Supporting Industry-based Benchmarking - Benefits of and Limits to the Application of Econometric Methods

T. Franz[1]*, F. Bertzbach[1], A. Schulz[2], S. Pfister[3] and J. Stemplewski[2]
[1] *aquabench GmbH, Ferdinandstraße 6, D-20095 Hamburg, Germany*
[2] *Emschergenossenschaft/Lippeverband, Kronprinzenstraße 24, D-45128 Essen, Germany*
[3] *hanseWasser Bremen GmbH, Schiffbauerweg 2, D-28237 Bremen, Germany*

**Torsten Franz, e-mail t.franz@aquabench.de*

ABSTRACT

Econometric methods can support and streamline benchmarking processes. They facilitate the consolidation of indicators of different performance areas and of structural characteristics into the analysis. Main benefit, herby, is not to calculate an overall score, used for ranking, but the potential to partly quantify and evaluate impact of structural characteristics. When using econometric methods, however, certain rules must be applied. From a scientific point of view, it is not admissible to use these methods at the company level. The tacit use of ranking orders is a highly questionable simplification that deviates from benchmarking principles. A balanced assessment of the steps that are necessary on the path towards sustainable and resource-efficient water management cannot be derived from the results of statistical calculations alone. Therefore, the inclusion of the expertise of local professionals is still required. In a cooperation project by several long-term benchmarking partners, process data from sewer construction, wastewater collection and transport and wastewater treatment were used for these purposes.

KEYWORDS: benchmarking, econometrics, economy, efficiency, performance indicators, sewer construction, wastewater treatment, wastewater collection and transport

INTRODUCTION

The voluntary benchmarking is a management tool developed by the water sector and it is used successfully (e.g. Bertzbach et al. 2012, Theuretzbach-Fritz et al. 2012, VEWIN 2013). Despite the proven and already achieved successes it remains a continuous challenge to support practicably the work on the original aim of benchmarking - performance improvement. A simplification of the performance indicator analysis and a clear inclusion and consideration of the various structural characteristics can help involved senior management to identify quicker and more easily the changeable factors and to assess their effects.

A previously in practice under-utilised – at least in Germany (see Oelmann and Growitsch 2011) – possibility for such a support and a slimming down of the benchmarking process (in particular to find more easily real performance differences caused by better practices) is the employment of so-called “econometric methods”. Using these methods based on procedures of statistics and linear optimisation several factors or performance indicators were consolidated mathematically. Therewith, both different performance indicators as well as structural characteristics can be consolidated in models.

In a cooperation project of long-term benchmarking partners and aquabench the following questions were put:

- Benefits: How can the benchmarking process be advanced efficiently by a simplified and meaningful assessment and quantifying of structural characteristics and performance aspects?
- Limits: With econometric methods the calculated differences between operators' values are by default interpreted as differences of efficiency. Which risks do arise by this interpretation with regard to the assessment of operators?

USING ECONOMETRIC METHODS WITH HIGH QUALITY BENCHMARKING DATA

Data

Since 1996 operators of the German water sector have been using benchmarking voluntarily as optimisation instrument (Schulz et al. 1998). In the meantime, alone with aquabench over 800 operators have carried out projects of which numerous have been repeated regularly. The detailed and comprehensive database built up does make it possible to assess constructively benefits and limits of the econometric methods. It is fundamentally that all applications with regard to the building of models and the data used must be checked critically, before one can follow any conclusions gained. Insofar it is also important that any already available knowledge is used in order to eliminate meaningless models.
In the project described, process data of sewer construction, waste water collection and transport as well as of wastewater treatment were used. From a database of more than 1,000 investigated objects some 100 were selected for the modelling (table 1).

Table 1. *Database of aquabench benchmarking projects*

	Plants, networks reaches of sewer	Variables	Assessment period
	(number)	(number)	(years)
Wastewater Treatment	272	44	2006 – 2010
Wastewater Collection and Transport	324	60	2008 – 2012
Sewer Construction	1,594	25	2002 – 2010

Econometric methods

Three econometric methods have been investigated, one method of regression analysis and two methods of so-called efficiency analysis. These methods represent different characteristics with regards to model building and consideration of model failures. They are used in existing regulated markets and dominate the national and international literature on the regulation of the water sector (e.g. Berg and Marques 2011, Konkurrence- og Forbrugerstyrelsen, 2014; CEPA, 2014). A graphical interpretation of the different methods is presented in Figure 5, for comparison also the approach of the classical performance indicator comparison, which is equally an econometric method (von Hirschhausen et al. 2009). Primarily for data preparation (relevance of characteristics, outlier analysis) other statistical methods were applied.

