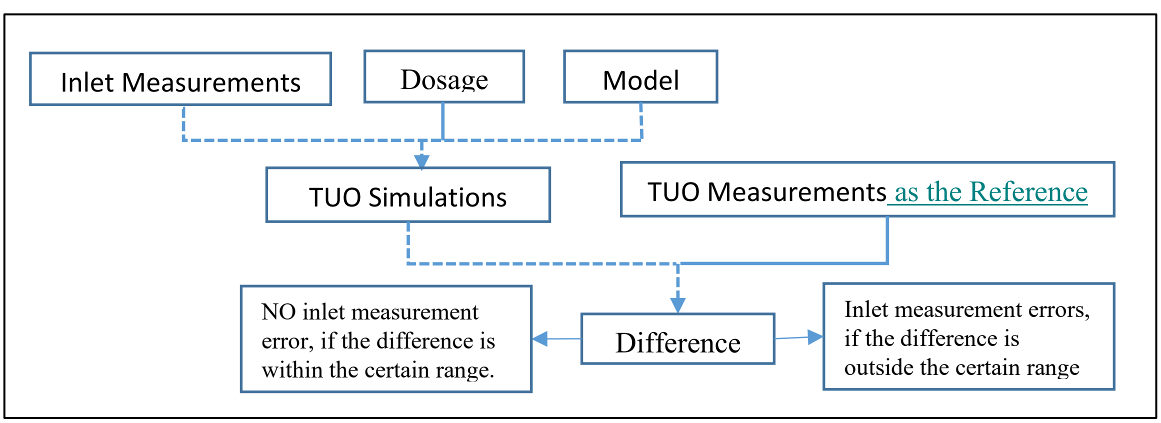
Equation (1)

Lo et al. (2016) pointed out that the process of sensor fault diagnosis can be classified into four steps: (1) detection (detecting whether there are any faulty sensors in the system), (2) isolation (determining which sensor(s) is (are) faulty), (3) identification (determining the type of faults that occurred) and (4) recovery (estimating the correct output of the faulty sensors). The model-based measurement error detection, what this paper proposes, belongs to the first step of the sensor fault diagnosis.

*1.2 Concept of the proposed error detection*

Reference is a key factor of the proposed error detection in this paper. Since outlet qualities measured by outlet sensors indicate the treatment performance, outlet sensors has importance role in coagulation process. Hence, this paper uses outlet sensor as the reference sensor of the proposed error detection. In practice, in order to ensure working status of outlet online measurements, online sensors are sometimes double or triple installed at process outlet so that difference between them can reveal the potential measurement error. Furthermore, in order to check the accuracy of outlet online measurements, laboratory measurements of grasping samples are often compared with online measurements. Moreover, much better working environment at the process outlet, where most particles settled down, reduces error possibilities of outlet sensors. Therefore, outlet sensors are some of the most reliable instruments in the dosage control system, which is able to function as the reference sensor of the error detection of inlet measurements. Since coagulation processes usually use turbidity meter to indicate outlet qualities, this paper considers the outlet turbidity sensor (TUO) as the reference sensor.

Based on the TUO software sensor as the result of authors' previous research (Liu and Ratnaweera 2017), it enables to simulate TUO at given inlet parameters and dosage. Hence, there are differences between TUO measurements and corresponding TUO simulations. This paper considers that these differences are caused by both



model errors and inlet measurement errors. Model errors are related to model accuracy, which causes unpredictable fluctuations of TUO simulation. Provided model errors repeat within a certain range and the related differences caused by the fluctuations are limited to the certain range, the proposed concept of the error detection is that when the differences exceed the certain range what model errors decide, the measurement errors of inlet sensors are considered to happen. Figure 1 describes the above concept, where dotted lines indicate sources of differences.

Figure 1. Concept of error detection of inlet measurements, dotted lines indicate sources of difference

*1.3 Objectives and research procedure*

Based on the TUO simulation model, this paper aims to i) define the difference range what is caused by model errors, ii) build up a detection criterion for inlet measurement errors, iii) compare the proposed method with a traditional method-normal variation range what the current dosage control system is using. The research is carried out by following procedure. Firstly, operation data is collected from a full-scale coagulation process, measurement errors of each parameter are defined by traditional method referring to the normal distribution and these errors of each parameter are marked in the dataset. Secondly, based on previous research on the TUO simulation model (Liu and Ratnaweera 2017), TUO simulation model is calibrated with the dataset including the marked errors. Thirdly, analyze the deviation of TUO simulations and define the range of model errors. At last, build up a criterion of the error detection and compare it with the traditional method.

