VATTEN – Journal of Water Management and Research 71:183–190. Lund 2015

QUICKER RESPONSE TO QUALITY CHANGES IN INCOMING WATER WITH DECISION SUPPORT FOR COAGULANT DOSAGE AT GÖRVÄLN DRINKING WATER PLANT

Beslutsstödssystem för koagulantdosering på Görvälnverket ger snabbare respons på kvalitetsförändringar i råvattnet

by S. NILSSON[1] and F. HALLGREN[2]
1 IVL Swedish Environmental research Institute, Box 21060, 10031 Stockholm, Sweden
e-mail: sara.nilsson@ivl.se
2 IVL Swedish Environmental research Institute, Box 53021, 40014 Gothenburg, Sweden
e-mail: fredrik.hallgren@ivl.se

Abstract

A decision support system (DSS) for control of the coagulant dosage at the Görväln drinking water plant has been developed and implemented. The goal with the DSS was to enable the transition from manual to automatic control of coagulant dosage. The DSS is based on a multivariate statistical regression (PLS) model mimicking the operators' manual dosage of coagulant and is based solely on UV-absorbance, colour, COD, TOC and conductivity in the raw water. By external validation with two years of historical data, the model was proven to provide a good estimation of the manual dosage. When the model was implemented for dosage control, the variation of the quality of the treated drinking water was significantly reduced as a result of quicker and correct response to changes in the raw water and at the same time the coagulant consumption was maintained. The results pave the way for future optimization of the coagulant dose, resulting in reduced coagulant consumption while still maintaining or even increasing the drinking water quality. The approach presented is expected to be able to give positive results on other drinking water plants as well.

Key words – Precipitation; coagulant; dosage; decision support system; multivariate; model

Sammanfattning

Ett beslutsstödssystem för styrning av koagulantdoseringen på dricksvattenverket Görvänverket har utvecklats och implementerats. Målet med beslutsstödssystemet var att möjliggöra en övergång från manuell till automatisk styrning av koagulantdosen. Beslutsstödssystemet baseras på en multivariat statistisk regressionsmodell (PLS-modell) som imiterar operatörernas manuella koagulantdosering med hjälp av värden på råvattnets UV-absorbans, färg, COD, TOC och konduktivitet. En extern validering med två års historiska data visade att modellen gav en god estimering av den manuella dosen. När modellen sedan implementerades för styrning av koagulantdosen minskade variationen i kvaliteten på det behandlade dricksvatten signifikant till följd av snabbare och mer korrekt respons på förändringar i råvattenkvalitén, samtidigt som kogulantförbrukningen var oförändrad. Resultaten banar väg för framtida optimering av koagulantdosen, och med detta en minskad förbrukning av koagulant samtidigt som dricksvattenkvaliteten bibehålls eller till och med ökar. Tillvägagångssättet som presenteras väntas kunna ge positiva resultat även på andra vattenverk.

Introduction

Precipitation control

A commonly used technique for water treatment in drinking water plants is precipitation. By adding a coagulant to the water, particles in the water form flocks. The flocks are then allowed to settle in sedimentation basins and can in that way be removed from the water.

The proper dose of coagulant depends on the composition of water to be treated. This is often determined by the operators who have developed a professional skill to tune the dose correctly for different compositions of the incoming water. One problem with the manual dose control is that the quality of the incoming water often varies greatly and quickly, resulting in too slow response to the changes. This might lead to changing drinking water quality and a general tendency to overdose to be on the safe side of the quality limits for the outgoing drinking water. A well-tuned automatic control of the coagulant dose could solve this problem.

Görväln drinking water plant

The Görväln drinking water plant is situated about 20 km northwest of Stockholm. The plant treats raw water from Lake Mälaren to produce drinking water for 500.000 users, as well as for companies, industries and hospitals. One of the process steps in their drinking water treatment is precipitation, which is manually controlled by the operators. The manual control makes it challenging to react to the rapid changes in the raw water.

