Reviewing and improving performance measurement systems: An action research

Renata Gomes Frutuoso Braz, Luiz Felipe Scavarda, Roberto Antonio Martins

A performance measurement system (PMS) is a vital part of a company’s managerial system (Neely et al., 2005). Despite the great attention of scholars and practitioners to designing a PMS, design is not the most important process. There are three other important processes: implementing, using, and reviewing/updating (Franco-Santos et al., 2007; Nudurupati et al., 2011).

Implementing a PMS stimulates managerial changes and promotes organisational learning by acquiring, storing, analysing, interpreting, and distributing data and knowledge about performance (Garengo et al., 2007). During implementation, many challenges arise due to fear, politics, and subversion (Kennerley and Neely, 2002).

If implementation succeeds, the challenge becomes how to properly measure and maintain the system’s relevance. When new measures are added, old measures are rarely deleted, which raises the PMS complexity (Neely, 1999; Kennerley and Neely, 2002), making the review process very important. The increase in complexity is not the only reason for outdated performance measurement systems; changes in strategy or in the external and internal environment of organisations can also contribute to the inability of the PMS to support decision makers (Bititci et al., 2001; Kennerley and Neely, 2002). The reviewing process aims to constantly update the PMS (Kennerley and Neely, 2002; Lohman et al., 2004).

In this context, the updating process should maintain alignment between the PMS and organisational strategy. The main steps in the updating process are the following: the review of targets and current measures, the development of new measures, and the challenge of strategy to improve the PMS (Bourne et al., 2000). In a recent literature review, Nudurupati et al. (2011) pointed out there is a need for longitudinal studies that explore and explain how a PMS within an organisation evolves in response to changes in the organisation’s internal and external operating environments. Other authors have highlighted the importance of reviewing, updating, and improving a PMS during the use stage (Kennerley and Neely, 2002, 2003; Nudurupati and Bititci, 2005; Bititci et al., 2006).

To address this gap, the following research question has been considered: how should a company accomplish the review/update process of its PMS? This article presents the empirical findings of our action research in reviewing and updating an existing PMS of a large multinational company in the energy sector and contributes to filling an identified gap in the literature.

The company’s decision makers and others involved in the process have criticised the PMS in the maritime transportation area. The studied company’s PMS is complex, with many performance measures, and it supports different decision makers at different hierarchical levels. These decision makers were willing to review and update the PMS.
To reach this target, the existing PMS was evaluated to identify, describe, and analyse the major performance measures used for the maritime transportation area and related to the performance of the fleet. An initial proposition to improve the studied PMS was presented based on a literature review, data collected from the company's information system, and interviews with those responsible for the PMS to analyse its current status. A PMS prototype in electronic spreadsheets was developed based on interviews, which were also used to evaluate and validate the improvement suggestion.

This article includes the following sections: PMS theoretical foundations, research design, analysis of the current PMS, suggestions for improvement, and conclusions.

2. Performance measurement system

Neely et al. (1995) define performance as the efficiency and effectiveness of actions within a business context. Effectiveness is compliance with customer requirements, and efficiency is how the organisation’s resources are used to achieve customers’ satisfaction levels.

Performance measurement is the process of quantifying efficiency and effectiveness. To do this, performance measures should be chosen, implemented, and monitored.

Performance measures are the metric used to quantify the efficiency and/or effectiveness of actions of part of or of an entire process or a system in relation to a pattern or target (Fortuin, 1988; Neely et al., 1996).

The performance measures should capture the essence of organisational performance (Gunasekaran et al., 2004). A lack of well-defined criteria to evaluate the performance of individuals and an organisation makes it difficult to plan and control the operations of an organisation and motivate its employees (Globerson, 1985).

A literature review illustrates a high correlation among the use of performance measures, objectives, and organisational strategy. This view considers these measures essential elements for planning and strategic control cycles (Neely et al., 1997). The use of such measures evaluates the strategy (Kaplan and Norton, 1996) and without them, decision makers cannot be certain whether their objectives were achieved (Goold and Quinn, 1990).

Other authors reinforce the importance of aligning performance measures with organisational strategy (Globerson, 1985; Eccles, 1991; Neely et al., 1996; Beamon, 1999; Bourne et al., 2000; Kaplan and Norton, 2001; Lambert and Pohlen, 2001; Coyle et al., 2002; Attadia and Martins, 2003; Lohman et al., 2004; Kaplan and Norton, 2006; Chen, 2008).

The Balanced Scorecard (BSC) is the most frequently applied framework used by companies worldwide to translate strategic objectives into a set of actions and performance measures. The BSC arranges the measures in four perspectives: (i) financial, (ii) customers, (iii) internal processes, and (iv) innovation and learning (Kaplan and Norton, 1993).