## 2. Materials and methods

The operation data is collected in Nedre Romerike (NRA) WWTP, located in Lillestrom Norway. The capacity is 110 000 P.E. and treatment process consists of a screen, a pre-sedimentation, a MBBR biological treatment and a coagulation treatment. Figure 2 shows schematic of the treatment process. The coagulation process includes two treatment lines in parallel. Inlet sensors mentioned above are installed in one of lines before the coagulant dosing pump, the PHO sensor lies in the flocculation chamber as a section of coagulation process, and one of two TUO sensors is installed at the end of one line and another is installed at the outlet of the process where these two parallel lines join. Online measurement signals are sent to supervisory control and data acquisition (SCADA, WWTP control system) at first and then transfer to the dosing control system by means of Modbus communication. These measurement signals are recorded at 15 minutes' interval and there is no time delay during the signal transferring. Normally, these online sensors are maintained once per week. The operation data was collected from January 2013 to December 2014.

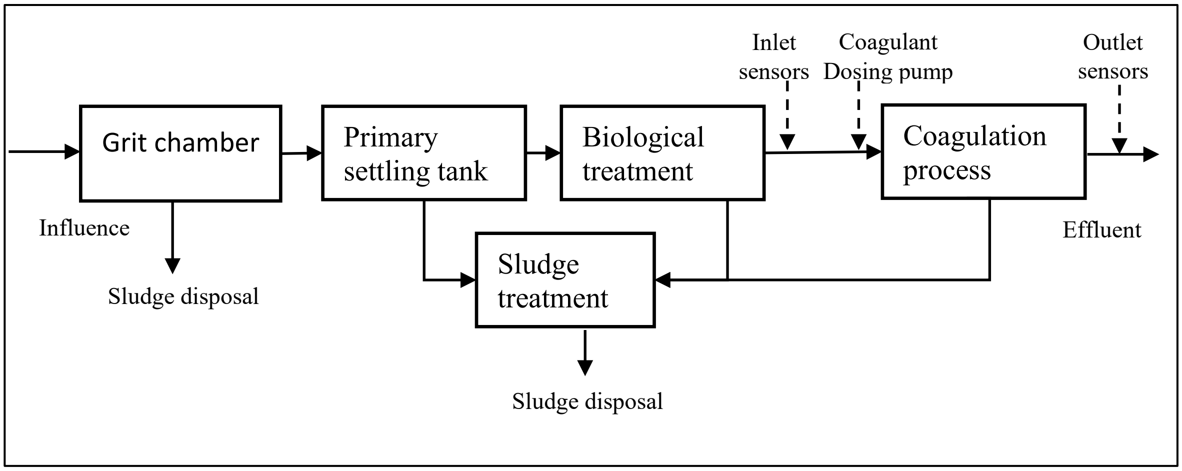



Figure 2, Schematic of treatment process in NRA WWTP, dash lines indicate the location of device and online sensors

The current dosage control system contains two criteria as the error detection of inlet measurements. The first criterion is to define whether inlet measurement values are within the normal range. Variation checking, as the second criterion, is to detect whether each sensor is active to work. Both two criteria belong to the univariate approach, which is used to compared with the proposed error detection method.

The normal variation range is determined by referring to standard deviation of measurements, mean value and historical observations. Based on the operation data, Figure 3 shows measurement distribution of each parameter. Total amount of samples is 67 872. Height of each bar indicates occurrence frequency within the total amount. Table 1 shows the normal range of each parameter. Some measurements with short bar, lying outside of the normal variation range, are identified as inlet measurement errors. Those errors defined by the normal variation range are marked in the dataset.