From manual to automatic dosage control

To tackle the challenge to dose the correct amount of coagulant in the rapidly changing raw water, the Görväln plant was looking for a way to implement a system for automatic dosage. The problem was addressed by creating a multivariate statistical regression model with the aim to predict the correct dose of coagulant based on parameters measured online in the incoming water. This paper presents the results from the modelling and the validation of the feasibility study model and the final model, and the first results from the implementation in the control system at the Görväln plant.

Materials and Methods

The aim was to develop and implement a PLS-model for automatic control of coagulant dosage at the Görväln drinking water plant. The plant is relatively well-instrumented and the process engineers provided data and process knowledge necessary for the project.

Görväln drinking water plant

Process outline

The raw water that is pumped from Lake Mälaren to the plant passes through a filter to remove larger particles. UV-absorbance (254 nm), turbidity, colour, temperature, TOC, COD, conductivity and pH are measured online in the water. Based on the values of the measured parameters, a manually controlled suitable concentration of the coagulant aluminum sulfate ($Al_2(SO_4)_3$) is added. In the following basin flocks are formed, a process which is also facilitated by the addition of sodium silicate. In the subsequent sedimentation basin, the flocks are separated from the water, and to remove the remaining flocks the water then passes through a sand filter. In the water leaving the sand filter UV-absorbance (254 nm), turbidity, colour, TOC and COD are measured online to monitor the quality of the treated water. The next step, the carbon filter, removes odour and taste from the water before the it passes through the UV aggregate where it is disinfected. The final step involves a pH-adjustment to prevent corrosion and addition of monochloramine to inhibit growth of microorganisms. An overview of the water treatment process can be seen in Figure 1.

Multiparameter probe

The values for UV-absorbance (254 nm), turbidity, colour, TOC and COD are measured by a Spectro::lyser from S::CAN, a multiparameter probe that uses UV-Vis spectrometry to determine the parameters mentioned.

Control system and data available

The control system at the Görväln plant is an ABB 800

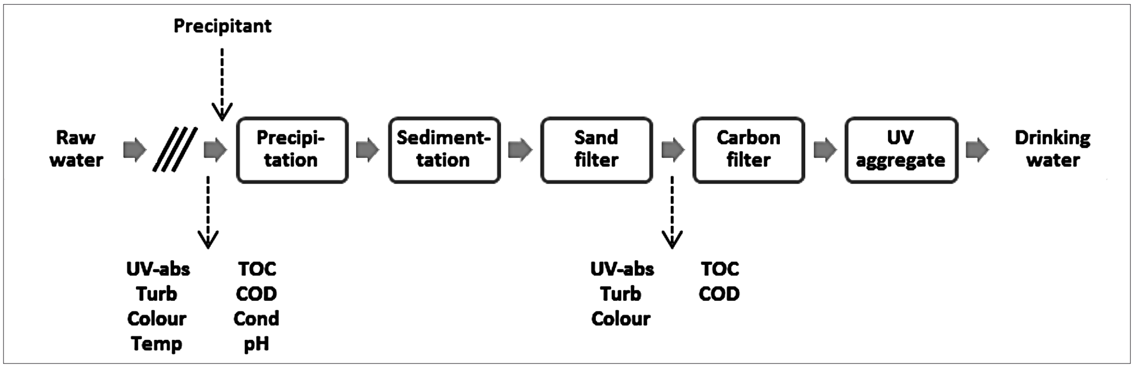

Figure 1. *Simplified schematic overview of the water treatment process at the Görväln plant. A number of parameters are measured before the precipitation to allow for control of the subsequent process steps, and after the sand filter to monitor the quality of the treated water.*

XA. It controls the process, and visualizes and saves process data. 5 minutes averaged data of sensor responses and process settings are saved in the historical database. This was the data available for the modelling.

Multivariate statistics

PLS

The multivariate statistical regression method used for calculation of the models for the precipitation was PLS, which is short for Partial Least Squares or Projection to Latent Structures (Geladi *et al.* 1986, Martens *et al.* 1989). With PLS, the aim is to establish a relationship between input variables (x), and output variables (y).