Performance measures should also be related to other hierarchical organisational levels, such as the tactical and operational levels (Lambert and Pohlen, 2001; Attadia and Martins, 2003) and should be linked to decision-making processes and provide control over the lowest levels of the organisational hierarchy (Gunasekaran et al., 2004).

Regarding to the aggregate level of the measures, a higher aggregation level implies a lower related cost. In contrast, it also suggests a decrease in reported accuracy and in the manager's ability to quickly detect the cause of the operational problem and to make an appropriate decision.

Disaggregating the performance measures generates a large number of performance measures. If the number is too large, the decision-making and control processes could be more difficult (Globerson, 1985). Therefore, the number of measures at the corporate level should be controlled through questioning the real need for the information generated (Martins and Costa Neto, 1998).

In contrast, any single performance measure will present such a myopic view that it will allow the managers to perform better without necessarily contributing to the organisation’s competitiveness. Therefore, managers should seek a good balance in terms of the performance measures that are employed (Kaplan, 1983).

Several performance measures were observed in the literature review. Shepherd and Günter (2006), for example, reported 132 measures, which were classified under three categories: (i) quantitative and qualitative, (ii) cost and non-cost, and (iii) quality, time, flexibility and innovation. Despite this categorisation, they indicated a lack of consensus in the literature on the best way to classify performance measures.

This reinforces the different classifications reported by Kaplan (1983), Neely et al. (2005), Fitzgerald et al. (1991), Kaplan and Norton (1993), Beamon (1998, 1999). Gunasekaran et al. (2001), De Toni and Tonchia (2001), Stephens (2001), Coyle et al. (2002), Chan (2003), Chan and Qi (2003), Hausman (2003), Huang et al. (2004), Lockamy and McCormack (2004), Lohman et al. (2004), Li et al. (2005), Neely et al. (2005), and Krakovics et al. (2008). Although distinct, the categories above can be used together. For details, see Shepherd and Günter (2006).

Some authors agree on the main characteristics considered for performance measures. Good performance measures are quantitative and have objective values instead of subjective ones. They should be straightforward and easy to understand to enable a quick identification of what is being measured and how it is being measured; practical with appropriate scales; consistent and maintain meaning over time; and clear on the objectives. Good measures also encourage adequate behaviour; are visible to everyone involved in the process; are understood and defined mutually; contain process inputs and outputs; measure only what is considered important; are multidimensional, demonstrating the existing trade-offs; and have a good cost–benefit relationship, prioritising collecting costs, and analysis over the expected benefits (Globerson, 1985; Fortuin, 1988; Neely et al., 1996, 1997; Coyle et al., 2002).

Another important consideration is defining performance measure attributes. Designing a performance measure involves more than just providing a complex formula. Issues, such as the meaning of the measure, the frequency of the measurement, and the source of the data, should be considered (Neely et al., 1997). Table 1 shows a list of attributes and their descriptions.

Franco-Santos et al. (2007) have surveyed the literature in search of PMS definitions. Rather than elaborating another definition, these authors provide the main characteristics of a PMS:

- features (properties or elements);
- roles (purposes or functions); and
- processes (series of actions that constitute a PMS).

The features are the performance measures and the supporting infrastructure, which can vary from manual to automated mechanisms, to acquire, collate, sort, analyse, interpret, and disseminate appropriate information to the decision makers. There are five roles: (i) measuring performance, (ii) strategy management, (iii) communication, (iv) influencing behaviour, and (v) learning and improvement. The processes can be grouped into five categories: (i) selection and design of measures, (ii) collection and manipulation of data, (iii) information management, (iv) performance evaluation and rewards, and (v) system review. The system review is the focus of this paper.
A PMS consists of a set of performance measures, such as software, a database, and procedures that thoroughly and consistently measure the performance of actions taken (Lohman et al., 2004). This system helps managers make decisions by collecting, compiling, analysing, and disseminating an organisation’s data and valuable information (Neely, 1998).

A PMS should be able to accomplish the following: to provide data to monitor past performance and plan future performance; to provide a balanced accounting of the organisation; to demonstrate how results are related to decisions; to prevent the inclusion of conflicting measures; to reinforce organisational strategies; to be compatible with the organisational culture and the available reward systems; and to provide data for external comparison (benchmarking) (Neely et al., 1996).

There are three stages in developing a new PMS: design, implementation, and use. The design stage identifies the customers and stakeholders’ needs and considers business objectives and a framework for adequate performance measures and their attributes (Bourne et al., 2000). When there is an existing PMS, the use stage is the starting point for any change, and then it is followed by the reflect, modify, and deploy stages. There are key organisational capabilities in each stage, except for the use stage, which are grouped into: process for reviewing, modifying, and deploying the measures; the people who are skilled to use, reflect, modify, and deploy the measures; the culture that ensures the value of measurement; and the systems that enable the collection, analysis, and reporting of appropriate data (Kennerley and Neely, 2002, 2003).

