Application of the Morris method for screening the influential parameters of fuzzy controllers applied to wastewater treatment plants

M. V. Ruano, J. Ribes, J. Ferrer and G. Sin

ABSTRACT

In this paper, we evaluate the application of a sensitivity analysis to help fine-tuning a fuzzy controller for a biological nitrogen and phosphorus removal (BNPR) plant. The Morris Screening method is proposed and evaluated as a prior step to obtain the parameter significance ranking. First, an iterative procedure has been performed in order to find out the proper repetition number of the elementary effects ($r$) of the method. The optimal repetition number found in this study ($r = 60$) is in direct contrast to previous applications of the Morris method, which usually use low repetition number, e.g. $r = 10 \sim 20$. Working with a non-proper repetition number ($r$) could lead to Type I error (identifying a not-important factor as significant (false positive)) as well as Type II error (identifying an important factor as not significant (false negative)), hence emphasizing the importance of finding the optimal repetition number for each study in question. With the proper $r$ found, the Morris Screening helped identify the parameter significance ranking, thereby facilitating the calibration of fuzzy controllers, which contain many parameters that need to be adjusted for different wastewater treatment plant (WWTP) applications.

Key words | fine-tuning, fuzzy controllers, parameter significance ranking, screening method, sensitivity analysis

INTRODUCTION

Wastewater treatment plant (WWTP) models are used for many applications/purposes including plant design, optimisation and control. It is generally accepted that the modelling and simulation of WWTPs represents a powerful tool for control systems design and tuning. However, the model predictions are not free from uncertainty, as these models are an approximation of reality (abstraction), and are typically built on a considerable number of assumptions. In this regard, sensitivity analysis provides useful information for the modellers as this technique attempts to quantify how a change in the input model parameters affects the model outputs. Different strategies have been applied in the literature (see, for instance, Saltelli et al. 2000), which are typically classified in two main categories: global sensitivity analysis, where a sampling method is taken and the uncertainty range given in the input reflects the uncertainty in the output variables (Monte Carlo analysis; Fourier Amplitude Sensitivity Test (FAST), Morris Screening (1991)); and local sensitivity analysis, which is based on the local effect of the parameters in the output variables (Weijers & Vanrolleghem 1997; Brun et al. 2002).

Fuzzy logic based controllers have been successfully applied to wastewater treatment processes (see e.g. Ferrer et al. 1998; Serralta et al. 2002), since fuzzy sets theory offers an effective tool for the development of intelligent control systems (Zhu et al. 2009). Fuzzy control algorithms can be used to create transparent controllers that are easy to modify and extend because the fuzzy-rules are written in the language of process experts and operators (Yong et al. 2006). Although these control systems have been shown to be more robust than classical controllers (Manesis et al. 1998; Traoré et al. 2005), they usually contain quite a number of parameters, which complicates their calibration. So far, these control systems have been tuned by trial and error methods, based on technical knowledge on the process and controller performance (Chanona et al. 2006). Whatever optimisation method is applied, the fine-tuning of these controllers requires a previous selection of the most important parameters.
parameters to be adjusted in each particular application. Sensitivity analysis can be used to find the proper parameters for fine tuning of a fuzzy controller applied to WWTP plant operation. Ruano et al. (2010) proposed a systematic approach for fine tuning of fuzzy controllers based on model simulations that employs three statistical methods: (i) Monte-Carlo procedure: to find proper initial conditions, (ii) Identifiability analysis: to find an identifiable parameter subset of the fuzzy controller based on local sensitivity analysis and (iii) minimization algorithm. However, this methodology is based on local sensitivity analysis, and then requires an iterative procedure to confirm that the identifiable parameter subset does not depend on the local point in the parameter space where the identifiability study has been carried out. Global sensitivity analyses are proposed to overcome the problem of selecting the proper initial point in parameter space.

In this work, a global sensitivity analysis is applied aimed at screening the most influential parameters of the fuzzy-control systems to be used in the fine-tuning procedure. The Morris method of Elementary Effects (EEi) (Morris 1991) was used as sensitivity analysis method. In addition, the implications of using this methodology at the usual low repetition number of elementary effects \( r = 10 \sim 20 \) is assessed. This screening approach is proposed as a prior step to obtain the parameter significance ranking and the best initial point to carry out the aforementioned identifiability study.

**MATERIAL AND METHODS**

**WWTP and control system description**

The fuzzy logic-based control system was implemented to control the aeration in a nutrient-removing WWTP based on a modified University of Cape Town (UCT) scheme (see Figure 1).

