How do Initial Design Assumptions Determine Plant Sizing? Assessing Activated Sludge Process Design using Monte-Carlo Simulation and Global Sensitivity Analysis

Xavier Flores-Alsina, Lluís Corominas, Dirk Muschalla, Marc B. Neumann and Peter A. Vanrolleghem

model*EAU*, Département de génie civil et génie des eaux. Université Laval, 1065, Avenue de la Médecine. Québec G1V 0A6, QC, Canada

Abstract:

The main objective of this paper is to show the interest of performing Monte-Carlo (MC) simulation and global sensitivity analysis (GSA) to assess how major design variables of activated sludge plants are influenced by the initial design assumptions *i.e.* the inputs to the design process. The initial assumptions considered relate to i) the uncertainty of wastewater characteristics, ii) effluent requirements considered by the local regulators, iii) operator preferences and iv) the safety that the plant owner is envisaging. In the case study used to illustrate the approach, the widely recognized Metcalf & Eddy design guidelines are investigated. The Latin Hypercube sampling (LHS) technique is applied to generate samples from the input ranges, which are then propagated by MC simulation. Next, GSA is performed using Standardized Regression Coefficients (SRC) to reveal which initial assumptions influence the variation on the design variables most. The application of this approach has three advantages. Firstly, designers, regulators, operators and plant owners are provided with useful information about why, when and how design variables e.g. reactor volume, air demand or recycle flow-rate, may change if the initial design assumptions are modified. Secondly, process engineers can recommend to their clients how to prioritize their resources. Finally, the scenario analysis is able to answer questions such as what would happen if ... evaluating possible changes on the design, the relative importance of the initial assumptions and its associated variability when pre-defined conditions are changed.

Keywords: Latin Hypercube Sampling, Model-based Design, Nitrogen Removal Wastewater Treatment, Response Surface, Standardised Regression Coefficients, Uncertainty Analysis

INTRODUCTION

A major task during the sizing of an activated sludge unit is the specification of the initial design assumptions *i.e.* input to the design process. Given these assumptions, the selected design guideline (e.g. Metcalf & Eddy, ATV, Ten State Standards, HSA principles or Custom Models) will determine the values of the design variables of the plant. Initial design assumptions need to be made for wastewater influent characteristics, kinetics & stoichiometry, solids retention time, future operational conditions (e.g. controller set-points), safety factors, effluent requirements and settling characteristics of the bio-solids. These initial design assumptions are extremely important because they determine the future construction volumes, air blowers' characteristics or capacity of pump stations (Metcalf & Eddy, 2003). Hereby, the ranges in which the initial assumptions may vary represent either lack of knowledge or the possible degree of freedom of stakeholder choices. Making design assumptions can be regarded to be a complex decision problem (Clemen and Reilly, 2001) as the available information can be ambiguous, incomplete and uncertain.

To assess how the ranges in which the design assumptions are made affect the final design, the authors suggest the combined use of Monte Carlo (MC) simulation and global sensitivity analysis (GSA). Recent studies in water technology have applied these techniques for design (**Benedetti** *et al.*, 2006, Sin *et al.*, 2009), for the evaluation of control strategies (Flores-Alsina *et al.*, 2008) or to predict the overall process performance dependencies (Neumann *et al.*, 2007). However, all these

previous applications have focused on how lack of knowledge (epistemic uncertainty) of activated sludge models such as the influent fractions or the kinetic and stoichiometric parameters affect the performance prediction or designs. In this study, additionally to accounting for uncertainties we investigate how the degree of freedom in the decisions by various stakeholders, *i.e.* regulators, process engineers, operators and plant owners, determine the final activated sludge design and operation characteristics.

The main objective of this paper is to show the interest of performing MC simulation and GSA of the initial assumptions made during the design of activated sludge plants. In the presented work it is investigated how i) the uncertainty of the influent wastewater characteristics, ii) the possible effluent requirements considered by the local regulators, iii) the operator preferences and iv) the safety that the plant owner is envisaging, determine the final design variables.

The paper is organized as follows. First, the process design guidelines, the sources of uncertainty, the stakeholder choices and the systems analysis techniques are described. Next, the results from the MC simulations and the GSA are presented. Finally, the study is complemented with a scenario analysis.