The validation of these measures and procedures can be supported by the use of prototypes for visualising deviations in the results and tendencies that can be represented by a learning curve (Globerson, 1985) to allow the inclusion of requirements and changes (Davies, 1982; Lohman et al., 2004). It is also important to consider data availability in the design stage because it is useless to design a system in which data are difficult to obtain or are unavailable (Lohman et al., 2004). The design of a feedback loop that takes action based on discrepancies between the targets set and actual performance also takes place during this stage (Globerson, 1985).

The PMS implementation stage requires the development of procedures to collect, process, and disseminate data, allowing the measurement to be performed precisely (Neely, 1998; Bourne et al., 2000). The management information system supports the implementation stage (Garengo et al., 2007; Nudurupati et al., 2011). One of most important tasks in this stage is to inform employees and make them aware of the new performance measurement system (Ukko et al., 2007). The implementation of some performance measures before the design stage has finished explaining, in many situations, the overlap between the implementation and use of performance measures (Bourne et al., 2000).

The pace of the PMS implementation progress can be increased by early involvement of information technology (IT) experts, the use of tools for recovering and manipulating data, and resource allocation (Bourne et al., 2000). Management information systems are critical factors in the success of PMS implementation (Garengo et al., 2007; Nudurupati et al., 2011), particularly in data collection, analysis, presentation, and dissemination (Neely, 1999).

In the PMS use stage, the performance measures change according to the internal and external environment; therefore, the PMS evolves and should be updated and improved (Lohman et al., 2004). Only a few organisations have processes to manage the evolution of a PMS (Neely, 1999). Most organisations use specific redundant measures, mainly because they rarely feel comfortable eliminating obsolete measures (Neely, 1999; Kennerley and Neely, 2002).

To improve the PMS during the use stage, it is necessary to accomplish the following tasks: (a) develop a procedure to periodically review the entire set of measures according to changes in the competitive environment and in the strategic approach (Wisner and Fawcett, 1991); (b) include an effective mechanism to review measures and attributes for continuous improvement and to agree on actions, such as arranging regular meetings for directors and managers responsible for the performance being measured (Bourne et al., 2000); (c) develop measures as the performance and circumstances change; and (d) be useful in challenging strategic suppositions (Bourne et al., 2000).

Managing the evolution of performance measurement systems consists of a number of stages that have, so far, received little attention from researchers. There is also little discussion on what to do when the change is necessary to avoid redesign of the entire current performance measurement system (Kennerley and Neely, 2002, 2003). Moreover, in the literature on longitudinal studies, there are claims that investigate how to review and update a PMS (Nudurupati and Bititici, 2005; Bititici et al. 2006; Nudurupati et al., 2011). The next section presents the research design used to develop a proposal for changing the PMS of an energy company, and the change process is presented below.

3. Research design

There are many different methods that can be applied in operations management, including surveys, case studies, action research, and modelling and simulation (Forza, 2002; Voss et al., 2002; Coughlan and Coghlan, 2002; Bertrand and Fransoo, 2002). In this article, the choice was based on the research question, the state of the art in the field, and the characteristics of the study.

The main characteristics of an action research study are as follows: it investigates more than actions; it is participatory; it
occurs simultaneously with the action; and it is a sequence of events and approaches used to solve problems, i.e., a longitudinal study. Furthermore, the action research is appropriate when the research is related to understanding the process of change, or its improvement to learn about it (Coughlan and Coghlan, 2002). Therefore, an action research design is the most appropriate research method because it develops a longitudinal study to review and update a company's PMS. The major research steps are presented next.

3.1. Step I: evaluation of the current measurement system

The semi-structured interviews and the data collected from the information systems, technical reports, and internal documents of the company were the information sources used to identify and describe the current PMS, as indicated by Davies (1982), Globerson (1985), Kaplan and Norton (1993), Lohman et al. (2004), Wouters and Sportel (2005) and Schmidberger et al. (2009). Based on these results, structured interviews to analyse the PMS were conducted.

A total of eleven employees involved in the maritime transportation area at the studied company were interviewed. All the employees were involved with the development, implementation, and use of the system, and they included the transportation director and four managers. As shown in Table 2, the interviews were guided by questionnaires consisting of relevant questions reported in the literature for the PMS analysis.

3.2. Step II: improvement proposition

The non-structured interviews and the workshops held for improving the current PMS took place with the same interviewees included in the system's description and analysis stages. The choice of performance measures and attributes was based on Neely et al. (1997), Neely et al. (2002), and Lohman et al. (2004). The process of choice was interactive, and the measures and the measurement procedures were developed and adjusted according to the availability of information regarding strategy, customers, and processes. The interviewees validated the measures and measurement procedures. The proposed measures were stored in an electronic spreadsheet that included company historical data, as suggested by Globerson (1985), which enabled the inclusion of requirements and changes indicated by Davies (1982) and Lohman et al. (2004). The availability of data was considered according to Lohman et al. (2004).