This control system was previously developed by the research group and it has been applied in several full scale WWTPs (Ribes et al. 2007). The developed control system consists of two hierarchical control layers: the supervisory control and the process control layer. In the process control layer, several independent fuzzy logic-based control loops were developed for two process variables: dissolved oxygen in each one of the aerobic zones and blowers that supply air and maintain pressure in the air pipelines. The first controller manipulates the air valve opening according to dissolved oxygen (DO) concentrations, while the rotational speed of the blower is manipulated by a frequency converter to reach the blowers’ pressure set point. Since each control valve is governed by an independent DO controller, the air pressure controller is implemented in order to enhance the overall control system performance, especially when there is more than one control valve in the same air pipeline system. This controller aims at obtaining the ideal situation where the movement of one valve, which is governed by its own DO controller, does not affect the air flow rate through the other valves in the same air pipeline system. In the supervisory control layer, the optimum set point of the blowers’ pressure is established according to the control valves’ opening and the DO concentrations. In this way, the discharge pressure level remains at the minimum level required to maintain the DO concentration at the set points in all the aerobic zones. Thus, this control system allows different DO set points to be fixed for different reactors, which makes it possible, for instance, to achieve simultaneous nitrification-denitrification, as was found in Denia WWTP (Ribes et al. 2007). However, it should be mentioned that only one aerobic zone has been considered in this case study (see Figure 1) in order to reduce the computational cost of the sensitivity study. Thus, the aeration control system is simplified to one DO controller and one air pressure controller.

For the DO controller, the input variables are the oxygen error \((OE)\) and the accumulated oxygen error \((AOE)\) and the output variable is the increment/decrement of the air valve opening \((IV)\). For the air pressure controller, the input variables are the pressure error \((PE)\) and the accumulated pressure error \((APE)\) and the output variable is the increment/decrement of the rotational speed of the blower \((IB)\), which is governed by a frequency converter. Both controllers are fuzzy logic-based controllers, which consist of five stages. Figures 2(a) and (b) show these five stages of the dissolved oxygen controller and the air pressure controller, respectively.

The first stage and the last stage are the same for both controllers. Stage 1 is the input stage (the measured or calculated variables are input to the controller) and stage 5 is the output stage. In stage 2, so-called fuzzification, the data collected from on-line sensors are converted into linguistic variables (fuzzy set), represented by membership functions (Gaussian shape in this study). In stage 3, a set of rules (so-called inference engine) are applied to the fuzzy set obtained in stage 2. The output linguistic variables are obtained in this stage by the Max-Prod operator, following the Larsen’s fuzzy inference method (Larsen 1980). These linguistic variables are converted into numerical control actions in stage 4, which is called defuzzification. In order to obtain a single output
value from our fuzzy linguistic set, the Height Defuzzifier method was employed (Mendel 1995), which only uses the centre of the Gaussian defuzzification membership functions. The total number of parameters of both controllers comes from the different stages of the fuzzy logic-based controllers, mainly derived from the defuzzification and fuzzification steps (Ruano et al. 2010). Both controllers use three Gaussian membership functions to fuzzify each input (“High Negative”, $HN$; “Low Negative”, $LN$; and “Low Positive”, $LP$), which gives a total of 12 membership functions from both fuzzification stages; and four to defuzzify each output (“High Negative”, $HN$; “Low Negative”, $LN$; “Low Positive”, $LP$; and “High Positive”, $HP$), which gives a total of eight membership functions from both defuzzification stages. As each Gaussian curve is defined by two parameters (centre, $c$, and amplitude, $a$), the control system has a total of 24 parameters corresponding to the fuzzification stages. In contrast, for the defuzzification stages, only the centres are used, giving a total of eight parameters. Including the response time of the control system ($RT$), i.e., time interval between two control actions, a total number of 33 parameters need to be adjusted. In order to identify the parameters of this control system, acronyms for each parameter have been used. These acronyms are constructed as follows: “abbreviation of input variable” + “c/a” + “fuzzification/defuzzification membership function abbreviation”. For instance, the acronym $OEaHN$ means the amplitude of the High Negative membership function for the input variable Oxygen Error; and the acronym $IVcLN$ means the centre of the Low Negative membership function for the output variable Increment air Valve opening.

