

## Performance assessment of urban utilities: the case of water supply, wastewater and solid waste

H. Alegre, E. Cabrera Jr and W. Merkel

### ABSTRACT

The importance of performance assessment in the water industry is demonstrated by the increasing number of initiatives taking place around the world. Since the first initiatives, which started 15 years ago, the topic has evolved significantly although it still presents challenges to be tackled.

Five years ago, a project on performance assessment of urban infrastructure services was submitted to the European Union under the COST programme. As a consequence, the COST C18 Action was created. This paper presents the main conclusions and future key research areas from the project. In the solid waste sector, existing practical experience is less advanced, but methodologies and conclusions will be equally applicable in performance assessment projects.

**Key words** | COST Action C18, infrastructure asset management, performance assessment, solid waste, wastewater, water supply

#### H. Alegre

LNEC—National Laboratory of Civil Engineering,  
Urban Water Division,  
Avenida Brasil 101, PL-1700-066, Lisboa,  
Portugal

#### E. Cabrera Jr (corresponding author)

Universidad Politecnica de Valencia,  
Instituto Tecnológico del Agua,  
Apdo 22012, E-46071, Valencia,  
Spain  
E-mail: qcabrera@ita.upv.es

#### W. Merkel

IWW Water Centre,  
Moritzstrasse 26, D-45476, Mulheim an der Ruhr,  
Germany

### INTRODUCTION

Performance assessment has been one of the hottest topics in the water industry for the past decade. The use of performance indicators and benchmarking techniques has become a common practice during this period. There are many good reasons behind this success. Water services are provided in a monopolistic environment, and in the absence of market forces it is hard to find motivation for efficiency. All the stakeholders in the business have come to realise that, by assessing the performance of the services in a systematic way, utilities are driven to continually improve their performance, with the consequential benefits for all those involved.

The motivation to assess performance in other urban infrastructure services is similar, although practice shows that development is a step behind. For instance, the case of solid waste could be quite similar to water services. As a matter of fact, all the good reasons suggesting the need to assess the performance of drinking water and wastewater services also exist in the case of urban solid waste services. However, the truth is, regardless of the reasons, that the initiatives in this area are much scarcer.

Five years ago, a project on performance assessment of urban infrastructure services was submitted to the European Union under the COST programme. Founded in 1971, COST is an intergovernmental framework for European Co-operation in the field of Scientific and Technical Research, allowing the co-ordination of nationally funded research on a European level. As a consequence, the COST C18 Action was created, having run between March 2004 and March 2008.

COST C18 had, as its objectives, to increase the knowledge and to promote the use of effective, robust and well-devised methodologies for decision-making for urban infrastructure services based on performance indicators.

The Action was also created to demonstrate that the principles and concepts of the approach developed for the water sector can be easily adopted in other urban infrastructure services. The case of solid waste was chosen to explore the synergies of considering these three types of urban infrastructures together, in the belief that the methodologies and conclusions would be easily translatable. This compatibility proved to be true

(Sjøvold *et al.* 2008) and, as a matter of fact, the conclusions resulting from the C18 Action are applicable to water and solid waste services alike, unless stated otherwise.

COST C18 proved to be a wonderful vehicle to discuss key topics on performance assessment. The meetings allowed for the gathering of experts from around the world. This paper summarises the knowledge, ideas and proposals that resulted from such work.

### THE COST C18 ACTION

The COST Action C18 spanned four years, bringing together over 50 experts from 19 European countries. A total of 10 meetings were held, including 5 technical seminars with 33 presentations that can be freely downloaded from the Action's website (<http://www.costc18.org>) (see Figure 1). Seven short-term scientific missions of young professionals contributed to networking the involved research organisations. The final conference of the Action, Pi08, was attended by over 100 delegates from 29 countries. The presentations of the 44 papers submitted can also be downloaded from the project's website.

The main deliverables from the action were:

- a website, with all relevant Action materials publicly available,
- international conference Pi08,
- Pi08 Proceedings, published by IWA Publishing (Cabrera & Pardo 2008) (Figure 2),
- final scientific report, including a case study compilation (Sjøvold *et al.* 2008),
- discussion and survey on research priorities,
- proposal to be submitted to the 7th Framework Programme of the European Union,
- executive summary document,
- dissemination papers.

### THE ROAD LEADING TO COST ACTION C18

#### Performance assessment in the 1990s

The major pioneering developments in the scope of performance assessment of urban infrastructure services started in the water supply and wastewater sectors. The first steps were undertaken by the initiative of regulating bodies<sup>1</sup>: however, succeeding initiatives were equally observed in

<sup>1</sup> To be precise, the Water Companies Association in the UK organised a conference on the topic in the early 1980s, but the proceedings remained private.



