# Evaluating Data Envelopment Analysis as a means of measuring van fuel efficiency

*Hubert Paul Bernard VIRTOS*

*A thesis submitted to the University of Huddersfield in partial fulfilment of the requirements for the degree of Doctor of Philosophy*

*The University of Huddersfield in collaboration with Masternaut Three X LTD*

*December 2010*

**Page left intentionally blank.**

# Copyright Statement

# Notes

All specific terms that are mentioned in this thesis are generally explained the first time they appear in the text. However, all the specific terms are also listed and defined in the 'Glossary of Terms and Abbreviations'. Most of these definitions specialise on a specific subject and a certain degree of necessary tacit knowledge is assumed to read the thesis. While this is fine where the subject is very specific, this is rather less so when the thesis brings knowledge from different specialities. Because this study is cross-disciplinary – spanning a gap between transport operations and performance measurement mathematics methods – it was thought best to explain most notions discussed in this thesis to avoid any potential confusion.

# Acknowledgment

Although the worth of this thesis still needs to be appraised, it would be unfair not to thank and mention all the people who have contributed to its writing and without whom I would perhaps have never completed this work. Thus, while trying to be concise but without forgetting anyone, I would like to first thank my parents who always supported me both financially and morally and without whom I would not be the man I am now; I would also like to thank Sara, my friend now wife, for her consistent support and words of encouragement during the ebb and flow of this thesis writing; thanks to Cruz Reyes for his useful comments in proof-reading this thesis; thanks also to Masternaut the company which trusted me enough to offer me a sponsorship for this thesis; thanks as well to Professor Alan Slater for his cryptic but constructive, smart and cunning comments – useful wisdom for those who can listen and decipher; thanks to Will Maden, who, although often late for reading, was always here when it really mattered; and finally, thanks to my supervisor Chris Savage who has always helped and supported me since the very beginning.

# Abstract

In order to improve fuel efficiency, fleet managers need methods to accurately measure fuel performance. Miles per gallon – the industry fuel efficiency standard measure – has several limitations. These relate to some aspects of fuel efficiency not reflected in the measure but also to the fact the measure cannot be interpreted without knowing some external factors (such as vehicle weight). This research addresses some of these limitations through the application – within three companies – of a Data Envelopment Analysis (DEA) model to van fuel efficiency measurement. In order to use the fuel information obtained from the fuel cards statements, it was necessary to develop a cleansing and smoothing algorithm which ensured that the data could be safely used in the models. The model results indicate that DEA provided a better and more comparable fuel efficiency measure while effectively addressing some key limitations of the mpg measure. The originality of this research comes from the limited amount of published literature on fuel efficiency measurement in road transport operations. Effectively, only a limited number of papers can be found on the measurement of road operations efficiency using DEA and, with the exception of this study, none could be found on van operations or fuel efficiency measurement. Debriefing discussions confirmed that the fleet operators appreciated the measure and also suggested that more research on fuel theft could be useful. Finally, the recent success of driver competitions seems to indicate there is a latent need in the industry for accurate driver performance measurement, which suggests that methods such as the one developed in this study could be of greater use in the near future.

# List of Contents


1. Introduction ........ 19
1.1. Background ........ 19
1.2. Hypothesis ........ 22
1.3. Aims & Objectives ........ 23
1.3.1. Aims ........ 23
1.3.2. Objectives ........ 24
2. Literature Review – Van Fuel Efficiency Measurement ........ 26
2.1. Explaining the Focus on Van's Fuel Efficiency ........ 26
2.2. On the Ways to Improve Fuel Efficiency ........ 29
2.2.1. Improved Scheduling and Network Optimisation ........ 29
2.2.2. CANbus Technology & Driver Training ........ 32
2.2.3. Fuel Card Management ........ 35
2.2.4. Traditional Fuel Saving Interventions ........ 40
2.2.5. A Word on the Ways to Improve Fuel Efficiency ........ 46
2.3. Explaining the Focus on Fuel Efficiency Measurement Based on Fuel Cards ........ 47
2.3.1. Improved Scheduling and Network Optimisation ........ 47
2.3.2. CANbus Technology & Driver Training ........ 48
2.3.3. Traditional Fuel Saving Interventions ........ 49
2.3.4. Fuel Cards Management and the Limitations of CANbus ........ 49
3. Literature Review – Performance Measurement ........ 53
3.1. Performance and Performance Measurement ........ 53
3.1.1. Key Concepts and Definitions ........ 53
3.1.2. Purpose of Measuring ........ 58
3.1.3. Type and Classification of Performance Measures ........ 62
3.2. Around the Performance Measure ........ 67
3.2.1. Data Gathering ........ 67
3.2.2. Possible Risks and Issues in Measuring Performance ........ 70

3.2.3. Performance Measurement Recommendations ..................................... 73
3.3. Performance Measurement Methods ........................................................ 78
3.3.1. Traditional approaches ............................................................... 80
3.3.2. Pair-wise and outranking methods.............................................. 85
3.3.3. Efficient Frontier Analysis ........................................................... 94
4. Methodology ............................................................................... 105
4.1. Reasons for this study to use Data Envelopment Analysis ............................ 105
4.2. Introduction to DEA.................................................................... 110
4.2.1. Performance Ratio.................................................................... 111
4.2.2. Single Input, Single Output ......................................................... 112
4.2.3. Single Input – Two Output Case .................................................. 116
4.2.4. A brief introduction on DEA computational process .......................... 120
4.2.5. A brief introduction to DEA's main models....................................... 121
4.2.6. Summary of Introduction to DEA .................................................. 123
4.3. Case Study Protocol.................................................................... 125
5. Case Study and Results .................................................................. 131
5.1. Case Study Theoretical Background .................................................. 132
5.2. Research Process....................................................................... 138
5.2.1. Population Selection.................................................................. 138
5.2.2. Fuel Efficiency Model General Considerations.................................... 140
5.3. Data Cleansing.......................................................................... 149
5.3.1. Cleansing Algorithm.................................................................. 150
5.3.2. Theft detection ........................................................................ 156
5.3.3. Smoothing Algorithm................................................................. 158
5.4. The fuel efficiency model............................................................. 167
5.4.1. Computational aspects and research process ................................... 167
5.4.2. Model with Fuel & Miles Alone ..................................................... 169
5.4.3. Adding the Cost ....................................................................... 177
5.4.4. Adding the Weight.................................................................... 192
5.4.5. Adding the Age ....................................................................... 222
5.4.6. Sensitivity Analysis .................................................................. 228

5.5. Summary of Results ........ 235
5.5.1. Verification, validation and testing ........ 235
5.5.2. Multi-companies benchmark ........ 238
5.5.3. Communicating the results ........ 239
6. Summary of Results and Discussion ........ 243
6.1. On the results and their usefulness ........ 243
6.1.1. Brief summary of the case study chapter ........ 243
6.1.2. Usefulness of the Results ........ 247
6.2. Limitations of the Results ........ 249
6.2.1. Limitations in relation to data cleansing and fuel cards ........ 249
6.2.2. Limitations in relation to DEA ........ 251
6.2.3. Limitation of the study itself ........ 252
6.2.4. Aspects not studied ........ 256
6.3. Contribution and Applicability ........ 259
6.3.1. Contributions ........ 259
6.3.2. Applicability ........ 261
7. Conclusion ........ 263
7.1. Research Summary ........ 263
7.2. Findings ........ 268
7.3. Limitations of the Research ........ 270
7.4. Potential for Further Research ........ 272
7.5. Contribution to Knowledge and Concluding Remarks ........ 276
8. Appendices ........ 279
8.1. Appendix 1: The CANbus technology ........ 279
8.1.1. A bit of history ........ 279
8.1.2. Available information on vehicles ........ 280
8.1.3. Technical description ........ 283
8.1.4. CAN-bus on vehicles ........ 287
8.2. Appendix 2: From Econometrics to the Charnes Cooper and Rhodes model ........ 294
8.2.1. Transforming the Fractional Problem ........ 295
8.2.2. Optimal Weights ........ 299

8.2.3. Production Possibility Set ........................................................................ 301
8.2.4. Dual Problem and computational aspects ...................................................... 303
8.2.5. Model Orientation ..................................................................................... 307
8.3. Appendix 3: Other DEA models ........................................................................ 309
8.3.1. BCC Model ............................................................................................. 310
8.3.2. SBM Model ............................................................................................. 313
8.3.3. Summary ................................................................................................ 316
8.4. Appendix 4: The Charnes Cooper transformation ............................................... 318
8.5. Appendix 5: Smoothing Algorithm Calculations Example ...................................... 320
8.6. Appendix 6: Details on non-controllable and non-discretionary models ....... 323
8.7. Appendix 7: Subsidiary information about the companies ..................................... 325
9. Glossary of Terms and Abbreviations ...................................................................... 326
10. References ........................................................................................................ 332

## Word count: 62,064 (excluding appendices)

# List of Tables


Table 3.1: Multi-criteria matrix ........ 90

Table 3.2: Matrix of concordance subsystems $J^c$ ........ 91

Table 3.3: ELECTRE concordance index matrix ........ 92

Table 3.4: ELECTRE concordance matrix ........ 92

Table 4.1: Performance Measures comparison table ........ 110

Table 4.2: Single Input, Single Output example data ........ 113

Table 4.3: Relative efficiency ........ 115

Table 4.4: One input – two outputs example ........ 117

Table 5.1: Case study design tests ........ 137

Table 5.2: Participating companies details table ........ 139

Table 5.3: Impact of weight on the traditional mpg measure ........ 143

Table 5.4: Factors to include in the fuel efficiency model ........ 149

Table 5.5: Smoothed mpg results ........ 165

Table 5.6: Fuel and Miles model results ........ 172

Table 5.7: Comparison of average distance to frontier based on vehicle type ........ 174

Table 5.8: CCR results for the Fuel, Cost -> Miles model ........ 180

Table 5.9: SBM results for the Fuel, Cost -> Miles model ........ 185

Table 5.10: SBM – CCR model results comparison for the Fuel, Cost -> Miles model ....... 189

Table 5.11: CCR model results comparison for Scenario 1 and 3 ........ 190

Table 5.12 Ranking discrepancies between score and mpg in Scenario 3 ........ 191

Table 5.13: Van gross weight category colouring ........ 192

Table 5.14: Impact of weight on the traditional mpg measure ........ 194

Table 5.15: SBM-NC-I model results ........ 205

Table 5.16: SBM-ND-I model results ........ 210

Table 5.17: Manual verification of SBM-ND-I score calculation ........ 211

Table 5.18: Processed data for weight treatment ........ 212

Table 5.19: SBM-ND-I results with treated weight ........ 214

Table 5.20: The impact of 'Vehicle Age' on the results ........ 224

Table 5.21: One drawback of incorporating 'Vehicle Age' .............................................. 225
Table 5.22: Score differences when adding 'Vehicle age' as a non-discretionary variable227
Table 5.23: Hj(w) at 1% .......................................................................................... 233
Table 6.1: Table summarising model applicability in regards to fuel information ........... 254
Table 8.1: Maximum CAN length at different speeds .................................................... 283
Table 8.2: Two inputs – one output weight example ..................................................... 299
Table 8.3: Two inputs – one output optimisation results ............................................... 300
Table 8.4: Summary table of models' properties .......................................................... 317
Table 8.5: Fuel Transactions ........................................................................................ 320
Table 8.6: Smoothed mpg results .................................................................................. 322

# List of Figures

Figure 2.1 Costs Distribution for Typical Transport Operations ........................................ 27

Figure 3.1: The S curve ..................................................................................... 65

Figure 3.2: Data Sampling ................................................................................. 69

Figure 3.3: Schematic representation of the performance management process ............. 77

Figure 3.4: The 3 levels AHP structure ............................................................... 87

Figure 3.5: The pair-wise comparison matrix ......................................................... 88

Figure 3.6: ELECTRE dual thresholds model with indifference, weak and strict preference zones........................................................................................................ 91

Figure 3.7: ELECTRE ranking with C = 0.50 .......................................................... 93

Figure 3.8: ELECTRE ranking with *C* = 0.75 ........................................................ 93

Figure 3.9: Efficient frontiers.............................................................................. 95

Figure 4.1: Graph comparison of depot's efficiency ................................................ 113

Figure 4.2: Improvement of Depot A ................................................................... 116

Figure 4.3: One input – two outputs example......................................................... 117

Figure 4.4: Graph of improvement in a single input – two output case ......................... 119

Figure 4.5: Illustrating different DEA models ........................................................ 122

Figure 5.1: Fuel Efficiency model – fuel used ........................................................ 141

Figure 5.2: Fuel Efficiency model – cost spent on fuel ............................................. 142

Figure 5.3: Fuel Efficiency model – vehicle weight .................................................. 144

Figure 5.4: Fuel Efficiency model – age ................................................................ 145

Figure 5.5: Spot missing transactions .................................................................. 154

Figure 5.6: Tank level ........................................................................................ 159

Figure 5.7: How not refilling to the top of tank affects measurement ........................... 160

Figure 5.8: Measuring fuel efficiency over a period of time ....................................... 161

Figure 5.9: Inaccuracy at the period start ............................................................. 162

Figure 5.10: Inaccuracy at the period end.............................................................. 162

Figure 5.11: An accurate mpg between first and last refills........................................ 164

Figure 5.12: Regression between the SBM scores and mpg......................................... 173

Figure 5.13: Graphical illustration of fuel -> miles data ................................................ 174
Figure 5.14: Graphical illustration of BCC scores ......................................................... 176
Figure 5.15: Fuel Efficiency model – cost spent on fuel ................................................ 179
Figure 5.16: Graphical illustration of DMU 10 and 29 ................................................... 181
Figure 5.17: Score difference spread when adding cost – CCR...................................... 181
Figure 5.18: Explaining why score difference is low ...................................................... 182
Figure 5.19: Rank difference spread when adding cost – CCR....................................... 183
Figure 5.20: Score difference spread when adding cost – SBM...................................... 186
Figure 5.21: Rank difference spread when adding cost – SBM ...................................... 187
Figure 5.22: Fuel Efficiency model – vehicle weight ...................................................... 192
Figure 5.23: Frontier for each weight category ............................................................. 193
Figure 5.24: Relation between Smoothed mpg and vehicle weight ............................... 195
Figure 5.25: The NCN (non-controllable) model............................................................ 199
Figure 5.26: The NDSC (non-discretionary) model ........................................................ 200
Figure 5.27: The SBM-NC model in its linear form......................................................... 201
Figure 5.28: The SBM-NC-I model.................................................................................. 202
Figure 5.29: The SBM-ND model in its linear form ........................................................ 202
Figure 5.30: The SBM-ND-I model ................................................................................. 203
Figure 5.31: Graphical results of the SBM-NC-I model ................................................. 206
Figure 5.32: SBM-NC and the need to allow for slacks ................................................. 208
Figure 5.33: The SBM-ND-I model ................................................................................. 211
Figure 5.34: Plot of SBM-ND-I results with treated weight........................................... 215
Figure 5.35: Explaining the disparities between SBM-ND and SBM-NC.......................... 217
Figure 5.36: SBM-ND-I results with K = 4,000................................................................ 219
Figure 5.37: SBM-ND-I results with K = 30,000.............................................................. 220
Figure 5.38: Fuel Efficiency model – vehicle weight ...................................................... 223
Figure 5.39: The Chebychev norm model ...................................................................... 229
Figure 5.40: The SBM-ND-I model in its dual form ........................................................ 231
Figure 5.41: SBM-ND-I results with data modified at 1%............................................... 233
Figure 5.42: Multi-companies benchmark graph............................................................ 239
Figure 5.43: Reverted results graph............................................................................... 241