4. Analysis of the current performance measurement system

The strategic objectives related to maritime transportation on the company's Balanced Scorecard were the following: to have competitive transportation costs, to reach excellence for logistic services, assuring reliability in terms of timely delivery and quality; and to achieve standards of excellence in terms of social and environmental responsibility, complying with the corporate requirements of safety, environment, and health.

To verify whether the maritime transportation area was contributing to the business, the following performance measures were designed based on the analysis of the performance of the company fleet.

The Maritime Transportation Unit Cost (MTUC) measures the cost of a ton-mile produced. It is computed by dividing the total amount spent on maritime transportation by the total produced in thousands of ton-miles. The data are gathered with Enterprise Resource Planning (ERP) software and an information system developed in-house. A lower value achieved by the measure leads to better performance.

The Operational Availability Index (OAI) measures the operational reliability of the fleet. It is calculated by the average of the total duration percentage of the contracts of each on-hire ship. Hire can be defined as the daily rent paid by the ship's charterer. During the time the ship is off-hire, all expenses are the responsibility of the owner. When the ship is available, it is on-hire. The measure target is reviewed annually, considering that the ships are allowed 72 h off-hire annually for repair, excluding the scheduled dry dockings.

The Ship Performance and Reliability Index (SPRI) measures the performance of the fleet based on events related to reliability and operational health and safety- (HSE) related aspects of the ships. The events of non-compliance with schedules, breakage or damage are related to the reliability of the ship. Crash, pollution, contamination, and loss of cargo are related to HSE. The performance of the ship is related to the operational aspects (consumption, speed, and size). This measure starts with a maximum score (100) and decreases according to the occurrence of each event.

The Maritime Transportation Efficiency Index (MTEI) measures the efficiency of the maritime transportation regarding dead freight. It is represented as the relationship between the ton-mileage accumulated by the ships and its navigation potential within this period. The total ton-mileage performance of the ships is divided by the ideal total ton-mileage performance of the ships (distance multiplied by 95% of the deadweight tonnage of the ship) that corresponds to the trips made by those ships.

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Relevant questions for the performance measurement system evaluation.</th>
</tr>
</thead>
<tbody>
<tr>
<td>References</td>
<td>Relevant questions for the evaluation of performance measurement systems</td>
</tr>
<tr>
<td>Beamon (1999)</td>
<td>(a) What is being measured? (b) How frequently is the measurement performed? (c) When and how are the measures reevaluated?</td>
</tr>
<tr>
<td>Neely et al. (1995), Neely et al. (2002), and Neely et al. (2005)</td>
<td>(a) Which performance measures are used? (b) What are they used for? (c) What benefit do they provide? (d) Are the measures related to the business unit's objectives? (e) Are some measures used for benchmarking?</td>
</tr>
<tr>
<td>Kennerley and Neely (2003)</td>
<td>(a) Does the measure definitely assess what it is supposed to assess? (b) Does the measure assess only what it is supposed to assess? (c) Is the measure consistent regardless of when it is performed or who performs it? (d) Can the data be promptly communicated and easily understood? (e) Is there any possibility of ambiguity in data interpretation? (f) Is it possible to take actions based on the data? (g) Can the data be analysed quickly enough so that actions can be taken? (h) Are collection and analysis worth the cost?</td>
</tr>
<tr>
<td>Neely et al. (2002)</td>
<td>(a) Is there any measure that should be discontinued?</td>
</tr>
</tbody>
</table>
The Product Spill Index (PSI) measures the performance of the fleet regarding the environment. It is measured monthly, and the target is to achieve zero volume of product spill. Table 3 shows the synthesis of the major results obtained through the current performance measurement analysis. The first column shows the relevant questions used in the interviews, which were based on the questions from Table 2. The other columns present the results for each performance measure.

The analysis of the data in Table 3 indicates that the performance measure PSI should be maintained without changes. It is in accordance with both the characteristics reported in the literature for a good performance measure and with the objectives of the company’s maritime transportation area. With respect to the OAI, this measure should undergo changes only in the actions concerning the data measured. Nevertheless, the analysis of the MTUC, SPRI, and MTEI indicated the need for adjustments in the measurement method or in the measure design. This also meant changes in the PMS, largely concerning the use of information technology. Next, details for these last three measures are provided.