Figure 1 | Home page of the COST C18 website (<http://www.costc18.org>).

non-regulated environments. The privatisation process carried out in the United Kingdom in 1989, the creation of a powerful economic regulator, the Office of Water Services (OFWAT) and the establishment by OFWAT of a direct relationship between tariff revision and the levels of service provided to the customers had a major leverage effect in this process. The levels of service set up by OFWAT in the early 1990s had some relevant characteristics, mostly still valid in the current version of this assessment system:

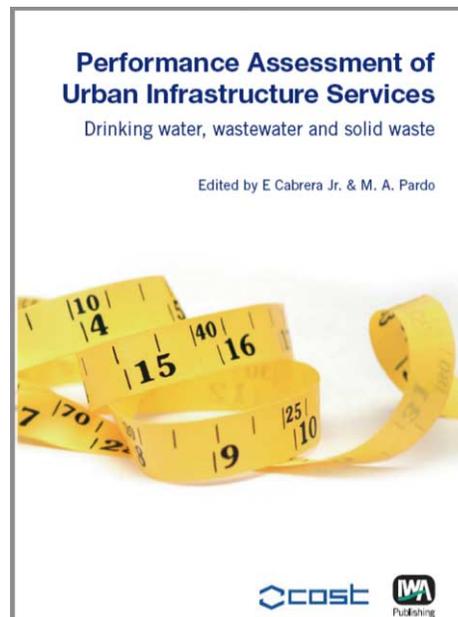


Figure 2 | Cover of the Pi08 Proceedings book (COST C18 final conference).

- a small number of measures were adopted,
- the emphasis of the assessment was put on the quality of the service directly perceived by the customers,
- the accuracy of the data used to assess the indicators had to be reported and audited,
- each indicator had a clearly defined assessment procedure,
- the results were made public.

The OFWAT system had—and still has—a major international impact. It is a key reference and source of inspiration.

Step by step, other initiatives started to pop-up worldwide during the 1990s. The World Bank, for instance, started to develop indicators to monitor project design and implementation (Janssens *et al.* 1994; Yeppes & Dianderas 1996, 1997). In 1992, the 13 larger Dutch water companies created a “contact club for water companies”. These companies agreed to define indicators to assess their own performance and compare the results achieved once a year. There were no goals set or yearly targets. Their motivation was to interpret the differences and draw conclusions for each company’s own benefit (van der Willigen 1997). The six-cities Group of the Nordic Countries (Stockholm, Gottebourg, Malmö, Copenhagen, Oslo and Helsinki) was also launched in this period, with similar goals and motivations (Adamsson 1997). Other relevant benchmarking processes started to be developed and implemented in other regions of the world, such as in Switzerland (Skarda 1997), Portugal (Faria & Alegre 1996), Malaysia (Malaysian Water Association 1996), North America (involving the US and Canada) and Africa (e.g. a process conducted by Umgeni Water, from South Africa).

### The IWA PI systems

In the early 1990s, the International Water Supply Association (IWSA) selected the topic “Performance Indicators” for one of its world congresses. No abstracts were submitted on this topic and it had to be cancelled. The subject did not seem to raise much interest. However, just three or four years later, the response to an inquiry sent by IWSA to about 150 senior members of water utilities from all over the world clearly showed that “performance indicators” and “unaccounted-for water” were by far the two topics of greatest interest in the scope of the water transmission and distribution systems. In 1997 a task force on “performance indicators” was established, and its final

output, the IWA PI systems for water supply services (Alegre *et al.* 2000, 2006) and for wastewater services (Matos *et al.* 2003) are likely to be the most widely used references in their field today.

Among many other applications, these systems were used for the establishment of the World Bank IBNET PI system, the regulatory PI system of ADERASA, South America, the framework for voluntary benchmarking of water supply in Germany, the quality of service regulatory system established in Portugal, the Japanese PI system of the Japanese Water Works Association and the water losses PI of the American Water Works Association. The international standards ISO 24500 (ISO 2007a,b,c), recently launched, also adopted the general principles recommended by IWA to set up and implement a PI system.

### Recent evolution and general trends

A significant evolution in methods, algorithms, tools and applications has taken place during the last 15 years of intense activity. In this time, systematic approaches for the design and implementation of performance indicator systems have been developed. Academics and researchers have also formulated alternative theories and more sophisticated algorithms to support the analysis of indicators and metric benchmarking. And, in parallel to the development of the tools, the number of experiences has grown considerably as associations and regulators have promoted the use of PI in different parts of the world.