METHODS

This section gives a general overview of the methodology used in this study. The ranges of values for the four initial assumptions are propagated (via the Metcalf and Eddy design equations) to the design variables by MC simulation. Their relative importance is identified by GSA on the MC results using Standardised Regression Coefficients (SRC).

Process Design Guidelines and Plant Setup

As case study the Metcalf & Eddy guidelines are selected to design the activated sludge unit. A modified Ludzack-Ettinger (MLE) configuration is chosen where the initial contact of the influent wastewater and return activated sludge occurs in an anoxic zone (ANOX), which is followed by an aerobic zone (AER) (see process layout in Figure 1). The process relies on the nitrate - formed in the aerobic zone - being returned via an internal recycle to the anoxic zone. The aerobic zone volume (V_{AER}) is sized on the basis of the net specific growth rate of nitrifying organisms, the desired mixed liquor suspended solids (MLSS) concentration and the total mass of solids that has to be removed to maintain the required sludge residence time. Next, the required internal recycle flowrate (Q_{INTR}) is calculated through a mass balance which includes the nitrate produced in the aerobic zone, the nitrate in the return activated sludge and the desired nitrate level in the effluent. The anoxic volume (V_{ANOX}) is designed by comparing the nitrate produced in the aerobic zone and the nitrate which potentially can be removed for a given hydraulic retention time. Finally the air flow (Q_{AIR}) is quantified based on the difference between the oxygen required (for carbon removal and nitrification) and the oxygen saved (by nitrate reduction). The wastewater to be treated has the same default profile and composition like the one used in the Benchmark Simulation Model No 1 (BSM1) (Copp 2002). In this case study, the design variables are represented by [X] and include V_{AER} , V_{ANOX} , Q_{INTR} and Q_{AIR} . To obtain [X] the 20 non-linear implicit algebraic equations of the Metcalf & Eddy guidelines were implemented as an m-file in MatLab. Further information about the design procedures can be found in Metcalf & Eddy 2003.

Monte Carlo (MC) Simulation

The MC procedure is commonly used for evaluating variations in the predictions of simulation models. It involves 3 steps: (1) Specifying ranges for the input factors, (2) sampling from the input factor ranges and (3) propagating the sampled values through the model to obtain a range of values for the output. In this case study the range of values considered for the initial assumptions [A] during the activated sludge plant sizing are the input factors.



Figure 1. Sources of uncertainty (influent fractions) and stakeholder choices (safety factors, operational conditions, effluent requirements) that determine the final activated sludge plant design

The range of values of the initial assumptions [A] are characterized using uniform probability distributions (**Table 1**). These distributions are assumed to characterize either the lack of knowledge by the design engineer (influent fractions) or the ranges of values that various stakeholders are considering for their decisions (safety factors, operational conditions, effluent requirements) (**Figure 1**). While the total organic load is considered to be constant and known, the different biodegradable and non biodegradable fractions are considered to be uncertain. The different effluent requirements, safety factors and operational conditions refer to decisions to be made by owners, regulators and future operators *i.e.* choices. The authors are aware of other parameters with strong impact on the future plant such as design temperature or the organic and nitrogen loads. Some of those are further investigated in the scenario analysis while the others are assumed to be constant.

Initial assumption [<i>A</i>]	symbol	Default value	Lower value	Upper value	units		
Influent fractions							
Inorganic soluble	Su	0.09	0.05	0.14	-		
Organic soluble	S _B	0.16	0.08	0.24	-		
Inorganic particulates	$X_{U, \text{inf}}$	0.12	0.06	0.18	-		
Organic particulates	X _B	0.52	0.35	0.72	-		
Heterotrophic biomass	Хоно	0.11	0.06	0.17	-		
Effluent requirements							
Effluent ammonium	S _{NHX}	2	0.5	6	gN⋅m⁻³		
Effluent nitrate	S _{NOX}	6	5	10	gN⋅m⁻³		
Safety factors							
Aerobic section	SFAER	1.25	1	1.5	-		
Anoxic section	SFANOX	1.25	1	1.5	-		
Operational conditions							
Dissolved oxygen in the aerobic zone	S _{O2}	2	0.5	4	(-gCOD)⋅m ⁻³		

 Table 1. Range of values of initial assumptions [A] expressed as uniform distributions characterised by default, upper and lower values

We apply the Latin hypercube sampling (LHS) method to generate 1000 samples from the space of initial assumptions in **Table 1** (McKay *et al.*, 1979; Iman *et al.*, 1981). The LHS method is a stratified sampling technique that enables covering the entire sampling space with a lower number of samples compared to random sampling. For each sample of initial assumptions [A] the different design variables [X] are computed with the Metcalf & Eddy equations: [X] = f([A]). The average time to run this analysis (1000 MC simulations) on a regular PC is around 2 minutes.

