SK-DSSy: how to integrate the YouTube platform in WWTP cooperative decision support?

Dario Torregrossa[1], Ulrich Leopold[1], and Joachim Hansen[2]

[1] Luxembourg Institute of Science and Technology (LIST), 41 Rue du Brill, 4422 Sanem, Luxembourg
dario.torregrossa@list.lu;orcid.org/0000-0002-5863-1628
[2] Université du Luxembourg, 6 rue Richard Coudenhove-Kalergi, L-1359 Luxembourg

Abstract. In wastewater treatment plants (WWTPs) domain, the decision support tools are nowadays necessary in order to efficiently process the large databases generated with on-line sensors. In this paper a cooperative decision support system (DSS) is presented: SK-DSSy. This DSS uses a KPI-based fuzzy logic engine to analyse the plant performance and identify the operational conditions that occur in the plants. Then, it associates the dominant operational conditions with YouTube pages in which videos are uploaded to explain the plant condition and propose suggestions. The YouTube platform is then used to share and validate suggestions trough the comment functions and the 'likes'. This approach is innovative, free of costs, and useful for plant managers that can rely in a user-friendly platform.

Key words: decision support system, wastewater treatment, fuzzy logic, YouTube, Shared Knowledge Decision Support System (SK-DSS)

1 Introduction

The energy management of Waste Water Treatment Plants(WWTPs) is generally considered a complex task because of: high amount of information, complexity of phenomena, and multi-parameter optimization requirements [1]. The present paper proposes a decision support system for WWTP energy optimization that combines the fuzzy-logic analysis with the popular YouTube platform (www.youtube.com). This work is an extension of the Shared Knowledge Decision Support System (SK-DSS), the cooperative decision support system presented in [1]. Since this new tool incorporates YouTube, it has been named SK-DSSy in which 'y' stands for YouTube. Before presenting the methodology, in this introduction, the reader can find a short introduction on WWTP energy efficiency issues (subsection 1.1), a short description of the first SK-DSS (subsection 1.2), the novelty and the added value of this new SK-DSSy (subsection 1.3).

1.1 Energy efficiency in WWTP

In the WWTP domain, several researchers have shown an interest for energy optimization because of a high energy efficiency potential associated with a relevant energy efficiency consumption [2]. For example, in [3], the authors investigated the efficiency of WWTPs and estimated a relevant potential energy saving potential (up to 25%). In [4], the authors shown that in NW-Europe, the WWTP are mostly not energy-efficient. The magnitude of WWTP energy consumption is also relevant since, for example, in Europe, it has been estimated in 27 TWh/year [5].

1.2 SK-DSS: a cooperative platform for decision support in WWTPs

One of the opportunities to save energy is the optimization of WWTP management. In order to do this, many authors proposed methodologies to reduce the energy consumption in these facilities. For example, Krampe proposed a benchmark analysis to efficiently identify inefficiencies [6]. Panepinto [7] used remote sensors for the on-line monitoring of electromechanical devices. Poch was one of the first authors to propose an environmental decision support system to optimize the WWTP performance [8]. In line with the Poch's work, recently, we have proposed SK-DSS [1], i.e. a plant generic decision support system that enables the plant operators to cooperate exchanging information. SK-DSS aims to perform analysis on a daily base using the SCADA systems installed in the plants; these systems generally produce a really large amount of information that a human operator cannot satisfactorily use in the decision-making process [1]. SK-DSS analyses the plant performances using a KPI-based fuzzy logic engine and uses the fuzzy scenario analysis to propose solutions. SK-DSS presented several novelties such as its plant-generic profile or the high-frequency, multi-parameter and on-line plant assessment. An extended dissertation about SK-DSS can be found in [1]; for the purposes of this paper, it is necessary to remark a specific characteristic: SK-DSS enables the operators to enlarge the list of case-based suggestions by directly upload solutions in the software databases. Then, SK-DSS automatically processes the information retrieved by the cooperative platform to provide a robust set of solutions. At the date of the first SK-DSS publication, the cooperative platform has not been fully developed yet and the operators had to charge their suggestions in an on-line PostgreSQL database. This solution was effective but not user-friendly, and some issues were still to be fixed such as the evaluation of the proposed solutions. These aspects are improved in this paper.