The models created using econometric methods are always a limited copy of reality. Therefore the findings of so-called efficiency analyses (“efficiency of x %“) are also always to be interpreted as theoretical values. Limits of modelling, e.g. with the selection of parameters considered or the taking into account of factors which cannot be expressed in figures, are to be kept in mind (see chapter on limits). The findings of a statistical analysis are always linked with statistical errors (“confidence interval“).

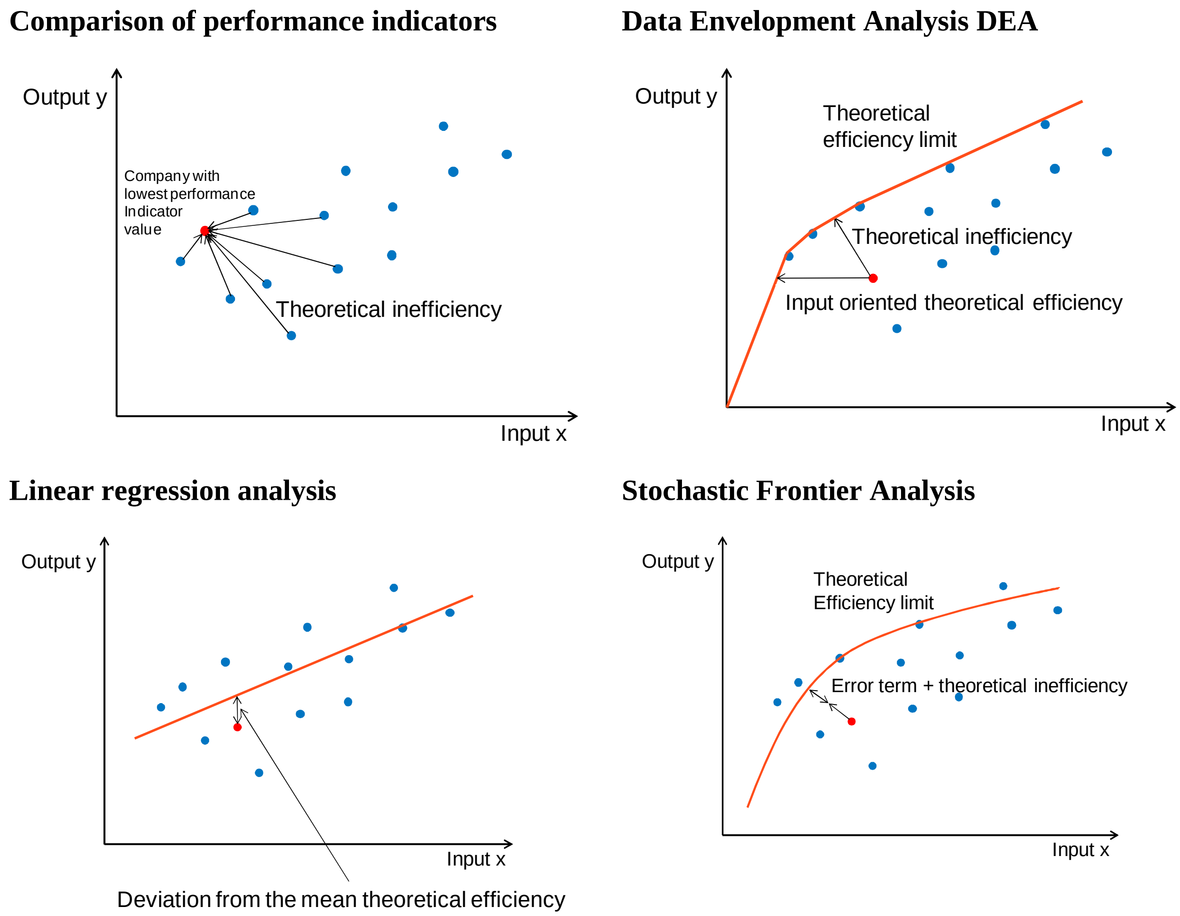

Figure 1: *Different theoretical inefficiencies by using different methods*

Linear regression analysis. The quantifying of the influence of structural characteristics on the economic performance indicators and the determination of their contribution to the level of the performance indicators takes place using regression analyses. Regression analyses are statistical methods, which determine relationships between a dependent and one or more independent characteristics. Here, the special case of multiple linear regression have been used: several independent characteristics are linked together using regression coefficients in the first power. The resultant function represents average relationships or an average theoretical efficiency.

Data Envelopment Analysis (DEA). Data Envelopment Analysis is a method for the measurement of theoretical efficiency. Every object examined is characterized by inputs (e.g. expenditure, working hours, revenue) and outputs (e.g. quantity of treated wastewater, discharge concentration). By means of the DEA several inputs and outputs can be considered simultaneously.
The theoretical efficiency value calculated for each object measures the distance to an efficient margin based on observing in- and outputs. This efficient margin or the theoretical efficiency limit is the envelope around all objects and is determined using linear optimisation.