### 4.1. Maritime transportation unit cost (MTUC)

Most interviewees noted the MTUC as a clear measure, despite it being difficult to understand and impractical. The difficulty is that the measure considers the various ship sizes jointly, which prevents the simple and direct identification of the reason for the deviations between real and planned performance. As reported by Globerson (1985), this aggregation level makes the analysis complex, as a superficial analysis can lead to incorrect conclusions and the need to justify variations that may not reflect the inefficiencies of a specific class of ship.

To compute the measurement, costs are allocated in different action periods (before and after), which can lead to misinterpretations. For example, the payment for fuel is made 30 days after filling the tank, and the freight payment is made at the end of the trip. This results in the allocation of costs in a period different than when the event occurred. Because the measurement frequency is monthly, distortions are generated due to cost accounting.

In addition, the ton-miles are only accounted for at the end of the trip, which also generates distortions. Therefore, a trip that begins in a certain month and ends in a different month will have its ton-miles only at the end of the trip. Generally, this takes place up to three months after the trip has begun, leaving the ton-mile measurement field blank for the previous months.

### 4.2. Ship performance and reliability index (SPRI)

The interviews verified that most interviewees consider the SPRI to be a difficult measurement to understand, and most of them classified the formula used to calculate the ship’s performance as complex. The use of different types of variables in the same measure and the aggregated result of all such variables per ship do not allow identification of which specific variable had the highest weight in the performance in a certain month.

This can lead to ambiguity when analysing the results. For example, if the ship loses points solely due to pollution or accident, then the ship could be regarded as being within the acceptable limits and treated as a ship that had lost the same number of points due to other minor events. However, in the information system, only the aggregated result per ship is available.

The subjectivity in determining the points related to each event occurrence and the lack of proportion among the points attributed to each event was also observed. Therefore, it can be stated that the scales and units are not appropriate. The data are gathered manually and involve several people. The task is not integrated into the information system, so the collection and analysis of the data are not worth the cost.

### 4.3. Maritime transportation efficiency index (MTEI)

Like the MTUC, the MTEI considers in the same measure various sizes of ships, making it difficult to identify inefficiencies and take action based on performance measures. Generally, the data cannot be analysed fast enough for the firm to take action in an appropriate time. Due to this aggregation, most interviewees considered the measure relatively unclear and difficult to understand, as Globerson (1985) argues.

### 5. Improvement proposal for the current performance measurement system

This section proposes a new PMS for the studied company's maritime transportation area. The scope of change is the set of performance measures in some of their attributes, and the

---

**Table 3**

Synthesis of the analysis of the current performance measures.

Source: elaborated by the authors.

<table>
<thead>
<tr>
<th>Are the measures clear?</th>
<th>MTUC</th>
<th>OAI</th>
<th>SPRI</th>
<th>MTEI</th>
<th>PSI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td></td>
<td></td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Are the measures practical?</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Are the measures easy to understand?</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Are data collection and analysis worth the cost?</td>
<td>Yes (collection)/no analysis</td>
<td>Yes</td>
<td>No</td>
<td>Yes (collection)/no analysis</td>
<td>Yes</td>
</tr>
<tr>
<td>Do the measures definitely assess what they are supposed to?</td>
<td>Not the way it is structured</td>
<td>Yes</td>
<td>No</td>
<td>Not the way it is structured</td>
<td>Yes</td>
</tr>
<tr>
<td>Are the objectives of the measure aligned with the objectives of the maritime transportation area?</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Is it possible to take actions based on the data?</td>
<td>No</td>
<td>Not always</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Are the scales and units appropriate?</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Is there any possibility of ambiguity in results interpretation?</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Can the data be analysed quickly enough so that actions can be taken?</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Is the measure consistent no matter when it is performed?</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Is the frequency of measurements adequate?</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Can the data be promptly communicated and easily understood?</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>
introduction of a mechanism for updating the system continuously. First, the changes in the performance measures of the system and in the measurement process are presented.

5.1. The operational availability index (OAI)

With respect to the OAI, the system should generate an alert for the operations area manager, warning about ships with an OAI lower than the target. This will enable the identification of ships with low reliability, which is not possible by analysing only the OAI average in aggregate.

IT personnel will need to be involved to implement the alert in the information system. The measurement process will include a procedure where the operations area manager will receive a warning about ships with an OAI lower than the target.

5.2. Maritime transportation unit cost (MTUC)

One of the most significant criticisms about this measure was the level of aggregation. An ideal proposition, though difficult to implement, and a viable proposition, which is easier (more realistic) to implement, are presented for this measure.

Both proposals, the ideal and the viable, consider the creation of a disaggregated measure per class of ship as the sum of the total cost incurred and the total ton-mileage produced that is unable to allow the identification of the most troubled classes. Therefore, all classes will be given the same weight, so these propositions allow verification of whether each class is being well managed. It is worth mentioning that disaggregating the measure per ship was not in the proposal because of the difficulty of setting targets per ship due to dynamics of the routes. In addition, the interviews indicated that this level of detail is not necessary. Regarding the level of aggregation, these propositions are aligned with Globerson’s (1985) recommendations.

