Identification of Alternatives Strengths and Weaknesses during the Conceptual Design of Environmental Systems

Xavier Flores Alsina

Manel Poch

Laureano Jiménez

Rene Bañares-Alcántara

Ignasi Rodríguez Roda

Identification of Alternatives Strengths and Weaknesses during the Conceptual Design of Environmental Systems

Xavier Flores[1], Manel Poch[1], Ignasi Rodríguez-Roda[1], Laureano Jiménez[2] and Rene Bañares-Alcántara[3]

[1] *Laboratory of Chemical and Environmental Engineering, University of Girona, Montilivi Campus s/n 17071. Girona, Spain.*

[2]*Department of Chemical Engineering, University of Barcelona Martí i Franquès 1, 08028 Barcelona, Spain.*

[3]*Department of Engineering Science, University of Oxford, Parks Road, OX1 3PJ Oxford, United Kingdom.*

Abstract: The objective of this paper is to present a systematic procedure to support environmental engineers during the evaluation of alternative options, identifying both, the strong and weak points of their decisions. The evaluation of alternative options in the design of environmental systems is complex because several objectives must be considered simultaneously (e.g. environmental, technical, economical, legal), *i.e.* the problem is multicriteria. The systematic procedure presented consists of three steps: 1) multicriteria decision analysis (MCDA), where the best alternative is selected among the most promising options by evaluating the degree of accomplishment of several design objectives; 2) sensitivity analysis, where multidimensional response surfaces are generated to represent the variation of the selected option with respect to the relative importance of the design objectives; and 3) data analysis through the application of classification trees. The resulting set of rules highlights the relationship amongst the design objectives and the selected alternative without requiring a detailed examination of data values. The capability of this new procedure is demonstrated with a case study where the bioreactor of a wastewater treatment plant is designed to achieve simultaneous carbon and nitrogen removal. The results demonstrate how this new procedure supports the systematic evaluation of alternative options during the conceptual design of environmental processes.

Keywords: wastewater, environmental systems; multicriteria decision analysis; conceptual design, modelling, machine learning.

1. INTRODUCTION

From an economic and environmental point of view the early stages of the design process (conceptual design) must be considered very carefully. Mc Guire and Jones (1989) reported that up to 80% of the capital cost of any process is committed during this design stage. In addition, conceptual design is a complex task, because it is often necessary to evaluate alternative options in detail based on uncertain data and information.

In view of this complexity, there is a major need for tools that support environmental engineers in the appropriate selection of the most suitable system satisfying the design objectives and required process performance.

A novel systematic procedure to identify alternatives strengths and weaknesses during the conceptual design of environmental processes is presented in this paper. This procedure combines both, multicriteria decision analysis (see for example Vincke, 1992; Belton and Stewart, 2002) and machine learning techniques (ML). As a result, the relationship between options and design objectives are codified in a set of rules, thus avoiding the need of detailed examination of numerical values. The procedure is demonstrated with a case study where the conceptual design of a bioreactor to treat wastewater is carried out. Several design options are proposed and their strong and weak points are identified by rules.

2. IDENTIFICATION PROCEDURE OF DECISION MAKING STRENGTHS AND WEAKNESSES

This section details the proposed procedure to identify the strengths and weaknesses of the options being evaluated during multicriteria decision making.

In step 1 [Flores *et al.,* 2005], possible design solutions are represented as alternative options [$A = \{A_1,\ldots,A_m\}$]. Different criteria [$X= \{X_1,\ldots,X_n\}$] are used to measure the satisfaction of the design objectives [$OBJ = \{OBJ_1,\ldots,OBJ_p\}$] and weighting factors are assigned to determine the relative importance of these objectives [$w = \{w_1,\ldots,w_p\}$]. Weights are normalized to add to 1 and distributed through the evaluation criteria. The quantification of a design option A_j with respect to criteria X_i is indicated by $x_{j,i}$. Thus each option can be represented as a n-dimensional score profile [$A_j= (x_{j,1},\ldots,x_{j,n})$].

Value functions [$v(X_i)$] map the score profiles of each design option in a normalized value (1 to 0) in step 2. The 1 and 0 values are associated to the best (x_i^*) and worst (x_{i*}) situations respectively whilst a mathematical function is used to evaluate the intermediate cases. The collection of the best [$x^* = (x_1^*,\ldots,x_n^*)$] and worst [$x_* = (x_{1*},\ldots,x_{n*})$] scores for all criteria determine the best [$v(x^*) = (v(x_1^*),\ldots,(x_n^*)) = 1$] and the worst profiles [$v(x_*) = (v(x_{1*}),\ldots,(x_{n*})) = 0$].

