Performance evaluation of Groundforce Portugal luggage management process

Beatriz Medeiros Melo

Department of Engineering and Management, Instituto Superior Técnico, Universidade de Lisboa

May, 2014

Abstract

The luggage management process is one of the most important operations for handling agents, in which mishandled bags imply high costs for airlines and lead to dissatisfaction from passengers. Aiming to foster continuous improvement of operations, Groundforce Portugal (GF), as a handling agent, assesses the performance of this process using a set of performance indicators established in accordance with the company's strategic goals.

This paper has two main goals: to determine whether the performance indicators used by GF are relevant to improve the luggage management process, or if there are other measures, complementary or substitute, which better fit the objectives of the company; and to create a system that allows monitoring the operating performance of the company, at the baggage management level.

This work develops a causal map that structures the problem and from which are defined the key performance indicators (KPIs), which are used in a multi-criteria evaluation model created to measure the aggregate attractiveness allowing analyzing the company's global performance in its luggage management process. In order to improve the monitoring of the process performance is built a dashboard, using the traffic lights methodology.

Keywords: handling, luggage management process, performance evaluation, KPIs, multi-criteria analysis, dashboard.

1. Introduction

As an integral process in air traffic, transporting baggage is a challenge for the industry in that it faces the problem of mishandled bags. A bag is considered mishandled when it is lost or damaged, when not arrives on time at destination, or when it is stolen. The main causes for mishandled bags are transferring between flights (53%) at an annual cost of around 1.3 billion dollars, problems associated with loading luggage on airplanes (15%) and also problems inherent safety, error ticketing and baggage exchange among passengers (13%) (SITA 2012).

If mishandled, it is the airline which carried the luggage and its respective ground handling agent that bear the costs and make the logistics process for the recovery of the luggage. Although the values are alarming, in the last seven years there has been a decline in the number of mishandled bags due to the effort and investment in new technology and processes by aviation industry. It is in this context that Groundforce Portugal (GF), provider of ground handling services, invests in mechanisms for continuous improvement through analysis of indicators in order to reduce baggage loss and consequent cost reduction.

This work aims to analyze the indicators currently used at GF and to investigate their application and relevance in the luggage management process at Lisbon Airport.

The remainder of this paper is organized as follows: section 2 presents the company that is used as case study; section 3 reviews the methods and approaches to address the problem; section 4 describes the case study, including the KPIs used and the multi-criteria evaluation model developed for GF; section 5 presents the dashboard developed for GF and section 6 concludes the paper with some final remarks and a proposal of future work.

2. Case Study: Groundforce Portugal

GF is a leader in the provision of ground handling services and operates in Lisbon, Porto, Funchal and Porto Santo airports, ensuring

handling services (i.e., passenger service, board service, cargo service, maintenance of ground equipment, custom assists, lounges, load control, aircraft cleaning, ground transportation, airport representation, flight operations and crew management, airport security, management facilities, advice and general assistance, and delivery of luggage) to more than 150 airlines. In general, the company ensures all ground operations required by airlines and has a team of employees to ensure the needs of the market and its customers, fulfilling the motto "We Handle, You Fly" (Groundforce 2013).

This work fits the key process PC-04 - "handling the plane, move the luggages safely and on time" because there was the need to ascertain the relevance of the performance measurement indicators for this process, since the GF created a monthly report of luggage, which contains several performance indicators. That report serves to inform about the baggage operations performance, identify problem areas and develop strategic solutions to solve problems in order to better use the resources available, reduce costs and improve operations in a continuous base. The report is produced taking into account the data of airline TAP.

2.1. PC-04: Performance Indicators

The monthly report of the performance evaluation for luggage management process is based on information about passengers and baggage and provides the following family indicators (non-financial measures) for the industry:

- Number of left behind bags (AHL – Advise in Hold);
- Number of damage bags (DPR – Damage Property Report);
- Number of unclaimed bags (OHD – On Hand Bags);
- Number of bags edited by Edit Team (responsible for reconciling luggage left behind with the respective passenger);
- Number of bags handled by the Tail to Tail team (responsible for carrying transfer luggage from one plane to another, without introducing the volume in the baggage sorting system);
- Number of baggage searchers on flights;
- Luggage delivery time;
- TAP punctuality on arrivals and departures.

3. Literature Review

3.1. Causal Maps

According to Axelrod (1976), causal maps are a network composed of nodes representing a concept, which are connected by arrows that indicate the type of connection: causality, influence or implication. Connections in network are assigned a positive or negative sign indicating, respectively, a positive or negative influence between a concept and another. It is also possible to connote each connection with its strength, e.g., strong, moderate or weak. (Montibeller et al. 2007)

It is important to define the notions of concept and link: concept – short phrase that expresses an idea; link – expresses a relation between two concepts.