This is done by reducing the multidimensional data set to lower dimensions by calculating so-called principal components that describe the data. A PLS model is calculated in such a way that it describes as much variance as possible in the data, while at the same time maximizing the covariance between the x-variables (e.g. the parameters measured in the raw water) and the y-variables (e.g. coagulant dose). The final result is an equation expressing y as a linear combination of the x-variables:

$$y = a_0 + a_1x_1 + a_2x_2 + \ldots + a_nx_n \tag{1}$$

where n is the number of x-variables.

External validation

A relatively reliable way of validating the predictive ability of a model is by external validation. When externally validating a PLS-model, data that has not been involved in the calculation (calibration) of the model is used. The external validation data set consists of the same x- and y-variables as the calibration data set, but with data points that are new to the model. The PLS-model is fed with the values of the x-variables only and is allowed to calculate (predict) the corresponding y-values. The predicted y-values can then be compared with the corresponding "real" y-values to give an estimate of the predictive ability of the model.

Measures of model quality

The quality of a PLS-model can be represented in several ways. Quality measures used in this paper are:

- $\mathbf{R^2}$ is the part of the variance explained in the calibration data, thus, it is a measure of how well the model fits the calibration data. Note that it does not give information about model performance for new data points. If R^2 is 1, the model explains the data perfectly, if R2 is zero it is as good to guess a random number as to use the model.
- $\mathbf{Q^2}$ is an estimate of the predictive ability of the model and is calculated by cross-validation. If Q^2 is 1, the model predicts the data perfectly.
- **RMSEE** (root mean square error of estimation) is a measure of the fit of the model and has the same unit as the y-variable. The lower the value, the better the model fit the y-data.

$$RMSEE = \sqrt{\frac{\sum_i (y_i - \hat{y}_i)^2}{n - 1 - A}} \tag{2}$$

  $y - \hat{y}$ refers to the fitted residuals for the data points in the model calibration set, n to the number of samples and A number of principal components in the model.
- **RMSEP** (root mean square error of prediction) is a measure of the predictive power of a model and has the same unit as the y-variable. It is calculated similarly to a standard deviation and can be used roughly as a standard deviation of predictions. Thus, the lower the value, the better the prediction.

$$RMSEP = \sqrt{\frac{\sum_i (y_i - \hat{y}_i)^2}{n}} \tag{3}$$

  y is the reference value and $\hat{y}$ is the predicted value, hence $y - \hat{y}$ refers to the predicted residuals for the data points in the external validation data set.

Software

The software used for the modelling was Simca 13.0 from Umetrics.

Model development

The goal with the model development was to come up with a model that can predict a correct coagulant dose based on values of parameters measured online in the water entering the precipitation process. The model development work was carried out in three steps.

Data selection

The engineers at the Görväln plant were asked to select 33 periods of approximately one day each for the model development. The selection criteria for the periods were 1) the plant performed well, and 2) the combination of the selected periods should cover a representative variation in raw water composition.

Feasibility study model

Every third period were excluded from the data set before the model development. A PLS model for coagulant dose was then calculated with the remaining two thirds of the data. The starting point was to use all parameters measured in the water entering the precipitation as x-variables (UV-absorbance, turbidity, colour, temperature, TOC, COD, conductivity and pH). During the development of the model, x-variables that did not contribute positively to the model were excluded. The model was then externally validated with the 11 periods that were not involved in the calculation of the models.

Final model

A final model was developed with data from all 33 selected periods and a more extensive external validation was performed. The parameters selected as x-variables were the same as in the feasibility study model. The external validation was made with data from 12-07-01 to 14-06-12. Values that were clearly incorrect had been manually excluded from the validation data set before it was used.

Implementation

If the final coagulant dosage model performed well in the external validation, it would be implemented directly in the ABB 800 XA controllers for evaluation and dosage control.