As the output (connected number of inhabitants and population equivalents, operated network kilometres) by wastewater disposal operators cannot be changed arbitrarily, the special case of the input DEA has been investigated, i.e. with the theoretical efficiency value determination the output is held constant.

Stochastic Frontier Analysis (SFA). The Stochastic Frontier Analysis is based on regression processes. In the findings, using SFA, one obtains a statistical model and/or a production function which defines the efficient margin or the efficiency limit taking into account an error term (statistical error).

BENEFITS FOR BENCHMARKING

Placement of the econometric methods in the benchmarking process

Benchmarking is defined as "…a tool for performance improvement through systematic search and adaption of leading practices"(Cabrera et al. 2011), i.e. the aim of a benchmarking project is ultimately optimising operations. The German water sector has developed a voluntary benchmarking approach based on this understanding (DVGW and DWA 2008).

Performance improvement is not aimed at economic improvement alone. Alongside economic efficiency, the so-called "five pillar model" defines quality, safety, customer service and sustainability as optimisable performance areas. The responsibility of operators themselves for cause analysis and determination of potential improvements is an essential factor for the successful application and broad acceptance of benchmarking (Bertzbach et al. 2012). Cause analysis and determination of potential improvements of participating operators, i.e. the search for that which is changeable, is based on a classical comparison of performance indicator values. The difference between own performance indicator values and the minimum values needs to be analysed in benchmarking. This difference is called the "search area", since this area needs to be explained in order to find better practices (Figure 2).

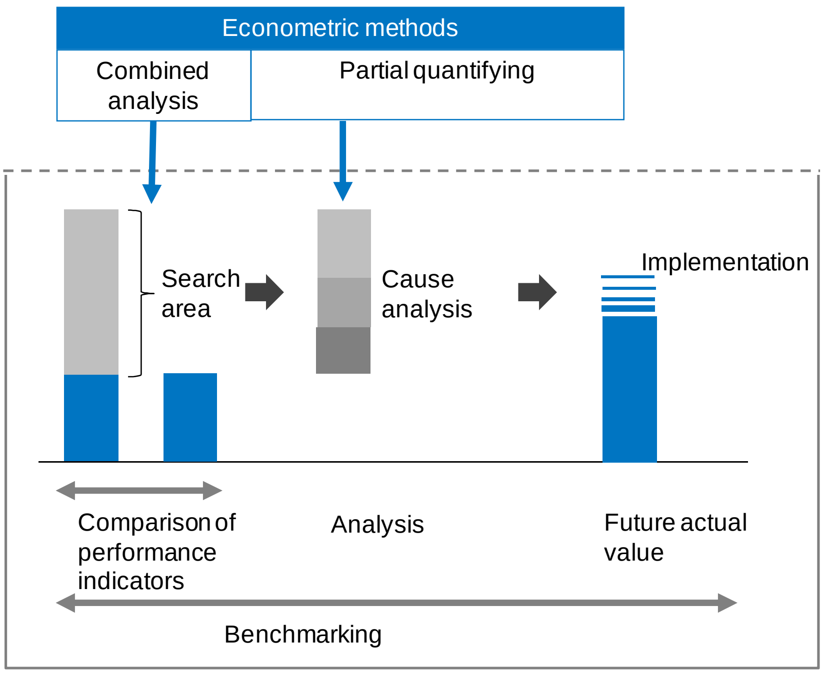

Figure 2: *Econometric methods supplement the benchmarking process*

The as a rule non-changeable structural characteristics are included in various ways into the benchmarking process (DVGW and DWA 2008, Cabrera et al. 2011):

- They are made transparent in order to give partners the opportunity of being able to assess their conditions in a comparison group.
- They are used for clustering.
- They are used in individual analyses in order to explain the values of performance indicators. With this, performance indicators from the various performance areas are qualitatively interrelated and discussed.

This easily comprehensible, transparent and tried and tested procedure requires considerable specialist input and therefore manpower. The actual influence of structural characteristics and the relationship between the performance indicators cannot be represented in a focused manner in two-dimensional analyses and, in part, are also only qualitatively assessable. Here, the appliance of econometric methods may improve the knowledge about values, relationships and processes. The search area analysis may then be supplemented by combined analyses from several performance areas and structural characteristics and by a quantifying of possible influencing characteristics (comp. Figure 1).