3.2. Multi-Criteria Decision Analysis

The multi-criteria decision analysis has a number of formal approaches that seek to take into account several criteria in order to help individuals (decision-makers) to explore alternatives that really matter (Belton & Stewart 2002). There are several methods that produce a quantitative measure that aims to represent performance attractiveness, for example: AHP, SMART, MACBETH. (Belton 1999; Belton & Stewart 2002). However, AHP has several known issues (see, e.g. Dyer 1990; Bana e Costa & Vansnick 2008), which is a sufficient reason to not use this method. The other two methods are in line with multi-attribute value theory (MAVT) (Belton 1999; Belton & Stewart 2002), thus both can be applied. However, MACBETH only requires qualitative judgments from a decision-maker to assess value scales and to weigh the criteria, whereas SMART requires quantitative judgments; so the former can be a better choice when dealing with decision-makers with higher fluency (Fasolo & Bana e Costa 2014). Another alternative, for decision-makers with higher numeracy (Fasolo & Bana e Costa 2014), is to use the numerical techniques usually applied in SMART, such as the direct rating method (Edwards 1977) or the bisection method (Von Winterfeldt & Edwards 1986) to create value scale and the swing weighting

procedure (Edwards & Barron 1994), or, despite not being used in SMART, the trade-off procedure (Keeney & Raiffa 1976). Given that MACBETH also has a companion software tool (M-MACBETH –Bana e Costa et al. 2012)) that validates the consistency of decision-maker' answers, we opted to use this approach.

3.2.1. The MACBETH Approach

MACBETH (Measuring Attractiveness by a Category-Based Evaluation Technique) is an interactive process to support multi-criteria decision analysis used to help build a quantitative value model based on qualitative value judgments. The objective of this technique is to measure the attractiveness of options, from a non-numeric comparison method that asks the decision-maker about the difference in attractiveness between two options (or two performance levels), e.g. "for *x* and *y* such that *x* is preferred to *y*, the difference in attractiveness between *x* and *y* is: very weak, weak, moderate, strong, very strong or extreme?" (Bana e Costa & Vansnick 1999; Bana e Costa et al. 2012). While other methods require quantitative judgments, MACBETH only requires qualitative judgments about attractiveness differences between each two alternatives or levels of a performance descriptor in order to generate value in a range of criteria, and requires only qualitative judgments about differences between fictitious alternatives to weigh the criteria. Being a MAVT approach, the construction of a MACBETH multi-criteria model encompasses three phases: structuring, evaluation and testing.

Structuring

- Criteria definition: criteria allow comparing alternatives according to the fundamental points of view (FPV) of the decision-maker..

- Performance descriptors definition: is an ordered set of plausible performance levels in terms of a FPV and is intended to serve as a basis for describing the effects of the alternatives regarding to a FPV.

Evaluation

- Value functions constructions: the goal of this step is to transform performance into value, using simple additive model, as shown on Equation 1 (Bana e Costa et al. 2012):

$$V(a) = \sum_{j=1}^{n} w_j \, v_j(a) \quad (1)$$

With $\sum_{j=1}^{n} w_j = 1, w_j > 0$ and

$$\begin{cases} v_j\,(good) = 100 \\ v_j\,(neutral) = 0 \end{cases}, j = 1, \dots, n$$

Being:

$V(a)$ – overall value of alternative *a*

w_j – weight of criterion *j*

$v_j(a)$ –partial value of alternative *a* on criterion *j*

$v_j\,(good)$ e $v_j\,(neutral)$ – assigned values to "good" and "neutral" performance levels, respectively, for each criteria *j* (which in this case are defined arbitrarily as 100 and 0, respectively).

- Criteria Weighting: the weighting coefficients required for the alternatives evaluation in multi-criteria models can also be obtained by assessing qualitative judgments from decision-makers.

Testing

To test the model and its outputs, should be done a requisite analysis, in order to define if the model form and content are sufficient to solve a particular problem (Phillips 1984).

3.3. Dashboards

Few (2006) defines a dashboard as a visual presentation of the most important information needed to achieve one or more pre-defined goals. Also argues that the information available from a set of KPIs should be consolidated and organized under the combination of text and graphics in a single panel so they can be monitored at a glance. Andra (2006) suggests the following steps for the development of a dashboard:

a. Define the dashboard goals;
b. Understand the organization's business model;
c. Define the key areas that must be measured and included in the dashboard;
d. Identify the benefits;
e. Define KPIs;
f. Select tools and methodology;
g. Proactive management for the results.

4. Case Study

4.1. Structuring

The problem was laid out through a causal map. The map was created in collaboration with

members of the three entities involved in the luggage management process (GF, ANA and TAP). The following question was posed to each one of them: “which are the factors that affect the performance of the luggage management process and why?”