Implementation for monitoring

The first step in the implementation process was to monitor the values given by the model in real-time. This was important to make sure that the implementation worked correctly in the control system and that the model provided reliable values, but also to give the operators a chance to get acquainted and comfortable with the idea of an automatic model based dosage control.

Implementation for control

The final step was to use the model for controlling the coagulant dosage. Before the model was implemented for control, it was decided to introduce limits for the x-variables and the dose. To ensure that the x-data (i.e. parameter values) fed to the model are valid, limits for the sensor data were set. The limits for each parameter were chosen according to the maximum and minimum parameter values that the model was built on. Limits were also set for the predicted coagulant dose, which ensure that if it is out of range it will not be used by the controller. By using the existing controller no additional hardware or software was needed to implement the dosage model. The existing control system was also used for visualization of predicted dosage.

Results and Discussion

Model development

The aim with the model development was to create a PLS-model that could predict the appropriate coagulant dosage (y) in real-time based on values from on-line sensors measuring parameters in the water entering the precipitation (x).

Data selection

The first step in the model development was the selection of representative data to build the model on. 33 approximately one day long periods were selected. They were distributed over a time period of about 21 months, from 2012-07-03 to 2014-03-23. The periods contained data for UV-absorbance (254 nm), turbidity, colour, temperature, TOC, COD, conductivity and pH in the water entering the precipitation, the coagulant dose and UV-absorbance (254 nm), turbidity, colour, TOC and COD for the water leaving the sand filter. The periods were selected in such a way that they represented a broad range of different raw water compositions, but that the treatment results still were satisfying, i.e. that the UV-absorbance in the water leaving the sand filter was below 0.45 (Table 1). The total number of data points was 9228, each having values for the above mentioned parameters.

Feasibility study model

After selection of suitable data, a first model was calculated to investigate the possibility of developing a model for prediction of correct coagulant dose. UV-absorb-

Table 1. *Properties of each parameter during the 33 periods selected for model development in terms of minimum value (Min), maximum value (Max), mean value (MV) and standard deviation (SD).*

	Water entering the precipitation step								Dosage	Treated water after sand filter			
	Temp *°C*	pH	Turb *FNU*	Cond *mS/m*	UV-abs *abs5cm*	Colour *mgPt/L*	TOC *mg/L*	COD *mg/L*	Dose *mg/L*	UV-abs *abs5cm*	Colour *mgPt/L*	TOC *mg/L*	COD *mg/L*
Min	7.3	0.6	1.3	17.5	1.20	23.5	7.8	7.0	36.0	0.37	4.0	4.1	2.6
Max	8.3	11.7	7.5	27.3	1.72	43.4	11.8	9.2	74.0	0.45	7.3	5.2	3.2
MV	7.7	5.1	3.8	22.6	1.37	30.2	9.2	7.7	52.4	0.42	5.6	4.7	3.0
SD	0.2	3.6	1.3	3.1	0.16	5.4	1.2	0.7	11.0	0.02	0.5	0.2	0.1

Table 2. *Properties of the feasibility study model.*

Measure	Value
R^2	0.946
Q^2	0.946
RMSEE	2.6 mg/L
RMSEP	2.6 mg/L

Table 3. *Properties of the final model.*

Measure	Value
R^2	0.952
Q^2	0.952
RMSEE	2.5 mg/L
RMSEP	2.9 mg/L

ance, turbidity, colour, temperature, TOC, COD, conductivity and pH in the water entering the precipitation were used as x-values and coagulant dose as y-value in the model.

The PLS-model was calculated with 22 out of the 33 periods, corresponding to 6060 data points. Temperature, pH and turbidity did not significantly contribute to the model and were therefore excluded. The feasibility study resulted in a model based on two principal components describing the relationship between five parameters measured in the water entering the precipitation (UV-absorbance, colour, TOC, COD and conductivity) and the coagulant dose set by the operators. Based on the high R^2- and Q^2-values, the relatively low RMSEE-values (Table 2) and visual inspection of the correlation between the dose predicted by the model and the dose that was actually used (not shown), it was concluded that the model fitted the data excellently.