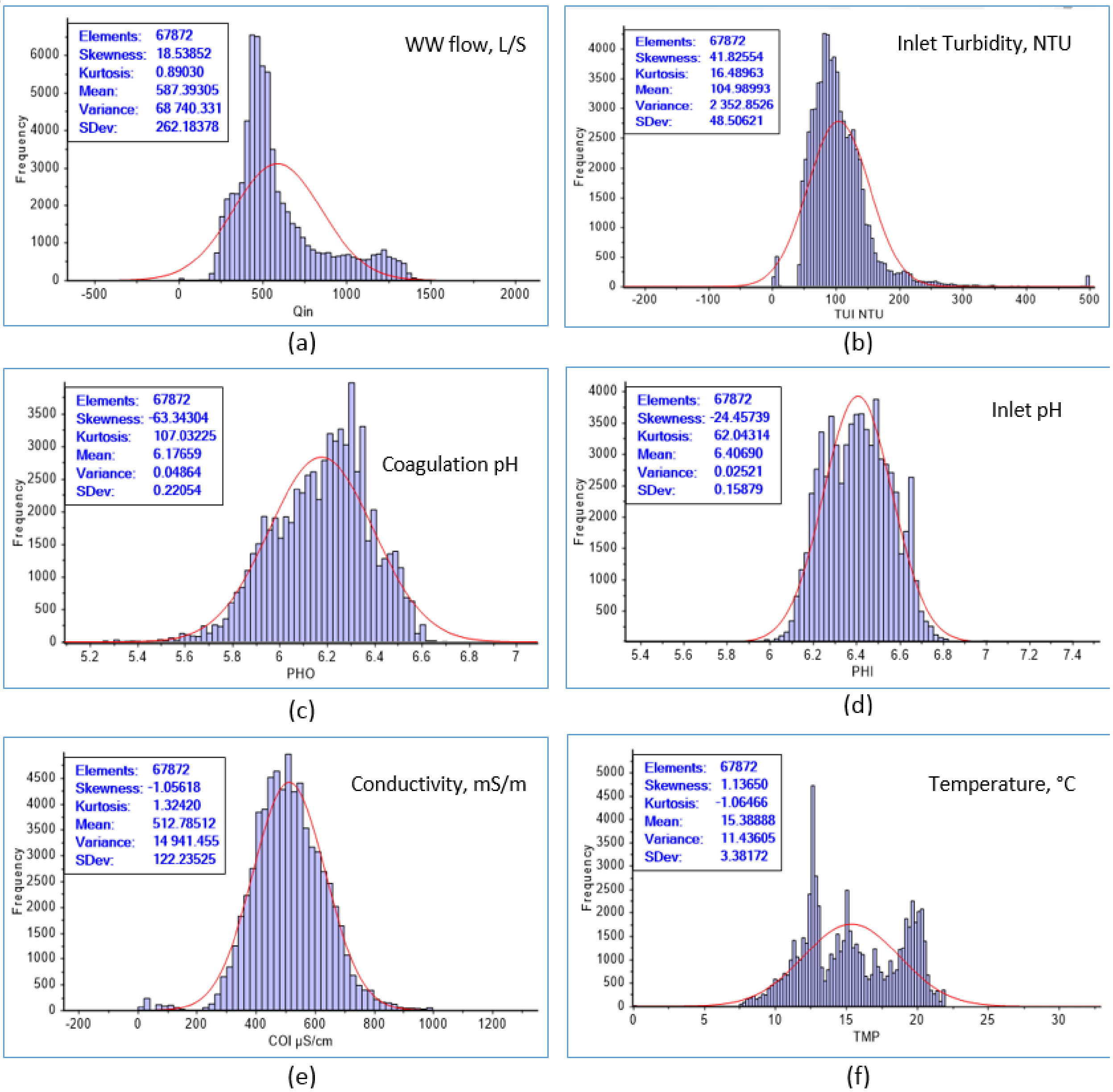



Figure 3. Online measurement distribution of each parameter during data collection. (a) distribution of WW flow, (b) distribution of inlet turbidity, (c) distribution of coagulation pH, (d) distribution of inlet pH, (e) distribution of conductivity, (f) distribution of temperature (features of the each distribution show in upper left of each figure, sample number=elements and SDev=standard deviation)

Table 1. Normal variation range of each parameter

| Parameter | Qin, L/s | PHI | TUI, NTU | CNI, mS/m | TMP, °C | PHO |
|---|---|---|---|---|---|---|
| Mean | 587 | 6.41 | 104 | 512 | 15.3 | 6.18 |
| standard deviation | 262 | 0.16 | 48 | 122 | 3.0 | 0.22 |

| | | | | | | |
|---|---|---|---|---|---|---|
| Low limitation in the variation | 50 | 6.00 | 50 | 200 | 5.0 | 5.80 |
| High limitation in the variation | 1400 | 7.00 | 300 | 900 | 25.0 | 6.80 |