In order to decrease the computational demand in this case study, we simplified the control system to 17 parameters, assuming symmetric behaviour of the membership functions defined for each fuzzy variable. This symmetric behaviour involves that the amplitude for the three Gaussian membership functions is the same and their centres are equidistant. This control structure is the simplest one that can be implemented in a WWTP, as the oxygen concentration is controlled in only one reactor. When more aerobic reactors are to be controlled, the number of parameters will increase proportionally. As an example, this control system has recently been implemented in Denia WWTP (Denia, Spain) (Ribes et al. 2007), with nine different oxygen variables and one pressure (as the same group of blowers is used to aerate the whole system). Hence, reducing the number of tuning parameters in this kind of controller is essential.
The fuzzy controller and the WWTP model were implemented and simulated using the WWTP simulation software DESASS (Ferrer et al. 2008). This software includes the plant-wide model Biological Nutrient Removal Model no 1 (BNRM1, Seco et al. 2004). The simulation strategy consisted of a steady-state simulation to obtain proper initial conditions followed by 28 days of dynamic simulations. The last 14 days were considered to evaluate the performance of the control system. The standardised influent file for dry weather proposed by Copp (2002) was used in this study. The Integral Absolute Error (IAE, integral of the absolute value of the time dependent error function) for each controller (Oxygen and Pressure) was selected as an output measure. So, in this study IAEO and IAEP are the output variables.

**Sensitivity analysis**

**Morris screening**

The method of Morris (1991) evaluates the so-called distribution of Elementary Effects (EE) of each input factor to model outputs, from which basic statistics are computed to derive sensitivity information. This distribution function is denoted as \( F_{jk} \), which stands for the distribution of the effects of the \( j \)th input parameter on the \( k \)th output. The \( EE_{jk} \) attributable to each input parameter is obtained from the following differentiation of model output, \( s_y \), with respect to the input, \( \theta \):

\[
EE_{jk} = \frac{sy_k(\theta_1, \theta_2, \theta_j + \Delta, \ldots, \theta_M) - sy_k(\theta_1, \theta_2, \theta_j, \ldots, \theta_M)}{\Delta}
\]

where \( \Delta \) is a predetermined perturbation factor of \( \theta_j \), \( sy_k(\theta_1, \theta_2, \theta_j, \ldots, \theta_M) \) is the scalar model output evaluated at input parameters \( (\theta_1, \theta_2, \theta_j, \ldots, \theta_M) \), while \( sy_k(\theta_1, \theta_2, \theta_j + \Delta, \ldots, \theta_M) \) is the scalar model output corresponding to a \( \Delta \) change in \( \theta_j \). Each input parameter, \( \theta_j \), can only take values corresponding to (a predefined set of) \( p \) levels within its range. The calculation of the elementary effects, \( EE_{jk} \), is replicated \( r \) times, at randomly sampled points in the input space, leading to a distribution of \( EE_{jk} \) used to infer a global sensitivity measure. To do that, an effective one-factor-at-a-time (OAT) design has been developed (Morris 1991). In this case study, the input uncertainty for each control parameter was fixed at 50% of the default value.

To compare the sensitivity measures obtained for different model outputs, the scaled elementary effects \( SEE_{jk} \) proposed by Sin et al. (2009) were applied. The resulting elementary effects of each output variable (IAEO and IAEP) show the sensitivity of each controller separately (oxygen and pressure controller, respectively). However, both controllers must be tuned as a global MIMO (Multiple Input Multiple Output) control system. Thus, sensitivity analysis aiming at selecting the most influential parameters in this MIMO system was carried out giving an equal weight contribution of the scaled elementary effects obtained from both output variables. The scaled elementary effect of the \( j \)th input parameter on the weighted contribution of both output variables was calculated as follows:

\[
SEE_j = \frac{EE_{jIAEO} \sigma_{IAEO}}{\sigma_{IAEO}} + \frac{EE_{jIAEP} \sigma_{IAEP}}{\sigma_{IAEP}}
\]

where \( \sigma_{IAEO} \) and \( \sigma_{IAEP} \) are the standard deviations of the corresponding output variables. Once the distribution of the scaled elementary effects is obtained, the sensitivity measures mean (\( \mu \)) and standard deviation (\( \sigma \)) of each \( F_{jk} \) are determined. In order to identify the influential parameters, these sensitivity measures were then interpreted using the graphical approach proposed by Morris. In this approach, the values of \( \mu \) and \( \sigma \) obtained for all the \( F_{jk} \) distributions are displayed together with two lines corresponding to \( \mu_i = \pm 2 \text{SEM}_i \), where the \( \text{SEM}_i \) represents the standard error of the mean that can be estimated as \( \text{SEM}_i = \sigma_i/\sqrt{r} \). Parameters with low \( \mu \) and low \( \sigma \) are deemed as non-influential (Morris 1991).