Figure 8.1: Diagram of a CAN network ........................................................................ 284
Figure 8.2: Diagram of a CAN data frame ................................................................... 285
Figure 8.3: ISO/OSI 7 Layer Reference Model ............................................................. 288
Figure 8.4: J1939 FMS definition of engine speed ........................................................ 289
Figure 8.5: Inputs and outputs data matrices .............................................................. 296
Figure 8.6: The virtual ratio........................................................................................ 296
Figure 8.7: The CCR model in its fractional form ......................................................... 297
Figure 8.8: The CCR model in its linear form ............................................................... 298
Figure 8.9: Two inputs – one output weight example graph ......................................... 300
Figure 8.10: CCR model in the multiplier vector form .................................................. 303
Figure 8.11: CCR model in dual form ........................................................................... 304
Figure 8.12: CCR model and slacks .............................................................................. 304
Figure 8.13: Computation phases of $CCR_{DLPO}$ ........................................................... 305
Figure 8.14: CCR output oriented model ...................................................................... 307
Figure 8.15: Illustration of the CCR and BCC frontiers .................................................. 309
Figure 8.16: CCR and BCC efficiencies ......................................................................... 310
Figure 8.17: BCC output efficiency............................................................................... 311
Figure 8.18: The BCC production possibility set ........................................................... 311
Figure 8.19: The BCC model ........................................................................................ 312
Figure 8.20: The dual of the BCC model ....................................................................... 312
Figure 8.21: The BCC model in its output oriented form ............................................... 313
Figure 8.22: The SBM model ....................................................................................... 313
Figure 8.23: The SBM model with scalar t.................................................................... 314
Figure 8.24: The SBM model in its linear form .............................................................. 314
Figure 8.25: The SBM model in its input oriented form................................................. 315
Figure 8.26: The SBM model in its output oriented form .............................................. 316
Figure 8.27: The CCR fractional problem ..................................................................... 318
Figure 8.28: Defining the variable *t* ............................................................................ 318
Figure 8.29: Multiplying the ratio by *t* .......................................................................... 318
Figure 8.30: Transforming the CCR problem ................................................................ 319
Figure 8.31: The CCR problem in its linear form ........................................................... 319

**Figure 8.32: The SBM-NC model in its fractional form.................................................. 323**
**Figure 8.33: The SBM-ND model in its fractional form .................................................. 324**

# List of Formulas


Formula 3.1 The AHP option's score formula........ 89
Formula 3.2 Generic element of the matrix $JC$........ 91
Formula 3.3 Concordance matrix formula........ 92
Formula 3.4 An expression of the ALS Stochastic Frontier model........ 96
Formula 3.5: The DEA virtual performance ratio........ 97
Formula 3.6: Advised minimum number of DMUs in DEA (1)........ 99
Formula 3.7: Advised minimum number of DMUs in DEA (2)........ 99
Formula 4.1: Efficiency ratio........ 111
Formula 4.2: Total factor productivity ratio........ 112
Formula 4.3: Relative efficiency ratio........ 114
Formula 4.4: Efficiency in a single input – two output case........ 118
Formula 4.5: Output efficiency in a single input – two output case........ 118
Formula 5.1: Incorrect mpg formula........ 163
Formula 5.2: Accurate mpg formula........ 164
Formula 5.3: Smoothed volume formula........ 164
Formula 5.4: Factorised smoothed volume formula........ 164
Formula 5.5: Smoothed mpg formula........ 165
Formula 5.6: The SBM – I objective function........ 185
Formula 5.7: Manual SBM-ND score calculations........ 211
Formula 5.8: Normalised variable formula........ 213
Formula 5.9: Introducing relation around vector *w*........ 230
Formula 5.10: Relation between efficient and inefficient DMU with *w*........ 230
Formula 5.11: Introducing relation around vector *w* for SBM-ND-I........ 231
Formula 5.12: Relation between efficient and inefficient DMU with *w* for SBM-ND-I........ 231
Formula 6.1: Assurance region weights constraints........ 257
Formula 8.1: Production possibility set........ 302
Formula 8.2: Relation between *v*, *u* and the $DLP_O$........ 305
Formula 8.3: Reference set........ 306

**Formula 8.4: Possible improvement formulae** .......... **306**

**Formula 8.5: Projection formulae** .......... **307**

**Formula 8.6: Relation between $DLPO_O$ and $DLP_O$** .......... **307**

**Formula 8.7: Relation between $DLPO_O$'s optimal solution and $DLP_O$'s** .......... **307**

**Formula 8.8: Relation between $DLPO_O$'s slacks and $DLP_O$'s** .......... **308**

**Formula 8.9: Relation between $DLPO_O$'s weights and $DLP_O$'s** .......... **308**

**Formula 8.10: BCC efficiency** .......... **310**

**Formula 8.11: CCR efficiency** .......... **311**

**Formula 8.12: The BCC reference set** .......... **312**

**Formula 8.13: The BCC projection formula** .......... **312**

**Formula 8.14: SBM projections** .......... **315**

**Formula 8.15: SBM Reference set** .......... **315**

**Formula 8.16: Bad mpg calculation** .......... **320**

**Formula 8.17: Smoothed Volume formula** .......... **321**

**Formula 8.18: Smoothed mpg formula** .......... **321**

# 1. Introduction

## 1.1. Background

The transport industry is a very competitive environment constrained by ever complex regulations (e.g. corporate manslaughter law (Health and Safety Executive, 2010)) and smaller profit margins.

In such an environment, measuring performance is essential to fleet managers in order to ensure resources are best used so that the organisation can stay competitive. Due to fleets' complex operations, performance can be improved in many different ways. Freight Best Practice (FBP, 2005) mentioned that fuel expenditure for commercial vehicles operators – intrinsic to any industry using vehicles – could be as high as 30 to 40% of all their expenditures. Fuel has also been shown to be a highly variable budget on which improvements are generally possible (Wilson, 1987). In addition, McKinnon (1993) stated fuel consumption can be improved in several ways. Consequently, it seems potentially easier and more beneficial to concentrate first on improving companies' fuel efficiency rather than other operational areas.

Finally, because vans have a bigger market share than HGVs (DfT, 2009, p.130) and that van fuel efficiency measurement is rather different from HGV's, this study will primarily focus on fuel efficiency improvement in the van sector.

Improving the design of a supply chain can, for example, have huge repercussions on fuel consumption. However, potential fuel savings resulting from an optimised supply chain might be outweighed by other costs (e.g. warehousing), thus optimising a supply chain for fuel saving is unlikely to be practical and ideal. Conversely, many different fuel saving interventions exist. Amongst these are diesel or oil additives, energy efficient tyres and aerodynamic kits. These interventions have a direct impact on fuel efficiency through improved fuel combustion, reduced friction or better aerodynamics. Technologies like CANbus (Controlled Area Network Bus, a bus on the vehicle – the electronic equivalent of a motorway – which allows different electronic units to share information such as rpm, distance travelled or fuel used) can also provide an accurate driver's mpg along with detailed information on each driver's behaviour. Although CANbus cannot alone lead to improvement in fuel efficiency, the accurate information and measurement it provides can help fleet managers make better informed decision which could ultimately lead to improvements in fuel efficiency. Although most of the interventions listed above can demonstrate a Return On Investment (ROI), they all represent an initial investment which some companies may not be able to afford.

On the other hand, fuel cards – electronic cards which drivers can use to buy fuel – are omnipresent in the transport industry. Thus, in a similar manner to fuel efficiency measurement based on CANbus information, improving fuel efficiency measurement based on fuel card data could indirectly improve fuel efficiency without requiring any extra investment. Besides, mpg, the industry-standard fuel efficiency measure, has several limitations which potentially impede the measurement of fuel efficiency and improvements that could be realised. These limitations should thus be addressed and this study will consequently concentrate on improving fuel efficiency measurement based on fuel card data.

This thesis will first introduce the hypothesis followed by the aims and objectives. Chapter 2 will then review the different alternatives that can improve fuel efficiency and justify the particular focus on fuel efficiency measurement based on fuel card data.

Chapter 3 reviews the background theory in regards to performance measurement. Most basic concepts in relation to performance measurement are described and explained in this chapter. Section 3.3 will review some relevant performance measurement methods or techniques available. The techniques considered range from the traditional benchmarking approach, to pair wise comparison techniques or Data Envelopment Analysis – another benchmarking technique with unique characteristics.

This study's methodology will be discussed in chapter 4. This chapter will justify why Data Envelopment Analysis is retained for this study. DEA key concepts will be introduced so that the remainder of the thesis is accessible even to individuals with no previous experience of DEA. Appendices 8.2 and 8.3 will provide more explanations on DEA and DEA models and should provide more detailed information on DEA. Because these appendices are technical sections which are not essential to understand the thesis, they were not included in the main body of the thesis. However, reading them could help understand some technical details of the 'Case Study and Results' and 'Summary of Results and Discussion' chapters. Chapter 5 will finally detail the protocol followed in this research.

'Chapter 5, Case Study and Results', will first briefly review case study background theory to then discuss the research in greater depth. The data cleansing and smoothing algorithms, which are essential to ensure the fuel card data is appropriate, will then be introduced and their use explained and demonstrated. This will be directly followed by a detailed description of the fuel efficiency DEA model with all the variables to be tested. As with many modelling approaches, each variable (including fuel used, vehicle weight, vehicle age and mileage) will be added to the model one by one in order to measure the impact each one has on fuel efficiency. Because of this step by step approach, each step's results will be discussed directly in this chapter as appropriate.

Chapter 6 will summarise the results and discuss them appropriately. This includes examination of the results and their usefulness, limitations, contribution and applicability. Chapter 7 will then conclude this thesis and give potential for further research. Appendices can be found in Chapter 8, the 'Glossary of Terms and Abbreviations' in Chapter 9 and the list of references in Chapter 10.

## 1.2. Hypothesis

The research hypothesis is as follows:

***It is possible to develop a form of vehicle fuel efficiency measurement that gives a fleet manager more relevant information than currently available***

It is interesting to observe that the hypothesis uses two important notions which will require adequate defining. These are:

- ⇨ Efficiency,
- ⇨ Measurement.

Although a definition of the terms relevant to the study can be found in the Glossary of Terms and Abbreviations, these two terms as well as a few important others will be defined and further discussed in section 3.1.1 Key Concepts and Definitions. It is also interesting to observe that the hypothesis does not suggest which performance measurement approach or which type of fuel data should be used.

## 1.3. Aims & Objectives

### 1.3.1. Aims

The research aims are as follows:

1. To analyse the main fuel performance measurement methods used in the transport industry.

2. To evaluate the limitations of these measures and discuss the consequent impact on fuel efficiency measurement in transport businesses.

3. To develop an advanced performance measurement method in order to produce a more effective measure and to assess its usefulness as a better measure.

4. To apply this advanced fuel efficiency performance measurement method to selected companies which operate vans.

5. To evaluate the extent to which this methodology is of operational value to transport businesses.

### 1.3.2. Objectives

The objectives of this research are as follows:

- ⇨ To demonstrate the relevance of fuel efficiency to transport operations
- ⇨ To critically review the factors and techniques which can have a positive impact on fuel efficiency
- ⇨ To demonstrate the relevance of improving the fuel efficiency performance measurement
- ⇨ To review the existing literature on performance measurement & performance measurement methods
- ⇨ To evaluate the applicability of some appropriate performance measurement methods
- ⇨ To demonstrate the relevance of DEA as a suitable performance measurement method
- ⇨ To identify the companies relevant to the study and collect the appropriate information
- ⇨ To develop a new fuel efficiency measure and appropriate (DEA) performance measurement models
- ⇨ To apply the developed model to this selection of companies
- ⇨ To evaluate the model results in collaboration with the participants
- ⇨ To iteratively improve these models with the participants feedback
- ⇨ To analyse the results
- ⇨ To critically analyse the results in comparison with traditional measurement methods

⇨ To appraise the applicability, usefulness and limitations of the new fuel efficiency performance measure

These objectives will be carried out in this research and the conclusion will review each objective and reference the section in which the objective was addressed.

# 2. Literature Review – Van Fuel Efficiency Measurement

This chapter will describe fleets' main operational costs and justify why fuel efficiency is an ideal area to seek performance improvements. The main potential methods that can be employed to improve fuel efficiency will be discussed elsewhere. The fuel efficiency improvement methods to be reviewed are: scheduling, driver behaviour management using CANbus information, fuel card mpg analysis, and some other traditional fuel saving interventions. The research interest in improving the mpg performance measurement using fuel cards will then be justified in the section 2.3 'Explaining the Focus on Fuel Efficiency Measurement Based on Fuel Card'.

## 2.1. Explaining the Focus on Van's Fuel Efficiency

Regardless as to whether it is caused by harsh competition or scarce resources, one intrinsic aim behind performance measurement is the need to improve performance. Due to fleets' complex operations, fleets' performance can be improved in many different ways. Coopers (1987, p.26) studied the cost distribution for typical vehicle fleets (the study focused mainly on fleets running 7.5 tonnes vehicles) and observed that the biggest costs are generally associated with 3rd party contractors expenses, wages, maintenance and fuel. This is illustrated in the Figure 2.1 Costs Distribution for Typical Transport Operations (Anon, 1985).