This trend has been followed by the industry. During these years, many utilities have started to assess their performance in a systematic way, by either using a pre-defined PI system or developing their own. And in such effort, it is remarkable to see how utilities have followed a similar evolution in different parts of the world.

Typically, utilities or associations start with the collection of statistical data. They recognise the importance of assessing the utilities’ performance, but the objectives of these exercises are not always clear, and the impact of the results in terms of service improvements tends to be limited.

In a more mature stage of development, utilities implement performance assessment systems as part of their management tools. In general, performance assessment is a part of broader approaches, such as quality

management programmes, risk management, infrastructure asset management, etc. Comparisons with the outside world are considered important, to allow the organisation to determine its position related to its peers.

It is then when regional or national metric benchmarking projects of some sort are undertaken, and results usually show managers that other utilities perform better in certain business areas. Competitiveness appears and the comparisons become periodic to determine the evolution from one year to the next. Experience shows that, even under merely monopolistic circumstances, benchmarking provides sufficient market forces to push a general development of a national sector. This process is often pushed by some leading companies, defending their leading position. Only an open question remains: how to deal with those companies left behind and not capable or willing to improve their performance.

Once utilities realise (at the national or international level) that someone is doing better in some areas, they feel the need to understand the reasons behind that fact (so they can improve and reach at least a similar level). This leads to process benchmarking initiatives which, depending on their origin, can range from a local club to large international projects encompassing utilities from several continents. To this date, process benchmarking seems to be satisfying participating utilities, although once a project is successfully completed, the topic is usually changed for the next periodic project.

During the course of these 15 years of experience, there are typical misunderstandings that may be encountered again and again in practical applications of performance indicators. It is important to be aware of them in order to minimise repetitions. These include:

- *Lack of engagement of the organisation's CEO.* Performance Indicators systems are useless if data are not reliable or if results achieved are not used to support improvement measures within the organisation. Any of these key success factors, which are inter-dependent, require a true engagement of the top managers of the organisation the indicators refer to. Data requested by an external body with no practical consequences for the organisation do not motivate corporate decision-makers. Additionally, isolated initiatives within the utility, led by any enthusiastic employee who found performance assessment a fashionable issue, are also doomed to fail.
- *Incorrect selection procedure.* The procedure recommended by IWA and in the ISO 24500 standards for the implementation of a PI system starts with the definition of objectives, followed by the establishment of assessment criteria and only then by the selection/definition of performance measures (set up according to these objectives and criteria). However, practice shows that this procedure is not often applied. For instance, many organisations willing to start using PI systems do so by seeing how many of the IWA indicators they would be able to calculate. The variables for which there are available data are selected and the related indicators are calculated. The result is an incoherent set of measures, almost useless.
- *Temptation of going from zero to a "PI heaven".* When an organisation starts to select and implement performance indicators, there is often the temptation that every aspect of the management should be covered. It is fundamental to ensure that a balanced solution is found, and that the number of indicators is kept as small as possible. In other words, that the effort required for data collection, validation, archiving and processing is cost-beneficial and that all the indicators selected are relevant and effectively used to support management decisions.
- *Temptation to reinvent the wheel.* People are imaginative by nature, and like to be original and provide a personal flavour to what they produce. Performance assessment is a field where this has proven to be the rule. Many organisations feel that they are unique and therefore will need to develop their own performance measures and establish their own systems, disregarding other sets of indicators. This is partially positive and understandable. Systems such as the IWA PI proposal resulted from the contribution of many experienced people worldwide and were tested in a good number of different situations. It is important to take the benefit of these proposals when possible, bearing in mind the specific environment of the organisation. The use of existing PI systems recognised as international references has the obvious added advantage of allowing comparisons with other organisations adopting the same platform.
- *Misuse of concepts.* The IWA PI, the technical committee who prepared the ISO 24500 and, last but not least, COST C18 worked hard in clarifying what is meant by performance measures (indicators, indices or levels),

variables, explanatory factors, context information, etc. These basic notions are well documented and easily available. However, misuse of concepts is still a very typical mistake found in performance indicator reports.