The ideal proposal provides the correct allocation of the costs, solving the problem related to the right time to make the allocation of costs and the apportionment of the ton-mileage produced by each ship among the number of days for each month of the trip. To correctly allocate the costs and apportion the ton-mileage produced, it would be necessary to restructure the company’s payment systems. Profound changes would be required to render the data adequate. According to the validation interviews, the cost of gathering the data would be prohibitive. Lohman et al. (2004) argue that designing a system with data that are difficult to gather or even unavailable is useless.

The viable solution considers the disaggregation of the measure per class of ship using a three-month moving average (the last three months) to measure the monthly cost and ton-mileage. The moving average lessens the effect of the method used to determine the ton-mileage produced, which is accounted only in the month the trip ends. It also mitigates the issue related to the proper time to allocate the costs.

The use of the three-month moving average can be justified by the fact that most trips begin in a given month and end up to two months later. This period of time is compatible with the question related to the right time to allocate costs. The interviews regarding validating the changes indicated the disaggregation per class of ship and the use of the three-month moving average were satisfactory. It is worth highlighting that the use of the two- and four-month moving averages were also studied, but the results were not satisfactory.

The MTUC value disaggregated per class of ship should be weighted by the ton-mileage produced by each ship, allowing for the same class of ships, but with different productivities, to be treated with different weights. The MTUC of a certain class of ship will correspond to the weight by the ton-mileage of the MTUC of that class of ships. Similarly, the MTUC of a ship corresponds to the weighting by ton-mileage of the MTUC of the routes of this ship. The targets set for each disaggregated measure will use the same logic used to set the target for the aggregated measure for the fleet, but the variables will be for the respective data of each class.

Disaggregating the MTUC increases the managers’ ability to quickly identify the source of the operational problem and respond adequately, as stated by Globerson (1985). At the executive level, maintaining the current formula of the aggregated fleet’s MTUC is necessary. The proposal to maintain the existing aggregation is in accordance with the organisational level, which is hierarchically higher, as reported by Attadia and Martins (2003) and Lohman et al. (2004).

To implement a viable proposal, IT must enable the automatic generation of the three-month moving average in the information systems, prepare the new data input procedure and allow the new measures to be visualised in the system.

5.3. Ship performance and reliability index (SPRI)

Based on both a critical analysis of the empirical findings and a literature review, the proposal for this measure is to separate the events to reduce the redundancy of some events and to highlight others.

The event of breakage or damage has already been considered for the calculation of OAI as well as the events of pollution, measured by PSI, which provides more thorough information because it also includes the amount of product spilled. So, there is no need to maintain these events in another measure. This action reduces redundancies in the PMS. These redundancies originated when measures were created in different periods without analysing the whole set, which could have eliminated an existing variable, as pointed out by Waggoner et al. (1999).

The analysis of speed and consumption is already performed monthly by the ship operators due to the performance clauses in the contracts and based on which ship is penalised. Therefore, it is not necessary to maintain these data in the performance measures.

For accidents, a measure per ship, denoted as NACCIDENTS, was proposed to measure the number of accidents per ship. Similar measures, such as non-compliance with the schedule and cargo contamination or loss, were disaggregated to create performance measures denoted as TERM and CARGO DAMAGE respectively.

The targets for the new measures NACCIDENTS and CARGO DAMAGE were determined as zero, as these events represent a direct impact on the company’s image. For the TERM measure, the target was based on an analysis of the time series and the validation interviews, and an annual reduction of the target, aiming for continuous improvement similar to that reported by Fortuin (1988), was proposed.

The proposal also highlights events that should be recorded directly into the company’s information system by the ship operators and not into electronic spreadsheets that record daily events (current method). By doing so, there should be fewer registration errors and more data consistency. The involvement of IT is crucial to properly include specific fields for this data in the information system. A change of procedures is needed, as the ship’s operators should directly include data in the information system.

5.4. Maritime transportation efficiency index (MTEI)

For the MTEI, disaggregation in classes of ships was proposed, which enables the identification of the classes with inefficiencies. To obtain the MTEI of each class, the same measure of each ship in
a class should be weighted by the ton-mileage produced by each ship in the period studied.

The MTEI for each ship in a certain period should be weighted by the ton-mileage produced by the route the ship made. This requires changing the calculation formula to incorporate these data. The new targets were set based on the company’s maritime transportation planning and on the time series. The inclusion of yearly improvement indexes for continuous improvement was also considered. Again, the involvement of the IT area is critical for modifying the information system to provide reports indicating both the ideal ton-mileage and the ton-mileage produced by each ship on each route.