Main results

The actual application for the procedures investigated has shown that some econometric methods support the initial questions regarding inclusion of performance indicators and structural characteristics:

- The quantifying of parts of the search area and combined analysis of performance indicators is possible: With the application of regression analyses on the data sets considered, the influence of structural characteristics and of other performance indicators on economic figures could be verified and, in part, quantified. With this, it can be expressed which constraints and which extra-service allow which range of additional or reduced costs to be expected. An example of such results is given below.
- Additionally, results can be used for cost-planning. E.g. applying regression models on base of sewer construction data has significantly reduced uncertainties of cost planning of future investments.
- Other authors stress the capability of DEA and SFA to produce efficiency objectives. For example, Alegre et al. (2009) even consider the advantage of these methods “of being fair to all those involved (the assessment is the same for everyone)”. Discovering the huge range of results for one utility depending on the selection of models and depending on data quality (see next chapter), this benefit can only partly be seen by authors of this study. Nonetheless, the ranges of the findings based on sensitivity analyses and the analysis of the differences of models can nevertheless be made usable in the benchmarking process (Franz et al. 2013). For this type of learning process various interactive models exist which have been developed in Danish benchmarking in order to be able to compare own objectives, conditions and demands individually with benchmarking partners (Bogetoft and Otto 2011).

Application example: Quantifying of parts of the search area is possible

With the aid of regression analysis, for example in the process benchmarking “sewer construction”, the influence of structural characteristics on the construction work costs [EUR/m] are quantified. The quantification is performed via regression coefficients.

Table 2. *Monetary impacts of structural characteristics in sewer construction for projects "Replacement in open trenches DN < 1200 taking into account type of shoring and groundwater drainage"*

Influencing factors (Structural characteristics)		Monetary impacts of recorded factors	
1	Project length	-0.78	€/m * m
2	Nominal diameter	1.75	€/m * mm
3	Average depth	132.61	€/m * m
4	Number of manholes per 100m	62.72	€/m
5	Construction trench length	0.64	€/m * m
6	Soil replacement	1.54	€/m * %
7	Type of material		
	Ductile cast iron	490.44	€/m
	Plastic	175.06	€/m
	Others	210.59	€/m
	Reinforced concrete	251.88	€/m
	Vitrified clay	148.30	€/m
	Concrete	0.00	€/m
8	Type of soil		
	Class 3 (easily removable)	0.00	€/m
	Class 4 (removable with moderate difficulty)	36.97	€/m
	Class 5 (removable with difficulty)	-20.76	€/m
9	Project location		
	Free area	-448.03	€/m
	Average	0.00	€/m
	Complex	20.06	€/m
10	Branch coordinated construction		
	Co-ordinated	128.77	€/m
	Independent	0.00	€/m
11	Road construction coordinated construction		
	Co-ordinated	-82.94	€/m
	Independent	0.00	€/m
12	Type of shoring		
	Element	-191.58	€/m
	Standard shoring	125.74	€/m
	Timber piling	358.09	€/m
	Berlin system	300.53	€/m
	Others	-194.82	€/m
	Pile chamber	0.00	€/m
13	Groundwater drainage		
	None	50.30	€/m
	Open	-82.49	€/m
	Closed	00.0	€/m

Model "Replacement in open trenches DN<1200 (with type of shoring and GW drainage)", 228 Data sets, $R^2 = 0.50$

They indicate how the dependent parameter changes if an independent one changes by one unit. It thus applies for example, on the basis of a specific model for the construction type "Replacement in open trenches DN<1200" and within the framework of the selected data sets, that an increase in depth of one metre generates additional costs of 133 EUR/m. Thus, in the process benchmarking ”sewer construction” 13 influencing factors are quantified in their monetary impacts (Table 2). Additionally, several influencing factors could also be identified as statistically insignificant.

Based on deviations between the actual construction work costs and those with regression function modelled costs of an individual participant, the following statement is now possible: How large is my individual search area with those for me characteristic construction site conditions? With this new differentiated search area the structural characteristics are taken into account and thus, in comparison with the pure performance indicator comparison, are to a large extent eliminated.

An example is presented in Figure 3. For this participant the search area amounts to 741 EUR/m compared with the minimum value. This search area, taking into account the modelled structural characteristics (Table 2), can be attributed to 69 % (equates to 513 EUR/m) to individual influencing factors. Non-modelled structural characteristics (regional construction market, construction standards not covered in the benchmarking or other factors) also influence the search area. The analysis of the search area is now significantly simplified. The participant can directly identify changeable factors and at the same time assess their monetary impacts.