1.3 Novelty and Added value of this paper

In the present paper, SK-DSS is modified to incorporate the YouTube functions as the cooperative platform. For this reason, this evolution is called SK-DSSy. To the best of our knowledge, a decision support system that couples YouTube and fuzzy logic is new not only in the WWTP domain but also in the decision support domain in general. At the current date (12-Nov-2017), a research with key words 'YouTube' AND 'Fuzzy Logic' shows only three records in Scopus database; these papers are not connected to decision support. No records are avaible with a Scopus research with this key-word set: 'YouTube','Fuzzy Logic','Decision Support'. Consequently, we are confident that this approach is fully original in the decision support domain.

SK-DSSy has added values when compared to his first version:

- it is user-friendly, because the YouTube platform is one of the most used over the world;
- the plant managers have not to deal with PostgreSQL to insert their suggestions;
- the knowledge is shared in a video-format; consequently video-description could be extensively adequate;
- the YouTube 'like' system can be used to give a score to the suggestions;
- the YouTube video-suggestions can be commented;
- it is possible to access the suggestions from each device connected to internet (smartphone, tablet, pc);

Moreover, SK-DSSy inherits all the characteristic of the first version: i) it is specifically oriented to energy saving, ii) it is plant generic, iii)it is able to produce daily analysis reports, iv) it provides case-based suggestions.

2 Data and Methodology

In order to provide the reader with a practical example, this paper proposes a decision support methodology to optimize the energy consumption of blower system installed in a WWTP biological stage. This is a relevant topic because the blower energy consumption can cover up to 60% of the global energy consumption of WWTPs.

2.1 Data description

The data used have been generated using the software SIMBA# (please refer to www.inctrl.ca/software/simba/). This software was used to simulate a WWTP designed for 120k population equivalent, equipped with blowers, pump and biogas generation stage. In order to show many operational conditions, the amount of air provided by the blower changes during the simulation process; according to the simulation model, these changes affect the pollution removal performance and the energy consumption.

2.2 KPI calculation

The software is used to simulate a 200-days time-series of several parameters such as: wastewater inflow $[m^3/day]$, COD concentration at the inlet $[mg/l]$, blower energy consumption $[kWh/day]$, dissolved oxygen in biological stage

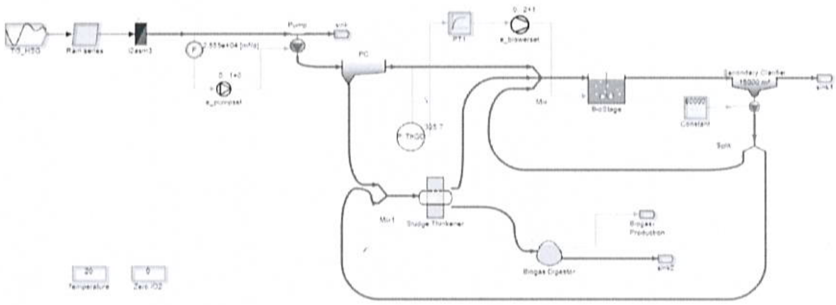

Fig. 1. Simba# simulator: screenshot

$[mg/l]$, COD concentration at the outlet of WWTP $[mg/l]$. For each day and for parameter, the software generates 130 values. As described in [5], these values can be aggregated in order to generate the following KPIs: population equivalent (PE), specific blower energy consumption $[kWh/year/PE]$. Moreover, as explained in [1], the SK-DSSy dynamically calculates the benchmarks taking into account the operational conditions: blower energy consumption benchmark $[kWh/year/PE]$, dissolved oxygen benchmark [fixed equal to 4 mg/l but customizable].
SK-DSS fed the fuzzy logic engine with KPIs and benchmarks, in order to be able to simultaneously work with several WWTPs [1]. Now, SK-DSSy uses benchmarks and KPIs to calculate index that are processed by the fuzzy logic engine. In particular SK-DSSy calculates the following indices:

- BloEA_Index; equal to the ratio between the specific blower energy consumption and the blower energy consumption benchmark.
- BioDO_Index; equal to the ration between dissolved oxygen in biological stage and the dissolved oxygen benchmark.