Finally, a weighted sum (see eq1) is calculated to obtain a single value for each option $s(A_j)$ by adding the products of each normalized criterion $v(x_{j,i})$ times their corresponding weight w_i.

$$s(A_j) = \sum_{i=1}^{n} v(x_{j,i})\cdot w_i \qquad (1)$$

The option with the highest weighted sum is the one recommended for implementation.

Once the best alternative is selected, in step 2, a sensitivity analysis of the alternative options with respect to the design objectives is made, the weighted sum is recalculated (eq1), and multidimensional response surfaces are generated to represent the variation of the selected option with respect to the design objectives.

Finally, in step 3 all the data generated in the previous step are processed to extract qualitative knowledge by means of classification trees [Quinlan, 1993]. The set of rules extracted are useful to identify both the strong ($x_{j,i}^*$) and weak points ($x_{j,i}^*$) of each alternative option (A_j), avoiding detailed examination of the data generated in the previous step.

3. CASE STUDY

This case study shows an application of the proposed procedure during the selection of the bioreactor in a wastewater treatment plant. All the steps of the procedure, including numerical details, are discussed and described hereafter.

3.1 Step 1. Multicriteria decision analysis (MCDA) of the design options

Three activated sludge configurations are evaluated. The first is (A_1) the IWA denitrifying simulation benchmark plant [Copp, 2003]. This plant is based on the modified Ludzack-Ettinger configuration. It is comprised of five reactors in series (tanks 1 and 2 are anoxic with a total volume of 2000 m^3, while tanks 3, 4 and 5 are aerobic with a total volume of 4000 m^3) linked by an internal recirculation from the 3rd aerobic tank to the 1st anoxic tank and a settling tank (with a total volume of 6000 m^3). In the second option (A_2) there are two consecutive units, each one with a settler tank (6000 m^3 and 4000 m^3). The first unit, a conventional plug flow reactor, is approximated using the IWA nitrifying simulation benchmark plant [Copp 2003], i.e. five aerobic reactors in series (tanks 1 and 2 have a total volume of 2000 m^3 and tanks 3, 4 and 5 have a total volume of 4000 m^3) while the second is an anoxic denitrifying reactor (tanks 6 and 7 have a total volume of 1000 m^3) with an additional carbon source in the 6th reactor. Finally, the third option (A_3) is an oxidation ditch. This configuration consists of an oval shaped channel with anoxic (tanks 1, 2 and 3 with a total volume of 4000m^3) and aerobic (tanks 4, 5 and 6 with a total volume of 4000m^3) zones present in the same tank and a secondary settler (total volume of 6000 m^3).

These options are evaluated using five design objectives: minimize environmental impact (OBJ_1), minimize economical costs (OBJ_2), maximize technical reliability (OBJ_3), comply with the limits fixed by the European Directive 91/271/EC (OBJ_4) and minimize land occupation (OBJ_5). We assume equal importance of all the control objectives ($w_i = 0.20$, i = 1 to 5).

A single criterion is proposed, X_1, to measure the satisfaction of OBJ_1, i.e. the global treatment efficiency [Copp, 2003]. Construction costs (X_2, EPA, 1982) and operation costs (X_3, Vanrolleghem and Gillot, 2002) are used for OBJ_2. Plant robustness (X_4) and flexibility (X_5), control performance (X_6) and the sensitivity to separation problems (X_{7-1}= foaming risk; X_{7-2} = bulking risk and X_{7-3} = rising risk) measure the satisfaction of OBJ_3. A detailed description of X_4 and X_5 can be found in Flores *et al.* (2005), while X_6 and X_7 are reported in Stephanopoulos (1984) and Comas *et al.* (2005) respectively. X_8-X_{11} reflect the percentage of time that the concentration of the pollutant exceeds the legal limits, i.e. the time in violation (TIV) for TSS, COD, BOD_5 and TN [Copp, 2003]. Finally, land occupation is approximated by X_{12} [EPA, 1982].