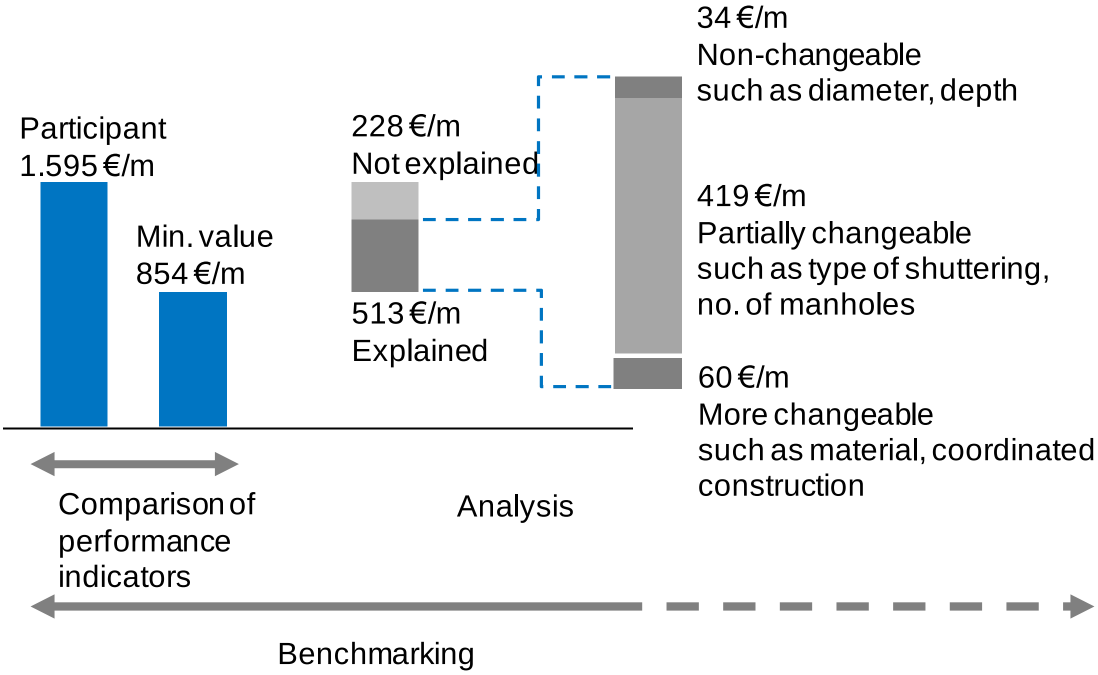

Figure 3: *Quantification of the search area for a participant, sewer construction process benchmarking, model “Replacement on open trenches DN < 1200 taking into account type of shoring and groundwater drainage”*

In a similar way main structural characteristics and performance indicators influencing the cost of wastewater collection and transport and wastewater treatment were identified and – with regard to the uncertainties - quantified.

LIMITS TO THE APPLICATION

General shortcomings

Shortcomings of the methods are recognized already in international benchmarking literature (Alegre et al 2009, Cabrera et al 2011):

- Selection of model can "bias" the results
- Models have risk of just ranking utilities instead of understanding improvement potential
- Methods are less transparent and more obscure than simple ratios
- Good results require good data quality

The here presented study stresses additionally:

- from a scientific point of view, it is not admissible to use these methods at the company level
- The findings of so-called efficiency analyses ("efficiency of x %") are also always to be interpreted as theoretical values.

Therefore, when using econometric methods, certain rules must be applied.

Focussing on sub-processes

The modelled dependent parameters of the regression analyses as well as the output parameters of the DEA and SFA must match both the processes considered and the economic figures, i.e. must be clearly associated cost drivers. Modelling is only possible for coherent processes. The selection of the respective parameters must be based accordingly on recognised findings on relationships between expenditure and performance from operational and benchmarking practice (Berg 2010, DWA 2007).

If expenditures and performances are analysed across processes, this leads, with high probability, to errors as non-existent cause-and-effect relationships between expenditure and benefits are assumed. Furthermore, with sub-processes, different economies of scale apply, e.g. with wastewater treatment the specific expenditure basically falls with the plant capacity, with wastewater collection and transport it rises. With a modelling at company level these scale effects will be balanced and helpful deductions for the operators are not to be expected.

The further-reaching idea, to sum up findings at sub-process level (e.g. in Danish regulation, Konkurrence- og Forbrugerstyrelsen, 2014) is associated with methodical imperfections: Efficiency analyses, due to statistical errors and necessary sensitivity analyses always deliver results with large ranges and not all impacts can be included (see e.g figure 3). With the summing of individual findings this information is lost and errors are propagated. Therefore the strength of the methods in the water sector has an effect at sub-process level, only.

Data basis must be analysed and observed – it determines the findings and limits the applicability

Compliance with the minimum number of data sets. A minimum number of data sets, which is dependent on the number of examined characteristics, is required for authoritative findings. A rule of thumb for the minimum number of data sets is that at least ten times as many data sets should be available as independent characteristics in the complete model (Held et al. 2010). For individual processes in the water sector such as, for example, the investigation of wastewater treatment plant sub-processes, this means that the processes currently are not authoritatively applicable within the German context.

