Identifying Techniques to determine the Overall Benchmarking Best Practices

James R. Oldham, M.S., Terry R. Collins, Ph.D., P.E., Manuel D. Rossetti, Ph.D., P.E and Heather L. Nachtmann, Ph.D.
Department of Industrial Engineering
University of Arkansas
Fayetteville, Arkansas 72701

Abstract

This research study investigated the use of multiple attribute decision theory (MADT) as a more robust mechanism to determine the best-in-class performer when conducting a benchmarking analysis involving inventory record accuracy within a public sector warehouse. The traditional gap analysis technique only identifies the absolute best-in-class performer for a single performance metric. The MADT method uses multiple criteria to identify the best overall performer for a selected set of critical metrics. Overall, this paper evaluated various decision-based methodologies to introduce a more scientific approach to benchmarking gap analysis.

Keywords

Benchmarking Analysis, Multiple Attribute Decision Theory, Warehouse Operations

1. Introduction

This paper presents the research efforts conducted towards the completion of a funded research project by a team of researchers (UA team) from the Industrial Engineering department at the University of Arkansas, and the thesis requirements for Mr. James Oldham. The research project entailed two phases which examined the policies and procedures involving inventory record accuracy within a public sector warehouse operation. The focus of the study was to discover if current industry practices were applicable for increasing overall warehouse accuracy levels.

The first phase involved an in-depth review of three warehouses in the public sector and identified the critical performance metrics for each facility. The second phase involved a process benchmarking study focused on sampling other warehouses across multiple industries to identify best-in-class performance metrics for warehouse management. A gap analysis was performed during phase two that compared general industry results to the public sectors' current operating procedures to identify opportunities for adaptation. The purpose of the benchmarking study was to develop recommendations for specific policy changes that would lead to higher accuracy levels.

However, new research questions were identified once the phase two of the study was completed. Both the sponsoring public sector warehouses and the UA research team desired a more scientific analysis tool than classical benchmarking analysis provided. This prompted the investigation of new methods for their application to benchmarking studies. The UA team decided to investigate the use of decision-based methodologies for the benchmarking portion of this research. These methods would allow comparisons to be made over a broad range of criteria, and allowed the sponsoring organization and the researchers to define specific priorities within the research.

The motivation for this continued research stemmed from the process benchmarking study previously conducted by the UA research team. The UA team used one warehouse as the "home" processes and nineteen other warehouses as potential "best-in-class" organizations. Most benchmarking questions reported specific warehouse procedures. The UA team evaluated the results using traditional benchmarking analysis tools. After the study concluded, the sponsoring organization and the UA team found three areas that were not completely described from the previous analysis. These points were as follows:

- Only one metric was used to identify the best-in-class for a particular set of questions. No stratification of the data was performed for judging best practices.
- Subjectivity was not analyzed during the discovery of the best-in-class performer. Suggestions were based solely on the best-in-class performer for each individual metric.
- No sensitivity analysis was performed to investigate what effect marginal changes in each critical metric had on the identification of the best-in-class performer.

Several works by Edward H. Frazelle and Steven T. Hackman addressed improving warehouse performance by benchmarking [8,9,10]. However, these methods analyzed resource input and the resulting output. The actual policies of conducting inventories and warehouse practices were not present in these models. This prompted the investigation of using an alternate method of analysis for this benchmarking study.

2. Alternative Methodologies

The UA team desired a different method of evaluating the benchmarking data. AHP and its associated mathematical formulations were the most robust for discovering best practices [15], but the method only allowed for pair-wise comparisons between alternatives and required a hierarchy structure to be developed. The UA team felt this form of an AHP structure was not warranted and prompted the investigation for an alternative method. Based on the positive studies with AHP and its reliance on mathematical theory for discovering best practices, other methods of multi-attribute decision theory (MADT) were examined for their applicability to benchmarking.

2.1 Multi-Attribute Decision Theory

Multi-attribute decision theory (MADT) provides an easily understood, yet comprehensive set of quantitative and qualitative approaches to justify a decision between alternatives [3]. Identifying the best-in-class performer from this benchmarking study is, in fact, a decision based on information gathered during the study. A specific type of MADT, called multi-attribute utility theory (MAUT), is examined for its applicability to this benchmarking study. Utility theory takes into account a range of the consequences of a particular decision and the risks of this decision, just as probability theory does for uncertainty. The following sections outline the basis for MADT and the eventual decision to use MAUT for analysis purposes.