When the index values are higher than 1, the observed values of the parameter are higher than the benchmark. When the index values are lower than 1, the observed values of the parameter are lower than the benchmark. For example, when $BloEA\_Index = 1$ the blower energy consumption corresponds to the expectations; when $BloEAIndex > 1$ the blower energy consumption is larger than a normal value; when $BloEA\_Index < 1$ the energy consumption is lower than its reference value. This changes has been done in order to improve the SK-DSSy usability; here the parameters are normalized values really easy to understand.

In conclusion at the end of the process, SK-DSSy produces the set of information necessary to feed the fuzzy-logic engine.

2.3 A plant generic fuzzy-logic for scenario analysis

The hearth of SK-DSSy is the fuzzy logic engine. The reader can find an extensive mathematical explanation of fuzzy logic methodology in [10, 11]. In [1, 2], the reader will find two fuzzy-logic applications in WWTP domain, alongside the advantages to use this methodology in environmental domain. In particular, the fuzzy logic well performs with databases affected by uncertainty and with high-complex system.

Table 1 shows the fuzzy rules used by SK-DSSy. These rules are obtained by processing authors' knowledge and the available literature. Each rules describes an operational scenario. For each day, the fuzzy logic engine uses these rules to analyse the indices and provide 2 outputs:

- a overall performance score in a range 0-100; higher the score value, better the plant performance;
- a truth degree (TD) for each rule in the range 0-1; the TD expresses the likelihood that in the WWTP the scenario described by the rule is verified. If, for a rule, TD=1, then the rule describes accurately the phenomena. When TD=0, the condition described by the rule is not verified in the WWTP.

The sum of all the truth degree for each day is equal to 1. The truth degrees of all the rules are used to identify the operational condition for given day. For example, if for one day, rule 2 has a TD=0.45, rule 3 TD=0.4, the sum of the remaining rule TD=0.15; in this case, the energy consumption is low and the oxygen concentration between medium and high. A potential suggestion, in this case could be to reduce a bit the air inflow to reduce energy. When the analysis is performed over a day, the interpretation of truth degree is really clear: by set-up, it is not possible that that a parameter is 'low' and 'high' at the same time and consequently, in the worst case, there are 2 coherent dominant rules.

The same kind of reasoning could be done by analysing the average truth degree of the rules over a period. In this case, this value should be interpreted as the expected values and the logic-coherence cannot be anymore imposed : in a period, for example, it is possible that for some days the energy consumption is low and for other days that it is high. If the analysis of a period produces two or more conflicting dominant rules, it is suggested to reduce the interval in order to have a more detailed assessment of operational conditions.

In order to use the rules, SK-DSSy needs to attribute the appropriate adjective to each parameter. For example, SK-DSSy need to know if (or better 'how much') it is true that BloEA_Index is Low. This operation is called fuzzification [10, 11] and it is performed trough fuzzification functions (fig. 2).

2.4 From scenario analysis to decision support: using the youtube platform

At this point, SK-DSSy explained the operational conditions occurring in the plant. This result can be coupled with a list of potential case-based solutions.

Table 1. Fuzzy logic rules

Rule	BloEA_Index	BioDO_Index	Score
1	Low	Low	30
2	Low	Medium	100
3	Low	High	80
4	Medium	Low	30
5	Medium	Medium	50
6	Medium	High	70
7	High	Low	0
8	High	Medium	20
9	High	High	40

The first column identifies the rule number. The columns 2, 3 and 4 identify the inputs. The last column identify score associated with the scenario. This table has to be translated in fuzzy logic language; for example, the rule 1 becomes:

IF BloEA_Index IS Low AND BioDO_Index IS Low THEN Score is 30

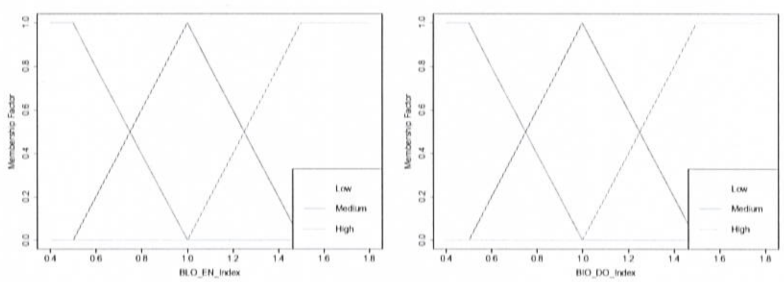

Fig. 2. Input membership functions

In SK-DSS [1], this was done with a PostgreSQL table that plant managers can use to visualize the results and upload new solutions.