From these interviews, we conclude that GF planning – correct dimensioning of resources for operations to be executed –, TAP punctuality and infrastructure capacity, as well as baggage handling system reliability are the three crucial elements for the performance of the luggage management system. Subsequently, the question “which factors affect GF’s performance regarding the luggage management process?” was posed to the GF decision-makers, so as to provide the detailed description in the map presented in Figure 1.

The causal map, made using the Decision Explorer software tool, allowed, in a first instance, to distinguish external from internal factors that affect the performance in study. Internal factors, which are GF’s responsibility, were marked with a circular border, so that they could be separated from the external elements.

In order to list the most important concepts in the map, so that all of them are taken into account when evaluating performance, a central analysis was made, which pointed out the ten most central concepts:

1. TAP delay on departure (non-compliance with the Service Level Agreement – contract for provision of services between the handling agent and the airline);
2. Dissatisfaction of the TAP/PAX customer;
3. Failure when handling luggage;
4. Establish GF responsibility / Others (RL Codes);
5. Rise in the number of irregularities in the operation;
6. Flaw when handling equipments for ground handling services;
7. Manage baggage flows in the Baggage Sorting System (BSS);
8. Real Time Control performance;
9. GF planning department performance;
10. Rise in the number of left behind luggage.

The concepts that are not directly connected to the luggage management process were excluded from further study: dissatisfaction of the TAP/PAX customer, rise in the number of irregularities in the operation, Real Time Control performance (software that adjusts

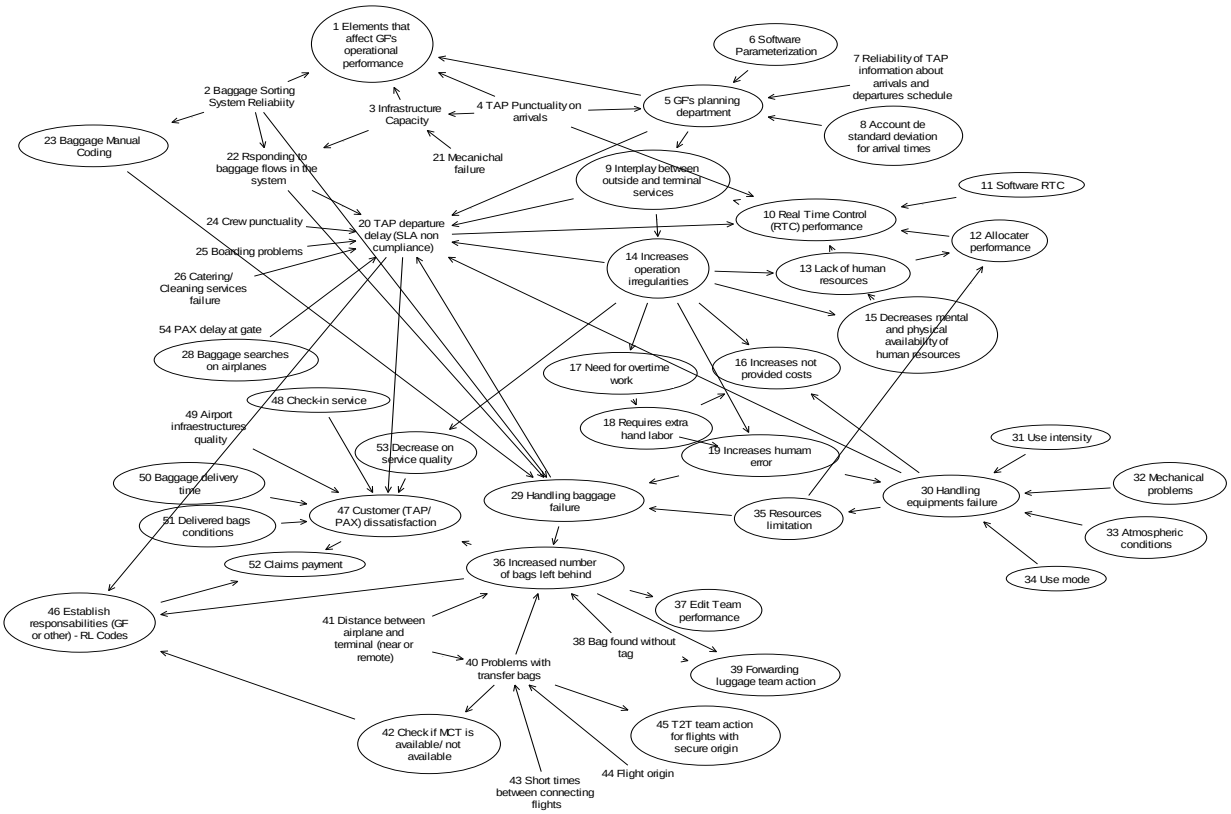

Figure 1. Causal map for luggage management process performance

operations planning to the daily reality) and GF planning department performance. Given the indicators monitored by GF and the FPV previously presented, we concluded that the following indicators should be added to the analysis of the luggage management process performance: lack of equipments, BSS capacity and performance of the Lost & Found department, which indicates the work developed in problem solving regarding left behind or damaged luggage.