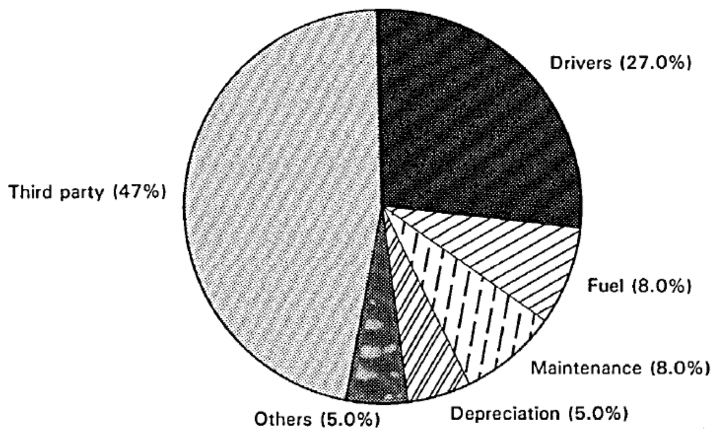


Figure 2.1 Costs Distribution for Typical Transport Operations

Wilson (1987, p.18) states that drivers' wages, third party costs, and maintenance costs represents the biggest budgets which are all quite stable and predictable. He also states that fuel cost is the budget that demonstrates the most variability. Freight Best Practice – a Department for Transport project – observed (FBP, 2005, p. 1) that fuel expenditures for commercial vehicles operators could represent a huge proportion of the total expenditures. The difference between the two figures (between fuel cost as shown in Figure 2.1 and FBP's fuel cost figure) can probably be explained by an increase in fuel costs since Wilson wrote his paper. Finally, the relationship in the road industry between profitability and fuel expenditures has quite logically been acknowledged by research (McKinnon, 1993). In his paper, McKinnon also recognises that there are a number of ways to improve fuel efficiency.

Due to the limited resources a fleet operator can spare to improve their operations, it is essential to concentrate efforts on areas that would bring maximum savings. It would be logical then to concentrate on the budgets that could be reduced the most. As detailed above in 'Figure 2.1 Costs Distribution for Typical Transport Operations', the biggest costs for fleets are drivers' wages and maintenance. These costs can be reduced mainly by cutting driver wages or by an improved use of scheduling techniques.

The use of fuel is also intrinsic to any industry using vehicles. Besides, as stated by McKinnon (1993), fuel consumption can be improved in many different ways which will be discussed in section 2.2 On the Ways to Improve Fuel Efficiency. Although fuel efficiency has been well studied, very little research has been conducted on the measurement of fuel efficiency itself and on addressing some of the key limitations traditional fuel efficiency measurement demonstrates (see section 2.2 On the Ways to Improve Fuel Efficiency). Because scheduling techniques have conversely been well researched (see following section 2.2.1 Improved Scheduling and Network Optimisation), concentrating on improving companies' fuel efficiency would consequently bring more originality than researching other operational areas. Because unnecessary fuel costs occur when the vehicles are badly driven but also when fuel is stolen, this study should consider both fuel efficiency and fuel theft. However, the issue of fuel waste due to tank contamination (FBP, 2007b, p. 3) is not discussed in this study (as the study will focus on the measurement methods rather than operational aspects of fuel management).

The DfT's Transport Statistics Great Britain document (DfT, 2009, p. 155 onward) shows there are more vans than LGVs and that the former's vehicle kilometres are also greater than the latter's (68.1 billion vehicle kilometres for vans and 28.7 for LGVs in 2008 - DfT, 2009, p.130 – because a huge proportion of vehicles potentially relevant to this study do not transport goods (e.g. engineering services vans) vehicle km is here more relevant to fuel efficiency than tonne km). Despite this lower vehicle kilometres figure, LGVs use in total more fuel than vans (DfT, 2009, p. 54). Therefore, it would seem more logical to tackle LGV fuel efficiency measurement first. However, the aim of this study is to demonstrate that it is possible to develop a form of fuel efficiency measurement that gives fleet managers more relevant information than currently available. Thus, it is more favourable to first tackle the simpler problem of van fuel efficiency measurement (in which issues such as vehicle load weight have a less significant impact) than the LGVs' fuel efficiency measurement problem. Consequently, this study will primarily focus on the measurement of fuel efficiency in the van industry. The applicability of the study to other industry sectors and vehicle types will however be discussed in chapter 6 'Summary of Results and Discussion'.

## 2.2. On the Ways to Improve Fuel Efficiency

Many different alternatives exist to improve fuel efficiency. This section will review the main ones.

### 2.2.1. Improved Scheduling and Network Optimisation

Transport operations implies that consignments or jobs are assigned to different vehicles and that the vehicles route themselves to their different job destinations or delivery / pick-up points. Research distinguishes routing problems (RP) from scheduling problems (SP) and from the problems combining the two (RSP) (Bagchi and Nag, 1991, p. 11). Routing relates to finding the most (or a more) advantageous route between two points (this can be for example the shortest or quickest path);

Bagchi and Nag (Bagchi and Nag, 1991) describe RPs as assignment problems. On the other hand, scheduling problems deal with the 'allocation of resources over time to perform a collection of tasks' (Bagchi and Nag, 1991, p. 10).

The complexity behind finding an optimal route for a list of deliveries or points was first observed by mathematicians (W. R. Hamilton and Thomas Kirkman) in the 1800s. Hamilton created his Icosian game which consists of linking all dots in a dodecahedron by a path visiting each vertex (dot) exactly once. Such paths are called Hamilton cycle or Hamilton path. Hamilton and Kirkam's work is discussed in Graph Theory (Biggs et al., 1976). This problem was later on considered further by mathematicians, notably Menger (1932, cited in Punnen, 2002, p. 1) who designate the 'messenger problem' as:

> 'The task of finding, for a finite number of points whose pairwise distances are known, the shortest path connecting the points. This problem is naturally always solvable by making a finite number of trials through the permutations of the given points. The rule, that one should first go from the starting point to the nearest point, then to the point nearest to this etc., does not in general result in the shortest path.'

Dantzig *et al* (1959) proposed a linear programming approach to tackle this problem (the problem in question was the Travelling Salesman Problem or TSP). Since then, academic interest in vehicle routing and scheduling problems has significantly increased.

In 1972, Richard Karp demonstrated that the TSP and many other RSPs were NP-hard (the hardest class of problems in complexity theory). Since the formulation of the TSP, RSP have grown in complexity. Tyagi (1983, cited in Slater, 2002) developed a method to minimise total fleet mileage. Other constraints were also studied to answer several problems observed in real life operations. Modern RSPs have several other constraints than the original TSP and VRP. These include:

- Delivery time windows (Rochat and Semet, 1994). Considering the TSP, this constraint implies that each city has to be visited during a period of time.
- 'Capacitated' Vehicle Routing Problem where vehicles are all located at a central depots and need to be routed to different customers with known demands and vehicle capacity constraints (and precedence constraints are possible (Malik et al., 2007)).
- Asymmetric model, i.e. a model in which some segments (link between two vertices) can be travelled one way only. This characteristic is crucial when doing real life RSP.
- Multi depots models. In this case, vehicles or engineers can belong to different depots. This is particularly useful when scheduling engineers as each engineer generally returns to his home. Vehicles can also leave a 'depot' and potentially return to another.
- Model in which capacity constraints are imposed. This can imply weight or volume constraints on different vehicles or depots.
- Skills constraints (e.g. drivers skills and competences, Punnen, 2002).

Most current commercial solutions incorporate all these constraints in their model. This list is not meant to be exhaustive but should provide a good overview of modern problems' complexity.

Several possible objective functions are possible. Traditionally the VRPs aim at minimising mileage, but it is also possible, amongst others, to optimise by minimising the time spent driving (Maden et al., 2010), or $CO_2$ emissions (Palmer, 2008).

By reducing the number of miles travelled, the time spent driving (Taniguchi and Shimamoto, 2004), or the amount $CO_2$ emissions, computerised scheduling can potentially reduce the amount of fuel to be used for a given list of deliveries. This can in turn have a beneficial impact on vehicles' fuel consumption (Baumgartner et al., 2008). In a similar manner, network optimisation at a supply chain level can – perhaps in an even greater manner – reduce the overall number of miles required for the whole supply chain operations and thus, the amount of fuel used. Supply chain optimisation can in a similar manner have a great impact on fuel consumption reduction.

### 2.2.2. CANbus Technology & Driver Training

In the last twenty years of the twentieth century, the automotive industry developed quickly and, following technological advances, more and more technology and electric and electronic devices were fitted and used on cars. The proliferation of wiring and wire looms all over the car was causing real problems throughout the manufacturing process, from car design to manufacturing and was also adding weight to the vehicle. Besides, excessive wiring was costly and did not provide good control over the vehicle’s electronics. Acknowledging this problem, R. Bosch started working on an in-vehicle network project as early as 1983 (Anon, 2008c); this became the Controlled Area Network, generally simply referred to as CAN.

This type of network allows different electronic devices such as brake or engine controllers to communicate through a common electronic network or central bus. The central bus reduces the need for wiring harness (Buchanan, 2000), thus reducing cost, vehicle weight – impacting positively on the fuel consumption – and improving the control over the vehicle’s electronic systems. This was, at the time, a revolutionary technology and it enabled an unhampered proliferation of electronic devices on the vehicles, which in turn has been associated with greater control and security (Anon, 2007b).

The original CAN bus specification (Bosch, 1991) described both the physical and data layers. The physical layer of the CAN protocol describes all the mechanical (cable, connectors, resistances…) and electrical aspects (signal level, bits, timing…) of the CAN. In short, CAN networks are made of a pair of twisted wires terminated by a resistor on both ends. The data layer on the other hand, describes the low level communication aspects between the different nodes connected to the CAN network (the electronic components that are connected to the network). CAN networks, communications are encoded by a voltage difference between the two cables (Anon, 2007b). Although other types of network are sometimes used instead of CANbus (e.g. FlexRay, LIN), CANbus is widely used for its robustness and low cost characteristics.

As explained previously, CAN networks transfer information broadcasted by the different Electronic Control Units (ECU) which equip the vehicle. Modern vehicles are equipped with ECUs that generally provide (at least) the following information:

- ⇨ Fuel used,
- ⇨ Distance travelled,
- ⇨ Speed,
- ⇨ Miscellaneous engine information such as rpm, throttle opening, speed pedal position,
- ⇨ Braking information

This information is generally available on most vehicles manufactured after 2000 although some vehicles will also have additional information available from the CAN (e.g. door opening, warning light). As the basic information listed above provides a key insight into accurate mpg analysis and driver behaviour, it is interesting for fleet managers to retrieve it so that they are in a better position to manage their operations. Many telematics tracking devices can now connect to vehicles' CAN, retrieve its key information and make it available to the fleet operator. Besides, most telematics solutions provide driver identification so that the CANbus information can be driver specific as well as vehicle specific (e.g. mpg performance by driver).

CANbus is one of the most accurate sources of information in term of vehicle information. The fuel consumption figure is obtained from the vehicle's fuel injectors, the distance travelled is the same as the one displayed on the odometer and more generally any information obtained from the CAN has the same digital accuracy. This makes CANbus one of the most accurate sources of information to obtain a vehicle or driver's mpg figure or driver behaviour information. More accurate mpg and driver behaviour information can in turn enable fleet operators to make better informed decisions that can have a greater positive impact on fuel efficiency. It is important to observe that retrieving CANbus information does not in itself improve fuel efficiency. It is rather the informed decisions (such as driver warning or training) based on more accurate information obtained from the vehicle's CAN that can in turn potentially lead to improvements in fuel efficiency (FBP, 2008a, p.27).

Further information on CANbus can be found in 8.1 Appendix 1: The CANbus technology.

#### 2.2.3. Fuel Card Management

As outlined by Baumgartner et al (2008) and FBP (FBP, 2008a), both improved scheduling and CANbus technology can have an indirect impact on fuel consumption. Scheduling can improve fuel consumption by reducing the number of miles required to complete a list of deliveries whilst CANbus can be useful in spotting drivers in need of driver training or for daily fuel performance monitoring and fuel theft detection. This section will show how fuel performance monitoring using fuel cards information can also indirectly improve fuel efficiency.

Fuel cards are cards which drivers can use to only buy fuel or sometimes other vehicle related commodities such as engine oil. Fuel card companies send reports to their customers which hold detailed information about all the fuel cards transactions that have occurred. These reports – combined with a distance figure – can then be used to measure the vehicles' mpg performance. In a similar fashion as for CANbus, decisions based on this mpg measure can indirectly lead to improvement in vehicle fuel efficiency. The concepts discussed in this section are only briefly introduced although more details and justifications on these will be given in the section 5.3.3 Smoothing Algorithm.

Although fuel cards can be used to buy commodities such as engine oil or food items at petrol stations, they are mainly used to buy petrol and diesel (Anon, 2009c). Fuel cards can only be used at the petrol stations which are part of the fuel card provider's network. The fuel card company then invoices the haulage company for the fuel bought (generally on a weekly or monthly basis). Fuel cards are obviously a necessity for most businesses as without them drivers would not have a convenient way to refuel their vehicles.

Fuel cards generally offer lower prices for diesel and petrol although prices vary from a fuel card providers and card types. Some fuel card providers tend to offer pump prices on a wide number of petrol stations while others will propose weekly set prices (generally lower than the pump price) at a limited number of petrol stations. Some more complex pricing scheme which depends on the type of petrol station the refill occurs at also exist (e.g. a bunker site where fuel will be cheaper).

Even though most fuel card offer discounted prices, de Kock (2009) warns that operators will have to balance out the extra cost of driving to compatible petrol stations as the extra miles driven might outweigh the savings at the pump. Cole

(2009) however mentions that fuel cards management is one way of reducing fuel, along with bunkering (having a fuel tank on site), bulk buying and careful fuel cost management. Similarly, Clarke (2008) mentions that although some fuel card companies still have a basic 'one card suits all' product, others fuel card companies propose a wide range of different cards for different types of operations, as well as 24/7 online accounting service, reports or advanced security features, etc.

Three main types of fuel cards, driver fuel cards exist – assigned to a unique driver, vehicle fuel cards – assigned to a unique vehicle, or vehicle-driver fuel cards – assigned to a unique vehicle and unique driver. This information (either driver name, vehicle registration, or both) is generally embossed on the card and will automatically be linked to each fuel transaction on the fuel card reports. However, the vehicle registration is generally spelt out to the attendant at the petrol station till (unless it is embossed on the card in the case of a vehicle or a driver-vehicle card). Similarly, the driver also needs to give the vehicle's odometer reading to the person at the till. This information will be made available on the fuel card reports and will generally be used to calculate the mpg performance measure.

As Paul Holland (cited in Cole, 2009) mentions, the 'whole raison d'être of fuel cards is control'. Indeed, fuel card data files provide an accurate picture of fuel expenditures thus enabling appropriate fuel cost micro management and control. Unlike CANbus technology, fuel cards management can also reflect the cost of theft and potential leaks between the tank and the injectors.