Criteria X_2 and X_{12} are calculated with the CAPDET model [EPA, 1982] and the rest of criteria through dynamic simulation. The ASM1 model [Henze, 2002] includes 13 state variables and describes the biochemical carbon removal with simultaneous nitrification and denitrification with 13 non-linear differential equations. The double exponential settling velocity of Takács *et al.* (1991), based on the solid flux concept, was selected as a fair representation of the settling process with a ten layer pattern. A PI control loop for the dissolved oxygen in the aerobic zone is included with a set point of 1 $g{\cdot}m^{-3}$. All the dynamic simulations are executed after a steady state simulation; this ensures a consistent initial point. Only the data generated during the last seven days are used to quantify the criteria.

Once the criteria are quantified the score profile for each option is obtained. The score profiles for each option are represented in Table 1.

Table 1. Score profiles for the three evaluated options

	A_1	A_2	A_3	units
$x_{j,1}$	81.92	85.78	87.21	%
$x_{j,2}$	$4.44{\cdot}10^6$	$4.8{\cdot}10^6$	$4.25{\cdot}10^6$	€
$x_{j,3}$	$8.15{\cdot}10^5$	$1.39{\cdot}10^5$	$5.94{\cdot}10^6$	$€{\cdot}y^{-1}$
$x_{j,4}$	12.36	11.29	9.05	-
$x_{j,5}$	17.36	26.19	12.48	-
$x_{j,6}$	$7.56{\cdot}10^{-6}$	$4.27{\cdot}10^{-6}$	$5.6{\cdot}10^{-4}$	$(gO_2{\cdot}m^{-3})^2{\cdot}day$
$x_{j,7-1}$	0	0	0	%
$x_{j,7-2}$	8	6.99	4.32	%
$x_{j,7-2}$	78.42	48.21	0	%
$x_{j,8}$	0	0	0	%
$x_{j,9}$	0	0	0	%
$x_{j,10}$	0	0	0	%
$x_{j,11}$	78.42	8.33	0	%
$x_{j,12}$	8.14	8.34	8.30	ha

The extreme criteria profiles used in this case study were based on expert judgments : $v(x)^* = v(x_1^*,\ldots, x_{12}^*) = 1 = (x_1^* = 100$, $x_2^* = 2.5{\cdot}10^5$, $x_3^* = 8{\cdot}10^5$, $x_4^* = 25$, $x_5^* = 25$, $x_6^* = 0$, x_7-$x_{11}^* = 0$, $x_{12}^* = 8)$ and $v(x)_* = v(x_{1*},\ldots, x_{12*}) = 0 = (x_{1*} = 0$, $x_2 = 5{\cdot}10^6$, $x_{3*} = 1.5{\cdot}10^6$, $x_{4*} = 10$, $x_{5*} = 10$, $x_{6*} = 1$;$x_{7-11*} = 0$, $x_{12*} = 9$).Then a lineal model between these extreme values is adjusted to calculate intermediate situations (*e.g.* for criterion X_1 the value function is: $v(X_1) = 0.01{\cdot}X_1$).

Finally, the weighted sum is calculated (eq1) using the normalized values for each criterion. The higher the value of the weighted sum, the better the satisfaction of the design objectives.

Table 2. Normalized criteria and weighted sums

		A_1	A_2	A_3	w_i
OBJ_1	$v(x_{j,1})$	0.83	0.86	0.87	0.20
OBJ_2	$v(x_{j,2})$	0.36	0.05	0.30	0.10
	$v(x_{j,3})$	0.98	0.16	1.00	0.10
OBJ_3	$v(x_{j,4})$	0.37	0.31	0.20	0.05
	$v(x_{j,5})$	1.00	1.00	0.37	0.05
	$v(x_{j,6})$	1.00	1.00	1.00	0.05
	$v(x_{j,7-1})$	1.00	1.00	1.00	0.017
	$v(x_{j,7-2})$	0.91	0.93	0.96	0.017
	$v(x_{j,7-3})$	0.26	0.52	1.00	0.017
OBJ_4	$v(x_{j,8})$	1.00	1.00	1.00	0.05
	$v(x_{j,9})$	1.00	1.00	1.00	0.05
	$v(x_{j,10})$	1.00	1.00	1.00	0.05
	$v(x_{j,11})$	0.22	0.92	1.00	0.05
OBJ_5	$v(x_{j,12})$	0.86	0.64	0.80	0.20
$s(A_j)$		**0.78**	**0.67**	**0.80**	

From the first analysis of the results (Table 2), the recommended option is A_3 (oxidation ditch) and the rejected options are A_1 (modified Ludzack Ettinger) and A_2 (double stage). Note that for this case study X_6, X_{7-1}, X_8-X_{10} have the same value, and therefore they are not useful to discriminate among alternatives.