Correct grouping (clustering). The processes investigated emanate fundamentally from continuous and monotonous processes of the modelled functions. This approach is not realistic for a combined consideration of (sub-) processes of all operators in the water sector.
Through the essentially parameter-conditioned structural characteristics such as, for example, concentration thresholds or applicable technologies, discontinuities arise, for example with the change from aerobic stabilisation to anaerobic sludge treatment or from unaerated to aerated plants. The ignoring of these discontinuities within the scope of an econometric modelling leads to false findings.
Here, a suitable grouping or clustering of participating operators appears to be sensible.

Consideration of influencing factors. With the modelling of performance indicators or output parameters all relevant and recordable influencing factors must be included or their influence eliminated by means of a harmonisation. If economically relevant influencing factors or structural characteristics are not taken into account, this leads inevitably to errors in the modelling and the subsequent evaluation. This concerns, for example, the use of revenues as main financial input as commonly practiced in a majority of academic surveys (e.g. Zschille et al. 2009). In contrary to costs revenues include non-efficiency-relevant components like disregarded reserves or different approaches for the determination of allowances for depreciation. Within the German context, the inclusion of these financial components alone can lead to a misinterpretation of the theoretical efficiency value of up to 20 % in addition to other influencing factors.

Elimination of outliers. Statistical outliers influence the findings of econometric methods massively. Therefore a comprehensive analysis of the data for outliers and a consequent elimination of these cases is required, in order to avoid erroneous findings. Extreme values are, however, a reality in the water industry due to the diversity of the plants/systems.

The determined theoretical efficiency values are always a relative dimension and vary dependent on the model's approach – sensitivity analyses are necessary

There is not just one econometric method and thus also not just one valid model for the determination of a theoretical efficiency value. It is therefore necessary to establish ever more models as determined by a sensitivity analysis with different methods and method assumptions and data sets.

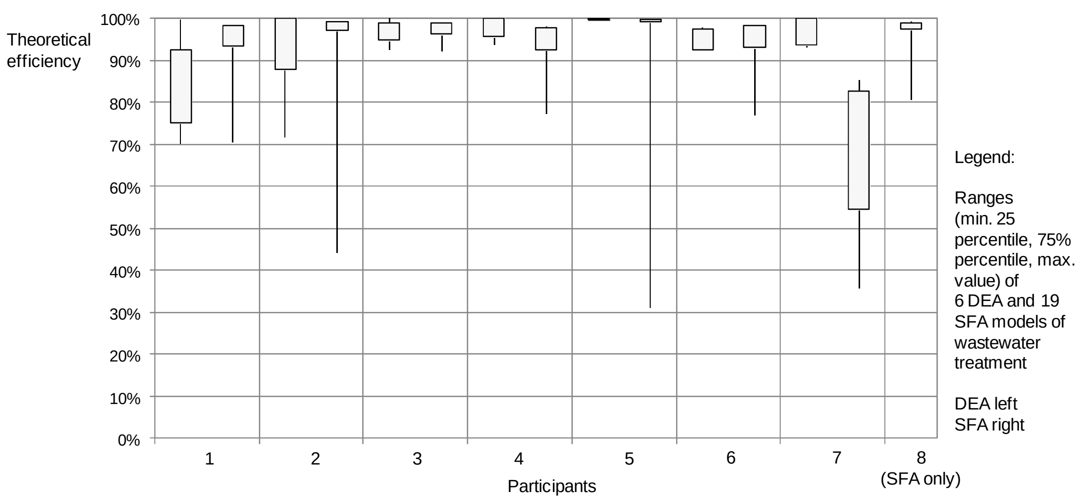

Figure 4. *Range of theoretical efficiency values from 25 efficiency models (DEA and SFA) in wastewater treatment for 8 benchmarking participants (Franz et al 2013)*

A sensitivity analysis for process data of wastewater treatment with 25 models, with which data and group membership have been varied very comprehensively, resulted in ranges of theoretical efficiency values of, on average, 24 % per participant (Figure 3). With this, the individual determined theoretical values are dependent not only on the selected model approaches but also on the available data bases.

Theoretical efficiency values cannot replace the balanced analysis of performance areas produced by local expertise

Using econometric methods selective findings of performance indicators or structural characteristics are made clear. The theoretical efficiency values determined here are, however, a long way away from actually representing all performance indicators and structural characteristics. On one hand, this would require a considerable number of data sets and, on the other hand, the performance areas of the water sector are not to be described completely through indicators.