The main limitation of this approach was that plant managers should be able to use PostgreSQL queries.

A not-addressed question concerned the evaluation of the solutions proposed by the common platform; according to the philosophy of this cooperative platform, each plant operator is authorized to upload solutions. This open-access approach does not guarantee the quality of proposed solutions. The quality of the end-user contribution could be decreased by several issues, for example: a limited comprehension of the variables, the upload of plant-specific solutions or the limited experience of the end users.

SK-DSSy proposes a solution to these issues by incorporating YouTube in the cooperative platform. In SK-DSSy, the fuzzy logic rules are connected to YouTube web-pages in which the solutions can be uploaded as video or as comment. SK-DSSy identifies the dominant scenario and lead the end-user to the right video-page. For each rule, the main video was inserted by the page administrator; it explains the operational scenario connected to the rule, some basic solutions and, in order be time efficient, it lasts less than 1 minute. The end-users can visualize the videos, comment it with a text, add videos or link to external resources (such as papers or other web-pages). Another important YouTube function is the 'like' command. Each user can mark with a 'like' the useful suggestions and with a 'dislike' the comments considered not useful. The comments are consequently scored and it is possible to sort them by popularity.

The advantages of this approach are: i) the knowledge can be shared in many formats (such as video, text or links), ii) the use of YouTube platform is extremely user-friendly, iii)the server and the maintenance of a part of SK-DSSy is externalized to Google-YouTube services without costs. iv) possible assessment of the proposed solutions.

3 Results

Table 2 reports a statistical summary of the parameters connected to blower energy consumption and, the dissolved oxygen concentration. In the period under investigation (200 days), the energy consumption of the blowers as well the dissolved oxygen values vary in a large range and comprise desirable and undesirable conditions.

Table 2. Summary of KPIs

	Spec. Blower En. cons. [kWh/year/PE]	**Blower En. Index [-]**	**DO [mg/l]**
Min	0.99	0.06	0.02
1st Quartile	7.51	0.49	2.76
Median	9.40	0.61	3.80
Mean	9.39	0.61	3.83
3rd Quartile	12.31	0.80	5.99
Maximum	17.98	1.17	6.17

The fuzzy logic engine, for each day, produces an overall performance score (fig. 3). For values higher than 80/100, the plant performance can be considered good, while a score value below 40/100 shows low performances of the WWTPs.

According to the SK-DSS methodology [1], it is possible to analyse single days or periods. In fig. 4, the analysis of 4 period is provided. During the first 50 days, the global index is quite high because the rule 2 is largely dominant. In the period between day 50 and day 100, the global index is in a normal range. Rules 2 and 6 show that the energy consumption is in a normal range, but it is still possible to save energy because in several day the oxygen concentration in the bio-stage can be reduced. In the third period, between day 100 and day 150, the global score is critical because we have simulated a failure of blower system. In this case, a wrong set-up of the system is detected by Rule1 that explains