- *Only best results welcome.* Utility leaders are human individuals and tend to easily accept good results (even without sufficient proof), while failing to adequately react to low performance. A common reaction is to invest a lot of effort to justify bad results instead of concentrating on the analysis of potential problems and countermeasures for improvement. A fault-positive culture in the utility is crucial for accepting bad performance results as a chance for doing better in future.
- *High short-term expectations.* Measuring the company or the sector performance will not provide automatically improved performance. A number of utility processes (e.g. investment processes or renewal strategies) need to be evaluated over some years to obtain sound and reliable data. Improvement measures often need some time to make an impact on the performance figures. Benchmarking is, by definition, a continuous process for which effectiveness cannot be evaluated after one period. Nevertheless, there are many examples of immediate benefits to performance measurements, presumably due to the fact that introducing performance thinking in a company automatically drives decisions towards higher efficiency.

The evolution described in the preceding paragraphs has been witnessed in several parts of the world. Public and private, small and large utilities in regulated sectors or non-regulated environments have all followed in different places, at different times, a similar path.

As an evolution of these separate approaches, with the objective to overcome specific shortcomings, some integrated approaches have been developed, such as the AWWA and WEF Qualserve program, the Scandinavian six-cities group and the WSAA National Performance Reports in Australia (formerly WSAAfacts).

These efforts are typically centred around a PI-based benchmarking system, including high-level PIs and more detailed process evaluation in dedicated workshops, but also include internal and external evaluation systems providing feedback and views from outside the utility. In order to add the customers' viewpoints, several utilities or national water

associations have also enlarged their benchmarking initiatives with focused customer and satisfaction polls: for instance, those performed by Water UK on the 2004 water price review, the VEWIN customer survey embedded in the Dutch Water Industry (Vewin 2006) or the annual customer survey of the German water industry by BDEW.

This constant change in needs and perspective requires similar developments in the tools used and their sophistication; developments which, in some cases, are still to come and constitute the current challenges of performance assessment.

#### Relevant milestones in the application of performance indicators

- **1989**—OFWAT is established as the economic regulator of the water services in England and Wales. A system of performance indicators, including data quality assessment and auditing, is established.
- **1995**—The cities of Copenhagen, Oslo, Helsinki, Stockholm, Gothenburg and Malmo decide to start a comparison of their performance by using indicators. The project has evolved through the years into new forms of benchmarking.
- **1998**—The Water Services Association of Australia (WSAA) starts publishing WSAA Facts (currently National Performance Reports) including audited data from the major urban utilities in the country. WSAA starts running international process benchmarking efforts in 2004.
- **2001**—A large metric benchmarking project in context with the IWA field-test establishes the concept of voluntary performance assessment in the German water supply system. A first major application of the IWA indicator system was a project in Bavaria (Germany) with almost 100 small and medium utilities.
- **2004**—Portuguese utilities begin reporting to the national regulator IRAR using a performance indicators' system based in the IWA proposal. The assessment does not include mathematical modelling and indicators are not only focused on the user perception but include assessment of the operators and of the environmental sustainability.

## AVAILABLE TOOLS AND APPROACHES

### Introduction

Performance indicators are the basis of most performance assessment systems. However, an increasing variety of methods and tools are available to support the interpretation and decision-making processes subsequent to the indicators' assessment.

The following subsections present an overview of these methods and tools that may be used in the scope of the utilities' self-assessment, of metric benchmarking and of process benchmarking. They do not intend to be exhaustive, but rather focus on the main alternatives available.

### Methods and tools for utilities' self-assessment

Self-assessment is the basic step of assessment. It refers to a specific system or utility. Interpretation is exclusively based on the comparison between pre-established targets and actual results and trend analyses over time. The main methods to assist in the interpretation and decision-making are the direct observation of the elements of the performance indicator system used and the use of indices.

#### Direct observation of performance indicator system

A basic procedure to interpret performance indicators is the direct observation and analysis of results. This includes not only the analysis of the PI results, comparing them with targets and with previous results, but also the observation of the variables used to assess them and their respective accuracy. The analysis is carried out with the support of tabular and graphical representations (e.g. Tables containing the PI values and the targets, time series graphs, spatial representations, spider graphs, etc.). This type of interpretation procedure may be used in the scope of management systems such as, for instance, implementations of balance score cards and Total Quality Management Systems (TQMS).

Advantages:

- Ease of representation.
- Transparency of results.
- Allow decision-makers to incorporate their experience in the interpretation process.

Shortcomings:

- When performance systems adopted are complex, the global performance of the service is not easy to perceive.

- Given that decisions for improvement should depend on sets of indicators, and not on individual detached ones, decision-making processes may become complex.
- As this direct interpretation includes the user experience, this may introduce bias into the analysis.