4.2. KPIs

The following KPI are taken into account to evaluate the performance of the luggage management process:

KPI 1: TAP Punctuality on arrivals – percentage of the flights whose difference between the scheduled time of arrival (STA) and the actual time of arrival (ATA) is less than 15 minutes.

KPI 2: TAP Punctuality on departures – percentage of the flights whose difference between the scheduled time of departure (STD) and the actual time of departure (ATD) is less than 15 minutes.

KPI 3: Standard LB Bags – industry standard value for the number of left behind baggage for each 1,000 passengers carried.

KPI 4: LB Bags GF Resp percentage – percentage of the total left behind baggage by GF's responsibility.

KPI 5: Standard LB Bags GF Resp – industry standard value for the number of left behind baggage, by GF's responsibility, for each 1,000 passengers carried.

KPI 6: TTL LB Bags RL 21 – total number of left behind baggage, by GF's responsibility, given the baggage was not loaded onto the airplane, in spite of being correctly labeled.

KPI 7: TTL LB Bags RL 51 - total number of left behind baggage, by GF's responsibility for the cases in which the passengers are redirected to other flights and their luggage remain in the original route.

KPI 8: TTL LB Bags RL 52 – total number of left behind baggage, by GF's responsibility, given that the transfer of luggage from one airplane to another, when they belong to different airlines, was not executed within the established MCT.

KPI 9: TTL LB Bags RL 55 – total number of left behind baggage, by GF's responsibility, given that the transfer of luggage from one airplane to the other, when they belong to the same airline, was not executed within the established MCT (Minimum Connection Time of 60 min).

KPI 10: Standard Bags OUT Resp – industry standard value for the number of left behind baggage, by others' responsibility, for each 1,000 passengers carried.

KPI 11: TTL LB Bags RL 56 – total number of left behind baggage by external responsibility, given that the transfer of luggage from one airplane to the other, when they belong to the same airline, was not executed once the MCT was reduced.

KPI 12: TTL LB Bags RL 62 – total number of left behind baggage by external responsibility, given that for extraneous reasons (e.g. meteorology) the luggage do not reach the airplane in time.

KPI 13: TTL LB Bags RL 26 – total number of left behind baggage by external responsibility, given that the luggage was not loaded onto the aircraft, once it did not have space or because it was overweight.

KPI 14: TTL LB Bags RL 76 – total number of left behind baggage by external responsibility, given that the luggage was found with no label.

KPI 15: TTL LB Bags RL 54 – total number of left behind baggage by other handling agents' responsibility, when they do not deliver the luggage in time to the handling agent responsible for the flight where the volume will proceed.

KPI 16: TTL LB Bags RL 53 – total number of left behind baggage by external responsibility, given that the transfer of luggage between two airplanes from different airlines was not executed once the MCT was reduced.

KPI 17: L&F LB Performance – Lost & Found department performance ratio concerning left behind baggage.

KPI 18: TTL Damaged/Pilferage Bags – total number of baggage that the passenger claims to be damaged and/or violated.

KPI 19: L&F DPR Performance – Lost & Found department performance ratio concerning damaged/violated baggage.

KPI 20: WTC opened cases for unclaimed baggage – total number of cases relating to unclaimed baggage opened in the World Tracer platform.