2.2 Multi-Attribute Decision Theory Methods

The first and most basic method of MADT discussed here is the use of comparison tools. These methods range from scorecard evaluation of weighing factors to more complex use of polar graphs and evaluation ratings [4]. A second method of MADT uses an analytic hierarchy process (AHP). This method allows quantifiable and intangible criteria to be broken down and measured in its simplest form [14]. The third example of MADT concerns utility models. This method uses a prescribed mathematical relationship to decide between alternatives, using either an additive or a multiplicative relationship between attributes [11]. For this research, each model was examined. Table 1 shows the relationship between these methods in a benchmarking context.

Table 1 – Comparison of multiattribute decision theory methods

	Comparison Tools (Clemen, 1996) [4]	Analytical Hierarchy (Saaty, 1980) [14]	Utility Models (Fishburn, 1970) [6]
Scaling Factors	Actual unit	Ratio; priorities	Interval; utiles
Evaluation of Preferences	Single gap	Pair-wise comparison	Trade-offs
Weighting	N/A	Normalized ratio from eigenvalues	Assigned
Synthesis	Varies	Additive, eigenvectors	Additive, multiplicative
Structure	Varies (majority are graphical)	Hierarchical	Matrix or tree
Feedback from method	Limited view of best-in-class performance	Synthesis; consistency Measure; technique produced weightings	Synthesis

The ideal MADT method would be applicable to benchmarking analysis and has potential for novel benchmarking analysis. Basic tools such as radar graphs and scorecards have already been used for benchmarking studies, which were also used during the previous analysis. The UA team felt the results obtained from these techniques did not contain robust suggestions for improvement for this study. The data did not fit into a AHP structure for evaluation, as the critical metrics did not need to be broken down into sub-elements for comparison purposes [13].

It was decided to apply the utility method for analyzing the benchmarking data. This method is known as multi-attribute utility theory (MAUT). This method was chosen due to its relative ease of both formation and computation. The MAUT approach enables the decision maker to incorporate preference and value trade-offs for each metric and measure the relative importance of each [11]. Though other MAUT studies have been performed, there still exists a need for documented applications of this type of analysis [16]. Bordley (2001) describes the use of MAUT to perform gap analysis for service research. The resulting gap analysis discounts the gap between performance and expectations, providing more empirical inferences than conventional gap analysis [2]. It is expected that MAUT applied to benchmarking will have the same benefit. The following section outlines the basic details of this method.

2.3 Multi-Attribute Utility Theory

The basic goal of multi-attribute utility theory (MAUT) is to substitute information with an arbitrary measure called utiles so that the information can be compared. The utile values range from a low of 0 to a high of 1, with intermediate values decided upon by the decision maker. That is, the critical metrics identified are plotted on a graph from 0, being the worst case, to 1, being the best case. Then, a utility curve is plotted to model the subjective value of each outcome [5]. To illustrate this, consider the case of being hungry and receiving a slice of pizza. The first slice is very attractive, and so may be the next slice. However, additional slices will become less attractive as the hunger dissipates. Figure 1 illustrates the utility curve for this scenario.

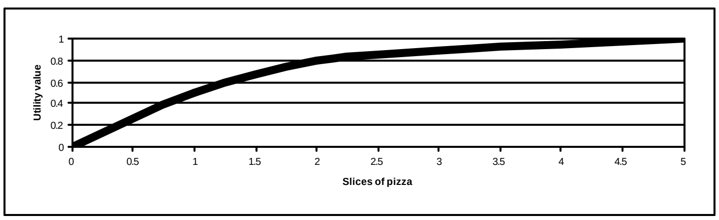

Figure 1 – Value of additional slice of pizza

Conversely, consider the case of a fire truck responding to an emergency call. The earlier the truck arrives, the more lives can be saved. As response time increases, the value of the truck arriving decreases. Figure 2 illustrates the utility curve for this scenario.