#### Performance indices

Performance indices are measures:

- resulting from the combination of more disaggregated performance measures (e.g. weighed average of performance indicators), aiming at aggregating several perspectives into in a single measure. An example of a well-known index is the Total Factor Productivity (TFP), which is the ratio of all output quantities (weighted by revenue shares) and all input quantities (weighted by cost shares).

Advantages:

- If properly defined, they are easy to understand and adequate to communicate results between stakeholders.
- As they synthesise sets of individual measures into single measures, they are easier to incorporate in the decision-making processes.

Shortcomings:

- Weight definition determines the result and therefore this is a potential source of bias.
  - Being synthetic, they lose information, preventing detailed analyses.
  - Their accuracy is not easily assessed and is not directly perceivable.
- derived from analysis tools (e.g. simulation models, statistical tools, cost efficiency methods).

Advantages:

- Can be used to assess future scenarios.
- Allow for the comparison of virtual scenarios.

Shortcomings:

- They are more subjective and less auditable than PI, sometimes turning into black boxes for the users.
- Their assessment may be complex.

### Frontier methods

Frontier methods are based on the concept that, in a given group, all companies should be able to operate at an optimal efficiency level which is determined by the most efficient companies in that group. These efficient companies are usually referred as the “peer firms” and determine the “efficiency frontier” (Eggen *et al.* 2001).

Advantages:

- Produce very synthetic results, allowing one to rank utilities.
- Produce an efficiency objective to the participating utilities.

Shortcomings:

- Require the definition of the input and output parameters, which can easily introduce bias to the analysis.
- If the users are not familiar with the method and its underlying assumptions, they do not accept and incorporate them easily into the decision-making processes.
- Statistical methods require good data quality, which is not often found in the water industry. Furthermore, once the data are processed, no information on the original data quality is found in the results.

### Data envelopment analysis

Data Envelopment Analysis (DEA) is the most-used non-parametric approach. Non-parametric methods do not impose any functional form of relationships between input and outputs. In DEA, the frontier is made up by linear combinations of the best performing companies in the sample (Eggen *et al.* 2001).

Advantages:

- Produce very synthetic results, allowing one to rank utilities.
- Do not require the specification of input and output parameters.
- Produce an efficiency objective for the participating utilities.

Shortcomings:

- A limited selection of the performance measures is likely to bias the output of the method.
- If the users are not familiar with the method and its underlying assumptions, they may not accept the

results and incorporate them easily into the decision-making processes.

- Statistical methods require good data quality, which is not often found in the water industry. Furthermore, once the data are processed, no information on the original data quality is found in the results.

### Metric benchmarking

Interpretation of assessment results carried out in the scope of metric benchmarking involves, by definition, the comparison between peers. These comparisons add a new requirement to the self-assessment: the differences in context that significantly affect performance results need to be taken into account.

As in the self-assessment applications, direct observation of performance indicator system is very important for metric benchmarking. The main difference is that it must also incorporate the analysis of the context. The advantages and the shortcomings are the same, added by the potential bias caused by the way differences in context are dealt with.

The use of performance indices is also relevant to assist metric benchmarking. As referred to, there are two main types of indices, but the aggregated measures are typically the type used for metric benchmarking. The advantages and shortcomings previously referred to remain valid. Another advantage is the ease of ranking the utilities based on the values of the aggregated indices.

There are other methods available. The main ones are the frontier methods and the data envelopment analysis.

### Process benchmarking

Process, or Xerox benchmarking, is different in nature to metric benchmarking. While the output of metric benchmarking is merely information (a regulator is able to differentiate better and worse performers, a utility manager discovers in which areas utility performance can improve compared to one of its peers), process benchmarking is all about improvement. The success of a project benchmarking project should be measurable in terms of cost reduction or improved performance.

This conceptual approach opens the door to further alternatives in partner selection, allowing for greater

### Applying best in class practices

Learning and adapting are the pillars of process benchmarking. The main steps of best practice, for a specific project and a specific group of participants, are:

- *Identification of the best in class.* The most obvious way to do this is by measuring performance, and this is the step that quite often leads to confusion (and to mixing up metric and process benchmarking).
- *Study of the process itself.* This study is not solely numerical, and the process should be assessed from a methodological, conceptual and even philosophical point of view.
- *Identify improvement measures.* Not everything that a best-in-class performer does can be applied in another system. Sometimes, some of the techniques or sub-processes will not be convenient or applicable. In choosing what to implement and what not to implement, in other words, in adapting a process to specific particular needs, lies the success of a process benchmarking project.