KPI 21: TTL Bags without label – total number of baggage found without label.
KPI 22: BSS capacity – SSB processing capacity in baggage per hour.
KPI 23: TTL LB Bags RL 64 – total number of left behind baggage by responsibility of the BSS managing entity, given that the treadmill was out of service.
KPI 24: TTL Bags RL 59 – total number of left behind baggage by responsibility of the BSS managing entity, given that the luggage is not authorized to be loaded once the information given to it at the source (BSM – *Baggage Source Message*) is not recognized by the system in LIS.
KPI 25: Edit Team Recovery Ratio – analyses the *Edit Team* performance.
KPI 26: T2T ratio above MCT – analyses the T2T team performance for longer periods than 60 minutes.
KPI 27: T2T ratio below MCT – analyses the T2T team performance for shorter periods than 60 minutes.
KPI 28: Baggage searches – percentage indicative of the number of baggage searches in TAP flights when departing from Lisbon.
KPI 29: Average time per searched baggage – amount of time, in minutes, GF took to find the baggage in a certain flight.
KPI 30: 1st Bag Delivery to NB/ Remote Stand – percentage indicative of the number of times that the delivery time of the first baggage to a *Narrow Body*-type aircraft parked in a stand remote from the luggage terminal met the established time of delivery (≤ 25 minutes).
KPI 31: 1st Bag Delivery to NB/ Near Stand – percentage indicative of the number of times that the delivery time of the first baggage to a *Narrow Body*-type aircraft parked in a stand near the luggage terminal met the established time of delivery (≤ 20 minutes).
KPI 32: Last Bag Delivery to NB/ Remote Stand – percentage indicative of the number of times that the delivery time of the last baggage to a *Narrow Body*-type aircraft parked in a stand remote from the luggage terminal met the established time of delivery (≤ 30 minutes).
KPI 33: Last Bag Delivery to NB/ Near Stand – percentage indicative of the number of times that the delivery time of the last baggage to a *Narrow Body*-type aircraft parked in a stand near the luggage terminal met the established time of delivery (≤ 25 minutes).
KPI 34: 1st Bag Delivery to WB/ Remote Stand – percentage indicative of the number of times that the delivery time of the first baggage to a *Wide Body*-type aircraft parked in a stand remote from the luggage terminal met the established time of delivery (≤ 25 minutes).
KPI 35: 1st Bag Delivery to WB/ Near Stand – percentage indicative of the number of times that the delivery time of the first baggage to a *Wide Body*-type aircraft parked in a stand near the luggage terminal met the established time of delivery (≤ 20 minutes).
KPI 36: Last Bag Delivery to WB/ Remote Stand – percentage indicative of the number of times that the delivery time of the last baggage to a *Wide Body*-type aircraft parked in a stand remote from the luggage terminal met the established time of delivery (≤ 45 minutes).
KPI 37: Last Bag Delivery to WB/ Near Stand – percentage indicative of the number of times that the delivery time of the last baggage to a *Wide Body*-type aircraft parked in a stand near the luggage terminal met the established time of delivery (≤ 40 minutes).
KPI 38: Lack of equipment – percentage indicating the absence rate of equipment directly used in handling operations.

4.3. Performance Aggregation Measures

In order to evaluate the luggage management process performance in several criteria, one-by-one, and produce a set of aggregate measures to determine which group of indicators is critical for the performance, the following multi-criteria evaluation model was built using MACBETH. The value tree in Figure 2 shows (in red) the KPI that were defined. A continuous performance descriptor was associated with each KPI, with five benchmark levels (including the reference levels “good” and “neutral”, which are essential for weighing the criteria), as may be seen in the example presented in Table 1.

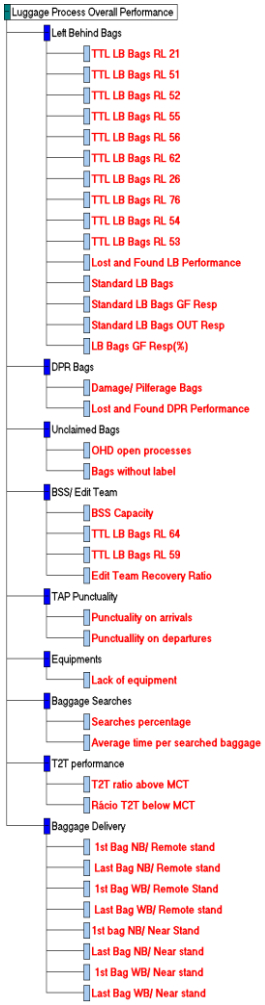

Figure 2. Value tree "Baggage Process Overall Performance"

Table 1. Performance descriptor and reference levels for KPI "Edit Team Recovery Ratio (%)"

KPI "Edit Team Recovery Ratio"		
Performance descriptor	Reference Levels (%)	
Represents the percentage of baggage with edit problems which were analyzed by Edit Team	L1	95
	L2 – "Good"	80
	L3	65
	L4 – "Neutral"	50
	L5	35

For each KPI was built a value function using the MACBETH method. The decision-makers (DM) were asked to sort the descriptor levels in descending order of preference, and were asked about the difference in attractiveness between each two levels, to which the DM answered qualitatively using the MACBETH semantic scale: extreme, very strong, strong, moderate, weak, very weak or null. For example, to build the value scale for the KPI "Edit Team Recovery Ratio" we questioned the DM about "the difference in attractiveness between the most preferred level 95 and the second most preferred level 80", to which the DM replied "moderate or strong" (see "mod -strg " in the cell indicated in Figure 3). The questions about the difference in attractiveness between every two levels of the descriptor were repeated. Sometimes the M-MACBETH software detected inconsistencies between judgments expressed by DM, just after being inserted into the software, which were shown to DM and then solved. With the MACBETH judgments matrix completed without inconsistencies, the software proposed a scale, which was shown to DM for validation and adjustment if needed, by comparing the value ranges between performance levels. Figure 4 shows the final value function for KPI "Edit Team Recovery Ratio", which resulted from the judgments shown Figure 3. This process was repeated for all KPIs, allowing building a value function for each KPI.