Partial Least square Regression (PLSR), as method of model calibration, applies to build up the relationship among TUO, inlet parameters and dosages. The model structure is shown as Equation 2, where interactions among variables mean to product of two inlet parameters (e.g. Qin*PHO) and squares are responsible for the non-linear relationship. PLSR is carried out by Software Unscrambler® (version 10.3). Considering retention time of sedimentation tank, TUO measurements are not results of inlet parameters and dosages, which are recorded at the same time and written in the same row of dataset. Thus, before inputting the collected dataset in the software, TUO measurements are shifted and matched to inlet parameters considering the hydraulic retention time.

$$\textit{TUO prediction} = f(\textit{QIN, TUI, PHI, PHO, CNI, TMP, dosage, interaction among variables, variables squares})$$

Equation (2)

3. Results and discussions

The TUO simulation model is calibrated by the above concept and method, which enables to simulate TUO with given inlet parameters and dosages. Since a random sample indicates that each individual sample has same possibility to be selected, it has advantage of avoiding classification interruption such as sampling time and data duration (Yates et al. 2008). Hence, in order to check stability of model errors, certain number of random samples are selected from total samples (67 872) for the model calibration. Table 2 shows statistics results of five models. Four of them are calibrated with different 5000 samples that are randomly selected from total samples, whereas one of them is calibrated with total samples. The results show that there are small variations among these five models. This not only indicates that the relationship among inlet parameters, dosage and TUO is quite constant when models are calibrated with different small amount of data, but also results of root mean square error (RMSE) reveal that model errors are so stable that the proposed method of error detection can rely on. Due to the small variations among these five models, one of them does not have obvious difference over others. Hence, model 4 is selected randomly by authors to display the performance of TUO simulations, which is shown in figure 4 and plotted with TUO simulations and TUO measurements. Equation 3 represents the correlation line (black line) between TUO simulations (TUOs, shown as predicted Y on Y-axis in figure 4) and TUO measurements (TUOm, shown as reference Y on X-axis in figure 4), which is used for detection criterion in chapter 3.2.

Table 2. Statistics of outlet turbidity simulation models with different sample selections

| parameter | Model 1 with 5000 random-selected sample | Model 2 with 5000 random-selected sample | Model 3 with 5000 random-selected sample | Model 4 with 5000 random-selected sample | Model 5 with whole samples |
|---|---|---|---|---|---|

| | | | | | |
|---|---|---|---|---|---|
| $R^2$ | 0.817 | 0.823 | 0.823 | 0.817 | 0.814 |
| slope | 0.817 | 0.825 | 0.817 | 0.823 | 0.814 |
| Offset | 0.426 | 0.404 | 0.406 | 0.406 | 0.431 |
| RMSE[1] | 0.727 | 0.716 | 0.732 | 0.733 | 0.732 |

[1] RMSE is abbreviation for Root Mean Square Error. Model with 5000 random-selected samples is able to show constant results, comparing with 50 or 500 random-selected samples.