An external validation of the model with the remaining 11 periods (3168 data points) gave a relatively low RMSEP value and showed that the correlation between predicted and actual dose was very good (Figure 2). It was concluded that it had the potential to be used for automatic dosage control.

Final model

After the promising results from the feasibility study, a final PLS-model was calculated with data from all 33 periods. It was based on the same five x-variables as the feasibility study model and had the purpose to predict an accurate enough coagulant dose to be implemented in the control system at the Görväln plant. Like the feasibility study model, the final model was based on two principal components, and the R^2-, Q^2- and RMSEE-values (Table 3) and the correlation between the dose predicted by the model and the dose that was actually used (not shown) was also comparable to the feasibility study model. When studying the time series plot it was also obvious that the dose was not changed very often when it was manually controlled compared to the dose that was recommended by the model.

The final model was then externally validated with continuous historical data from about two years, from 12-07-01 to 14-06-12 (195444 data points). The cor-

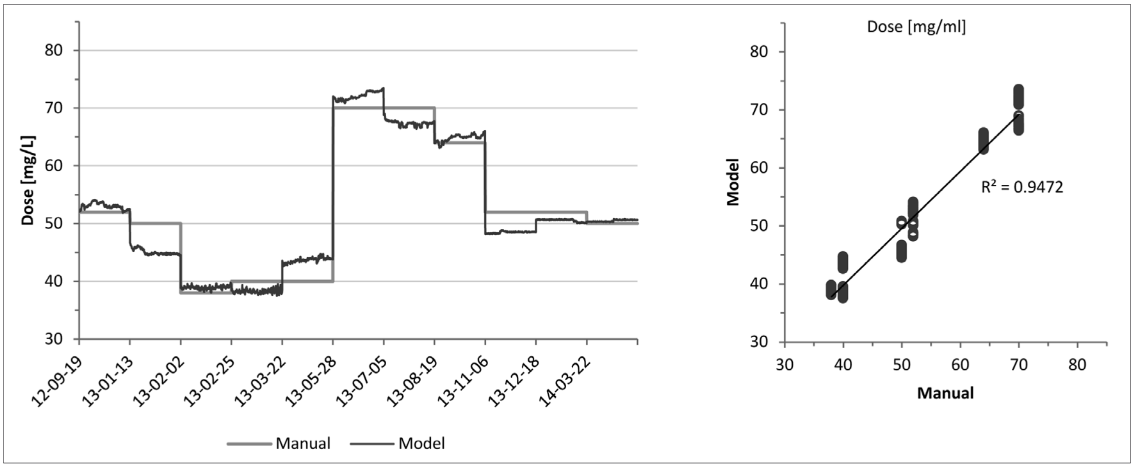

Figure 2. *External validation of the feasibility study model. The time series plot (left) and the scatter plot (right) illustrates the relationship between the dose actually used ("Manual") and the dose predicted by the feasibility study model ("Model") during the 11 periods used for external validation.*