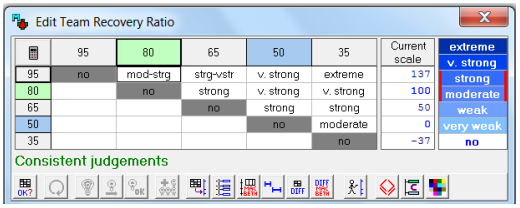

Figure 3. Judgments matrix for KPI "Edit Team Recovery Ratio"

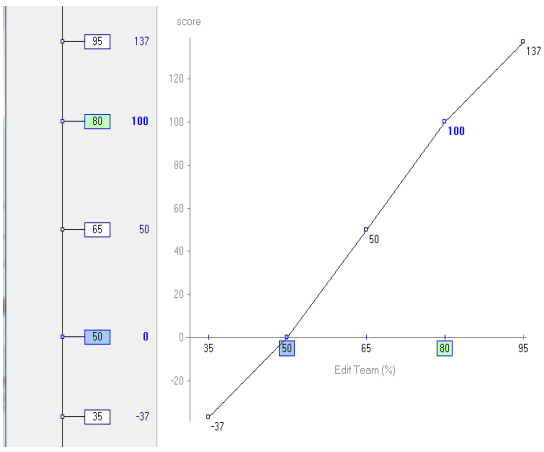

Figure 4. Value scale for "Edit Team Recovery Ratio"

Subsequently, and in order to weigh the KPI, the DM were asked to rank the KPI from the

most to the least attractive neutral-good swing, through the following question: "if you had a performance with impacts on the neutral level on all of the criteria and if you had the chance to improve the impact of that performance to the good level in only one of those criteria, which would it be? After ranking the criteria swings, the DM were asked to judge the increase of attractiveness in a given KPI that is 'upgraded' from the neutral to the good level, in order to complete the rightmost column. Finally, the difference of attractiveness between swings, two-by-two, were also judged, to fulfill the first row and the diagonal of the matrix. Figure 5 shows the technique used to help the DM when weighing the KPI. The judgments matrix in Figure 6 and the histogram in Figure 7 illustrate the weights obtained for each KPI of the "BSS/ Edit Team" family.

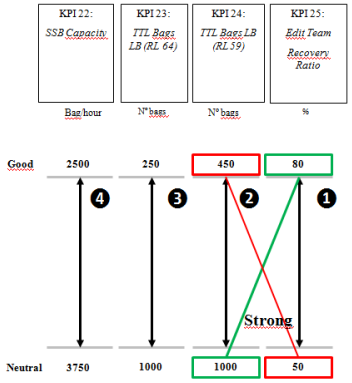

Figure 5. Used technique for weighting KPI

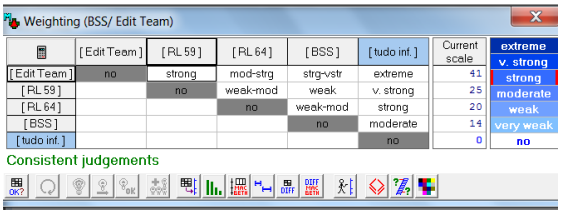

Figure 6. Weighting judgments matrix for the "STB/ Edit Team" family

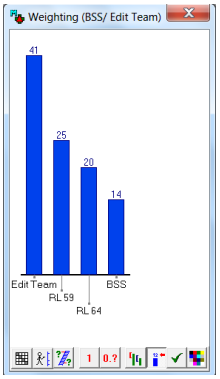

Figure 7. "BSS/ Edit Team" KPI's weights

The process of KPIs weighting illustrated above was repeated for each family of KPIs in the value tree. The KPIs weighting coefficients were applied to calculate, by simple additive model, the aggregate performance value of each KPIs family: TAP Punctuality, Left Behind Luggage, Damaged/Pilferage Luggage, Unclaimed Luggage, BSS/ Edit Team, T2T, Baggage Searching, Baggage Delivery and Equipment.

In order to calculate the overall performance value of the luggage management process was necessary to consider the hierarchical additive model, which corresponds to apply simple additive models bottom-up. To obtain the weights of the families of KPIs we used the KPI with the highest weight within each family: Bags LB Resp GF, Damage/Pilferage Luggage, OHD Processes, 1st bag Delivery NB/Remote Stand, Lack of Equipment, TAP Punctuality (Arrivals and Departures), Average time per searched baggage, Edit Team Recovery Ratio and T2T ratio below MCT. The referred KPIs where weighed according to the process described above, resulting the following matrix (see Figures 8 and 9) and respective family's weights (see Figure 10).