Fuel cost management with fuel cards has a number of disadvantages as it relies on the mpg measure calculated from the fuel consumption shown on the fuel card reports. This mpg measure can show the vehicle's fuel performance (both in terms of driver behaviour or potential leakages). The mpg measure can also indicate whether any fuel was misused or stolen (either via siphoning, by refuelling somebody else's vehicle or filling a jerry can). To calculate the vehicle's mpg, both the fuel consumption and the vehicle odometer information are used. However, and as described before, the odometer information and sometimes the registration are keyed in at the petrol station. This manually entered information is prone to errors and is consequently often incorrect or inaccurate.

Inaccurate information can have a dramatic effect on the vehicle's mpg measure. As an example, missing a single refill (e.g. because of a registration misspelling) in a month for a vehicle that does 500 miles a week at 35 mpg would give a 20% inaccuracy on the calculated mpg measure. Due to the dramatic impact potential misspellings can have on the final mpg measure, it is consequently crucial to cleanse the data before the mpg figure is calculated. Odometer readings can also be discarded where blatantly wrong although evaluating this can be complicated (as in some cases it might not easy to determine whether it was the refuel which was not made up to the top of the tank or whether an approximate odometer reading was given at the till).

Alternatively operators using telematics system will be able to retrieve the vehicle mileage from these systems as they provide a more consistent distance figure. As telematics devices also provide location information, a telematics system can also help confirm that a vehicle was really at a petrol station at the time of the transaction; information particularly useful to correct registration misspelling and also check on theft (i.e. vehicle not at the petrol station at the time of refill). Without telematics, cleansing registration will only be possible with educated guesses and with some knowledge of the operations.

Operators frequently want to know the average mpg for each vehicle of their fleet across a period of time (often monthly). As fuel card information is generally the only fuel consumption information available to the operators, they tend to calculate the vehicle mpg by totalising the fuel bought during the measurement period and by taking the distance travelled during that same period. However, this method can be highly inaccurate and this research will appropriately address this limitation. This will be further detailed in section 5.3.3 Smoothing Algorithm.

As already explained in this section, fuel cards are a necessity for most businesses to run their operations. They also bring several advantages such as the possibility to buy cheaper fuel and to better control cost in a way which CANbus alone cannot. However, fuel cards management relies on the mpg figure to be calculated and this requires data to be thoroughly cleansed before use. Despite these limitations, fuel cards analysis is still an essential tool that allows fleets to control their fuel performance and cost.

### 2.2.4. Traditional Fuel Saving Interventions

CANbus and fuel card analysis can both provide the fuel consumption information which can be used to calculate the mpg performance of the vehicle or drivers (for CANbus). However, measuring mpg does not in itself improve the fuel consumption. The mpg measure is merely a descriptive performance measure which only shows the current performance level (see Type and Classification of Performance Measures for more information on this). However, the mpg measure provides key information that enables informed decisions to be taken and it is the results of these decisions that can potentially improve fuel consumption performance. This concept is discussed in 'Fuel Efficiency Trials Research' (FBP, 2008a, p.27) and in 'In fleets trials of fuel saving information' (FBP, 2005, p.4). Several alternatives exist, called fuel saving interventions, which can have a more direct impact on fuel efficiency. The alternatives will be discussed in the following section.

#### 2.2.4.1. Fuel additives

Fuel additives can potentially improve the vehicle's fuel efficiency by cleaning the engine injection system, the engine cylinders and the filters. Three types of fuel additives exist, the deposit removal additives, the bacterial growth preventive additives (FBP, 2005, p.5) and diesel emulsifiers.

1. The first cleans the deposits that form on the injectors, cylinders and valve as these deposits can potentially affect the combustion process. The benefits of this type of additives vary with the age of the vehicle and the state of the engine.

2. The second prevents the bacterial growth in diesel and water contaminated petrol (diesel additive cannot be used on petrol and vice versa). Bacteria, when present in fuel, can clog filters which can impact the combustion process.

3. Diesel emulsifiers are also claimed to improve fuel combustion process although incorrect proportions will not produce any benefit and can potentially cause severe damage to the vehicle's engine.

Finally, some fuel brands claim that they produce better fuels which allow vehicles to achieve more miles for a given volume of fuel (Bearne, 2010). Although it is possible that enhanced fuels deliver a better combustion performance, operational complexities (see 2.2.5 A Word on the Ways to Improve Fuel Efficiency) makes it really difficult to quantify the proportion of the improvement caused by the deemed superior fuel. Furthermore, many companies use different fuel brands and there is no indication on the level of performance that would be delivered by a mixture of fuel brands in the tank. For these reasons, the study will not consider fuel brands. This decision has little or no impact on this study's results as the three companies used in the case studies (see section 4.3 Case Study Protocol) use a wide range of fuel brands (which implies that the fuel in the tank is a mixture of different fuel brands).

#### 2.2.4.2. Lubricating Oils and Additives

Lubricating oils and additives can also potentially improve fuel efficiency by limiting friction in the engine, gearbox and drive axle. FBP (2005, p.5) states that because oil can be used in different parts of the vehicle but also because studies evaluating the efficiency of lubricating oils and additives do not generally mention which parts of the vehicle are concerned, it is hard to estimate whether the potential savings in terms of fuel outweigh the cost of oil. Labeckas and Slavinskas (2005) have however demonstrated theoretical fuel saving by using a combination of oil and additives (a reduction of 7.3% of brake specific fuel consumption which is the rate of fuel consumption divided by the power produced).

#### 2.2.4.3. Using Hydrogen in the Air / Fuel Mix

Mixing hydrogen with the air before combustion can potentially improve the quality of the combustion process itself and thus reduce the amount of diesel used for similar output power. GSE Haulage and Dodd's Transport which have tried Hydrogen Injectors devices from Oil Drum have claimed fuel savings as high as 11% (Milnes, 2009). In turn, Saravanan and Nagarajan (2008) have conducted laboratory experiments on mixing hydrogen with diesel and air prior to combustion. They concluded that best results were attained from a 30% hydrogen mix – a limit above which the use of hydrogen becomes detrimental. They also concluded that particulate matter was reduced and that energy consumption decreased with an increase in hydrogen percentage for the whole range of operations.

#### 2.2.4.4. Energy Efficient Tyres

Part of the energy that is transferred from the engine to the wheel is converted into heat because of the contact of the tyre with the road. Tyre heat increases friction thus has a negative impact on the vehicle's mpg. Modern fuel efficient tyres have shallower treads which enable them to better dissipate heat thus reducing friction and improving the fuel consumption. Because of their lighter design, energy efficient tyres will also not be adequate for harsher conditions such as those experienced in the aggregates industry.

Due to their shallower treads, fuel efficient tyres also wear down more quickly than traditional tyres thus it will be necessary to balance out fuel savings against added wear and tear cost. FBP (2005, p.8) mentions typical 2% fuel savings and 75% lifetime of the life of a standard tyre. The benefits from using energy efficient tyres will consequently depend on the business' operations. Following some trials, Moy Park LTD reports fuel savings as high as 8% with Michelin energy efficient tyres (Michelin Energy), while on the other hand Turners Rolls declare that it found energy efficient tyres to be beneficial on international operations only and not in UK's (FBP, 2005, p.9).

Tyre manufacturers which propose energy efficient tyres are now battling for a fuel efficiency performance eco-rating (Anon, 2005b). Consequently, fuel efficiency indexes start to appear slowly as with the NHTSA's proposition of a tyre label (Anon, 2009b). Quite surprisingly in regards to the operations interest in reducing cost and to the potential savings that well inflated (Anon, 2007c), energy efficient tyres can bring, a recent study still found that 76% of transport operators did not consider fuel efficiency characteristics when buying tyres (Brown, 2008).

#### 2.2.4.5. Aerodynamic Kits

A vehicle at rest contains inertia. For the vehicle to move, energy generated by the engine is transferred to the wheels. At a given speed, the vehicle then contains momentum (the product of the mass and velocity of an object) and the vehicle only needs to compensate for friction, potential incline and aerodynamic drag to maintain a constant speed.

As a vehicle's speed increases, so does its rolling resistance (friction) and aerodynamic drag. The former is proportional to the vehicle weight so the aerodynamic drag – which is not – will be proportionally more important on a partly laden or unladen vehicle. On the other hand aerodynamic drag increases exponentially as a vehicle speeds up resulting in being the major factor for fuel consumption (Weatherley, 2009).

Vehicles need to limit the three following types of aerodynamic drag:

- ⇨ Form drag (relates to how well air flows around the object's overall shape). This is the most influential form of drag
- ⇨ Surface friction (caused by air viscosity)

- ⇨ Interference drag (caused by projections of an object's angles creating vortices; e.g. the end tip of a standard shaped wing (i.e. without a winglet)).

Aerodynamic kits are designed to reduce the aerodynamic drag of the vehicle. As the three main types of trucks (articulated, rigid and drawbar) react in different ways to wind, the aerodynamic kits will be different for each of these vehicle types (Daimler Chrysler reports fuel savings as high as 6% (on LGVs, Anon, 2007a)). Due to vehicles' large frontal area and a poor aerodynamic design, the potential for savings is greater when the vehicles are travelling fast. Consequently, LGVs – due to their size and type of operations – generally have a greater scope for fuel saving from aerodynamic kits than vans (FBP, 2007a).

#### 2.2.4.6. A Word on Fuel Saving Interventions

Although these interventions are claimed by manufacturers to improve fuel efficiency see (see FBP, 2005, p.5) and that experiments also demonstrate some potential fuel savings, FBP recommends having a performance management system in place to appraise – before the intervention is used throughout the whole fleet – the potential fuel saving benefits of any intervention. This is crucially important as the potential savings from fuel saving interventions generally vary depending on how well the fleet is already performing in regards to fuel efficiency. Carlsberg and Somerfield case studies illustrate the importance of measuring the potential benefits of an intervention, as these two companies have declared some impressive fuel savings by measuring the performance (especially in relation to mpg) before the purchase of any equipment (Tonkin, 2009a).

### 2.2.5. A Word on the Ways to Improve Fuel Efficiency

This section has reviewed some interventions that have a direct impact on fuel consumption (e.g. fuel additives, hydrogen added to the air/diesel mix, aerodynamic kits...), and other which have a rather indirect impact such as scheduling or CANbus technology (with adequate following-on driver training). By providing key information such as distance, location or CANbus information, telematics technology enables transport companies to make better informed decisions in regards to their operations thus could potentially be also considered as a fuel saving intervention (FBP, 2009). Fuel savings from telematics solution can be achieved for example from reductions in private mileage or from a better understanding of driving practices. Anti-theft devices which can be placed on the tank's aperture to hinder any fuel siphoning can also be considered as fuel saving interventions. Similarly, engine automatic cut off at traffic light could also be regarded as fuel saving interventions although these systems are generally found on top range luxury cars and are generally not retro-fitted. Finally, it is important to observe that regardless of which fuel interventions are used on a truck, several other variables can have an impact on fuel efficiency. These are mainly:

- ⇨ correctly inflated tyres (mentioned in section 2.2.4.4 Energy Efficient Tyres),
- ⇨ driver behaviour (mentioned in section 2.2.2 CANbus Technology & Driver Training),

- ⇨ properly done maintenance,
- ⇨ topography,
- ⇨ weather,
- ⇨ traffic and infrastructure and
- ⇨ vehicle speed.

## 2.3. Explaining the Focus on Fuel Efficiency Measurement Based on Fuel Cards

There are many different methods that can be used to improve vehicle's fuel efficiency. These sections will review the methods mentioned above and will justify why fuel efficiency performance measurement based on fuel card information was selected as the main focus for this study.

### 2.3.1. Improved Scheduling and Network Optimisation

Section 2.2.1 explained how improved scheduling could reduce the number of miles which would indirectly benefit fuel consumption. It was also stated that network or supply chain optimisation (e.g. depot location) could have an even greater impact on the number of miles required to conduct operations, thus on the overall fuel consumption.

To ensure that fuel consumption is reduced via scheduling, the optimisation function needs to either reduce the $CO_2$ emissions or the fuel consumption (there is direct linear relation between fuel used and $CO_2$ emissions). However, some companies' main concern might not be to optimise fuel efficiency but instead reaching their customer on time – the cost of customer dissatisfaction being possibly greater than any fuel saving. Depending on the optimisation criteria, improving fuel efficiency through scheduling would not be appropriate in these cases.

Similarly, network or supply chain optimisation needs to consider the total cost (e.g. cost of different depot locations, cost of average transport, number of drivers necessary, number of depots necessary, etc). In this case potential fuel savings might be outweighed by other savings (e.g. cheap depot location). Thus, optimising a whole network or supply chain just considering only fuel cost is likely to be a mistake.

### 2.3.2. CANbus Technology & Driver Training

CANbus is probably the most accurate source of information in regards to vehicle fuel consumption and driver behaviour information (providing accurate mpg, engine rpm, throttle position, etc). CANbus' extreme accuracy – which can, depending on the telematics solution, sometimes be tied in to each individual driver – makes it an ideal solution to target inefficiencies and thus improve fuel consumption through driver management. However, CANbus technology still remains expensive (generally found from £5.00 to £15.00 per vehicle-month on top of most telematics offers (already between £5.00 and £15.00 depending on options) although some telematics providers' offer this as part of the standard package). Furthermore, some telematics providers might not cover all vehicle makes and models, and some vehicles do not even have a CAN – i.e. it will not be possible to retrieve CAN information on these vehicles. In addition, some CAN installation can be problematic as fuel used needs to be calibrated against fuel card data in order to report accurate fuel consumption.

This means that despite the benefits of using CAN technology, this option is generally costly and occasionally on some vehicles, a difficult or inapplicable option.

### 2.3.3. Traditional Fuel Saving Interventions

The section Traditional Fuel Saving Interventions has reviewed the main fuel saving interventions ranging from fuel additive, hydrogen-diesel mix, to aerodynamic kits. All these techniques have a direct impact on fuel efficiency although each will generally have different fuel saving potentials.

It is consequently possible to improve fuel efficiency by enhancing an existing traditional fuel saving intervention or creating a new one. This option would have the advantage of directly improving fuel consumption, making improvements more tangible and more easily measurable. Nevertheless, all the aforementioned fuel interventions listed above relate more to engineering, chemistry or physics rather than transport. These limitations make fuel efficiency improvement via traditional fuel saving intervention not an ideal subject for this transport research.

### 2.3.4. Fuel Cards Management and the Limitations of CANbus

The relation between fuel cards management and improvement in fuel efficiency is similar to the relation between CANbus and fuel efficiency improvement; these two methods do not have a direct impact on fuel efficiency and it is rather the informed actions based on the measurement information that will improve fuel efficiency.