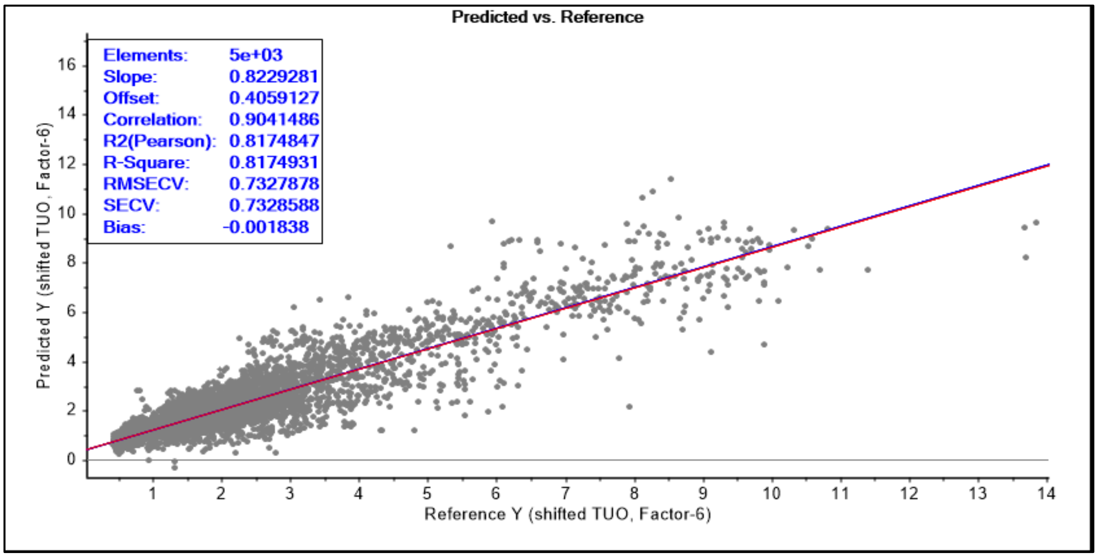



Figure 4. Correlation between simulations and measurements of outlet turbidity, black line presents the correlation line. (Sample number=elements, Slope, Offset, $R^2$= coefficient of determination and RMSE=root mean square error as main parameters of the correlation line list in the upper left of the figure)

$$TUOs=Slope*TUOm + Offset \tag{3}$$

*3.1 Defining the range of model errors*

Since model errors and the measurement errors cause differences between TUO simulations and TUO measurements, the plotted points deviate from the correlation line. Hence, a distance from a plotted point to the correlation line indicates model error and measurement error. The feature of model errors is similar to random errors, which repeats constantly and is caused by unknown and unpredictable changes (Taylor 1999). This section is to define the range of model errors by the distance. Hence, the distances from every point to the correlation line in Figure 4 are calculated. The distance distribution is shown in Figure 5, where length of each bar presents sample percentage. In probability theory, the normal distribution presents there are 68% of possibility for a sample lying in the range of m±s (m: mean value and s: standard deviation), 95% of possibility for a sample lying in the range of m±2s and 99.7% of possibility for a sample lying in the range of m±3s. According to the mean distance=-0.007 and standard deviation of distance=0.508 of distribution results, this paper chooses m±3s (±1.5) as the distribution range. The Figure 5 displays that there are 98.4% of possibility for a sample lying the range of ±1.5, which is close to the ideal value 99.7%. This distribution range (±1.5) is considered as maximum range what model errors cause. If a sample lies outside of the distribution range, this sample has big possibility to relate to the inlet measurement error. Thus, the distribution range serve as a boundary between model errors and inlet measurement errors. Based on this boundary, the differences caused by model errors and inlet measurement errors can be separated.

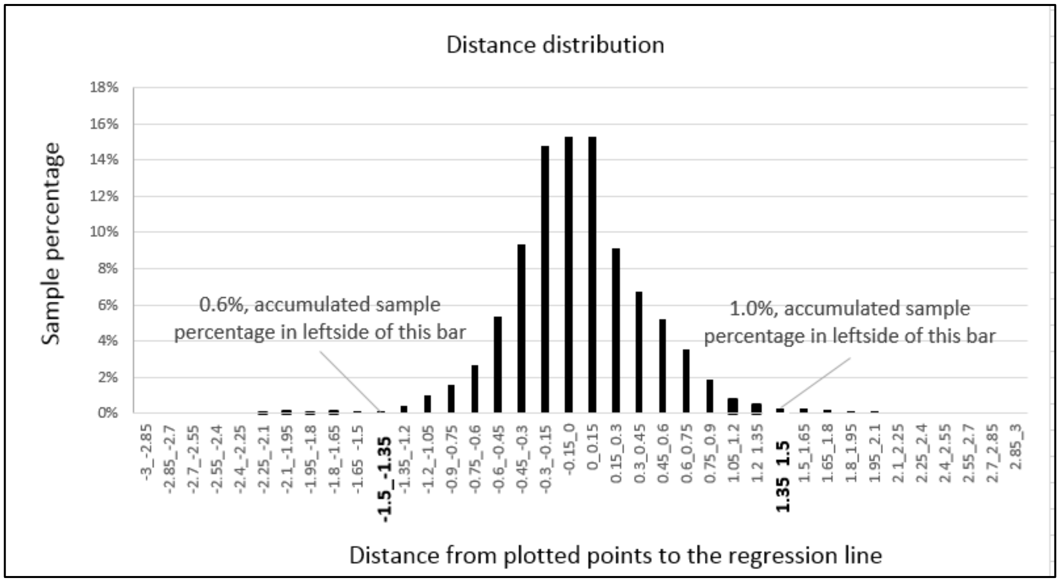



Figure 5. Defining the range of model errors based on the distance distribution. The total random sample number is 5000. Mean distance=-0.007 and standard deviation of distance=0.508. The range of model error is defined as mean± 3*standard deviation (±1.5).