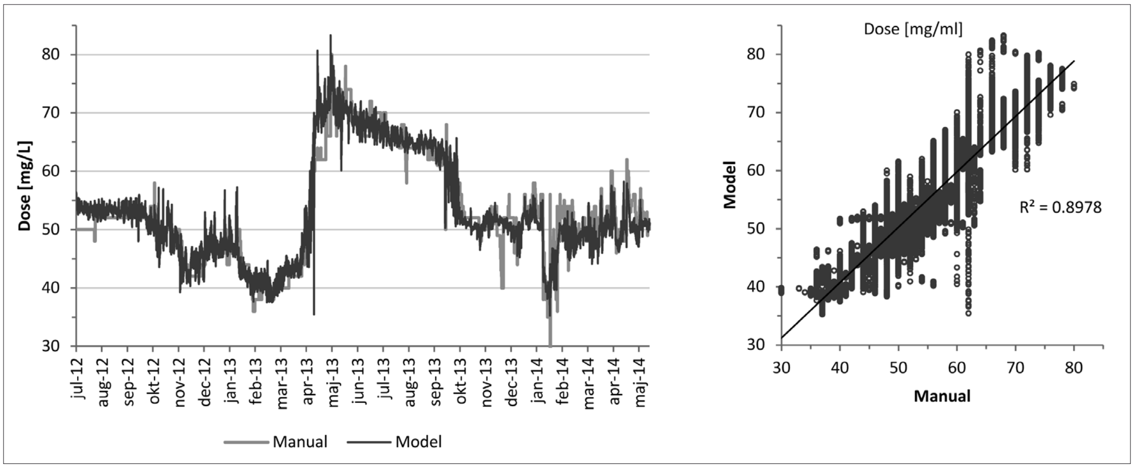

Figure 3. *External validation of the final model. The time series plot (left) and the scatter plot (right) illustrates the relationship between dose actually used ("Manual") and dose predicted by the final model ("Model") during the 2 years of historical used for external validation.*

relation between predicted and actual dose was high (see Figure 3) and the RMSEP (Table 3) was relatively low, i.e. the predictive ability of the model was good.

Worth noticing when comparing the predicted dose with the actual dose is that the dose that was actually used not always was the "optimal" dose. There might have been periods when an overdosing or underdosing has occurred due to for example rapid changes in raw water composition. Also, on many occasions when there is a deviation between predicted and actual dose, an effect on the treated drinking water could be noted (not shown). The deviations in some cases indicate that the operators might have had to under- or overdose coagulant to facilitate optimal performance of the rest of the plant. The operators might for example increase the coagulant dose when the flow limit for optimal performance of the UV-treatment is exceeded. This is done to decrease the contaminant concentration and in that way increase the UV-treatment efficacy of the UV-aggregates.

Implementation

Implementation for monitoring

The external validation of the final PLS-model showed that it seemed to be reliable, and it could therefore be implemented in the ABB 800 XA control system at the Görväln plant in September 2014 with the purpose to evaluate its performance online. After solving some initial problems originating from the control system, it was shown that the model performance was similar to its performance in the external validation. The next step was therefore to use the model for real-time control of the coagulant dose.

Implementation for control

On April 1st 2015 the final PLS-model was implemented for control of the coagulant dose. To evaluate the effect of the implementation, the first 26 days of the implementation, i.e. April 1–26 2015, was compared with the corresponding period during 2014, when the coagulant dose was manually controlled. The two periods were concluded to provide a good comparison evaluation based on the fact that they were from the same time of year and that they had similar level and variation of the composition of the water entering the precipitation. This is illustrated in Table 4 by mean value and standard deviation for UV-absorbance (254 nm) in the water entering the precipitation (1.34 and 0.026 respectively for April 2014 and 1.35 and 0.026 respectively for April 2015) and in Figure 4.

It was concluded that the treatment result was significantly more stable during April 2015 than during April 2014. This can be exemplified by the UV-absorbance in the water leaving the sand filter as a measure of contaminants in the treated water. The standard deviation of the UV-absorbance in the water leaving the sand filter was much lower in April 2015 than in April 2014 (0.005 compared to 0.012, see Table 4) and by visual inspection of Figure 4 it is clear that the UV-absorbance varies less in 2015. Furthermore, Figure 4 also shows that the variation of the treatment result was drastically increased on the occasions when the model based control was disconnected during April 2015.

Table 4. *Standard deviations and mean values for dose, UV-absorbance (254 nm) in the water entering the precipitation (UVin) and UV-absorbance (254 nm) in the water leaving the last sand filter (UVout) are shown for a comparison between April 2014, when the dosage was manually controlled, and the periods during April 2015 when the model was used for dosage control.*

	Average			Standard deviation		
	Dose (mg/L)	UVin	UVout	Dose (mg/L)	UVin	UVout
Manual (2014)	51.5	1.34	0.410	2.2	0.026	0.012
Automatic (2015)	49.4	1.35	0.418	1.5	0.026	0.005

The average dose was in the same range during the two time periods, but the standard deviation was lower in April 2015, i.e. approximately the same amount of coagulant was used. The average treatment results were also comparable during the time periods, which is exemplified by the small difference in the mean values for UV-absorbance (0.008).