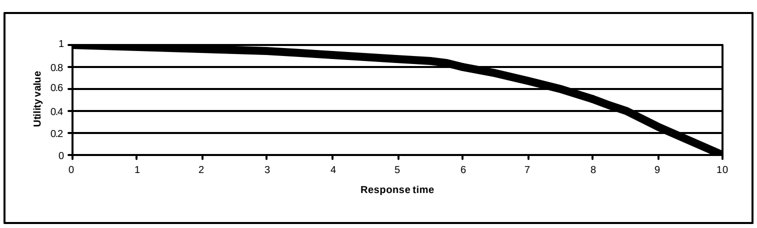

Figure 2 – Value of increasing fire truck response time

The most common method of curve fitting involves the decision maker identifying the lower and upper bounds for each curve, then identifying intermediate points on the curve. That is, other metric values between these extreme points are decided upon by the decision maker, which further defines the shape of the curve. This process continues until the decision maker is comfortable with the overall curve structure. The curve can then be interpreted by hand or by common spreadsheet software [5]. Figures 3-6 show graphical representation of different types of utility curves. The predetermined points are marked on each curve.

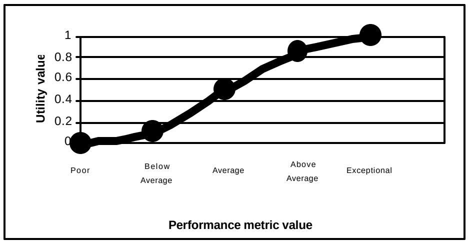

Figure 3 – S-curve relationship

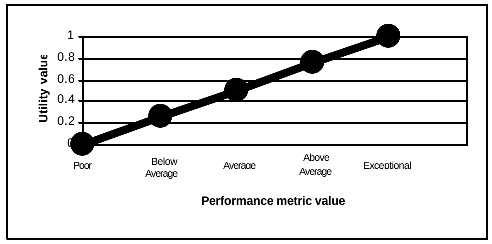

Figure 4 – Linear relationship

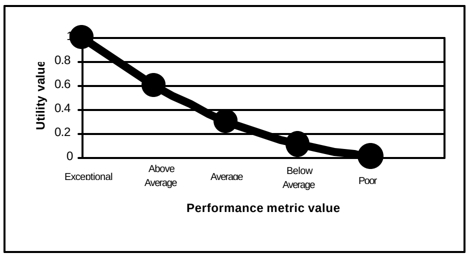

Figure 5 – Negative curve relationship

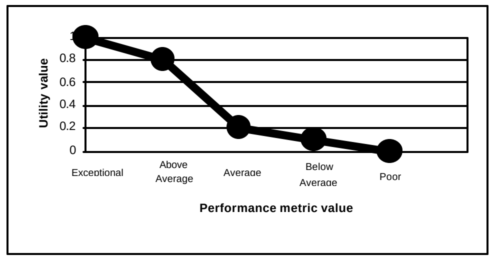

Figure 6 – Linear relationship, multiple slopes

The end result of MAUT is simply to maximize the combined utility value [11]. That is, each metric is assigned a utile value and is combined with other utile values to assess an aggregate utility value according to set mathematical procedures. These procedures are explained in detail in the next paragraph. MAUT allows the decision maker to develop reasonable preference criteria, determine which assumptions are most appropriate, and assess the resulting utility functions [12].

Two types of multi-attribute utility theory are common in current literature: additive and multiplicative utility theory. For i alternatives with j attributes, the additive utility model is expressed as:

$$U(x_i) = \sum_{j=1}^{n} k_j * u_j(x_{ij}) \tag{1}$$

$$\sum_{j=1}^{n} k_j = 1.0 \tag{2}$$

where:

1. k_j is a relative weight factor of the *j*th attribute.
2. $u_j(x_{ij})$ is the utility of the outcome x_{ij} for the *j*th attribute.
3. All attributes are independent of each other.

For i alternatives with j attributes, the multiplicative utility model is expressed as:

$$U(x_{ij}) = \frac{\prod_{j=1}^{n}[K * k_j * u_j(x_{ij}) + 1] - 1}{\text{K}} \tag{3}$$

$$\sum_{j=1}^{n} k_j \neq 1.0 \tag{4}$$

where:

1. k_j is a relative weight factor of the *j*th attribute.