### *3.2. Detection criterion for inlet measurement errors*

According to the boundary between model errors and inlet measurement errors, this section is to definite detection criterion for inlet measurement errors, which aims to display the boundary in the Figure 4. Two lines with the same slope as the correlation line are added and the distances to the correlation line equal to 1.5 (m+3s), which is shown in the Figure 6. These two lines expressed by the Equation 4. According to the slope=0.82 and distance=1.5, K in Equation 4 equals to 1.9 as a side of the right triangle (solid line in the Figure 5). Offset ± K are intercepts of the two additional lines. These two lines can function as detection criterion of inlet measurement errors. Namely, if a plotted point lies between these two lines, which are described by the Equation 5, then there is an insignificant difference between TUOs and TUOm and hence inlet online sensors are working normally. Otherwise, the plotted points beyond this range indicate measurement errors of inlet online sensors. When implementing the proposed method of error detection, TUO simulations are generated continuously and compared with TUO measurements by Equation 5. If the comparison result displays the significant difference, measurement errors are considered to happen to inlet sensors and message of claiming sensors maintenance will generate.

$$TUOs=Slope*TUOm + Offset \pm K \qquad (4)$$

Detection criterion: -K+ Offset< TUOs-Slope*TUOm<K+ Offset (5)

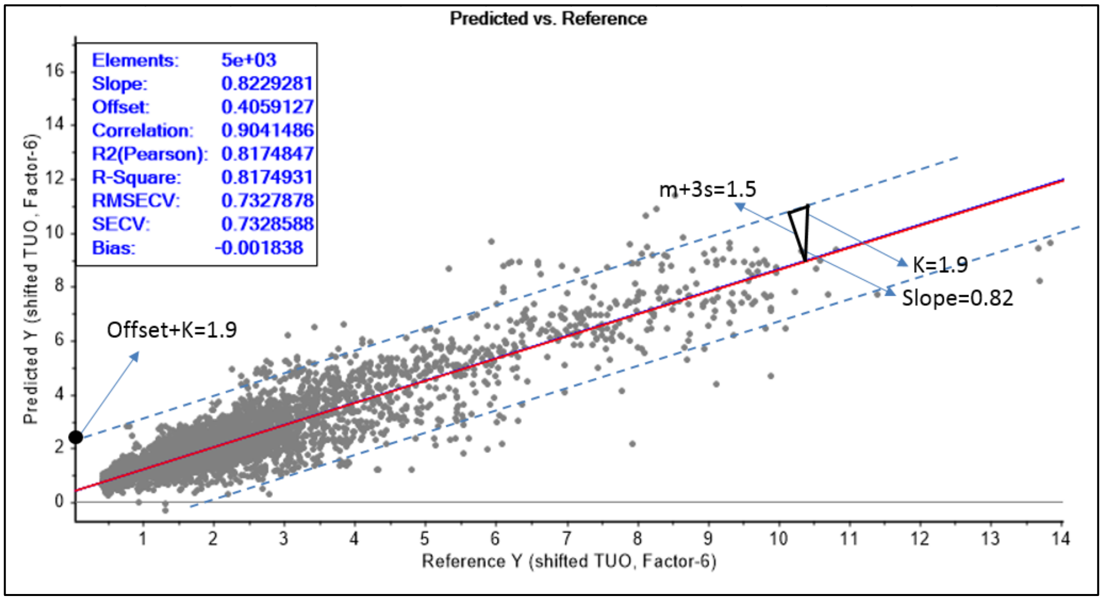



Figure 6. Detection criterion of inlet measurement errors, Solid line presents correlation line between simulations (X-Axis) and measurements (Y-Axis) of outlet turbidity. Based on the slope and the offset of the correlation line and defined range of model errors, two dotted lines parallel with the correlation line are defined as proposed detection criterion. Offset of additional lines=offset of correlation line ± K)

### *3.3. Testing the detecting criterion*

As a new method of error detection, it is necessary to comparing with the traditional method in order to validate working efficiency. The outliers of inlet measurements have been defined previously by a traditional method-the normal variation range, which is already shown in Table 1. These outliers are assumed to influence TUOs and to generate significant differences between TUOs and TUOm. This section is to test the proposed method with these outliers. Since previously marked in the whole dataset, these marked outliers are also randomly selected during the model calibration. In Figure 7, the plotted points related to random-selected outliers are marked. The plotted points with various shapes represent different inlet parameters. It can be seen that these random-selected outliers do not always result in the obvious differences. Taking square points (related to inlet turbidity outliers) as an example, some square points close to the correlation line indicate small influence on TUOs while some square points far from the line generate the large influence. Other points with different shapes also have such situation. Therefore, the results of proposed detection criterion are not fully in accordance with the traditional method-the normal variation range.