Both CANbus and fuel card management rely on measuring the fuel efficiency performance to uncover potential gaps in fuel performance. Because the information recorded from the CANbus generally only measures the fuel that passes through the injectors and other driver's behaviour related parameters, it is best at uncovering inefficiencies caused by poor driving behaviour. However, and unlike fuel cards, CANbus alone fails to show potential fuel thefts as having no knowledge of how much fuel was bought. In comparison, fuel cards truly reflect both the cost of fuel expenditures and of fuel theft.

The mpg measure which is calculated with fuel cards data has several limitations however:

1. The measure **does not include factors which are important or essential to its interpretation**. These are mainly:

    ⇨ *Vehicle weight* (heavier vehicles are likely to use more fuel),

    ⇨ *Vehicle type* (some vehicle types are less efficient than others in terms of fuel efficiency – e.g. drawbar vehicles have poorer aerodynamics qualities than articulated thus would be likely to show worse performance in terms of fuel efficiency),

    ⇨ *Type of operations* (urban driving uses – for the same distance – more fuel than motorway driving for example),

    ⇨ *Vehicle age* (older engines are likely to show lower performance in terms of fuel efficiency as being of an older generation but also because of wear and tear).

    Including these factors would enable the measure to reflect performance level by itself instead of relying on extra information to appraise the performance.

Furthermore, and when considering benchmarking approaches, this would enable vehicles of different sizes (and potentially: types, type of operations, and different ages) to be compared together in a fair and potentially unbiased manner.

2. The mpg measure **does not include cost** which is another key aspect of fuel performance (i.e. pence per mile measure). It is however conceivable that a vehicle could be mpg efficient but not fuel-cost efficient (and vice versa). Because both aspects of performance are relevant to fuel efficiency, it is important to appraise a vehicle which is cost efficient yet mpg inefficient as a fuel efficient nonetheless (and vice versa).

3. The measure **is generally not used correctly** in fuel trials (i.e. vehicle not always refilled at the beginning and the end of the measurement period – this is explained in greater details in the section 5.3.3 Smoothing Algorithm).

These limitations potentially hinder an easy interpretation and use of the mpg measure.

Fuel cards, as a whole, have several attributes that make them an ideal choice for research on fuel efficiency improvement. Not only their use is nearly universal for companies that run road operations, but they also are more suitable than CANbus in term of cost control and are definitely cheaper than this advanced technology. Finally fuel efficiency measurement using fuel cards also demonstrates several limitations which would be interesting to address (as explained in section 2.2.3). As mentioned earlier, improving the measurement can only have an indirect impact on performance, i.e. any improvement on fuel efficiency measurement using fuel cards would only have an indirect impact on fuel efficiency and would be relying on adequate driver management. Despite this indirect relation to fuel efficiency, and in view of all the qualities fuel cards have, this study will focus on improving the fuel efficiency measurement based on fuel card data. The details of how this data is used can be found in chapter 5 Case Study and Results.

# 3. Literature Review – Performance Measurement

The previous chapter has explained why this study will focus on improving the mpg measure calculated from fuel card data. This current chapter will define most of the key terms used throughout this study. It will also review the core theory around performance measurement, to finally discuss the main aspects related to measuring performance. The performance measurement methods relevant to this study will be described and discussed in the following chapter 3 Literature Review – Performance Measurement.

## 3.1. Performance and Performance Measurement

### 3.1.1. Key Concepts and Definitions

This section will list the definitions of several notions which are intrinsic to this study. The terms are not listed in absolute alphabetical order as some notions need to be introduced to explain others.

This list of definitions is not exhaustive and some might criticise the absence of some terms (e.g. there is no definition of effectiveness whilst efficiency is defined). These were not included in the list as they are not deemed essential to the purpose of this study. Most terms relevant to this study – including all those deemed non-essential to the understanding of the study – can be found in the 'Glossary of Terms and Abbreviations' section.

**Performance**

The Oxford English Dictionary (1989) defines ‘performance’ as:

> Performance. The action of performing, or something performed… The carrying out of a command, duty, purpose, promise, etc.; execution, discharge, fulfilment. Often antithetical to promise… The accomplishment, execution, carrying out, working out of anything ordered or undertaken; the doing of any action or work; working, action (personal or mechanical); spec. the capabilities of a machine or device, now esp. those of a motor vehicle or aircraft measured under test and expressed in a specification… The observable or measurable behaviour of a person or animal in a particular, usu. experimental, situation… The action of performing a ceremony, play, part in a play, piece of music, etc…

**Measurement**

The Cambridge Dictionary (2008b) defines ‘measurement’ as both:

(noun)

1 [C or U] the act or process of measuring [or]

2 [C] the size, shape, quality, etc. of something, which you discover by measuring it:

Whilst ‘measuring’ (from the verb “to measure”) is define as:

(verb) [L only + noun; T]

To discover the exact size or amount of something, or to be of a particular size

**Performance Measurement**

It is important to observe that the definition of performance can apply to any activity; i.e. there are as many types of performance as there are occasions to perform. On a discussion on performance, Bourne *et al* (2002) provide a different definition of performance in a business context. Performance is there defined as: "the efficiency and effectiveness of a purposeful action". Bourne *et al* used the example of better effectiveness that could lead to a better product, thus better customer satisfaction whilst efficiency might lead to greater profits through cost reduction. 'Performance measurement' could consequently be defined as follow:

> 'The qualification and/or quantification of a purposefully executed action'

Similarly, Harbour (2009) defines 'performance measurement' as follows:

> The 'process of measuring actual outcomes or the end goal of performance, as well as the means of achieving that outcome as represented by in-process measures'

The term 'performance appraisal' is often used in sociology instead of performance measurement. This is probably intended to better reflect most sociologist's tendency to appraise (generally individuals') performance with some qualitative element – i.e. not solely quantitative as per explained by Longenecker and Ludwig (1990). The two approaches are essentially similar yet they both intend to appraise or evaluate the performance of a purposely executed action or entity. As this study's main focus is on measuring productive efficiency, this research will prefer the terminology

‘*Performance Measurement*’ (PM) (see both the definition of efficiency below, and the section ‘Type and Classification of Performance Measures’ for an explanation of the term ‘productive’).

**Efficiency**

Efficiency is defined by the Cambridge dictionary (2008b) as follows:

> ‘when someone or something uses time and energy well, without wasting any’
>
> Specialized: the difference between the amount of energy that is put into a machine in the form of fuel, effort, etc. and the amount that comes out of it in the form of movement.

For the sake of this research, the specialised definition will be retained. I.e. efficiency is the differential between outputs and inputs of a purposely conducted action or process.

**Fuel Efficiency**

Fuel efficiency is consequently the relation between the inputs and outputs used on a vehicle in relation to fuel efficiency. The inputs need to be both related to the ‘production’ of miles and relevant to businesses (e.g. fuel used, cost of fuel...). Conversely, the main relevant output is the number of miles travelled. This will be discussed in more detail in section 5.4 ‘The fuel efficiency model’. The ‘Traditional fuel efficiency’ measure aforementioned is here the ‘miles per gallons’ measure (mpg) used in UK. It is interesting to observe that this measure also links an input (i.e. gallons) to an output (i.e. miles).

**Improvement**

The Cambridge dictionary (2008) defines improvement as follows:

> 'When something gets better or when you make it better'

**Modern Performance Measurement Method**

The limitations of 'traditional' benchmarking – i.e. the comparison of one's performance against others' performance – have been well acknowledged by modern research. This research has developed new methods that have addressed some or all of the limitations of more traditional methods. These methods – mostly developed in the last 40 years – are referred to as 'modern performance measurement methods' in this research. An appropriate selection of these methods will be discussed in the section 3.3 'Performance Measurement Methods'.

**Van**

A van is an independent small vehicle on which the driver's cab and the load carrying compartment are mounted on the same (rigid) chassis. The term van is used by official bodies such as DfT (Anon, 2005a) or in research such as in the paper from Brackstone *et al* (2009). The size of a van greatly varies, ranging from small vans with a gross weight slightly greater than 1000 kg to large vans weighting up to 3.5 tonnes. Vehicle weighting between 3.5 tonnes and 7.5 tonnes are sometimes called heavy vans in opposition to light vans (from 1 to 3.5. tonnes). This study will concentrate on vehicles up to 3.5 tonnes only.

**Fuel cards**

Fuel cards are special cards given to drivers or employees which allow them to purchase fuel or goods at petrol stations or to access fuel at company or associate fuel tanks.

### 3.1.2. Purpose of Measuring

Performance measurement has been of interest from times as far back as Aristotle (Landy and Zedeck, 1983). This need for improving a given action or process can be explained, amongst others, by common constraints such as scarce resources, harsh environment, competition or ecological motives. Halachmi (2002a) and Harbour (2009) also state that it is not possible to know whether the performance has improved or worsened without appropriate and adequate measurement. This section will thus discuss in more detail the need to measure performance.

Harbour (2009, p.1) pinpoints, in a very clear manner, one of the main reasons to measure performance when stating: 'You can't understand, manage, or improve what you don't measure' and 'a critical enabler in achieving desired performance goals is the ability to quantitatively measure performance'. This author sees performance measurement as not only the process of quantifying an actual outcome or an end goal, but also as the means of achieving that outcome. In this respect, performance measurement can help a business or structured organisation to (Harbour, 2009, pp.5-6):

⇨ Determine where they are (baseline)

- ⇨ Establish goals on the basis of their current performance
- ⇨ Determine the gap or delta between a set of desired goals and current performance levels
- ⇨ Track progress in achieving desired performance goals and ensure that such goals are maintained
- ⇨ Compare and benchmark their competitors' performance levels with their own
- ⇨ Assess variation within a system or process and help control such variation within predetermined boundaries
- ⇨ Identify problem areas and possible causes,
- ⇨ Make more informed performance, cost, and fact based decisions
- ⇨ Allow better forecasts to be made

The key steps which are necessary to know (or prove) that performance has actually improved (or worsened) are to determine the actual performance (by measuring) before and after a modification to the method or to the process. Harbour includes acting on performance measurement results (i.e. 'Make more informed performance, cost, and fact based decisions' & 'Better plan for the future') as part of performance measurement. The relation between improvement and performance measurement will however be further discussed in the section 'Performance Measurement Recommendations'.

Conversely, Halachmi (2002a) gives a comprehensive list of reasons supporting the need for performance measurement which are detailed below:

- ⇨ if you cannot measure it you do not understand it;
- ⇨ if you cannot understand it you cannot control it;
- ⇨ if you cannot control it you cannot improve it;
- ⇨ if they know you intend to measure it, they will get it done. This point is also mentioned in (Eccles, 1995);
- ⇨ if you do not measure results, you cannot tell success from failure;
- ⇨ if you cannot see success, you cannot reward it (this point is again mentioned in (Eccles, 1995));
- ⇨ if you cannot reward success, you are probably rewarding failure((Eccles, 1995) p. 95);
- ⇨ if you will not recognize success you may not be able to sustain it;
- ⇨ if you cannot see success/ [or] failure, you cannot learn from it;
- ⇨ if you cannot recognize failure, you will repeat old mistakes and keep wasting resources;
- ⇨ if you cannot relate results to consumed resources you do not know what is the real cost;
- ⇨ if you do not know the actual cost you cannot tell whether or not you should do it or outsource it;
- ⇨ if you cannot tell the full/ real cost you cannot get the best value for money when contracting out;

- ⇨ if you cannot demonstrate results, you may undermine your ability to communicate with important stakeholders to mobilize necessary support because you provide value for money;
- ⇨ if you cannot document that the business process, material or people you use are the most suitable for achieving the sought after results your performance will be questioned;
- ⇨ if you cannot show that in comparison to the past or to another provider you are on a par or doing better there may be question about your accountability and;
- ⇨ if you do not have the data about who is happy/unhappy with your performance and why, you may change when you should not or, even worse, stay a course [which] on its face seems to be right but in fact is wrong.

In this compelling list of reasons why performance measurement is often essential, Halachmi covers a wide range of topics from the need to understand the business' operations to create an appropriate performance measurement system, accounting reasons (e.g. 'if you cannot relate results to consumed resources you do not know what the real cost' is), or sociological/organisational reasons (e.g. 'they know you intend to measure it, they will get it done' – a point well discussed by Hayes *et al* (1988)). The psychological aspects linked to performance measurement will be discussed with more details in the section 'Possible Risks and Issues in Measuring Performance'.

It is important to observe that the purpose of measuring – whether it is to ensure improvements occur, to keep performance levels under control, or to better control costs or people – often differs depending on who is concerned by the performance measure. Brady ((Brady, 1985) cited in (Longenecker and Ludwig, 1995)) discusses this issue by proposing his 'Janus-Headed' model (a god from the Greek mythology who had a head with two faces). Longenecker and Ludwig (1990) illustrate this Janus-Headed model concept by mentioning that the manager sees the (employees') performance measurement as a means whilst the Human Resource department might see it as an end.

Finally, performance measurement does not necessarily bring improvement to a business. It is rather the adequate and appropriate actions based on the performance measure results that can lead to improvement. Murphy and Cleveland (sociologists and researchers) see performance measurement as both the qualitative or quantitative measurement and the communication of these results. They state that considering performance appraisal strictly as a measurement instrument is unrealistic if the follow up communication is not taken into account (Murphy and Cleveland, 1995).

### 3.1.3. Type and Classification of Performance Measures

Research on performance has proposed many different classifications of performance measures and proposed an even larger number of characteristics or possible attributes to these measures (Christopher and Thor, 1993, Abbott, 1994, Harbour, 2009, Gass and Prince, 1993). This section will introduce one classification proposed by Harbour (2009) on performance measures' types and families. Numerous dichotomous measure's characteristics which regularly appear in the literature (e.g. qualitative vs. quantitative) will also be introduced. The characteristics which families of measure should possess (e.g. cost, productivity, quality, timeliness...) relate more to performance management and will consequently be discussed in section Performance Measurement Recommendations below.

The previous section highlighted the need to establish the baseline performance prior to an improvement exercise. This was necessary to evaluate whether an improvement actually occurred. In view of this, Harbour proposed the following three performance measurement categories:

- ⇨ Descriptive measures: measures which describe what has happened;
- ⇨ Diagnostic measures: helping to understand what caused a good or bad performance;
- ⇨ Predictive value: a value that helps forecasting what will happen based on what has been measured.

**Descriptive Measures**

Descriptive measures, also called lagging-indicators (Anon, 2009a) and (Erikson, 2009), describe what is happening or has happened. These measures generally include baseline and trending performance measures.

Baseline measures show the current levels of performance. They are essential to any performance measurement system as without a baseline, it would be impossible to know whether improvement has occurred or not. A trending measure is essentially the same but considers performance over time so trends can be shown (and improvement demonstrated). An illustration of a descriptive measure could be the mpg (miles per gallons) of a vehicle. A baseline measure could be let's say a mpg of 7.5 for a HGV and its corresponding trending measure the plot of the HGV's mpg performance over a year; the latter showing clearly whether performance is constant, worsening, or improving.