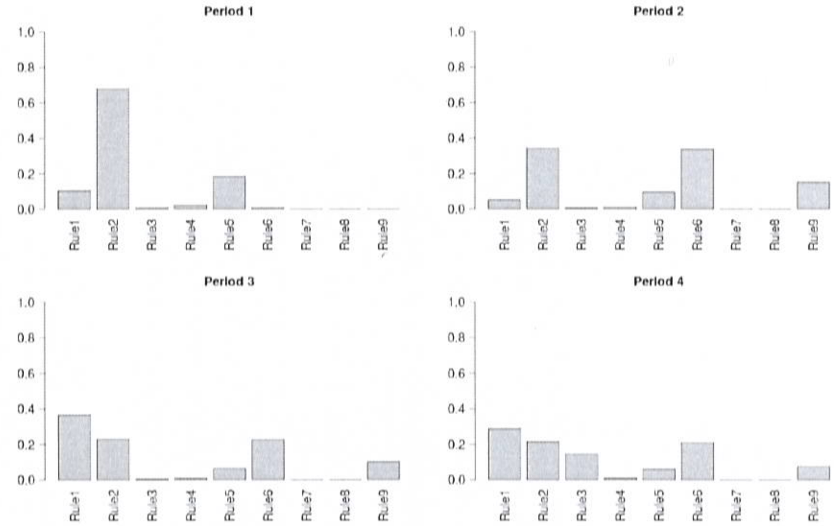

Fig. 4. Scenario analysis for many periods; Period 1, between day 0 and day 50; Period 2,between day 50 and day 100; Period 3,between day 100 and day 150; Period 4,between day 150 and day 200

Table 3. Link to Solution

Rule	Link
rule1	Link to rule store 1
rule2	Link to rule store 2
rule3	Link to rule store 3
rule4	Link to rule store 4
rule5	Link to rule store 5
rule6	Link to rule store 6
rule7	Link to rule store 7
rule8	Link to rule store 8
rule9	Link to rule store 9

I do not understand that? Necessary!?

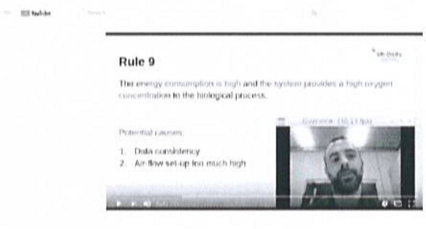

Fig. 5. Screen-shot of intro-video

in the answer shown by fig. 6, a link to an external book is provided. The 'like' system is used to score the answer and the contributions: the contributions with more 'like' can be sorted and visualized on the top of the page.

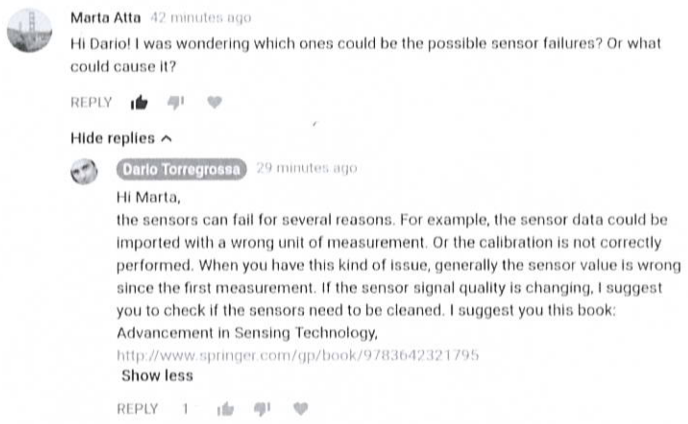

Fig. 6. Screen-shot of cooperative use of YouTube platform for decision support

In summary, with SK-DSSy, the plant operators have a decision-support system able to perform the on-line analysis of WWTPs and to share knowledge. The contributions of the network can be voted and ranked. Moreover, the sharing-knowledge platform relies on YouTube platform that is well-know and generally considered user-friendly.

4 Discussion

This paper presents an extension of the decision support system presented in [1]. The decision tools are important for plant operators that are required to deal with large databases (almost 300k records/day). The cooperative decision support tools enable the plant operators to cooperate and consequently the set of proposed solution is robust, self-updating and peer-reviewed. To the best of our knowledge, in water domain, cooperative DSSs represent a novelty. According to the Scopus research

4.1 Field of application of this methodology

This methodology has been applied in WWTP domain, but it is suitable to be applied in each domain in which it is necessary to analyse a large amount of data for decision support.

4.2 Advantages of the methodology

This methodology has several advantages:

- it enables the on-line analysis of complex systems; in this case, a demonstration has be performed on wastewater treatment plants;
- it provides a robust and self-updating set of solutions by enabling the cooperation between experts;
- it uses YouTube, i.e. an user-friendly, cost-free, popular, and flexible platform;
- the maintenance of this cooperative platform is externalized to YouTube service providers;
- the fuzzy logic input is expressed as an index set in order to make the comprehension really easy for the end user

4.3 Limitations of the methodology and potential new developments

Despite the advantages exposed before, the use of YouTube platform has a big limitations: it is not possible to know and/or influence the evolution of the platform during the next years. This could be a problem if YouTube administration remove functionalities (such as the comments). Nevertheless, given the popularity of YouTube, we expect that the service will be active at least for several years.