This inconsistency is analyzed below. According to the TUO simulation model shown as Equation 2, any single parameter has two ways to contribute TUOs. Firstly, the direct contribution is carried out by the single parameter. Secondly, interactions among variables are combined contributions with two parameters (e.g. PHO *dosage in equation 2). This combined contribution could become insignificant when one of two parameters is varying in low level, whereas the combined contribution could be significant when both parameters are varying in high level. Hence, a difference between a TUOs and a TUOm depends on whether a parameter with the measurement error has a significant influence. If an insignificant influence leads to the unobvious difference between TUOs and TUOm, then it is unnecessary to identify the inlet measurement errors under such situation. Therefore, the traditional method of normal variation range focuses on sensor itself and could work universally but not as efficiently as the proposed method. The proposed detection method is only to identify the measurement error that tends to deviate TUOs from TUOm.

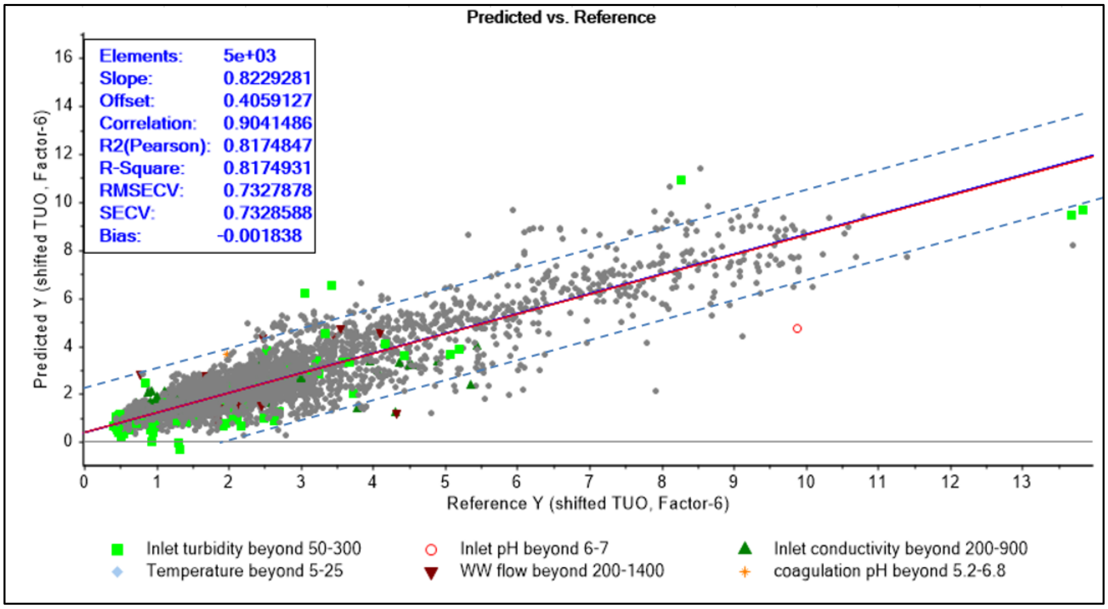



Figure 7. Comparison of proposed detection method and the traditional method (this figure bases on the correlation between simulations and measurements of outlet turbidity, solid line presents the correlation line and two dotted lines indicate proposed detection criterion of measurement errors, various points marked by difference shapes are measurement errors defined by the traditional method)

## 4. Conclusions

In order to detect potential inlet measurement errors for the coagulant dosage control system, a concept of model-based measurement error detection is proposed in this paper. This concept is implemented by the differences between measurements and simulations of outlet turbidity. This paper considers that the differences are caused by model errors and inlet measurement errors. If the differences are outside of the range what model errors decided, then inlet measurement errors are considered to happen. The concept proved to be effective based on results. And following conclusions can be drawn.