Global Sensitivity Analysis (GSA)

GSA involves performing a linear regression on the output of the MC simulation (here 1000 simulations), revealing the (linear) relationships between the inputs [A] and the outputs [X]. The regression that is conducted for each design variable is represented in the following equation (**Eq1**)

$$\widehat{X} = b_0 + \sum_{k=1}^{ns} b_k A_k$$
 Eq1

where \hat{X} is the regression model prediction, b_0 is the offset, b_k are the slopes and *m* is the number of inputs [A]. The standardized regression coefficients (SRC) are obtained by \hat{X} normalisation: $SRC = b_k \sigma_{A_k}$

 $SRC = b_k \sigma_{A_k} / \sigma_{\hat{X}}$. According to Saltelli *et al.*, (2004) the SRC are a valid measure of sensitivity if the coefficient of determination $R^2 > 0.7$. The higher the absolute values of the SRC, the stronger the influence of the corresponding input [A] on determining the output X. The absolute values of the regression coefficients are then ranked and categorized in strong, medium and weak influence by k-means clustering (Hair *et al.*, 1998).

RESULTS

Monte-Carlo (MC) simulation and Global Sensitivity Analysis (GSA)

Table 2 summarizes the ranges of the design variables [X] obtained from propagating the initial assumptions [A] from **Table 1** in the MC simulation. The ranges are characterised by average value, maximum and minimum values as well as the first and third quartile. The ranges are considerable and show the large impact of the initial design assumptions on tank and pump sizing. For example, depending on the influent characteristics, safety factors, effluent requirements, the designed aerobic volume (V_{AER}) can vary between 3740 and 15670 m³.

	Design variable [X]						
	V_{AER} (m ³)	$V_{ANOX}(m^3)$	Q _{AIR} (m ³ min ⁻¹)	Q _{INTR} (m ³ day ⁻¹)			
Average value	6230	3080	200	88420			
Maximum value	15670	5140	260	14050			
Minimum value	3740	1620	170	55630			
Percentile 25 (Q1)	5040	2650	190	67760			
Percentile 75 (Q3)	6930	3450	210	106540			
Q3-Q1	1890	800	20	38790			

Table 2. Average values, maximum and minimum values, quartiles Q1 and Q3 and Q3-Q1 for the designvariables [X] obtained in the uncertainty analysis

Table 3 summarizes the calculated SRC from the GSA (**Eq1**) for the four calculated design variables. The coefficients of determination obtained for the linear regression conducted for the sensitivity analysis are all above $R^2 > 0.9$ indicating that the output from the MC simulation is well described by the linear model and the SRCs are a valid measure of sensitivity. The influence of the initial assumptions is categorized in strong, medium and weak after classifying the absolute values of the different SRC using 3 group k-means clustering. The input parameters with strongest influence on the output uncertainty are highlighted in bold.

Initial	Design variable [X]								
assumption [A]	V _A	V _{AER}		V _{ANOX}		Q _{AIR}		Q _{INTR}	
	SRC	rank	SRC	rank	SRC	rank	SRC	rank	
Su	0.08	7	0.03	10	0.04	9	0.00	8	
SB	0.13	6	-0.72	1	0.70	2	0.01	3	
$X_{U,inf}$	0.26	5	0.26	5	0.05	8	0.00	9	
X _B	0.28	4	-0.46	3	0.95	1	-0.01	4	
Х _{оно}	0.05	8	-0.23	6	0.49	3	0.00	10	
S _{NHX}	-0.53	2	-0.16	8	-0.31	4	0.01	5	
S _{NOX}	0.02	10	-0.29	4	0.24	6	-0.98	1	
SFAER	0.38	3	0.12	9	0.22	7	-0.01	2	
SFANOX	-0.02	9	0.60	2	0.02	10	0.01	7	
S _{O2}	-0.56	1	-0.17	7	-0.32	4	-0.01	6	