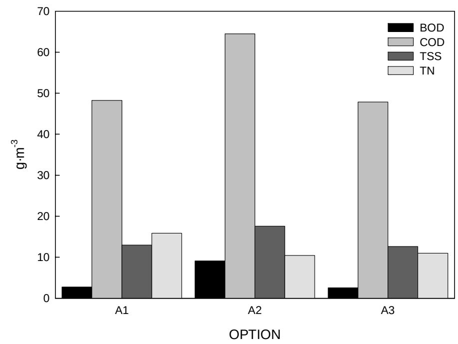
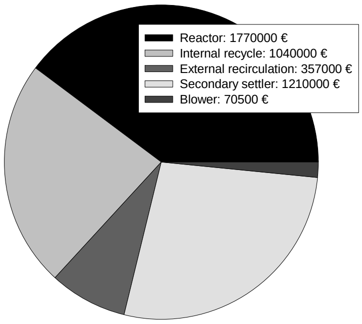
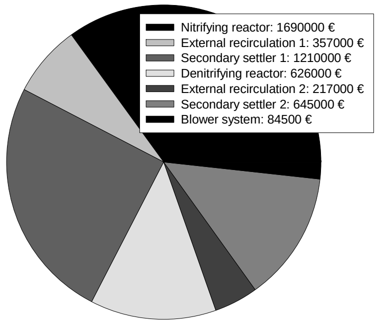


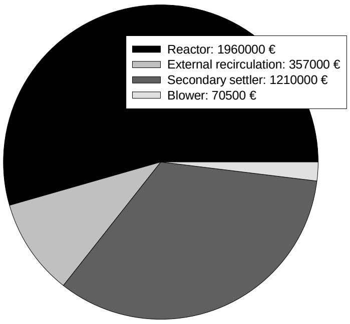
Figure 1. Average concentration of the effluent for the three options evaluated



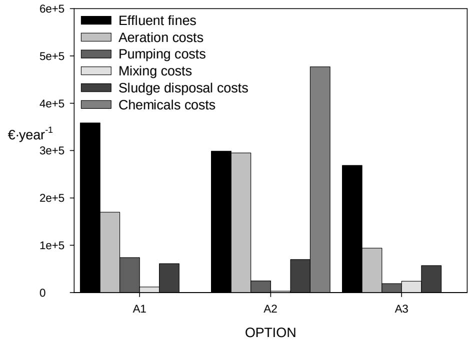
a)



b)



c)



d)

Figure 2. Breakdown of construction costs for options A_1 (a), A_2 (b) and A_3 (c) and of operation costs (d).

Option A_2 is the least favored despite its good performance with respect to OBJ_1, OBJ_3 and OBJ_4 due to its adaptability to long term perturbations (see Table 1 and 2) and high nitrogen removal efficiency (see Figures 1 and 3). Nevertheless, this option has the highest operation costs (see Figure 2d) resulting from its need of periodic purchase of chemicals for the post anoxic denitrification, the high aeration costs (because the entire BOD is completely removed in the aerobic zone), and its highest sludge production due to the addition of methanol (and the costs derived from its disposal). Moreover, construction costs (Figure 2b) are high because of the additional reactor and settler.

The second best option is the modified Ludzack Ettinger (A_1) because it has the lowest scores for OBJ_1, OBJ_2 and OBJ_3. The high construction cost (see Figure 1a) of this option is due to the internal recirculation. Furthermore, partial denitrification (Figures 1 and 3) in the anoxic section causes the plant to operate above the legal limits during most of the time, induces potential problems with rising sludge and damages the receiving water body.

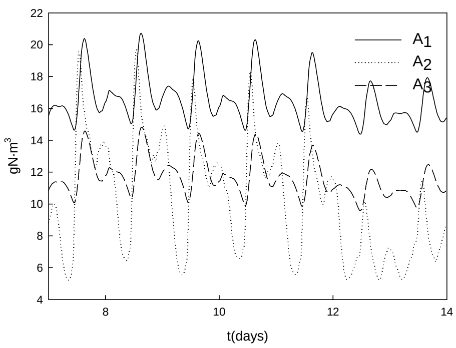


Figure 3. Dynamic effluent nitrogen concentration for the three evaluated options

Finally, option A_3 achieves the best punctuation because it scores higher in all the criteria used in OBJ_1 (Figure 1), OBJ_2 (Figures 2c and 2d) and OBJ_4 (Figure 3). This option is recommended, despite the fact that its technical reliability is not the maximum.