Another potential limitation is the absence of a score to motivate the end-users to cooperate. In the scientific domain, the impact factor is considered important for the profile of authors and it is also a stimulus to improve the quality and the quantity of contributions. A similar mechanism should be implemented to stimulate the SK-DSSy end users to cooperate. This could have a relevant impact above all if SK-DSSy is used by professional networks able to provide impulse for cooperation. For example, the authors of this paper are in contact with Croon and Vereniging van Zuiveringsbeheerders, the Dutch association of WWTP managers, that developed the Dutch national WWTP database that collects on-line information from more than 300 WWTPs; SK-DSSy could easily be applied to analyse these WWTPs and enable the cooperation inside the professional community also.

These two limitations will be overcome with a new YouTube like cooperative platform, that this group of research will develop in the next years.

5 Conclusion

In the wastewater treatment plant domain, decision support tools are required to efficiently manage the large amount of data produced by on-line sensors. In this paper, a you-tube based cooperative environmental decision support system has been presented. This cooperative decision support system combines the fuzzy logic analysis with the use of YouTube platform; this approach enables an efficient cooperation between plant operators and the evaluation of the proposed suggestions. Moreover, a research on Scopus database shows that this approach is extremely innovative in the water as well as ~~and~~ in the decision support domain.

Acknowledgements

The authors gratefully acknowledge the financial support of the National Research Fund (FNR) in Luxembourg (7871388-EdWARDS). The methodology, the script and the results presented in this paper are part of the PhD project of Dario Torregrossa (EdWARDs project). The authors thanks Marta Antonelli (LIST) for her support in testing the platform.

References

1. Torregrossa D, Hernndez-Sancho F, Hansen J, Cornelissen A, Popov T, Schutz G. Energy saving in wastewater treatment plants: A plant-generic cooperative decision support system. J Clean Prod 2017;167. doi:10.1016/j.jclepro.2017.08.181.
2. Torregrossa D, Hansen J, Hernndez-Sancho F, Cornelissen A, Schutz G, Leopold U. A data-driven methodology to support pump performance analysis and energy efficiency optimization in Waste Water Treatment Plants. Appl Energy 2017:111. doi:10.1016/j.apenergy.2017.09.012.
3. Castellet L, Molinos-Senante M. Efficiency assessment of wastewater treatment plants: A data envelopment analysis approach integrating technical, economic, and environmental issues. J Environ Manage 2016;167:1606. doi:10.1016/j.jenvman.2015.11.037.
4. Becker M, Hansen J. Is the energy-independency already state-of-art at NW-European wastewater treatment plants. Proc IWA-Conference Asset Manag Enhancing Energy Effic Water Wastewater Syst 2013:1926.
5. Torregrossa D, Schutz G, Cornelissen A, Hernndez-Sancho F, Hansen J. Energy saving in WWTP: Daily benchmarking under uncertainty and data availability limitations. Environ Res 2016;148:3307. doi:10.1016/j.envres.2016.04.010.
6. Krampe J. Energy benchmarking of South Australian WWTPs. Water Sci Technol 2013;67:205966. doi:10.2166/wst.2013.090.
7. Panepinto D, Fiore S, Zappone M, Genon G, Meucci L. Evaluation of the energy efficiency of a large wastewater treatment plant in Italy. Appl Energy 2016;161:40411. doi:10.1016/j.apenergy.2015.10.027.
8. Poch M, Comas J, Rodrguez-Roda I, Snchez-Marr M, Corts U. Designing and building real environmental decision support systems. Environ Model Softw 2004;19:85773. doi:10.1016/j.envsoft.2003.03.007.
9. Shi CY. Mass Flow and Energy Efficiency of Municipal Wastewater Treatment Plants. IWA Publishing; 2011.
10. Zadeh L a. Fuzzy sets. Inf Control 1965;8:33853. doi:10.1016/S0019-9958(65)90241-X.
11. Starczewski JT. Advanced concepts in fuzzy logic and systems with membership uncertainty. vol. 284. 2013. doi:10.1007/978-3-642-29520-1.

Version revised on: Monday 16th October, 2017- 12:29