**Diagnostic Measures**

These measures do not describe what is happening but help identifying why something has happened. Taking the example of a manufacturing product cycle time, a descriptive measure might be the overall product cycle time, and the corresponding diagnostic measure the cycle time of the individual processes. In this case the predictive measure is made of the diagnostic measure but this is not always the case (e.g. diagnostic measures for mpg would be any variables impacting mpg for example driver behaviour, load weight or weather conditions).

**Predictive Measures**

Predictive measures, often called 'leading indicator' (Chatchai et al., 2007), are used to extrapolate what is likely to happen in terms of performance – based on past observed performance. Developing good predictive measures is often difficult and most predictive measures will also require some degree of interpretation. Harbour

(2009) exemplifies this with the example of an S curve illustrated in Figure 2.1 Costs Distribution for Typical Transport Operations.

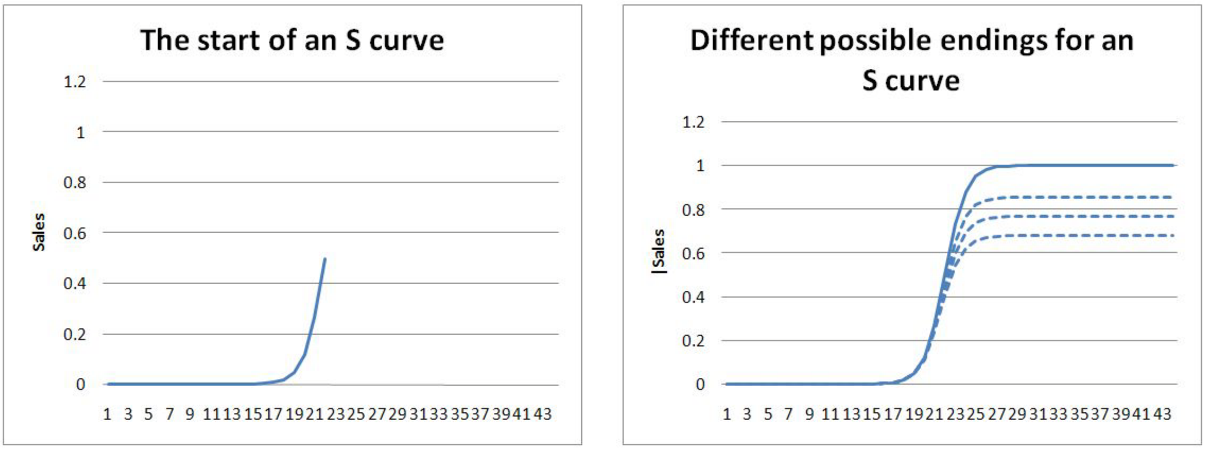


Figure 3.1: The S curve

Here, the predictive measure was the volume of sales over a period of time (trending measure). As illustrated by the graph on the right hand side, forecasting the sales volume based on the start of the 'S' curve (left hand side graph) will require expert knowledge (this is illustrated by the many different potential curve endings on the right hand side figure).

Aside from the framework introduced above, there are numerous dichotomous characteristics for performance measure which can be found in research papers. The most commonly used characteristic is probably quantitative versus qualitative. This attribute relates to the possibility to measure a variable using a numerical scale. For example, the distance an athlete can jump is a quantitative variable whilst how a person is appreciated by their colleagues (or not) is qualitative variable. Interestingly, qualitative variables are often 'encoded' using numerical scales (such as the Likert scale – the term 'Likert' being an eponym) in order to run statistics on the results (Harzing et al., 2009).

Another important distinction can be made on whether the performance measure is depicting efficiency or effectiveness, the former relating to the how much output is created in regards to the inputs, while the latter only considers how well the task or process has been completed (see the section Key Concepts and Definitions and the Glossary of Terms and Abbreviations for term definitions).

Zairi (1996, p.393) also proposed a two-way classification of performance measure by suggesting a process versus results approach which he refers to as in-process measures and output measures. He describes the process based measures as relating to a process and providing a swift performance feedback whereas result based measures relate to ‘broader issues or targets’ and are used more as ‘management information’. Conversely, Lawson ((1995) in (Walters, 1995, p.11)) distinguishes internal versus external performance measure. He argues that the most important measures are the external ones as those are visible to the customer (e.g. quality or delivery times).

This section has briefly introduced a 3 classes classification from Harbour (2009) to then describe several dichotomous classifications which are often used in the literature. The list presented above does not intend to be exhaustive in any way, but rather to give a quick outline on different performance measure classifications. The characteristics which performance measure families should have (e.g. cost, timeliness...) will be discussed in the ‘Performance Measurement Recommendations’ section as they relate more to performance management systems rather than mere taxonomic discussion.

## 3.2. Around the Performance Measure

Now that the key definitions, purpose and main taxonomies of performance measurement have been discussed, this section will delve in to more details regarding performance measurement. The aspects related to data and data gathering will be introduced in the first section. The possible risks and drawbacks associated with measuring performance will then be discussed. This section will finally end with a series of recommendations regarding performance measurement. These were collected from selected research papers. The notion of performance management will also be introduced at the end of this section.

### 3.2.1. Data Gathering

Research projects always require data to conduct analysis. The data collected can be of various forms and can for example consist of research papers (i.e. reviewing peer reviewed papers and discussing them), statistical data collected by different bodies (McNally, 2008), or in-field data. Probably due to this invariable need for data, the data collection methods and associated issues have been well researched. However, some distinctive differences between quantitative and qualitative collection methods exist. In light of the quantitative nature of this study, this section will focus mainly on general issues and the quantitative aspects of data collection.

Curwin and Slater (2002, p.31) stress the fact that the data collected need to be appropriate, adequate, and without bias. It is effectively not relevant to gather statistics about accident rates whilst the purpose of the research is fuel efficiency (unless there is reason to believe there is correlation and causation between accident rate and fuel efficiency of course). To know whether the data are adequate requires being clear about the problem boundaries. This can generally be answered by considering what needs to be achieved or what the customer's requirements are. Data collection can also be constrained by the time or resource limitations of the project. Another obvious property of collected data are that it needs to be representative (Curwin and Slater, 2002, p.14), e.g. no point in collecting LGV's data when the study is only concerned with vans.

Another concern when gathering data is their own accuracy, concern which is present in both qualitative (e.g. (Murphy and Cleveland, 1995, p.7)) and quantitative research (e.g. (Vincent, 1998, p.59 onwards)). The potential reasons for inaccuracies in the data are various. Murphy and Cleveland mention a potential leniency and other cognitive issues as potential causes behind data inaccuracy (1995, p.96). For instance a certain leniency is for example generally found in the rating of employees for administrative purposes in comparison with when the rating is for research purpose. The reasons for such behaviour being well detailed in the paper 'Ethical dilemmas in performance appraisal revisited' (Longenecker and Ludwig, 1990). On a more general basis, Stephen Vincent (1998, p.60) mentions normal error rates and annoyances which can be, amongst others, rounding inaccuracies, or inaccuracies due to the fact the operator not being aware of the accuracy requirements. The author illustrates the sampling and data collection processes in Figure 3.2: Data Sampling.

Curwin and Slater (2002, p.32) also describe potential biases in the collected data as another potential risk when gathering data. Biases can occur when the sample data misrepresent the population, thus biasing any inference made to the population studied. This issue is particularly important when the research attempts to prove a whole population's characteristic by studying a sample of this population.

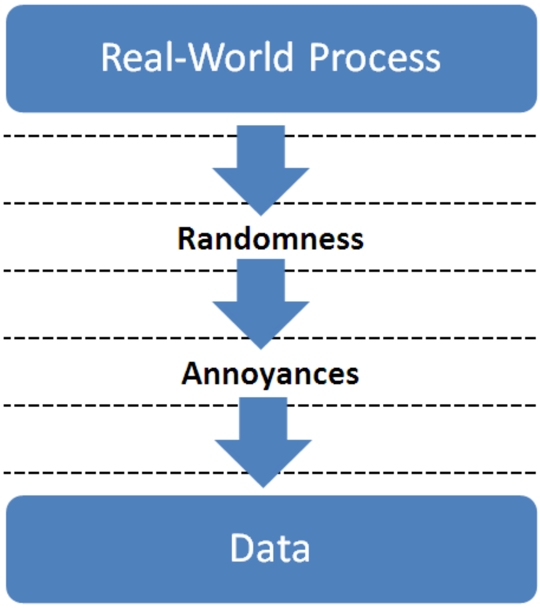


Figure 3.2: Data Sampling

In order to reduce the risk regarding potential inaccuracies and bias, several methods can be used. Axinn and Pearce (2006) suggest mix-methods as a solution offering better control. This methodology suggests using different methods (e.g. surveys, interviews, Delphi method, etc) in order to provide different orientations of the same problem and thus, limit bias. This approach is also shared by Saunders *et al* (2009) where the authors define mix-method strategies as methods which combine quantitative and qualitative for the data collection phase as well as for other phases of the research. Axinn and Pearce (2006, p.1) further add that mix-methods offer opportunities to use multi-sources of information which in turn reduce the risk of non-sampling error by 'providing redundant information from different sources and ensure that a potential bias coming from one particular approach is not replicated in alternative approaches'. It is also possible to further reduce the risk of random errors by using significance testing or bigger samples.

Finally, it is important, when using some methods (e.g. simulation, DEA), to assess the independence of the different variables and of any potential correlation. This can be done with both linear variables' correlation and scatter diagram techniques, or with formal tests. However, Vincent (1998, p.61) recommends the first two approaches when the assumption regarding the populations' distribution cannot be made.

### 3.2.2. Possible Risks and Issues in Measuring Performance

Section 3.1.2 'Purpose of Measuring' explained that measuring performance does not guarantee that it will systematically improve. This is especially true in this context as the drivers do not need to be aware of the measurement, thus human behaviour effect such as the one observed in the Hawthorne experiment are not relevant to this study (Franke and Kaul, 1978). This lack of connection between measurement and performance improvement might be due to already existing good performance levels allowing little room for improvement, to the fact that the performance measures are irrelevant, or that there is no action taken on the performance measurement results. Measuring performance also involves a certain number of risks; this section will discuss the most common ones.

Measuring performance is often believed to systematically lead to improvements in performance. Yet, Halachmi (2002b) points out that potential dysfunctions of performance measurement should not be ignored. He particularly highlights the risk that the cost of measuring could potentially exceed the benefits of measuring performance thus nullifying the benefits of performance measurement. It is also quite hard to estimate potential improvement although monitoring a few key performance indicators (KPIs) over time can help appraising the room for improvement. This relates to the 'S' curve, i.e. performance improves slowly at first until reaching a constant tangent, to finally finish with a slow slope once again (Harbour, 2009, p.76).

Section 3.1.2 'Purpose of Measuring' also explained that the purpose for performance measurement generally differs from the people or the department involved. Longenecker and Ludwig (1990) gave the example that performance measurement is generally seen as an end for the Human Resources department but as a means for the managers. This difference in purpose can lead to some deliberate performance measure manipulation as cited by Murphy and Cleveland (1995, p.103). Similarly, Eccles (1995) states that managers manipulating performance measurement figures and the consequent need to secure data relating to or used for measuring performance (e.g. employees manipulating sales data to collect greater bonuses). Although the reasons why performance measurement results are sometimes manipulated is beyond the interest of this research, it is important to cite Longenecker and Ludwig (1995) who have listed a compelling list of potential reasons for these manipulations to happen. Finally, Harbour (2009, p.31) mentions the importance of ensuring the performance measure chosen and the corresponding data cannot 'be easily manipulated to achieve desired results'.

Fenner also states that there is a risk associated with measuring performance if the decisions made are based on insufficient or inappropriate performance measures (2002). Curtis (1985) (cited in Eccles, 1995, p.62) observed this phenomena in most companies of the 1980's. At this time companies' performance measures were mainly focusing on economical or accounting aspects missing an important aspect of performance relating to quality and customer service. Similarly, King *et al* (in (2002) (cited in Halachmi, 2005)) also suggest that there is (still) a growing concern among scholars who fear executives may base their decisions on "sometimes-arbitrary performance measures rather than [for] improving public management per se". Eccles (1995) also stressed the risk that focusing excessively on short term performance measures could potentially impede potential but necessary long term investments from being made.

Furthermore, there are also some important legal aspects to performance measurement when it is used to make decisions about people. Murphy and Cleveland (1995, p.11) stress the importance of being able to explain the performance measurement mechanism. This is because any system used to make decisions about people is subject to court action on its accuracy and validity.

Since this observation, Cullen (1999) (cited in Halachmi, 2005)) has pointed out that the number of evaluation systems has proliferated so much that the 'selection of appropriate measures has become a difficult task' and that 'none of us can answer the question 'does performance measurement improve organizational effectiveness?' positively and confidently…'. Even though there are strong and natural incentives to measuring performance (see aforementioned list), more performance experts seem to agree that simply measuring performance without including performance measurement in a wider picture can fail to bring the expected benefits of measuring.

### 3.2.3. Performance Measurement Recommendations

Whilst the previous sections have mainly looked at performance measurement classifications and the risks associated to performance measurement, there have been few recommendations on how performance measurement should be approached. This section will introduce the concept of families of measures as well as detailing which essential characteristics performance measurement should have. The notions of performance measurement framework and performance management will then be briefly introduced.

The performance of most purposely executed actions has generally several facets. This might be best illustrated by a company that needs to be creative and innovative in design, while efficient in production and effective in selling the products it produces. This can also be illustrated with a vehicle that needs to be simultaneously mpg efficient and at the same time pence per mile efficient. Because in most situations a single performance measure is rarely enough to encompass all aspects of performance, families of measure are generally required instead (Harbour, 2009).

The concept of family is generally attributed to the work of Christopher and Thor (1993, p. 2–6.2). These authors define a family of measures as a group of performance measures that capture all important aspects of a given process or organisation performance. Harbour (2009, p.26) illustrates the concept of family of measures with an analogy to a vehicle dashboard. The vehicle dashboard has several measures (speedometer, fuel gauge, rpm...). All these measures capture the key attributes of the car and are all necessary. This is exactly the same principle with families of measures. Although families should encompass all aspects of performance, Harbour (2009) also recommends that only the vital few should be selected in order to avoid diluting what is really important.

Carl Thor (1993) discussed the attributes which measures of a family should ideally have. Harbour (2009, p.26) and Gass & Prince (1993, p4-8.3) further discussed this topic. A summary of their work is described below:

**Productivity**: this relates to the notion of efficiency or the relation between the outputs produced and the inputs used. An example of a productivity measure could be the number of deliveries per hour for a delivery van.