Advantages:

- Produces original solutions instead of just pointing out weaknesses.
- Improvements can be quantified in a cost/benefit analysis.
- Results may be directly incorporated in the decision-making process.

Shortcomings:

- Very demanding task. Even large utilities usually need an adaptation period before they obtain important benefits.
- Cannot be used for regulation, since it does not assess performance relative to peers. However, regulators in some parts of the world are favouring that utilities participate in process benchmarking efforts as a way to encourage performance improvement within their influence area.

opportunities and flexibility and lessening the impact of confidential data leaking out to direct competitors. *Out-of-the-box* benchmarking consists of looking for *best-in-class* performers in other industries and sectors, thus widening the scope of potential participants considerably.

### CURRENT CHALLENGES IN PERFORMANCE ASSESSMENT

In the 15 years that have lapsed since the first initiatives in performance assessment of urban infrastructure services, the evolution of tools, objectives and projects have led to new kinds of challenges. In the early stages, the focus was placed on the establishment of adequate measures and the precise definitions needed to obtain the correct data from the systems and calculate the indicators correctly.

Once the actual projects went on their way, the everyday implementation of a performance measurement system became crucial as data were not so easily obtained and it was not easy to integrate all information systems in a utility. Additionally, it became necessary to set up a clear structure differentiating simple measures from indicators and from purely contextual information.

Today, performance assessment faces the challenge of understanding better the results and making them more understandable to others. Mature projects have difficulties in the analysis stage, determining the real performance of a utility, whether it is evaluated on its own or compared to others. Furthermore, results often need to be communicated to the public or the press in a way which is understandable, an objective sometimes incompatible with the more sophisticated assessment tools.

These are some of the most relevant challenges presently faced by the water supply, wastewater and solid waste sectors regarding the assessment of performance:

- *Performance assessment tools for solid waste services.* Although many of the principles and concepts developed for other sectors (and particularly for the water and wastewater services) can be adopted by the solid waste services, there is the need to develop appropriate tools specific to this sector. The existing applications of coherent performance assessment systems are limited. The challenge is to develop a set of tools that is robust, flexible, customisable and as universal as possible in terms of its applicability.
- *Taking into account the data quality in the decision-making process.* Performance measurement systems are always the means to some sort of decision-making process. Unfortunately, those decisions are often

made based on poor data. A fact that is hidden behind the performance indicator figure, a number which hardly ever shows how it was obtained or where it came from.

Even in the PI systems in which some sort of data quality information is included, it is difficult to take such information into the analysis stage of the results. In other words, it is difficult to determine whether a good indicator value based on poor data should be better than an indicator below target obtained from accurate data. How to incorporate data uncertainty into the decision-making process is still a challenge.

- *Finding the driving factors.* When measuring performance, the *why*'s are as important as the *how much*'s. One of the key aspects of making performance assessment successful is to identify the driving factors behind a certain level of performance, and whether those factors can be corrected or not. The influence of context in performance (e.g. climate, size, geography, etc.) also needs to be assessed as precisely as possible.

Unfortunately there is no single answer to this question. Data are often insufficient to guarantee the results of complex mathematical analyses and simpler methods rely too much on the experience or intuition of those in charge of the results' interpretation.

- *Synthesising the results.* Whether they are to be presented to a board of directors, the press or the general public, performance assessment results often need to be simplified. However, both in metric benchmarking and in individual performance assessment, the big picture usually fails to deliver some of the most interesting details, and it can even lead to misunderstandings and wrong conclusions.

From the simple method of weighing indicators and aggregating them into an index, to the use of complex tools providing a measure of the efficiency, all the techniques used to aggregate results present some sort of shortcomings. Providing a reliable and transparent way of synthesizing the results of the assessment is a problem that still needs to be solved.

- *Applying the tool to small scale systems.* A problem which is not limited to performance assessment is how to apply best practices in the management of services in small scale systems. In the case of performance indicators, the available data in these systems is often scarce

or unreliable. The resources are also limited, increasing the difficulty of improving the situation. Finally, in such systems there are often impending problems of greater importance. How can performance measurement systems become useful in such systems?

- *Quantitative statistical models vs. qualitative assessments.* Regulators in the utilities' sector have been using performance indicators for decades. Many of them used approaches in which a complex mathematical model assesses comparative performance by establishing relationships between inputs and outputs or evaluating efficiency. Such methods have the advantage of being fair to all those involved (the assessment is the same for everyone) but the disadvantage that a deep knowledge of the equations driving the model is needed to avoid erroneous conclusions. Whenever this is not the case, the users do not trust and use the results, unless they "look good in the picture".