officiants (CDC) for the

As can be seen in Table 3 the effluent requirement for ammonium (S_{NHX}) and the oxygen concentration in the reactor (S_{02}) have a strong influence on determining the aerobic zone tank volume (V_{AER}). Also, an increase in the influent biodegradability (S_B and X_B) significantly increases the design airflow (Q_{AIR}) and decreases the anoxic volume (V_{ANOX}). Effluent nitrate requirements (S_{NOX}) have a strong influence on the internal recycle flow (Q_{INTR}) and a moderate effect on the anoxic volume (V_{ANOX}). The higher the degree of safety (SF_{AER} and SF_{ANOX}) in the aerobic and the anoxic zone, the larger the reactor volume (V_{AER} and V_{ANOX}) will be.

The results of this analysis also enable the creation of higher-dimensional response surfaces (one for each design variable), which represent the design variable X as a function of the initial design assumptions [A]. Figure 2 shows a 3D projection that displays the combined influence of the two "influential" assumptions for both aerobic (VAER) and anoxic (VANOX) volumes. As mentioned before, high nitrification rates (S_{NHX}) and low dissolved oxygen concentration (S_{O2}) in the reactor lead to high aerobic volumes (V_{AER}) and vice versa (see Figure 2a). On the other hand, high influent biodegradability (either S_B or X_B) lead to smaller anoxic volumes (V_{ANOX}) (see Figure 2b). The effect of the safety factors (SFAER and SFANOX), can be observed in both cases, giving smaller and larger volumes for the very same initial assumptions.

The creation of these response surfaces allows a "regional" instead of a "local" analysis of the design. In this way it is possible to study the variation of the different design variables (V_{AER}, V_{ANOX}, Q_{AIR} and Q_{INTR}) as function of the ranges of the initial assumptions allowing a better understanding of the design space.



Figure 2. Combined influence of influential initial assumptions for aerobic (a) and anoxic (b) volume

Scenario Analysis

In this section we analyse how the results from MC and GSA are affected by changing some of the constants. Scenario *S0* denotes the base case described above. Three further scenarios (*S1*, *S2* & *S3*) are suggested and studied. Scenario 1 (*S1*) reduces the design temperature from 15 to 10 deg C. Scenario 2 (*S2*) changes the influent condition by reducing the nitrogen load by 50 %. Finally, in Scenario 3 (*S3*) the design TSS in the reactor is increased from 3000 to 4000 gTSS·m⁻³.

Table 4. Results of the scenario analysis							
	Design variable [X]						
	V_{AER} (m ³)	$V_{ANOX}(m^3)$	Q_{AIR} (m³min⁻¹)	Q_{INTR} (m ³day⁻¹)			
S1: D	S1: Decrease of the design operating temperature down from 15 to 10 deg C						
Average value	8398	3760.	204	88420			
Q3-Q1	2390	869	869 18				
most influential							
parameters	S _{O2} & S _{NHX}	S _{B,} X _B & SF _{ANOX}	S _B & X _B	S _{NOX}			
S2: 50% reduction of the nitrogen influent load							
Average value	6185	747	177	29541			
Q3-Q1	1827	235	17	19345			
most influential	S _{O2} & S _{NHX}	S _{NOX} , S _{B,} X _B &	S _B & X _B	S _{NOX}			
parameters		SF _{ANOX}					
S3: Increase of the design operating MLSS 3000 to 4000 gm ⁻³							
Average value	4670	2311	201	81086			
Q3-Q1	1415	593	18	38790			
most influential	S _{O2} & S _{NHX}	S _{B,} X _B & SF _{ANOX}	S _B & X _B	S _{NOX}			
parameters							

The effects of the temperature (*S1*) and the MLSS (*S3*) lead to increase and decrease respectively the aerobic V_{AER} and anoxic V_{ANOX} volumes and their associated variability (see **Table 4**). Further, a reduction of the internal recycle Q_{INTR} can be observed in *S3* due to the higher biomass concentration in the reactor. On the other hand, when the nitrogen load is reduced (*S2*) the aerobic zone (V_{AER}) remains practically equal while the anoxic zone (V_{ANOX}) and the aeration (Q_{AIR}) and internal recirculation (Q_{INTR}) flow rate are reduced in both average value and variation.