**Quality**: the quality measures generally reflect quantity of undesirable outputs (scrap, waste, $CO_2$ emissions...) or (customer) satisfaction levels and repeat frequencies.

**Timeliness**: these measures are related to how well things are delivered when they should be.

**Cycle Time**: cycle time measure are generally used in manufacturing to represent the time it takes for goods to be processed through a machine or for raw products to be transformed in to a final product.

**Resource Utilisation**: these measure how well capability is used. An example of this could be the vehicle fill (percentage of volume filled, percentage of weight loaded against total possible weight).

**Cost:** these measures are especially useful when used with a comparator, e.g. cost per unit, cost per vehicle/mile...

**Safety:** these measures show level of safety associated with particular activities (e.g. number of accidents, number of nearly misses).

This list is obviously not exhaustive and it is not essential for all families to have measures in each of the categories listed above. Conversely, it is possible to have a family with categories other than those listed above. The key criteria in the performance measures selection is to always ensure that the measures selected help achieving the desired performance levels. Eccles (1995, p.3) crystallises this when writing: ‘businesses must ask themselves what measures truly predict their [company’s] long-term financial success’.

Harbour (2009, p.30) gives further essential characteristics which performance measures should demonstrate:

**Have a comparative basis**, i.e. information that offers the possibility to appraise whether the performance measure level is good or bad. For example, it is not possible to know whether a vehicle's mpg is good or bad without knowing which vehicle type the mpg is for (e.g. 25 mpg does not mean anything if one does not know whether the vehicle is a car, van, or rigid).

**Be timely**. Information has an expiry date and late information is often useless. For instance, there would be no need for a driver to know the vehicle's speed with a 5 minute delay.

Conversely, Bititci and MacBryde (2002) state the need for performance measures to be dynamic and change with the natural changes of the organisation. This point is also supported by Maskell (1992) cited in (Folan and Browne, 2005).

Most of the above recommendations for performance measures are applicable in most cases. Research has however developed many frameworks to help conducting performance measurement study. Folan and Browne (2005, p.665) define a framework as 'the active employment of particular sets of recommendations'. They also distinguish two types of frameworks. The structural frameworks (e.g. the Balanced Scorecard model from Kaplan and Norton (1996)) help deciding which aspects of performance should be measured. On the other hand, the procedural frameworks for example Wisner and Fawcett's framework, (1991) define the overall performance measurement and action procedure.

The management of the performance measurement in relation with the organisation and its strategy has given birth to the new concept of performance management. Amaratunga and Baldry (2002) define performance management as the use of ‘performance measurement information to effect positive change in organisational culture’. They describe performance management as the transition from measurement, which simply states what has happened, to management and analysis, which explains why it has happened and potentially gives direction on what should be modified or corrected. Smith and Goddard (2002) capture the relation between performance management as illustrated below in Figure 3.3: Schematic representation of the performance management process.

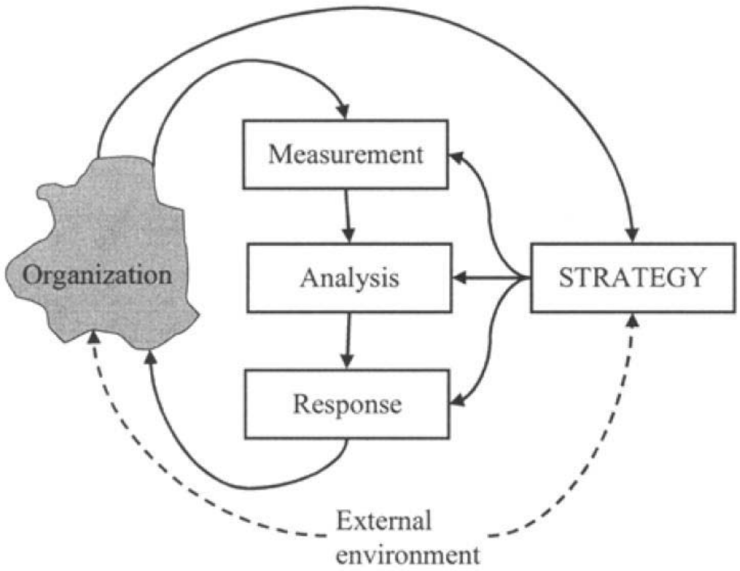


Figure 3.3: Schematic representation of the performance management process

Band (1990) cited in (Folan and Browne, 2005, p.665) also recommends that the performance measures have ‘top management support’. This point is also emphasised in Harbour (2009).

Research has suggested that it is necessary to consider the interactions between performance measurement and the organisation itself when implementing a performance management strategy (Halachmi, 2002a). This study will however focus chiefly on the technical aspects behind the actual measurement of performance rather than the interactions between those results and the organisations in question. Thus the technicalities of performance management will not be discussed further.

Section 3.1.1 Key Concepts and Definitions has introduced the key terms in relation to performance measurement. The purpose of measuring, performance measure taxonomy, and other issues related to performance measurement have also been described and discussed. All these aspects were briefly introduced and more could be said on all these subjects. The aim was solely however, to briefly introduce the key concepts related to performance measurement in regards to the study's objectives of improving fuel efficiency measurement. The concepts discussed in this section will be used in Chapter 5 Case Study and Results. The following section, 'Performance Measurement Methods' will first introduce some performance measurement methods of potential interest to fuel efficiency.

## 3.3. Performance Measurement Methods

Chapter 2 highlighted the potential interest in measuring van fuel efficiency using fuel card data. This previous section also listed several limitations which the mpg measure demonstrated and – in light of the importance of fuel performance measurement – suggested that these limitations should be addressed by modern performance measurement methods. These limitations were as follows (key summary of Section 2.3 Explaining the Focus on Fuel Efficiency Measurement Based on Fuel Cards):

- ⇨ The mpg measure does not include factors necessary to its interpretation (e.g. mainly vehicle weight and potentially, vehicle type, type of operations and vehicle age).
- ⇨ The mpg measure does not reflect fuel cost. It is conceivable that a vehicle can be mpg efficient but cost inefficient (e.g. by buying a more expensive fuel offering greater mpg performance). In industrial language, this means that mpg is not a Total Factor Productivity measure, i.e. a measure which accounts for all inputs and outputs.
- ⇨ The mpg measure is often misused in fuel trials.

Similarly, Chapter 3 introduced the basic concepts of performance measurement theory. This chapter provided recommendations about which characteristics a good performance measure should demonstrate and additionally warned against potential issues such as those linked with data gathering.

This section will first introduce several modern performance measurement methods that could be used to improve fuel efficiency measurement in the van industry. Their respective characteristics in terms of how well they address mpg's limitations and how they fit with the theory will be discussed. The appropriateness of each of these performance measurement methods is summarised in a table which can be found in section 4.1 Reasons for this study to use Data Envelopment Analysis.

The issue related to misuse of mpg measure in fleet fuel trials is not dependent upon the performance measurement method chosen. As a result, this issue will not be discussed in this section but in sections 5.3.3 Smoothing Algorithm, 5.3 Data Cleansing and chapter 6 Summary of Results and Discussion.

### 3.3.1. Traditional approaches

The most basic performance measurement method consists of using appropriate scales in order to appraise the performance of a process or action. This measurement is often realised through Key Performance Indicators (KPI) which are generally used to measure specific aspects of performance (e.g. for vehicle load: vehicle fill in percentage, weight expressed as percentage of maximum permitted load, etc). The two main KPIs used to reflect fuel efficiency are *miles per gallons* (mpg) and *pence per mile* (ppm).

KPIs are widely used throughout road operations and have been well documented in research (e.g. McClelland and McKinnon, 2004) and by the government (FBP, 2008b). KPI's calculations are generally straight forward (e.g. mpg calculation requires dividing the number of miles travelled by a vehicle by the number of gallons used to cover these miles). However, the data gathering and cleansing can generally prove to be more difficult (Curwin and Slater, 2002). When developing a performance measurement system, it is also relatively easy to develop KPIs for each of the descriptive, diagnostic and predictive category  of measures (as recommended by Thor (1993)). Taking the example of fuel efficiency on vehicles equipped with CANbus (see section 2.2.2 for more information on CANbus technology) this could be exemplified as follows:

- *Descriptive measure*: could be the mpg measure calculated directly from the CANbus distance and fuel used information. The mpg measure 'shows' the fuel efficiency performance.
- *Diagnostic measures*: could be over-revving, over-acceleration, or idling. These different KPIs help explain the performance level shown by the descriptive measure.
- *Predictive measure*: could be the mpg measure plotted on a daily or journey basis. This would show the fuel efficiency performance trend.

As with any performance measure, it is essential that care is taken in the appropriate selection of the different KPIs in order to truly reflect all aspects of performance and avoid using misleading measures (for further information, see section 3.2.3 Performance Measurement Recommendations).

Although KPIs used independently are ideal at describing, diagnosing and predicting performance, they remain an internal oriented approach; thus they fail to indicate what level of improvement is possible. When operational tasks are carried out by several similar units, it is however possible to compare – without bias – their respective performance on each KPI. This externally oriented approach, called benchmarking, can indicate which improvements are possible.

Benchmarking is a performance measurement and evaluation technique that compares an entity's performance against other entities' performance. One of the most quoted definition of benchmarking is 'Benchmarking is the search for the best industry practices which will lead to exceptional performance through the implementation of best practices' ((Camp, 1989), cited in (Anand and Kodali, 2008)). Benchmarking studies can be used to evaluate the comparative performance of people, business units, or entire companies. The origins of the technique are credited to Xerox in the United States which was at the time - and along with other printer companies in the US – threatened by Japanese competition. The technique has since then quickly developed in Europe in to a widespread management technique and which was quickly adopted by the rest of the world. Although originally mainly driven by cost reduction objectives and Total Quality Management programs, benchmarking is now used in a wider business context to identify best practices. By comparing performance within the same industry but also with other industries that share the same business processes, benchmarking can not only identify and quantify performance gaps but also uncover the practices that leads to competitive advantages (Dence, 1995).

Many benchmarking models have been developed (Anand and Kodali, 2008). The following categories are regularly quoted however (Dence, 1995), (Isoraite, 2005):

- ⇨ *Internal benchmarking*: is the type of benchmarking that is done internally, either between related divisions or departments, or between plants or equivalent business units. Because of the relative easy access to information and parties' cooperation, this is generally appraised as a good start for benchmarking activities.
- ⇨ *Functional benchmarking*: this relates to the performance comparison of functionally similar operations but in different organisations or

companies. For example this could be the benchmarking of customer service calls in a call centre company, a telematics company, or a heavy industry company. This is similar to *process benchmarking*.

- *Competitive benchmarking*: relates to the benchmarking of direct competitors within the same industry sector or indirect competitors in similar industries. This benchmarking is generally difficult to do directly as competitors might be diffident about providing their information. A third party is generally used to handle safely the confidential information. Solely limiting this benchmarking to general performance measures such as quality of service might identify best performers but not necessarily identify best practices.
- *Generic benchmarking*: relates to the benchmarking done between companies from different industries but all best in class for some of their operations.
- *Strategic benchmarking*: refers to the corporate benchmarking at a strategic level.

The specific requirements of each benchmarking study have resulted in the development of many different methodologies and approaches to benchmarking. These generally include different steps, usually ranging from 6 to 20 steps. The following steps seem however to be frequently included in most approaches (Dence, 1995):

- determine the key (aspects of) performance to measure

- ⇨ set the key standards and variables to measure
- ⇨ identify the most relevant competitors and best-in-class companies
- ⇨ measure regularly and objectively (quantitative)
- ⇨ analyse the best-in-class performance (qualitative)
- ⇨ specify programmes and actions to close the gap, and implement them
- ⇨ monitor on-going performance

As by nature external or competitive benchmarking techniques use performance indicator information from different parties and do so in order to provide information to other parties, legal and ethical issues need to be seriously taken into account when undertaking a benchmarking study. A third party or a system can handle the information in order to limit the risks and enable the different parties to focus on the benefits of information sharing rather than the associated risks. Confidentiality agreements can also be written in order to ensure legal acceptance of the confidentiality of the information. Depending on how the analysis is conducted, benchmarking could also require a minimal number of participants or a minimum dataset in order to provide valid results. Despite the challenges specific to the method, benchmarking generally provides new insights on business performance measurement by measuring performance not against historical data or future pre-defined levels but by comparing it against best performers. In light of this, it does not seem surprising that DfT's programme Freight Best Practice has launched a road transport benchmarking programme online called On-Line Benchmarking or 'OLB' (Anon, 2008a).

### 3.3.2. Pair-wise and outranking methods

Benchmarking using an external oriented approach can potentially make participants aware of possible changes that are of an order of magnitude beyond what they could have originally thought possible. When it comes to multi-criteria benchmarking, participants being best-in-class for all criteria are rarely observed however, thus a method is needed to address which are the best-in-class. Sharif (2002, p. 76) emphasises this issue by stating the following:

> "There is no performance management enterprise […] that will be best across all areas".

In the absence of a best-in-class performer, traditional benchmarking solves the multi-criteria problem with the creation of a synthetic indicator calculated by making a weighted average of each score (Laise, 2004). The performer with the best average score will be considered the best-in-class (e.g. as in Goh and Richards, 1997). This approach has several inherent problems however. Average is a measure of a central tendency that is a representative value when data have a low variability, which may not always be the case with benchmarking performance scores. This averaging issue has been addressed by several pair-wise and outranking comparison methods. The Analytic Hierarchy Process (AHP) is a powerful pair wise comparison tool which can help making decisions (by ranking different possibilities). AHP was first proposed by Thomas Saaty (1980) cited in (Sureshchandar and Leisten, 2006). Several outranking methods successively developed by Roy and Bouyssou and Pomerol and Barba Romero also offer an interesting alternative to AHP (Bouyssou and Roy, 1993), (Barba-Romero and Pomerol, 2000) cited in (Laise, 2004). This section will review and discuss how these methods address the averaging issue associated with traditional benchmarking.

#### 3.3.2.1. Analytic Hierarchy Process

As Sharif (2002) observed, best-performers in traditional multi-criteria benchmarking are rarely found; besides, using averages to find best performers has been shown not ideal (Laise, 2004). Pair-wise comparison such as the AHP method are particularly well suited to address complex decision making problems such as ranking performers based on their relative performance on several different characteristics. The decision problem is structured with the overall focus or objective at the top, the criteria at the middle and the decision variables at the bottom.

Although the AHP allows for multi-level decisions (e.g. with sub-criterions), it is easier to explain the concept with the 3 level AHP model. This model can be represented as in Figure 3.4 (Sureshchandar and Leisten, 2006).