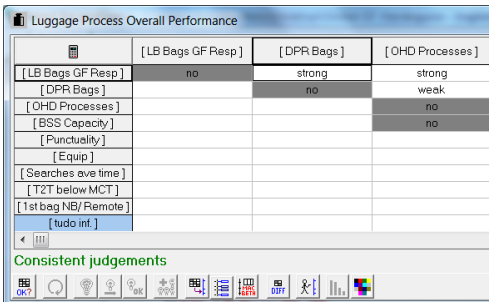

Figure 8. Judgments matrix for the KPIs with highest weights - part I

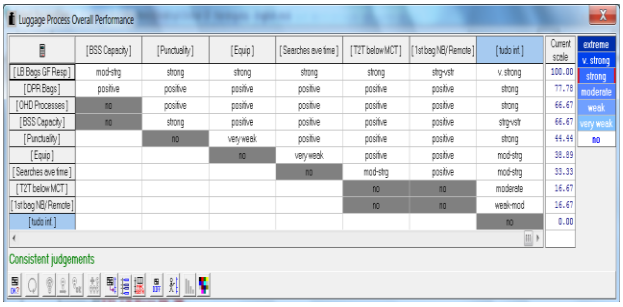

Figure 9. Judgments matrix for the KPIs with highest weights - part II

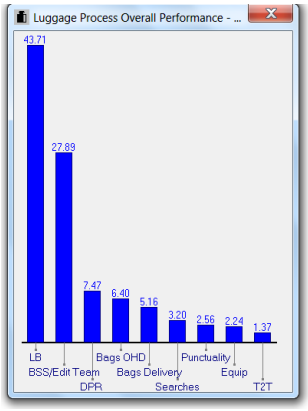

Figure 10. Family's weights histogram

Left Behind Luggage, BSS / Edit Team and DPR Luggage families are those that most influence the overall performance value, as are those with higher weights.

5. Dashboard

In order to monitor the new set of KPIs for assessing the baggage management process was programmed in Excel the digital dashboard presented in Figure 11, which has the following characteristics:

- Allows only to view the strictly relevant information to the process;
- Groups information by families in order to facilitate the identification of problem areas;
- Shows pre-defined conclusions from the application of the traffic lights methodology with three levels of colors (green, yellow and red) in order to demonstrate the performance (i.e. acceptable, investigate performance, or unacceptable, respectively) of a given indicator , thus helping to guide the user in data analysis;
- Shows information about the aggregate performance value of each family and the overall performance value of the luggage management process.
- Use graphs to display the group indicators Bags Left Behind and Baggage Delivery, focusing the attention of the user to the most relevant causes for luggage getting left behind and what the performance of GF in baggage delivery since it is an indicator which has a direct impact on passenger satisfaction.

6. Final remarks and future work

This work has met the goals previously established: determine whether the indicators used by GF are relevant to improving the process above, or if there are other measures, complementary or substitute, which better fit the objectives of the company; and creating a system that allows monitoring the operating performance of the company, at the baggage management level. For this it was necessary to build a causal map to support the creation of a MACBETH multi-criteria evaluation model and developing a dashboard to monitor the luggage management process.

The KPIs used are relevant to assessing the baggage management process performance and measures, properly, the process performance value, since the value functions translate