5.5. Prototype for evaluating the new performance measures

The historical data were applied to electronic spreadsheets to test the new performance measures and to validate the proposed changes. In this article, due to space restrictions, only the analysis of the MTUC of each class and fleet is presented as an example of the change process. The maritime transportation managers and the other interviewees validated the results.

The data concern two classes of ship, denominated Classes 1 and 2. These two classes were chosen because they accounted for approximately 70% of the total ton-mileage movement for the studied company. The three-month moving average was also applied for computing the cost and ton-mileage produced by each class. For confidentiality reasons, the values presented in Figs. 1 and 2 were changed, but the proportions were maintained so that the analysis performed would not be affected.

According to the current PMS, the values obtained for the MTUC of the fleet (aggregated level) during the analysed period were satisfactory, as they were below the maximum value set as the target for the entire period. Nevertheless, as illustrated in Fig. 1, the analysis made for each class of ship (disaggregated level) shows that Class 1 operated with MTUC values higher than the maximum set as the target for this performance measure during most of the studied period, indicating a poor performance. The disaggregation proposed per class of ship allows decision makers to quickly identify the source of the operational problem and respond adequately, as highlighted by Globerson (1985).

Despite the satisfactory global performance of the fleet, the Class 1 ships had a poor performance. Moreover, it can be seen...
in Fig. 1 that the use of the three-month moving average better represents performance, as it lessens the effect caused by concentrating cost accounting or ton-mileage in certain months.

From the data analysis of the current method, in some months, 60% of the ships of this class did not have their ton-mileage accounted for because they were travelling. This created distortions in the fleet's performance analysis. Fig. 1 shows this distortion for months 5 and 6. The analysis of the data also indicates that these two months presented similar costs, but month five showed a low ton-mileage figure, as few Class 1 ships finished their trips in that month. This results in extremely high values for MTUC in this class. In contrast, the majority of these ton-miles were accounted in month six, lowering the value measured that month. In view of this, there are considerable variations in the graphic result from the current method. Fig. 1 exhibits such distortions from the current performance measures.

Fig. 2 illustrates how aggregated performance measures, such as MTUC, can distort and hide performance inefficiencies. The graph allows the comparison of MTUC using the three-month moving average method for Classes 1 and 2 and for an aggregate measure of both classes. The behaviour of the measure proposed for the Class 2 MTUC during the analysed period is satisfactory, indicating that this class shows good performance. It can also be seen that the good performance of Class 2 offsets, to a great extent, the poor performance of Class 1 when both are associated, resulting in a more aggregated measure.

5.6. Summary of PMS changes

The measurement of the pursuit of the objectives of the maritime transportation area is grouped as follows: cost (MTUC), HSE (PSI and NACCIDENTS), and level of logistic service provided (OAI, MTEI, CARGO DAMAGE, and TERM). Table 4 summarises the changes proposed to the company's PMS. For each performance measure, both ideal and viable changes are presented where applicable and propose the necessary changes to implement a solution. The director, the managers, and the other interviewees validated these changes.

5.7. Procedures for continuously updating the PMS

The PMS change process in the company convinced the directors and managers of the importance of updating the company's PMS. Therefore, it is important to also propose a procedure to guarantee a continuous updating of the PMS to keep pace with the competitive environment.

The procedure proposed to the studied company consists of an annual review of the PMS by the director and managers involved