Thus, the first evaluation of the implementation showed a significantly reduced variation and maintained quality of the treated drinking water, without increasing the coagulant consumption. The decreased variation of the treatment results when using the model for dose control is most likely a result of a quicker and more correct response to changes of the raw water composition.

Prerequisites for accurate performance of the model based automatic dosage control

The fact that the model developed for the dosage control is based on sensor values makes it even more important to use reliable sensors, monitor their performance and to keep them well-maintained. If one or more sen-

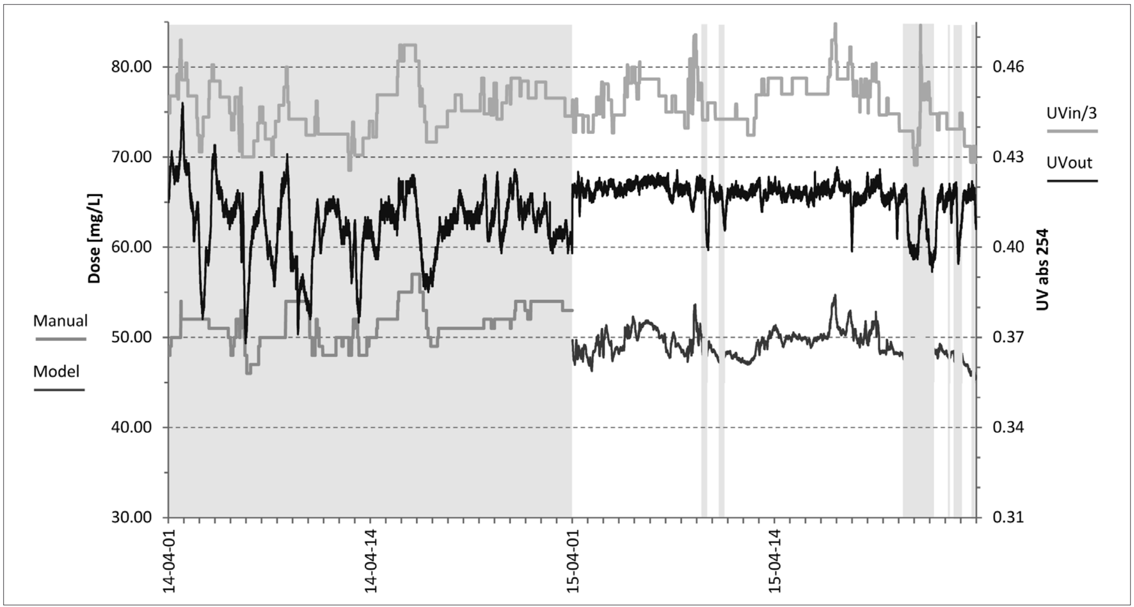

Figure 4. *The effect of the implementation of the model based dosage control illustrated by a comparison between April 1–26 2014 (manual dosage control) and April 1–26 2015 (mostly automatic dosage control). Manual dose during 2014 ("Manual") and automatic model based dose during 2015 ("Model") are shown in the lower part of the chart. The quality of the incoming water is represented by UV-absorbance in the water entering the precipitation step divided by 3 ("UVin/3") and the quality of the treated water is represented by UV-absorbance in the water leaving the sand filter ("UVout") are shown in the upper part of the chart. The time periods when the manual dosage control were used is shaded with grey.*

sors give incorrect values, the model will in its turn give the control system an incorrect estimation of which dose of coagulant to use, which may result in an undesirable treatment result.