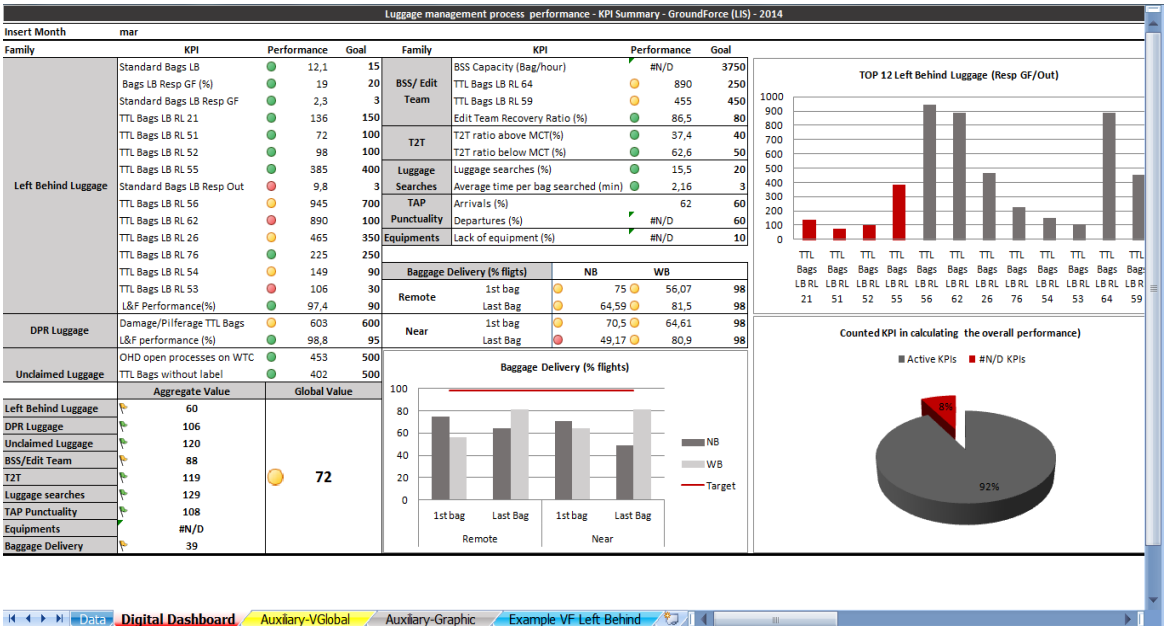

Figure 11. Dashboard for luggage management process - March 2014

correctly the value of each level of performance. However, and although, accordingly to the DM, the weights obtained represent properly the trade-offs between KPIs, it would be appropriate to conduct tests with historical data to verify the suitability of the model. This was not done by lack of information regarding to some KPIs. This aspect should not be neglected and so there is data, the model should be tested and, if necessary, calibrated by adjusting the model parameters. Also, the rules applied to the dashboard take into account that performance measures are compensatory, because it uses the additive model, but, in reality, some of these measures (e.g. baggage delivery for different aircraft types) may not be compensatory. Further studies should pay attention to this possibility.

At short term, the work developed for the Lisbon Airport should be replicated at other airports where the GF operates: Porto, Funchal and Porto Santo.

References

Andra, S., 2006. Action-Oriented Metrics for IT Performance Management. *Cutter IT Journal*, 19(4), pp.17–21.

Axelrod, R., 1976. *Structure of Decision: The Cognitive Maps of Political Elites.*, Princeton: Princeton University Press.

Bana e Costa, C.A. & Vansnick, J.-C., 2008. A critical analysis of the eigenvalue method used to derive priorities in AHP. *European Journal of Operational Research*, 187(3), pp.1422–1428.

Bana e Costa, C.A. & Vansnick, J.-C., 1999. The MACBETH approach: Basic ideas, software, and an application. In N. Meskens & M. R. Roubens, eds. *Advances in Decision Analysis*. Dordrecht: Kluwer Academic Publishers, pp. 131–157.

Bana e Costa, C.A., Vansnick, J.-C. & De Corte, J.-M., 2012. MACBETH. *International Journal of Information Technology & Decision Making*, 11(2), pp.359–387.

Belton, V., 1999. Multi-criteria problem structuring and analysis in a value theory framework. In T. Gál, T.J. Stewart, & T. Hanne, eds. *Multicriteria Decision Making Advances in MCDM Models, Algorithms, Theory and Applications*. Boston London: Kluwer Academic Publishers, pp. 12.11–12.32.

Belton, Valerie & Stewart, Theodor J., 2002. *Multiple Criteria Decision Analysis An Integrated Approach*, Kluwer Academic Publishers.

Dyer, J.S., 1990. Remarks on the analytic hierarchy process. *Management Science*, 36(3), pp.249–258.

Edwards, W., 1977. How to use multiattribute utility. Measurement for social decisionmaking. *IEEE Transaction on Systems, Man and Cybernetics*, 7, pp.326–340.

Edwards, W. & Barron, F.H., 1994. SMARTS and SMARTER: Improved simple methods for multiattribute utility measurement. *Organizational Behaviour and Human Decision Processes*, 60(3), pp.306–325.

Fasolo, B. & Bana e Costa, C.A., 2014. Tailoring value elicitation to decision makers' numeracy and fluency. Expressing value judgment in numbers or words. *Omega*, 44(0), pp.83–90.

Few, S., 2006. Clarifying The Vision. In *Information Dashboard Design: The Effective Visual Communication of Data*. O'Reilly Media, pp. 34–37.

Groundforce, 2013. Groundforce Portugal. Available at: http://www.groundforce.pt/ [Accessed February 20, 2013].

Keeney, R.L. & Raiffa, H., 1976. *Decisions with Multiple Objectives: Preferences and Value Tradeoffs*, New York: John Wiley and Sons.

Montibeller, G. et al., 2007. Reasoning maps for decision aid: an integrated approach for problem-structuring and multi-criteria evaluation. *Journal of the Operational Research Society*, 59(5), pp.575–589.

Phillips, L.D., 1984. A theory of requisite decision models. *Acta Psychologica*, 56(1-3), pp.29–48.

SITA, 2012. *SITA Baggage Report 2012*,

Von Winterfeldt, D. & Edwards, W., 1986. *Decision Analysis and Behavioral Research*, New York: Cambridge University Press.