Compared to the base case (*S0*) the results of the GSA practically remain unchanged. The main difference can be observed in *S2*, with a re-ranking of the initial assumptions with strongest influence with respect to V_{ANOX} : In this case, the most influential initial assumption is S_{NOX} while the influent biodegradability (S_B and X_B) moves to second rank. With a reduced influent nitrogen load, even with low contents of organic matter, the factor that will determine the total anoxic volume (V_{ANOX}) is the effluent requirement for nitrate and not the composition of organic matter.

The results generated in this study allow process designers learn about the design procedures, in this illustrative case study the Metcalf & Eddy guideline. Thus, it will be possible to deduce general properties of the design guidelines that can be applied to a wide range of cases as shown for S1 and S3 (they do not lead to changes in the importance ranking of the initial assumptions). Nevertheless, there are always special cases that will have to be treated separately *e.g.* S2 (which leads to a change in the rankings).

DISCUSSION

The results presented in this paper open the door to several discussions. Firstly, from the response surfaces generated in the previous studies, consulting engineers, plant owners, operators and regulators are provided with useful information about *why*, *when* and *how* the different construction volumes, air blower characteristics or pumping station capacities may vary for different initial design assumptions. In the presented case study, the *why*, *when* and *how* was depicted in the previously presented response surfaces (**Figure 2**). Specifically, **Figure 2a** identified that strict

effluent requirements (low S_{NHX}) and a low oxygen concentration (S_{O2}) in the bioreactor were the main *cause* of large aerobic volumes (V_{AER}). Similarly **Figure 2b** showed that poor organic matter biodegradability (low $S_B \& X_B$) increased the anoxic tank volumes (V_{ANOX}). Also, it was possible to identify the situations *when* the changes were more pronounced and *how* these changes occur. For example, the aerobic volume (V_{AER}) increased strongly when the effluent requirements (S_{NHX}) were lowered from 4 to 1 g m⁻³. On the other hand, the anoxic volume (V_{ANOX}) doubled or tripled if the organic biodegradability (S_B) or the safety factor (SF_{ANOX}) was changed from 0.05 to 0.3 and 1 to 1.5 respectively.

Secondly, process engineers can recommend to their clients where they *should best invest resources* during the design process. For example, the analysis brought to light the existing synergies and trade-offs between cost for construction (larger volumes) and equipment (smaller aeration system) and operation (lower aeration energy). For this reason a thorough analysis balancing construction costs, aeration costs and costs of possible effluent violations would be strongly encouraged in order to optimize the tank volume required for nitrification. Another case was the study carried out for the anoxic reactor. After the analysis one may conclude that it is useful to perform a detailed characterization of the influent organic matter, particularly if lower nitrate effluent concentrations are demanded by the regulators. The results revealed potential savings through reduced anoxic volumes and size of the aeration system.

Thirdly, thanks to the scenario analysis it was possible to complement the entire evaluation process and to answer what-if questions permitting to evaluate changes in the design and the relative importance of the initial design assumptions. For example, if the plant is constructed in a location with large differences between winter and summer temperatures (S1), higher averages and larger ranges have to be expected for both aerobic (V_{AER}) and anoxic (V_{ANOX}) volumes. Nevertheless, the results of the GSA showed that VAER and VANOX were still sensitive to SO2 and SB and XB. For this reason, if an additional investment is done in a good aeration system or if the occasional addition of an external carbon source is considered, the biological volumes (VAER and VANOX) could be significantly reduced. Another interesting example was scenario 2 (S2), where it could be seen that the default influent wastewater did not have a suitable C/N ratio. This fact could be suspected from the anoxic zone hydraulic retention times obtained in SO (HRT = 4 hours), which is higher than the values suggested in literature (HRT = 1-3 hours). The scenario analysis revealed that a substantial reduction in construction (V_{ANOX}) and operational costs (Q_{INTR} and Q_{AIR}) occurred at a lower nitrogen load. The results of this analysis can warn the plant manager not to accept some high Nstrength industrial influent, encouraging the implementation of source control measures. Finally, in scenario 3 (S3), it was possible to see a substantial reduction of the biological volumes when the design MLSS concentration was allowed to increase. Keeping this idea in mind, some designers may want to invest in a larger secondary clarifier, thus avoiding possible solids separation problems for such high MLSS concentration.