---

**Table 4**

<table>
<thead>
<tr>
<th>Current measure</th>
<th>Proposition/suggested change</th>
<th>Procedures necessary for implementation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MTUC</strong></td>
<td>Ideal proposition</td>
<td>- Maintain MTUC for the fleet (aggregated level).</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Create a MTCU for each class of ship (disaggregated level).</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Correct allocation of all costs in the formula.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Apportionment of the ton-mileage produced by each ship among the number of days related to each month of the trip.</td>
</tr>
<tr>
<td></td>
<td>Viable proposition</td>
<td>- Maintain MTUC for the fleet (aggregated level).</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Create a MTUC for each class of ship (disaggregated level); weighed by the ton-mileage produced by each ship.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Use three-month moving average analysis.</td>
</tr>
<tr>
<td><strong>OAI</strong></td>
<td></td>
<td>- Maintain OAI for the fleet (aggregated level).</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Create alert mechanism in the information system regarding the ships with OAI values lower than the desired level (disaggregated level).</td>
</tr>
<tr>
<td><strong>SPRI</strong></td>
<td></td>
<td>- Remove SPRI in an aggregate way eliminating redundancies already considered in other measures and creating new aggregated measures.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Create a new measure, NACCIDENTS, for the event of Accidents/pollution of the SPRI.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Create a new measure, CARGO DAMAGE, for the event of contamination/cargo losing of the SPRI.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Create a new measure, TERM, for the event of non-compliance with the due date of the SPRI.</td>
</tr>
<tr>
<td><strong>MTEI</strong></td>
<td></td>
<td>- Maintain MTEI aggregated for the fleet.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Create a MTEI disaggregated for each class of ship.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Change calculation formula of MTEI.</td>
</tr>
<tr>
<td><strong>PSI</strong></td>
<td></td>
<td>- Maintain without changes.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Restructure the information system enabling the apportionment of the ton-mileage.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Restructure the cost allocation system.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Implement the targets proposed for each measure.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Implement the IT area for enabling the automatic generation of the three-month moving average by the information systems, allowing the new measures to be fed into the system, and allowing such measures to be visualised in the system.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Implement the targets proposed for each measure.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Involve the IT area for the creation of an automatic alert mechanism to the manager of the operations area about ships with OAI values that are lower than desired.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Procedure Change: the manager of the operations area would receive the warning about ships with lower OAI values.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Involve the IT area for including this information in the information system.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Implement the targets proposed for each disaggregated measure per class of ship.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Involve the IT area for the report preparation to verify the ton-mileage produced and the ideal ton-mileage for each ship on every route to feed the new measure calculation formula.</td>
</tr>
</tbody>
</table>
with the performance of that area in accordance with Wisner and Fawcett (1991), Bourne et al. (2000), Kennerley and Neely (2002, 2003) and Lohman et al. (2004). This procedure formalises the change process, which represents the object of this action research study.

6. Conclusions

The focus of this article was a longitudinal study on the performance measurement change process of an energy company's maritime transportation area to fill the gap pointed out by Kennerley and Neely (2002, 2003), Nudurupati and Bititci, (2005), Bititci et al. (2006), and Nudurupati et al. (2011). The review and improvement proposal of the current PMS suggested a new set of performance measures based on the objectives defined by the company for this area. They were evaluated in terms of cost (MTUC), HSE (PSI and NACCIDENTS), and the level of logistic services provided (OAI, MTEI, CARGO DAMAGE, and TERM). The proposition also contemplates changes in the measurement process for procedures and information systems. The director, managers and users of the PMS for the maritime transportation area validated the proposal after the simulation using historical data on an electronic spreadsheet, as recommended by Globerson (1985).

The changes in the company's PMS considered concepts proposed and discussed in the literature reviewed, such as the aggregation and disaggregation of measures, redundancies in the PMS, data availability, and constant target updating for continuous improvement. As Kennerley and Neely (2002, 2003) pointed out, the use of PMS was the starting point for the entire change process.

The disaggregation of some measures in the classes of ships can improve the ability of decision makers to quickly identify the source of an operational problem and take appropriate action, in accordance with Globerson (1985). However, for top management, the use of aggregated measures related to the fleet is more appropriate due to the kinds of decisions made. This proposal is coherent with the high hierarchical level, as reported by Attadia and Martins (2003) and Lohman et al. (2004), and with the decision-making level, as reported by Gunasekaran et al. (2004).

The proposal aims to reduce PMS redundancies that originated from the fact that measures are created in different periods and without an analysis of the whole set, which could eliminate an existing variable, as highlighted by Waggoner et al. (1999).

The proposal also examines data availability. According to Lohman et al. (2004), it is useless to design a PMS where data are difficult to obtain or are unavailable. In light of this, some procedures and measurement formulas were adequate for the available data in the company's systems and operations. This led to a viable change rather than an ideal change.

To set the measures' targets, the values should be gradually reduced on an annual basis, as stated by Fortuin (1988), who highlighted that once the target is achieved, a new target that is more challenging yet realistic should be set. The values of the new targets should be set to improve the system during annual meetings where those responsible for performance are present.

Regarding implementation of the improvement proposal, changes in the procedures and the involvement of information technology to optimise the structure of the existing data in the database used by the maritime transportation team are necessary. Information systems are fundamental to implementing and using a PMS, as reported by Garengo et al. (2007) and Nudurupati et al. (2011).

The findings of our longitudinal study have shown the difficulty and complexity in reviewing and updating the PMS of an energy company's maritime transportation area. The difficulty is related to the involvement of users of the PMS, the assessment of performance measures, the establishment of targets, and data availability. Complexity is related to changes in information technology to implement the changes in procedures for computing the performance measures. Therefore, this contributes to a better understanding of the process of reviewing and updating an existing company's PMS. However, it is necessary to carry out more longitudinal studies on this issue to continue to fill the gap.

Acknowledgement

The authors gratefully acknowledge the management team from the energy company analysed here for their invaluable assistance with this research project, and MCT/CNPq (research project: 309455/2008-1) and CAPES/DFG (BRAECGRIM research project: 010/09). The authors are also very grateful to the anonymous referees for their constructive suggestions.

References