The alternative might not be any better. In other systems, the human factor is introduced to assess performance. Experts review the PI values and take into account context factors to assess, in a qualitative manner, the performance. Obviously, despite the process being totally transparent, such a person can introduce bias into the method which is strongly influenced by the experience and knowledge of the expert.

The challenge remains in how to achieve the advantages of both methods without facing the disadvantages.

## LESSONS LEARNT: COST C18 RECOMMENDATIONS

COST C18 experts studied in detail the available methods to assess, analyse and synthesise utilities' performance. The main lessons learnt from those efforts and the corresponding recommendations are:

1. *Individual PI analysis is fundamental.* As discussed in this paper, there are several methods that combine and aggregate indicators and provide an overall assessment of the performance. Regardless of how useful these methods may be, the analysis of each individual indicator and the reasons for its current value should never be neglected.
2. *Graphics help to interpret results.* One of the most effective ways to perform the analysis of results is by

means of appropriate graphical representations. The variety of graphs used in performance assessment projects around the world suggests that a single solution will not fulfil the needs of every project, and finding the appropriate graphical tools should be an important part of the analysis stage. A number of inspiring examples have been compiled during the COST C18 seminars.

3. *Always perform a full analysis of results.* Calculating performance indicators is an easy task if all the necessary data is at hand. Interpreting their values is never easy. The analysis of indicators should always take into account the following factors:

- a. *Quality of input data.* The uncertainty and reliability of the data used to calculate the indicator are necessary to assess the possible risk taken during the interpretation of the results.
- b. *Explanatory factors.* The analysis should include the study of those factors influencing the value of the indicator, whether they can be modified by management decisions in the short/medium term (driving forces) or are simply part of the local characteristics of the utility (explanatory factors).
- c. *Targets.* Indicators are only useful when compared to a reference value. Any analysis should take into account the original target set when the indicators were chosen.

4. *Be careful when simplifying!* Aggregation or simplification of performance assessment results is often necessary due to several reasons. In any case, the need to provide a brief, easy-to-understand result for a performance assessment project should never result in a simplistic view of the situation which may lead to wrong conclusions. Therefore, under all circumstances, performance assessment results should be:

- a. Easy to understand and interpret.
- b. Clearly specify whether the communicated results have been obtained after some data processing, weighing or qualitative analysis.
- c. Transparent and traceable.
- d. In avoidance of “black box” approaches.

In any case, and for public communication, simple and transparent aggregation criteria should be preferred to more

complex processes, to allow any user to understand the information and go into more detail if needed. Individual indicators should, in any case, be accessible, allowing it to be traced back to the original data.

#### **The road ahead—key research areas and future applications**

Based on the experience achieved, members of COST C18 discussed which are the key research problems related to the management of urban water infrastructures, currently not covered by on-going projects of the European Framework Program. R&D topics recommended are:

- Efficient management of small community systems,
- What makes a utility sustainable?
- Common framework for capital maintenance suitable to medium and small European communities,
- Performance assessment and decision-support systems for the urban solid waste services.

#### **CONCLUDING REMARKS**

Judging from the growing number of initiatives worldwide, performance assessment has been accepted by the industry as a key tool to drive efficiency and best practices. The IWA has played a crucial role in this development by creating the system which is currently the worldwide reference in the water and wastewater sectors, and by thoroughly disseminating such practices. The publication of the ISO 24500 standards will only contribute further in that direction.

Action COST C18 was created four years ago to try to take matters a little further and determine future challenges and research needs. From such activities, it is possible to say that, although the water industry has already begun to actively use performance assessment tools, important work is still needed to refine the practices in those projects. This action demonstrated that principles for establishing and implementing PI systems developed for the water sector are sufficiently mature and universal to be adopted by other urban infrastructure services. The development of sector-focused PI systems, based on these generally accepted principles, is a remaining challenge.

An important part of that work has to do with the need for the simplification of results. Analysing the performance of a utility is a complex task, and all attempts to simplify the

results lead to an incomplete picture that may even result in wrong decisions being made. Until better methods and/or data quality lead to more reliable results, a more complete picture of the assessment should always be available to gain a deeper understanding of the assessment.

In the coming years, the search for best practices by the industry, as well as the increasing number of national and regional regulators, will produce an increasing number of performance assessment and benchmarking projects, and also an increase in the maturity of such projects. It is quite likely that the need to produce better results from those studies will also drive the improvements still needed in the assessment methods.