The global description of a benchmarking object by means of only one figure is contrary to the objectives and practice of benchmarking. Different practice-related manuals and studies (e.g. Bertzbach et al. 2012, Cabrera et al. 2011, DVGW and DWA 2008) have elaborated the significance of the responsibility of the operators for the cause analysis and evaluation of the findings of statistical comparisons. This responsibility cannot be fulfilled alone with the employment of econometric methods. This, for the above given reasons, applies all the more if it is attempted to generate, with the aid of aggregated values, a single overall statement on the capability of a complete company. The balanced evaluation continues to require urgently the inclusion of the expertise of the local experts as well as the continuous balancing out of the operator's objectives in the performance areas of the water sector (Figure 4). A reproduction in a single statistical method cannot be imagined.

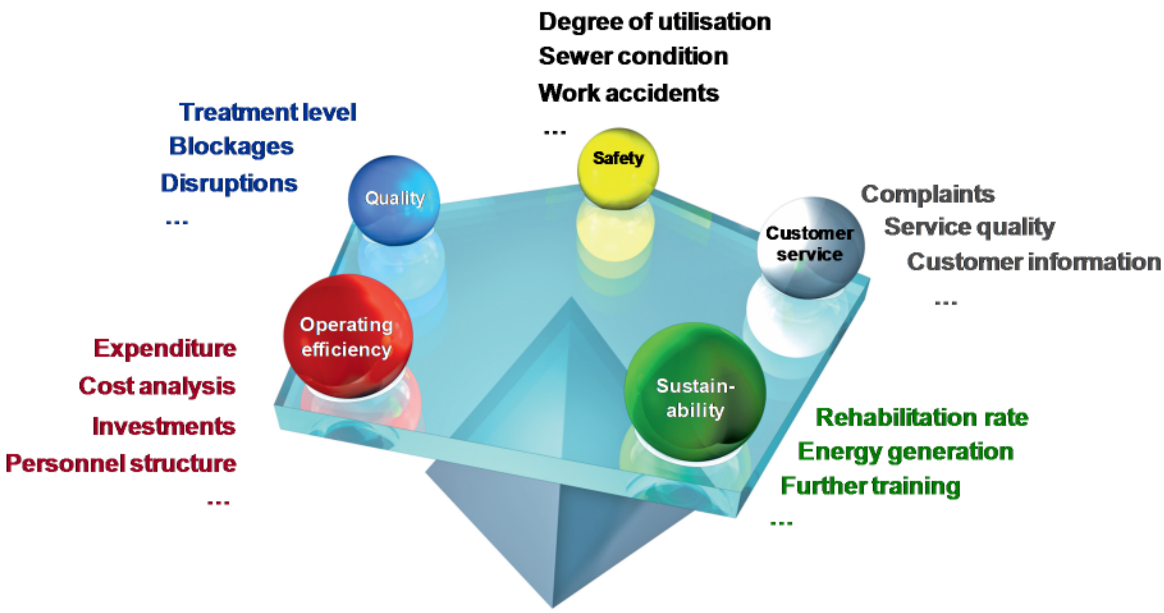

Figure 5: *The balanced analysis of performance areas (five pillar model of the German benchmarking) cannot be replaced by econometric methods (adjusted representation according to Schulz and Goebel, 2012)*

CONCLUSIONS

The actual performance of water utilities is modelled using the methods investigated – as they also modelled when using the classical comparison of performance indicators. The particular strength, that several characteristics and/or performance figures can be analysed combined is recognized in all tested methods. But especially the strength of regression analysis to quantify and show the influences of structural characteristics in a transparent way is seen as main benefit for benchmarking. Differences between partners are explained and changeable factors with great optimisation potential are identified more easily.

However, with the employment of econometric methods there are rules to be observed:

- An employment is only sensible with coherent processes with which clear cause-effect relationships exist. An application at company level in the water sector is therefore scientifically unacceptable.
- There is no "one" econometric method and no "one" model for the determination of a theoretical efficiency.
- Models are based on assumptions and data which, depending on analysis, can differ greatly and influence the model findings and/or limit the applicability. The data basis is to be analysed carefully and is to be taken into account with modelling. Above results have been possible on basis of a consistent data basis built up in years of benchmarking in Germany.
- The use of rankings without comment, which are derived from the findings of econometric methods, is a questionable simplification. It departs from the principles of the benchmarking. The balanced evaluation of necessary steps up to a sustainable and resource conserving water sector cannot alone be based on results of statistical calculations.

In line with those rules aquabench will use the methods targeted in chosen projects to increase the value of benchmarking for its participants.