In summary, it should be highlighted that the main objective of this paper is to present a new way of analyzing design guidelines by use of MC simulation and GSA. Even though the authors have used the Metcalf & Eddy guidelines to quantify the different design variables, other approaches such as the ATV rules, CAPDET or HSA can be subjected to such analysis.

CONCLUSIONS

The paper shows the effect of the initial design assumptions on the sizing of volumes, recycle & aeration flows for activated sludge processes. The authors suggest the use of MC simulation and GSA to assess the design process. The proposed approach combines MC simulations using Latin Hypercube Sampling with Standardised Regression Coefficients as measure of sensitivity. The paper contributes to the field of wastewater engineering with a method that allows a "regional"-

instead of a "point"-based design analysis by means of the creation of response surfaces. These response surfaces represent the variation of the different design variables as function of a set of ranges of initial assumptions representing lack of knowledge, engineering choices, regulator choices on effluent limits or preferences of the future operator. At the same time it was possible to assess the relative importance of these initial design assumptions in determining the design variables such as reactor size, air demand and the flow of the different recycles. Additionally, the results generated during the study allow the process designer to learn and deduce general properties of his/her design guidelines.

This approach has three advantages over current design procedures. Firstly, designers, regulators, operators and plant owners are provided with useful information about *why*, *when* and *how* the plant size may change when the initial assumptions are modified. Secondly, process engineers can recommend their clients where they have to invest their resources. Finally, the scenario analysis enables to explore *what-if* questions and test the robustness of results obtained in the MC and GSA.

ACKNOWLEDGEMENTS

This research is supported by the Canada Research Chair in Water Quality Modelling and a NSERC Special Research Opportunities grant as part of the Canadian contribution to the European Union 6th framework project NEPTUNE. Lluís Corominas benefits from the postdoctoral fellowship "Beatriu de Pinós" of the Government of Catalonia. The authors wish to acknowledge the support provided by the IWA Task Group on Design and Operations Uncertainty . Finally, the authors wish to acknowledge the invaluable help of Marie-Ève Boucher.

REFERENCES

- Benedetti L., Bixio D. and Vanrolleghem P.A. (2006) Assessment of WWTP design and upgrade options: balancing costs and risks of standards' exceedance, *Water Sci. Technol.* **54**(6-7), 371-378
- Clemen, R.T. and Reilly, T. (2001) Making Hard Decisions, Duxbury. USA
- Copp J.B. (2002) The COST Simulation Benchmark: Description and Simulator Manual, Office for Official Publications of the European Community, Luxembourg.
- Flores-Alsina X., Rodriguez-Roda I., Sin G. and Gernaey K.V. (2008) Multi-criteria evaluation of wastewater treatment plant control strategies under uncertainty. *Water Res.* **42**(17), 4485-4497.
- Hair J.F., Andersen R.E., Tatham R.L. and Black W.C. (1998) Black, Multivariable Data Analysis, London, UK.
- Iman R.L., Helton J.C. and Campbell J.E. (1981) An approach to sensitivity analysis of computer models, Part 1. Introduction, input variable selection and preliminary variable assessment. J. Qual. Technol. **13**(3), 174-183.
- McKay M.D., Beckman R.J. and Conover W.J. (1979) Comparison of 3 methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics* **21**(2), 239-245.
- Metcalf & Eddy. (2003) Wastewater Engineering: Treatment and Reuse, McGraw-Hill Higher Education, New York.
- Neumann M.B., von Gunten U. and Gujer W. (2007) Uncertainty in prediction of disinfection performance. *Water Res.* **41**(11), 2371-2378.
- Saltelli A. Tarantola S., Campolongo, F. and Ratto, M., (2004), Sensitivity Analysis in Practice. A Guide to Assessing Scientific Models, John Wiley & Sons. Chichester. UK.
- Sin G., Gernaey K.V., Neumann M.B., van Loosdrecht M.C.M. and Gujer W. (2009) Uncertainty analysis in WWTP model applications: A critical discussion using an example from design. *Water Res.* **43**(11), 2894-2906.