One issue of particular interest is the selection of the resolution, \( p \), and number of trajectories, \( r \). Cropp & Braddock (2002) pointed out that a good choice of the repetition number (\( r \)) is more critical for obtaining a good estimate of the effects than the resolution (\( p \)). In this case study, an optimal setting of \( r \) was searched with a constant resolution of \( p = 8 \). To this end, the number of repetitions of elementary effects calculations (\( r \)) for each distribution \( F_{jk} \) was increased until the influential parameters remained more or less stable, i.e., the Type II error was minimised (type II error: identifying an important factor as insignificant). Once the \( r_{\text{opt}} \) was found, the graphical Morris approach was used to find the significant parameters.

**RESULTS AND DISCUSSION**

The Morris method was applied to a different number of elementary effects calculations, \( r \), until the sensitivity of parameters remained more or less stable. Table 1 shows the
resulting sensitivity measures ($\mu$ and $\sigma$) for the different number of elementary effects calculated.

As can be seen in this table, increasing the number of runs, due to a higher $r$, clearly demonstrates a closer similarity between the sensitivity measures of the parameters. However, increasing the number of elementary effects from 50 to 60 does not show significant differences between the significance ranking of the parameters.
considered influential (see below for the identification of influential parameters). As a result, \( r = 60 \) was selected as the optimal number of repetitions for this case study. The overall model evaluation costs were therefore 1,080 simulations \((= r^k(x + 1), k = 17)\) (computational cost of one simulation was around 10 min obtained using a PC with 4 GHz Intel® Pentium® processor). These results are in contrast to previous applications of the Morris method, since most of these studies used a low repetition number, e.g., \( r = 10 \sim 20 \) (Campolongo et al. 2007). As our results indicate, in agreement with Cropp & Braddock (2002), \( r \) has a significant effect on the identified parameter sensitivity, particularly for a low value of \( r \), (lower than \( r = 30 \), in this case study). This fact implies the necessity of finding out the optimal repetition number for SEEJK calculations \((r)\). A non-optimal selection of \( r \) would lead to Type II error (false negative), failing in the identification of a parameter of considerable influence in the model and Type I error (false positive), as well, considering a factor as significant when it is not. For instance, in this case study, for a repetition number of \( r = 10 \), the parameter RT (response time) was characterised by a low mean and low standard deviation and, for a repetition number of \( r = 15 \), the parameter OEaHN (amplitude of the High Negative membership function of the Oxygen Error) was also characterised by a low mean and low standard deviation. In contrast, the results for the optimal repetition number showed that these parameters present a considerable effect in the output variables, as was expected from practical experience. On the other hand, for a repetition number of \( r = 40 \) the parameter AOEaHN (amplitude of the High Negative membership function of the accumulated oxygen error) was considered as influential compared to the results for the optimal repetition number, where it is non-influential.

Figure 3 shows the graphical Morris approach for the optimal number of repetitions obtained for the two output variables (Figure 3(a) IAEo and Figure 3(b) IAEP). The resulting influential parameters from both output variables are similar (only one parameter is different). Figure 3(c) shows the same graph for the weighted contribution of the elementary effects obtained from both output variables (Equation (2)). This figure was used to screen out the non-influential parameters of the control system (i.e. the six parameters that are not labelled in Figure 3(c)). From the 11 influential parameters, RT, OEcHN (High Negative centre of the Oxygen Error), APEaHN (High Negative amplitude of the Accumulated Pressure Error) and IVcHN (High Negative centre of the increment of the air valve opening) presented high mean and low deviation, lying outside of the wedge formed by the two lines corresponding to \( \mu_i = \pm 2SE_i \). Thus, the effect of these parameters on the output variables is expected to be linear and additive, which is desirable for parameter estimation based on optimisation algorithms, since they are characterised by a higher mean than the rest of the parameters and not too high standard deviation.

These results are in agreement with the experience-based knowledge. RT is one of the most important
parameters, which is in agreement with its importance in the controller actions. High $RT$ values would give a slow response to changes in the process variables and, inversely, low values could cause too fast actions, which would lead to instabilities when controlling a process with either a low response time or a high time delay. $IVeHN$ is also important in the control system because it determines the change in the manipulated variable, i.e. the air valve opening. The magnitude of the oxygen control actions will basically be determined by this parameter and the frequency of each control action, given by $RT$. In contrast, the rest of the parameters characterised by low mean and high standard deviation present an effect that is dependent on the values of the other parameters through interaction effects or a nonlinear effect.