REFERENCES

Alegre, H., Cabrera, E., Merkel, W. (2009): „Performance assessment of urban utilities: the case of water supply, wastewater and solid waste", Journal of Water Supply: Research and Technology – Aqua, 58.5, 2009, S. 305 – 315.

Berg S. V. (2010). Water Utility Benchmarking – Measurement, Methodologies, and Performance Incentives. IWA Publishing: London. ISBN 978-1-8433-9272-9.Berg S. V. and Marques R. (2011). Quantitative studies of water and sanitation utilities: a benchmarking literature survey. Water Policy 13(5), 591–606. IWA Publishing, DOI:10.2166/wp.2011.041.

Bertzbach, F., Moeller, K., Nothhaft, S., Waidelich, P., Schulz, A (2012). 15 years of voluntary benchmarking – how it supports the modernisation strategy of the German water sector. Water Utility Management International, December 2012, London.

Bogetoft P. and Otto L. (2011). Benchmarking with DEA, SFA, and R. Springer: New York, Dordrecht, Heidelberg, London. ISBN 978-1-4419-7960-5.

Cabrera E. Jr., Dane P., Haskins S. and Theuretzbacher-Fritz H. (2011). Benchmarking Water Services - Guiding water utilities to excellence. IWA-Publishing: London. ISBN: 978-1-843391-98-2.

CEPA Cambridge Economic Policy Associates (2014). OFWAT Cost Assessment – Advanced econometric models. Final report, 20th March 2014. http://www.ofwat.gov.uk/pricereview/pr14/pap_tec1402feederbasiccostappb.pdf

DVGW and DWA (2008). Advisory Leaflet W 1100/DWA M 1100 – Benchmarking in Water Supply and Wastewater Disposal. DWA: Bonn/Hennef. ISBN: 978-3-940173-50-8.

DWA (2007). Corporate Benchmarking as Component of the Modernisation Strategy - Performance Indicators and Evaluation Principles. DWA: Hennef. ISBN: 978-3-940173-64-5.

Franz T., Bertzbach F., Schulz A., Pfister S. und Stemplewski J. (2013). Unterstützung des Benchmarking-Prozesses – Nutzen und Grenzen der Anwendung ökonometrischer Verfahren. (Support of benchmarking – Benefits and limits of econometric methods). KA Korrespondenz Abwasser, Abfall 60 (2013), issue 12, 1067-1074. DOI: http://dx.doi.org/10.3242/kae2013.12.005.

Held L., Rufibach K. and Seifert B. (2010): Einführung in die Biostatistik [Introduction into Biostatistics]. Universität Zürich. 6th Edition.

von Hirschhausen C., Walter M. and Zschille M. (2009). Effizienzanalyse in der Wasserversorgung – Internationale Erfahrungen und Schlussfolgerungen für Deutschland [Efficiency analysis in water supply – International Experiences and conclusions for Germany]. GWF Wasser Abwasser, no. 2-3, 170-175.

Konkurrence- og Forbrugerstyrelsen (2014). Resultatorienteret benchmarking for 2015- Fastsættelse af individuelle effektiviseringskrav i prislofterne for 2015 [Result-oriented benchmarking for 2015 Determination of individual efficiency requirements of price caps for 2015], website: http://www.kfst.dk/~/media/KFST/Vandtilsyn/Benchmarking/Benchmarking%202015/Hoering%20af%20Benchmarking%202015/Resultatorienteret%20benchmarking%202015%20version%202%20050814.pdf, visited February 19th 2015

Oelmann M. and Growitsch C. (2011) Analysis of the current German Benchmarking Approach and Its Extension with Efficiency Analysis Techniques. In: Uhlig U. (Ed.): Current issues of Water Management, ISBN: 978-953-307-413-9.

Schulz A., Schön J., Schauerte H., Graf P. and Averkamp W. (1998). Benchmarking in der Abwasserbehandlung – ein Praxisbericht [Benchmarking in Wastewater Treatment – a Practical Report]. Korrespondenz Abwasser 12/1998 (45), 2297-2303, Hennef.

Schulz A. and Goebel H. (2012). Balanced Score Card Emschergenossenschaft / Lippeverband. Presentation DWA-Wirtschaftstage March 2012, Hamburg.

Theuretzbacher-Fritz H., Koelbl J., Friedl F., Fuchs-Hanusch D. (2013). How benchmarking triggers water industry improvements. Water Utility Management International, June 2013, 28-32, London.

VEWIN (2013). Reflections on Performance. Benchmarking in the Dutch drinking water industry. Vewin nr. 2013 / 119 / 6281.

Zschille M., Guder J, Kittlaus B., Moll R., Walter M. (2009). The Performance of German Water
Utilities: A Semi Parametric Analysis. Water Economics and Management Working Papers WP-H2O-11.