2. $u_j(x_{ij})$ is the utility of the outcome x_{ij} for the *j*th attribute.

3. K is a scaling constant found by:

$$1 + K = \prod_{j=1}^{n}(1 + K * k_j) \tag{5}$$

and must be found iteratively.

4. All attributes are independent of one another; -1 < K < 0 implies utility independence.

Based on the various types of decision-based methodologies the UA team chose the additive utility method primarily for the following reasons:

- Additive utility theory (AUT) provides a more practical methodology due to easier computational analysis.
- AUT is easier to understand and explain relative to multiplicative utility theory.

AUT allows the benchmarking party to assign priorities to certain metrics and allows stratification of all critical metrics. Also, AUT can be applied using common spreadsheet software, which are readily available in most business settings. No components of the formulation require complex iterative solutions. Therefore, it is appropriate to apply additive utility theory for the purposes of this research. This analysis method uses subjectivity in formulating the relative weight factors (k_i), which therefore requires sensitivity analysis to be conducted to ensure robustness of assessment.

3. Conclusions

In essence, a formal MAUT analysis forces the benchmarking party to clearly define its priorities and measure the attractiveness of a discovered best practice. This is especially crucial in benchmarking studies, as the effects of the study are far-reaching throughout the organization [7]. As benchmarking studies continue to become more complex, traditional benchmarking tools do not apply to new research [1]. The selection of the MAUT technique allowed the UA research team to address the three previously mentioned areas of further analysis of benchmarking data. MAUT allows a benchmarking team to evaluate each metric, weight it against other alternatives, and pick a best-in-class representative for making recommendations. The anticipated short-term benefit of this research is to add robustness for suggestions presented to the public sector warehouse sponsor. The anticipated long-term benefit of this research is the increased benefit of benchmarking methodologies through MAUT analysis.

4. References

1. Ammons, D., 1999, “A Proper Mentality for Benchmarking,” Public Administration Review, 59(2), 105-109.
2. Bordley, R. F., 2001, “Integrating Gap Analysis and Utility Theory in Service Research,” Journal of Service Research, 3(4), 300-309.
3. Canada, J. R., and Sullivan, W. G., 1989, Economic and Multiattribute Evaluation of Advanced Manufacturing Systems, Prentice Hall, New Jersey.

4. Clemen, R. T. T., 1996, Making Hard Decisions: An Introduction to Decision Analysis, International Thompson Publishing, Washington, D. C.
5. Daellenbach, H. G., 1994, Systems and Decision Making, Wiley, New York.
6. Fishburn, P. C., 1970, Utility Theory for Decision Making, Wiley, New York.
7. Forger, G., 1998, "Benchmark your Warehouse for Future Success, Modern Materials Handling," 53(12), 39-41.
8. Frazelle, E. H., 1992, "How to Benchmark Warehouse Operations," Distribution, 91(9), 60-64.
9. Frazelle, E. H., and Hackman, S., 1994, "The Warehouse Performance Index: A Single-point Metric for Benchmarking Warehouse Performance," MHRC Final Report, Atlanta, #93-14.
10. Hackman, S. T., Frazelle, E. H., Griffin, P. M., Griffin, S. O., and Vlasta, D. A.,2000, "Benchmarking Warehousing and Distribution Operations: An Input-Output Approach," Georgia Institute of Technology, Atlanta.
11. Keeney, R. L., and Raiffa, H., 1993, Decisions with Multiple Objectives: Preferences and Value Tradeoffs, Wiley, New York.
12. Lindley, D. V., 1985, Making Decisions, Wiley, New York.
13. Pandey, P. C., and Kengpol, A., 1994, Selection of an Automated Inspection System using Multiattribute Decision Analysis. Asian Institute of Technology, Bangkok.
14. Saaty, T. L., 1980, The Analytic Hierarchy Process, McGraw-Hill, New York.
15. Saaty, T. L., 1995, "Transport Planning with Multiple Criteria: The Analytical Hierarchy Process Applications and Progress Review," Journal of Advanced Transportation, 29(1).
16. Walls, M. R., 1995, "Integrating Business Strategy and Capital Allocation: An Application of Multi-objective Decision Making," The Engineering Economist, 40, 247-266.