The model error of the multi-parameter dosage control system can be quantified by the differences between simulations and measurements of outlet turbidity. Moreover, model errors proved to repeat constantly within the certain range.

The range of model errors can be defined by means of probability theory and is able to functions as detection criterion of inlet measurement errors. Proposed detection method of inlet measurement errors has a better efficiency than the traditional method-the normal variation range.

This paper suggests that the sensitivity of the proposed error detection method should be improved with shorter range of model errors, which can be achieved by improving model accuracy. Since treatment results cannot be measured until coagulated water go through subsequent sedimentation tank, there is a time delay for the proposed error detection. However, in practical, the proposed detection method could be more efficient to claim maintenance of online sensors than the weekly maintenance.

Acknowledgement: The authors appreciate the assistance provided by NRA WWTP, and DOSCON Co Ltd. (www.doscon.no) for providing access to the multi-parameter based dosing control system.

Reference

Edward N, Charles F (2014) *Essentials of Testing and Assessment: A Practical Guide to Counselors, Social Workers and Psychologists*. 3rd ed.; Brooks Cole: California, USA

Liu W, Ratnaweera, H, Song HP (2013) Better treatment efficiencies and process economies with real-time coagulant dosing control, 11th IWA conference on instrumentation control and automation, Narbonne, France.

Liu W, Ratnaweera H (2016) Improvement of multi-parameter based Feed-Forward coagulant dosing control systems with Feed-back functionalities, *Water Sci. Technol.*74(2):491-499

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

Lo C, Lynch P, Liu M (2016) Distributed model-based nonlinear sensor fault diagnosis in wireless sensor networks. *Mechanical Systems and Signal Processing*. 66-67: 470-484

MacGregor JF, Jaeckle C, Kiparissides C, Koutoudi M. (1994) Process monitoring and diagnosis by multi block PLS methods. *AIChE J*. 40: 826–838.

Maier HR, Jain A, Dandy GC, Sudheer KP (2000) Methods used for the development of neural networks for the prediction of water resource variables in river system: Current status and future directions', *Environ. Modell. Softw*. 25: 891-909.

Maier HR, Morgan B, Chow C (2004) Use of Use of artificial neural networks for predicting optimal alum doses and treated water quality parameters. *Environ. Modell. Softw*. 19: 485-494.

Maier HR, Jain A, Dandy GC, Sudheer K. (2010) Method used for the development of neural networks for the prediction of water resource variables in river system: Current status and future direction. *Environ. Modell. Softw*. 25: 891-909.

Misra M, Kumar S, Qin SJ. (2000) Seemann, D. Recursive on-line data compression and error analysis using wavelet technology. *AIChE J*. 46: 119-132.

Rathnaweera S. Modelling and optimization of wastewater coagulation process', PhD thesis, Norwegian University of Life Sciences, Aas, Norway, 2010.

Ratnaweera H, Fettig J (2015) State of the Art of Online Monitoring and Control of the Coagulation Process. *Water*. 7: 6574-6597.

Rieger L, Thomann M, Gujer W, Siegrist H (2005) Quantifying the uncertainty of on-line sensors at WWTPs during field operation, *Water Res*. 39: 5162-5174.

Rieger L, Vanrolleghem PA (2008) MonEAU: a platform for water quality monitoring, *Water Sci. Technol*. 57: 1079-1086.

Robinson RB, Chris D, Odom K (2005) Identifying Outliers in Correlated Water Quality Data,. *Environ. Eng. Sci.* 131: 651-657.

Thomann M, Rieger L, Frommhold S, Siegrist H, Gujer W (2002) An efficient monitoring concept with control charts for on-line sensors, *Water Sci. Technol*. 46: 107-116.

Taylor JR (1999) An Introduction to Error Analysis: The Study of Uncertainties in Physical Measurements. University Science Books. California, USA

Venkat V, Raghunathan R, Kewen Y, Surya NK (2003) A review of process fault detection and diagnosis Part I: Quantitative model-based methods. *Comput. Chem. Eng*. 27: 293-311.

Winkler S, Rieger L, Saracevic E, Pressl A, Gruber G (2004) Application of ion-sensitive sensors in water quality monitoring, *Water Sci. Technol*. 50: 105-114.

Wu GD, Lo SL (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*, 21:1189-1195/

Yates DS, David SM, Daren SS (2008) The Practice of Statistics, 3rd ed.; Freeman, New York, USA.