At the Görväln plant limits for the x-variables as well as for the dose was set to prevent this, but one weakness with that approach is that the combination of sensor values could be outside the valid range of the model even though each parameter value is within its set limits. An attractive solution could be to use multivariate diagnostics tools to complement or replace the limits. One disadvantage with the multivariate diagnostics might be that it is more complicated to implement it in the control system.

Exchange of sensors or changes related to the precipitation process might call for a model update. After maintenance of the sensors or the process it is recommended to more carefully monitor the model prediction to validate the results.

It is worth emphasizing the importance of acceptance of and commitment to the new concept at the plant. The engineers and operators on site are continuously both handling the sensors and monitoring the performance of the implemented model and will also be the first to notice if a re-calibration of the model is necessary.

Conclusions

This paper described the process of developing a PLS-model based decision support system for control of coagulant at the Görväln drinking water treatment plant, with the aim of enabling the transition from manual to automatic coagulant dosage control. The external validation with two years of historical data showed that the model provided reliable estimations of the coagulant dose, and the first results from the implementation in the control system demonstrated that it significantly reduced the variation of the quality of the drinking water without increasing the coagulant consumption.

The positive results came with a number of future possibilities. The Görväln plant could start using an automatic control of coagulant, which should result in reduced variation of the quality of the treated drinking water as an effect of the quicker and more correct response to changes in the raw water composition. This could be done without any investments in hardware or software, which is a major advantage.

The results pave the way for future optimization of the coagulant dose, resulting in reduced coagulant consumption while still maintaining or even increasing the drinking water quality.

The approach presented can most likely also be used with positive results at other drinking water plants. The next step will be to investigate if a re-calibrated version of the model, or a similar model, is valid for other types of raw water, coagulants and sensors by introducing and to further develop the decision support system at other drinking water treatment plants. Another future approach could be to develop a more basic system for smaller plants.

Aknowledgements

The authors wish to thank David Heldt, Kristina Dahlberg, Sofia Wängdahl and Toni Aigner at Görvälnverket for providing data, knowledge and feedback that made this project possible, and for implementing the final result in their plant. We also thank Norrvatten for financing the project.

References

Geladi, P., Kowalski, BR. (1986) Partial least-squares regression: a tutorial. Anal. Chim. Acta, 185, 1–17.

Martens, H., Naes, T. (1989) Multivariate calibration. John Wiley and Sons, Chichester.

Stödjande medlemsföretag

AquaP
Arboga Kommun
Brenntag Nordic AB
COWI AB
Falköpings Kommun
Golder Associates AB
Gryaab AB
Göteborgs Stad Kretslopp och vatten
Haninge kommun
Hudiksvalls Kommun
Hässleholms Vatten AB
Höganäs Kommun
Kalmar Vatten AB
Karlstads kommun
Kemira
Kristianstads kommun
Käppalaförbundet
Köpings Kommun
Ljusdal Vatten AB
Malmberg Water AB
Miljö och vatten i Örnsköldsvik MIVA AB
Mittskåne Vatten
Mälarenergi AB
Mönsterås Kommun
Norconsult
Nordiska Plaströrsgruppen
Norrköpings Kommun/Livsmedel
Norrvatten
NSVA
Ragn-Sells Avfallsbehandling AB
Sjöbo Kommun
Skellefteå Kommun
Sollefteå Kommun
Stockholm Vatten VA AB
Stora Enso Hylte AB
Strömsunds Kommun
SWECO
SYDVATTEN
SYVAB
Säffle Kommun
Tomelilla Kommun
Trelleborgs Kommun
Trollhättan Energi
Tyréns AB
Tyréns AB
UMEVA
Uppsala Vatten och Avfall AB
Urban Water Management Sweden AB
VA-avdelningen NVK
Vattenfall AB
Veolia Water Technologies AB
WSP Sverige AB
Vänersborgs Kommun
Västvatten AB
Växjö Kommun
Xylem Water Solutions Sweden AB
Yara AB/ Industrial
Ånge Kommun
Östersunds Kommun

Vill ni synas bättre i listan?
För 2000 kr om året kan ni få er logotyp i listan ovan.