## ACKNOWLEDGEMENTS

This paper presents the experiences and lessons learned during the four years of the COST Action C18. The authors, as chairpersons of the Action, acknowledge the COST Office for making this Action possible, as well as the working groups chairs and to all COST C18 members, who actively contributed to the success of the Action. Acknowledgements are also due to the invited experts who participated in the seminars, sharing their knowledge and experiences with the group.

## REFERENCES

- Adamsson, J. 1997 The Swedish experience and viewpoints. In: *IWSA Workshop on Performance Indicators for Transmission and Distribution Systems, Lisbon, Portugal*. IWSA, London.
- Alegre, H., Hirner, W., Melo Baptista, J. & Parena, R. 2000 *Performance Indicators for Water Supply Services*. IWA Publishing, London.
- Alegre, H., Melo Baptista, J., Cabrera, E., Jr, Cubillo, F., Duarte, P., Hirner, W., Merkel, W. & Parena, R. 2006 *Performance Indicators for Water Supply Services*. IWA Publishing, London.
- Cabrera, E., Jr & Pardo, M. A. (eds) 2008 *Performance Assessment of Urban Infrastructure services. Drinking Water, Wastewater and Solid Waste*. COST Office and IWA Publishing, London.
- Eggen, E., Büchner, J., Petro, K., Giesbertz, P., Fredriksen, O. & Ajodhia, V. 2001 Micro and macro benchmarking methods for estimating efficiency in electricity networks. In: *Proceedings of Distributech EuropeBerlin, Germany*.
- Faria, A. L. & Alegre, H. 1996 Paving the way to excellence in water supply systems: a national framework for levels-of-service assessment based on consumer satisfaction, The Maarten Schalekamp Award—1995. *AQUA* 45(1), 1–12.
- ISO 2007a *ISO 24510:2007—Service Activities Relating to Drinking Water and Wastewater—Guidelines for the Improvement and for the Assessment of the Service to Users*. Technical Committee TC224, International Organization for Standardization, Geneva.
- ISO 2007b *ISO 24511:2007—Service Activities Relating to Drinking Water and Wastewater—Guidelines for the Management of Wastewater Utilities and for the Assessment of Wastewater Services*. Technical Committee TC224, International Organization for Standardization, Geneva.
- ISO 2007c *ISO 24512:2007—Service Activities Relating to Drinking Water and Wastewater—Guidelines for the Management of Drinking Water Utilities and for the Assessment of Drinking Water Services*. Technical Committee TC224, International Organization for Standardization, Geneva.
- Janssens, J., Pintelon, L., Cotton, A. & Gelders, L. 1994 Development of a framework for the assessment of operation and maintenance (O&M) performance of urban water supply and sanitation. In: *IWSA Specialised Conference on “The Quality of Service (or Managing a Water Company)”*, Amsterdam, The Netherlands, IWSA, London.
- Malaysian Water Association 1996 *Performance Indicators for Water Supply, A Proposal from the Malaysian Water Association for the Consideration of Member Countries of ASPAC*. MWA, Kuala Lumpur.
- Matos, R., Cardoso, A., Ashley, R., Duarte, P., Molinari, A. & Shulz, A. 2003 *Performance Indicators for Wastewaters Services*. IWA Publishing, London.
- Sjøvold, F., Conroy, P. & Algaard, E. (eds.) 2008 *Performance Assessment of Urban Infrastructure Services. State-of-the-Art in the Drinking Water, Wastewater and Solid Waste Sectors*. Final scientific report of COST Action C18, COST.
- Skarda, B. C. 1997 The Swiss experience with performance indicators and special viewpoints on water networks. In: *IWSA Workshop on Performance Indicators for Transmission and Distribution Systems, Lisbon, Portugal*, IWSA, London.
- van der Willigen, F. 1997 Dutch experience and viewpoints on performance indicators. In: *IWSA Workshop on Performance Indicators for Transmission and Distribution Systems, Lisbon, Portugal*. IWSA, London.
- Vewin 2006 *Reflections on Performance 2006. Benchmarking in the Water Industry*. Vewin, Rijswijk, The Netherlands.
- Yeppes, G. & Dianderas, A. 1996 *Water & Wastewater Utilities Indicators*. Water and Sanitation Division, The World Bank, Washington, DC.
- Yeppes, G. & Dianderas, A. 1997 The role of indicators to guide unaccounted for water reduction programs. In: *IWSA Workshop on Performance Indicators for Transmission and Distribution Systems, Lisbon, Portugal*. IWSA, London.