Compared to the results obtained in the previous work (Ruano et al. 2010), in which local sensitivity analysis was used to find the significant parameters, the following could be said: (i) the significant parameters identified in the local sensitivity analysis agreed mostly with those identified by the method of Morris, e.g., $OEcHN$ and $RT$ parameters; (ii) only one parameter identified as influential in the local sensitivity analysis is found to be non-influential according to the Morris screening study ($AOEaHN$, i.e. High Negative amplitude of the accumulated oxygen error). Hence, these results demonstrate that global methods, in particular the Morris Screening, are more resilient to Type II error, i.e., identifying non-influential parameters as influential. Hence, it is recommended for use in identifiability problems.

It is worth pointing out that the significant reduction of the important controller parameters could suggest a simplification of the control system design, i.e., to eliminate some redundant membership functions. However, it must be highlighted that one characteristic of fuzzy control is that, when the knowledge-based rule system is formulated, some redundant membership functions are defined in order to represent in a general way all possible system behaviours. In this way, the fuzzy control systems become more flexible at describing system dynamics, but at the expense of a large number of parameters (around 17 in this case). This is certainly different from proportional-integral-derivative (PID)-type controllers, which use much fewer numbers of parameters (around three per control loop) to describe input-output dynamics of the system; hence there are comparatively fewer parameters to fine-tune. However, when the system to be controlled has strong non-linear response to input dynamics, as the one presented in this work, fine-tuning PID controllers becomes a challenging (if problematic) issue since the tuning procedure is done for a specific working point (thereby ignoring other operating points in the system). From this point of view, fuzzy control emerges as an interesting alternative to PID controllers, which otherwise may require gain-scheduling-type adaptive controller strategies.

Finally, one caveat has to be mentioned when using the Morris screening method, which has to do with the selection of a proper repetition number of $EE_i$ calculations ($r_{opt}$). Probably, a high value of the parameter $r$ will be needed when either a highly nonlinear model is used, or a large input uncertainty is defined (lower and upper values in uniform distributions). It, thus, comes with a slightly higher computational cost, which is increasingly feasible to do.

As a guideline to find the proper repetition number, the Morris method must be applied to a different number of elementary effects calculations (at least for two $r$, starting with e.g. $r = 10$). In general, we proposed that the parameter significance ranking (as defined by the graphical approach of Morris) obtained for two different “$r$” must be similar. The $r$ is optimal when the following holds: (i) those parameters, which are found non-influential, are the same in both iterations with different $r$, and (ii) those parameters, which are influential, maintain similar ranking. Since the two measures must be evaluated at the same time, the graphical Morris approach facilitates this analysis.

We recommend the Morris screening as an efficient and promising method for sensitivity analysis in the wastewater modelling community for applications ranging from model calibration to controller fine-tuning applications.

**CONCLUSIONS**

Fuzzy controllers, in contrast to PID controllers, contain a large number of parameters to describe system dynamics in all possible operating conditions. This is an issue for their application, since fine-tuning of fuzzy controller parameters becomes challenging. Hence, the Morris screening approach has been applied to help with fine-tuning of a fuzzy logic based control system of a WWTP. Firstly, an iterative procedure has been executed in order to find out the proper repetition number of $EE_i$ calculations ($r_{opt}$). The optimal repetition number found in this study is in direct contrast to previous applications of the Morris method, which usually use a low repetition number, e.g. $r = 10 \sim 20$. Working with a non-proper repetition number of elementary effects ($r$) could lead to Type I error as well as Type II error. This approach is proposed as a prior step to obtain the parameter significance ranking for the identifiability
methodology used in the previous work (Ruano et al. 2010).

As regards parameter significance ranking, compared to the local sensitivity analysis, one can conclude that both results are in fairly good agreement with each other with the exception of one parameter identified falsely by the local method.

Overall, the Morris approach provides a good approximation of a global sensitivity measure, thereby overcoming the drawback of the differential analysis method whose results are only locally valid. The downside of the Morris method is the computational cost, which is increasingly feasible to do. Overall, the results supplemented with process engineering knowledge provide valuable insights into the engineering problems at hand, thereby helping engineers in doing their day-to-day work, e.g., helping to fine-tune a controller, checking the robustness of plant design, among others.

REFERENCES