3.2 Step 2. Sensitivity analysis of the design objectives

Next, a sensitivity analysis with respect to the weights is made for the five design objectives (OBJ_1 to OBJ_5). The weighted sum is recalculated varying the value of each weight within a defined feasible region defined by ±0.10 its initial value (with a 0.05 interval) and the options are ranked again. The result of this analysis in a four dimensional space shows the influence of the weighting factors over the selected option.

Thus, for example, a simplified weight sensitivity analysis is made with OBJ_1, OBJ_3 and OBJ_5. The weights for OBJ_2 and OBJ_4 remain constant ($w_2 = w_4 = 0.20$), while the other weights must add up to 0.60 (because the sum of all the weights has to be 1) distributed between OBJ_1 (w_1), OBJ_3 (w_3) and OBJ_5 (w_5). The weighted sum for the three competing options is recalculated to obtain a rank of scores.

As shown in Figure 4, for each pair of objectives (OBJ_1, OBJ_3), the option with the best score is plotted to create a two dimensional response surface (note that OBJ_5 is not represented, as it depends on OBJ_1 and OBJ_3 according to the constraint $w_1 + w_3 + w_5 = 0.60$). This surface represents the variation of the selected option with respect to the relative importance (weight) of the design objectives. From these results we can state that high values of w_1 (minimize environmental impact) clearly favours option A_3. However, as the importance of w_3 increases (maximize technical reliability), the situation is reversed, with option A_1 becoming the best. Finally if w_5 (minimize land occupation) is increased, the selected option depends on the relative importance of w_1 and w_3 as shown in Figure 4.

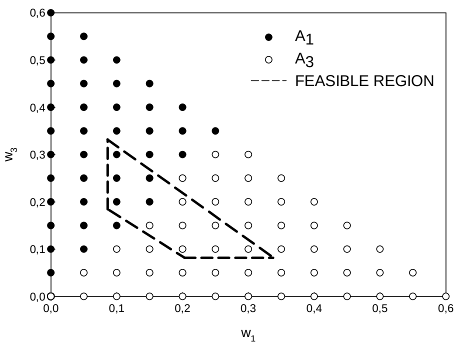


Figure 4. Response surface generated by the simultaneous variation of the weights for OBJ_1, OBJ_3 and OBJ_5

3.3 Step 3. Classification tree induction and extraction of rules

All the data generated in step 2 are processed and a set of rules is extracted using ML techniques (in this case study we use classification trees). A classification tree predicts the value of a discrete dependent variable (in this case the selected option) based on the values of a set of independent variables (in this case the values of the evaluation weights, within the limits fixed previously). The classification tree is generated by the C4.5 algorithm [Quinlan, 1993]. All the classification experiments are carried out using the WEKA software package and the rules are extracted using the PART decision list. PART builds a partial C4.5 decision tree [Quinlan, 1993] in each iteration and the "best" leaf is transformed into a rule [Frank and Witten, 1998]

A set of 12 rules were extracted. These rules were derived from a data set of 341 cases using five continuous variables (w_1 to w_5) that predict wheter the selected option is A_1, A_2 or A_3. These rules classify accurately 97.31% of the cases, which is a very good predictive capability. For illustration purposes some of the extracted rules for 253 of the 341 studied cases are shown:

Rule 1*: IF* $w_4 \leq 0.15$ *and* $w_3 > 0.15$ *THEN the selected option would be* A_1. [70]

Rule 2*: IF* $w_4 > 0.20$ *and* $w_3 \leq 0.25$ *THEN the selected option would be* A_3. [110]

Rule 3: *IF* $w_3 \leq 0.15$ *and* $w_4 > 0.15$ *THEN the selected option would be* A_3. [30/1]

Rule 4: *IF* $w_1 > 0.15$ *THEN the selected option would be* A_3. [14/1]

Rule 5: *IF* $w_1 \leq 0.15$ *and* $w_4 \leq 0.25$ *and* $w_5 > 0.2$ *THEN the selected option would be* A_1. [20]

Rule 6: *IF* $w_2 > 0.15$ *and* $w_3 \leq 0.2$ *THEN the selected option would be* A_3. [7]

The option preferred for the widest range of situations is A_3. The advantage of option A_3 is due to a better satisfaction of OBJ_1 (see rule 2), OBJ_2 (see rule 6) and OBJ_4 (see rule 4) that results in a clear advantage with respect to A_1. In spite of having the lowest risk of rising sludge (Table 1), option A_3 does not adapt well to short and long term perturbations. Moreover, the large bioreactor volumes used in this type of configurations increases land occupation and worsens the accomplishment of OBJ_5.

On the other hand, option A_1 adapts better to short and long term perturbations and requires less land occupation. This is the reason for its comparative advantage in OBJ_3 and OBJ_5 (see rules 1 and 5). Nevertheless, a lower denitrification capacity (main reason of its high values of rising) due to an insufficient anoxic retention time, and higher operation costs (due to large internal recycle flows) and construction costs (due to the pumping system) give a comparative advantage to A_3 in OBJ_1, OBJ_2 and OBJ_3.

4 CONCLUSIONS

This paper has addressed the problem of evaluating design options for complex systems, when several design objectives have to be taken into account. The paper contributes to solve this problem by proposing a systematic procedure that combines multicriteria decision analysis (MCDA) and machine learning (ML) techniques to support environmental engineers during the evaluation of alternatives.

The usefulness of this novel procedure has been tested with a case study where the bioreactor of a wastewater treatment plant is designed to achieve simultaneous carbon and nitrogen elimination. For this case study, three activated sludge configurations are evaluated and the best alternative is selected according to its degree of satisfaction of the design objectives.

Several sensitivity analyses are carried out and several response surfaces are generated and studied in order to identify the dependency of the selected design option with respect to the relative importance of the design objectives. Thus, the strong and weak points of the most promising options can be identified without having to analyse the complex behaviour of the multidimensional response surfaces.

To sum up, the proposed procedure supports the understanding of the overall design space, helping environmental engineers to identify the reasons behind the refusal or acceptance of each option.

5 ACKNOWLEDGEMENTS

The authors wish to thank the Spanish Ministry of Science and Technology for the financial support of project DPI2003-09392-C02-01. The authors also want to acknowledge Dr Ulf Jeppsson (IEA, Lund University) for providing the ASM1 benchmark MatLab/Simulink© code to carry out the simulations.

6 REFERENCES

Belton, V. and Stewart, T.J. Multiple Criteria Decision Analysis: an Integrated Approach. Kluwer Academic Publishers. Dordrecht, 2002.

Comas J, Rodríguez-Roda I., Poch M., Gernaey K., Rosen C. and Jeppsson U. Extension of the IWA/COST simulation benchmark to include expert reasoning for system performance evaluation. 2nd IWA Conference on Instrumentation, (ICA). 2, 907, Busan, Korea. 2005.

Copp, J.B. Respirometry in Control of the Activated Sludge Process: Benchmarking Control Strategies. IWA Scientific and Technical Report No.11. 2003.

EPA - Office of Water Program Operations. 1982. Process Design and Cost Estimating Algorithms for the Computer Assisted Procedure for Design and Evaluation of Wastewater Treatment Systems (CAPDET) Vicksburg, Mississippi 1982.

Flores, X., Bonmatí, A., Poch, M., Bañares-Alcántara, R. and Rodríguez-Roda I. “Selection of the Activated Sludge Configuration during the Conceptual Design of Activated Sludge Plants”. *Ind. Eng. Chem. Res*., 44, 3556 2005.

Frank E., and. Witten I.H. "Generating Accurate Rule Sets Without Global Optimization."In Shavlik, J., ed., Machine Learning: Proceedings of the 15th International Conference, Morgan Kaufmann Publishers. 1998

Henze M., Gujer W., Mino T., and van Loosdrecht M.C.M. Activated Sludge Models ASM1, ASM2, ASM2d and ASM3. IWA Scientific and Technical Report No. 9 IWA. London. 2002.

McGuire, M.L., Jones, K. Maximizing the potential of process engineering databases. *Chem.Eng.Prog*. 85(11), 78-83, 1989.

Quinlan, R. "C4.5: Programs for Machine Learning", Morgan Kaufmann Publishers, San Mateo, CA. 1993

Stephanopoulos, G. Chemical Process Control. An Introduction to Theory to Practice. Penitence Hall International. Inc: New Jersey, 1984.

Takács, I., Patry, G.G. and Nolasco, D. A dynamic model of the clarification thickening process. *Wat. Res*., 25(10), 1263-1271. 1991

Vanrolleghem, P. and Gillot, S. Robustness and Economic Measures as Control Benchmark Performance Criteria. Wat. Sci. Tech., 45(4/5), 117 2002.

Vincke, P. Multicriteria Decision-Aid. John Wiley & Sons. Inc., Chichester, 1992.