To compare different options in regards to several objectives, AHP uses the following:

- ⇨ The scores obtained by each option for each objectives and
- ⇨ Some objective weights (illustrating the objectives' importance).

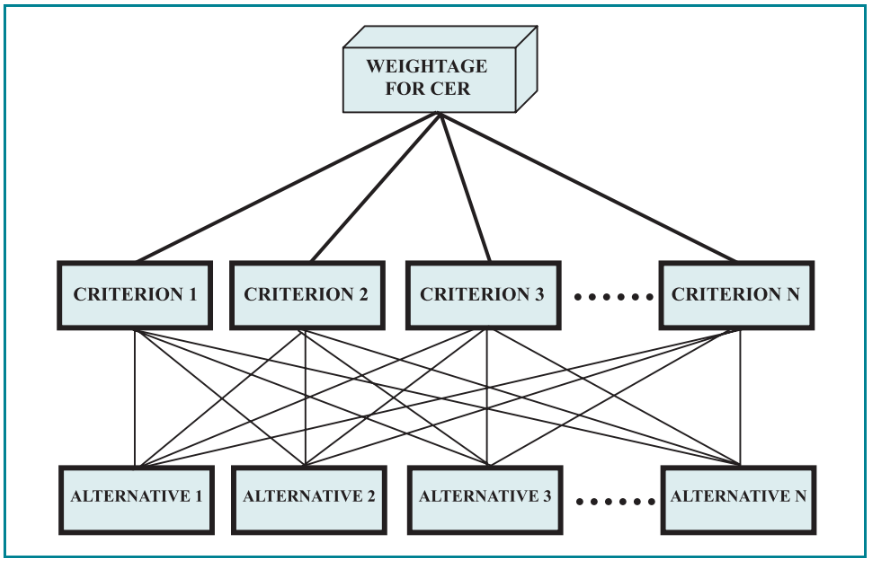


Figure 3.4: The 3 levels AHP structure

The scores are obtained via any kind of performance measurement system (e.g. KPI). The weights are however calculated by asking experts to evaluate how much more important a specific objectives is in comparison to another. Each option's score is then calculated using the options' scores on each objective and the corresponding calculated objective weights (Winston, 2004, p. 785).

AHP requires every objective to be compared with the others in a pair wise manner using a 1-9 scale in order to evaluate the dominance of each objective on the others. The decisions are entered in a pair-wise comparison matrix '**a**' (where entry $a_{ij}$ represents how much objective **i** is preferred over objectives **j**). The objectives' weights then correspond to the matrix's Eigen vector (Sureshchandar and Leisten, 2006, p.24). A similar approach is taken to find how much each option scores for each objective (Peters and Zelewski, 2008).

This process can be illustrated while trying to evaluate different job opportunities. In this scenario, the choice of the job offer could be made based on the following objectives:

- ⇨ Rent cost (RC),
- ⇨ Infrastructure quality (IQ),
- ⇨ Proximity to suppliers (PS),
- ⇨ Road network quality (NQ),
- ⇨ Nearby garages quality (GQ))

The pair-wise comparison matrix to obtain the objective weight (with a specific AHP scale generally ranging from 1 to 9) for the given problem could then look as in illustrated in Figure 3.5:

$$\begin{array}{c c} & \begin{array}{ccccc} RC & IQ & PS & NQ & GQ \end{array} \\ \begin{array}{c} RC \\ IQ \\ PS \\ NQ \\ GQ \end{array} & \begin{bmatrix} 1 & 5 & 2 & 4 & 5 \\ 1/5 & 1 & 1/2 & 1/2 & 1 \\ 1/2 & 2 & 1 & 2 & 2 \\ 1/2 & 2 & 1 & 1 & 2 \\ 1/5 & 1 & 1/2 & 1/2 & 1 \end{bmatrix} \end{array}$$

$$Where\ the\ rent\ cost\ (RC) is\ deemed\ 2\ times\ more\ important$$
$$than\ the\ proximity\ to\ suppliers (PS)\ (a_{13}).$$

**Figure 3.5: The pair-wise comparison matrix**

The pair-wise matrix above shows that the rent cost (*RC*) is judged two times more important than the proximity to suppliers (*PS*). Similarly, the proximity to suppliers (PS) is judged two times more important than infrastructure quality (*IQ*). This means that the rent cost (RC) should be logically 4 times more important than the infrastructure quality (*IQ*). However, the matrix shows that rent cost is perceived as 5 times more important than the infrastructure quality.

This is the type of inconsistency which is checked for in the weight calculations. Without detailing the inconsistency checks calculations, the matrix's Eigen vector corresponds to the objectives weight.

With these weights, it is now possible to use the options' objective score (the score attributed to each option for each of the evaluating objectives) to discriminate against each option. The options' scores are calculated using Formula 3.1.

$$\sum_{i=1}^{i=m} w_i(option\ j's\ score\ on\ objective\ i)$$
$$where:$$
$$m\ is\ the\ total\ number\ of\ objectives,$$
$$j\ is\ the\ option\ for\ which\ the\ score\ is\ calculated$$

**Formula 3.1 The AHP option's score formula**

In this example, this means a score would have been calculated for each depot and each of its parameters as per described above (RC, IQ, PS, NQ and GQ). The method used to calculate these scores is down to managerial decision. These scores will then be used, along with the weight calculated by the method described above, to calculate a unique score for each depot.

Although this method still weights each criterion, these are not arbitrarily chosen as with traditional benchmarking, but generated through a more robust process.

#### 3.3.2.2. ELECTRE methods

Methods such as the AHP or the SMART method (Edwards, 1977) cited in (Buchanan and Vanderpooten, 2007) are well suited for problems with a 'finite number of discrete alternatives'. ELECTRE methods – which are outranking techniques – differ from these traditional approaches by using a strict dominance approach (which can be relaxed), and by offering the possibility to add a fuzzy factor to the decision making process in order to illustrate the nature of decision making. For example, the ELECTRE III method considers whether the difference between two values is significant or not. ELECTRE techniques are also non-compensatory, i.e. a poor score in an area cannot be compensated by other scores in other areas.

ELECTRE (I) works by creating a matrix of concordance subsystems from a calculated multi-criteria matrix (i.e. a matrix of scores, Laise, 2004). For a list of 4 organisations which need to be evaluated on 5 different objectives, the matrix of organisations' scores on each objective could look as demonstrated in Table 3.1 (Laise, 2004).

| . | Objective 1 | Objective 2 | Objective 3 | Objective 4 | Objective 5 |
|---|---|---|---|---|---|
| **Org. 1** | 3.50 | 3.40 | 3.53 | 3.32 | 3.80 |
| **Org. 2** | 4.10 | 3.90 | 3.65 | 3.70 | 3.65 |
| **Org. 3** | 4.00 | 3.60 | 4.20 | 3.60 | 4.70 |
| **Org. 4** | 4.60 | 4.70 | 4.80 | 4.00 | 4.90 |
| **Weight** | 1/5 | 1/5 | 1/5 | 1/5 | 1/5 |

Table 3.1: Multi-criteria matrix

A matrix of concordance subsystems $J^c$ should then be computed from the multi-criteria matrix (this is a square matrix which lists for each organisation ***x***, the list of objective numbers where organisation ***x*** scores better than organisation ***y***). This is illustrated in Table 3.2.

| | Org. 1 | Org. 2 | Org. 3 | Org. 4 |
|---|---|---|---|---|
| **Org. 1** | | [5] | | |
| **Org. 2** | [1,2,3,4] | | [1,2,4] | |
| **Org. 3** | [1,2,3,4,5] | [3,5] | | |
| **Org. 4** | [1,2,3,4,5] | [1,2,3,4,5] | [1,2,3,4,5] | |

Table 3.2: Matrix of concordance subsystems $J^c$

Where the generic element $J^c(\text{Organisation}_i, \text{Organisation}_j)$ of the matrix $J^c$ is given by Formula 3.2.

$$J^C(Organisation_i, Organisation_p) =$$

$$[j \in J \mid Objective_j(Organisation_i) \geq Objective_j(Organisation_p)]$$

Formula 3.2 Generic element of the matrix $J^C$

ELECTRE II differs by differentiating low preferences from high preferences (Coello Coello et al., 2007). The ELECTRE III method further differs in that it adds a fuzzy component by specifying whether an objective *a* is strictly preferred to an objective *b*, weakly preferred, or is indifferent to the objective *b*. This is illustrated as in Figure 3.6 (Buchanan and Vanderpooten, 2007).

$aPb$ ($a$ is strictly preferred to $b$) $\Leftrightarrow$ $g(a) - g(b) > p$,
$aQb$ ($a$ is weakly preferred to $b$) $\Leftrightarrow$ $q < g(a) - g(b) \leqslant p$,
$aIb$ ($a$ is indifferent to $b$; and $b$ to $a$) $\Leftrightarrow$ $|g(a) - g(b)| \leqslant q$.

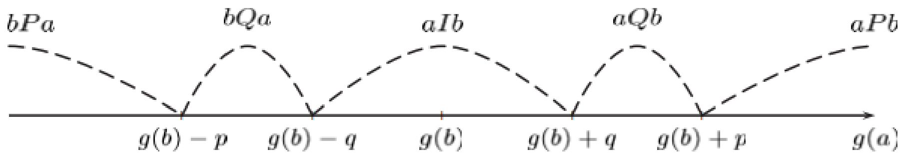


Figure 3.6: ELECTRE dual thresholds model with indifference, weak and strict preference zones

A concordance index matrix C is then calculated with Formula 3.3.

$$C(Organisation_i, Organisation_p) = \sum_{j \in J^c} K_j$$

Formula 3.3 Concordance matrix formula

Where $K_j$ is the weight attributed to objective j. This is illustrated as follows (in this example, all objectives have a same weight of 0.20) in Table 3.3.

An organisation is defined superior to another when its K value (the values in the matrix in Table 3.3) is greater than a concordance criteria *C*. *C* is generally chosen to be 0.50 although tighter concordance criteria (e.g. 0.75) will allow for a greater differentiation with traditional benchmarking (Laise, 2004).

| | Org. 1 | Org. 2 | Org. 3 | Org. 4 |
|---|---|---|---|---|
| Org. 1 | | 0.20 | | |
| Org. 2 | 0.80 | | 0.60 | |
| Org. 3 | 1 | 0.40 | | |
| Org. 4 | 1 | 1 | 1 | |

Table 3.3: ELECTRE concordance index matrix

A concordance criterion of 0.50 would give the concordance matrix illustrated in Table 3.4.

| | Org. 1 | Org. 2 | Org. 3 | Org. 4 | Count | Rank |
|---|---|---|---|---|---|---|
| Org. 1 | | | | | 0 | 4 |
| Org. 2 | ✓ | | ✓ | | 2 | 2 |
| Org. 3 | ✓ | | | | 1 | 3 |
| Org. 4 | ✓ | ✓ | ✓ | | 3 | 1 |

Table 3.4: ELECTRE concordance matrix

Which would consequently give the ranking Source illustrated in Figure 3.7 (Laise, 2004):

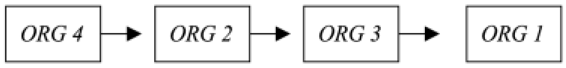


Figure 3.7: ELECTRE ranking with C = 0.50

A tighter concordance criterion of 0.75 gives the ranking illustrated in Figure 3.8 below (Laise, 2004).

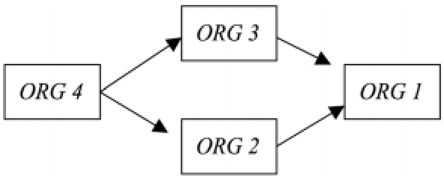


Figure 3.8: ELECTRE ranking with *C* = 0.75

AHP and ELECTRE methods are powerful tools which address the averaging limitation of traditional benchmarking through the use of outranking and pair-wise comparison techniques. Whilst AHP makes the most of matrix operations and provides a score, ELECTRE methods provide an interesting alternative with attributes such as moderation criteria for difference (fuzzy factor) and more subjective weights (Buchanan and Vanderpooten, 2007, see 5. Discussion). However, despite addressing some limitations of traditional benchmarking, these outranking methods do not address some aforementioned limitations of the mpg measure; notably the lack of method to satisfactorily include the criteria necessary to the interpretation of the mpg measure (or more generally of fuel efficiency). This problem is addressed by another class of performance measurement technique called frontier analysis.

### 3.3.3. Efficient Frontier Analysis

As Eccles (1995) mentioned, early performance measurement is used to principally focus on financial output performance disregarding other areas (such as production or customer service) or ignoring the concept of efficiency. The consequences of this omission prompted econometricians to rethink how conventional econometric analysis looked at production functions and how it dealt with variations in efficiency (Kumbhakar and Knox Lovell, 2000, p. 1).

Production functions, which model the structure of production, have been developed and refined over more than 80 years (e.g. by Cobb and Douglas (1928)). However, Kumbhakar and Knox Lovell (2000) point out that while conventional econometrics tends to use production, cost, and profit functions, they assume that producers allocate inputs and outputs efficiently and that producers operate on these functions apart from randomly distributed statistical noise. The authors state anecdotal evidence (p. 2) which suggests that producers are not always successful in solving their optimisation problems efficiently. This can be illustrated by inefficiently utilising the resources (inputs) in the production process (this is called technical inefficiency (Cooper et al., 2007)), or by poorly allocating resources and production targets (this is called mix inefficiency (Cooper et al., 2007)). Producers not solving their optimisation problem correctly were consequently not operating on the production functions used, up to then, to measure performance.

In the light of the clear limitations of traditional production functions, productivity analysis' focus moved towards production frontiers. The literature that directly influenced the development of frontier analysis methods began in the 1950's with the work of Koopmans (1951) who mentioned that a producer would be efficient 'if, and only if, it is impossible to produce more of any output without producing less of some other output or using more of some input'. Koopmans' original work prompted Debreu (1951) and Shephard (1953) cited in (Kumbhakar and Knox Lovell, 2000, p. 7) to develop models which associated the distance function with technical efficiency. This work was critical to the development of further literature on efficiency. Farrell (1957) applied for the first time these developments to measure technical efficiency in an agricultural context. This innovative work influenced the creation of two major frontier analysis techniques: Stochastic Frontier Analysis (SFA) and Data Envelopment Analysis (DEA). The concept of efficient frontier is illustrated Figure 3.9 below (Farrell, 1957, p. 258) for a one input, one output case. The best performers are on the frontier line [O, S] (the frontiers represented below are stochastic frontiers; DEA also identifies frontiers but assumes these are piecewise-linear (i.e. frontiers made of segments instead of curved lines)).

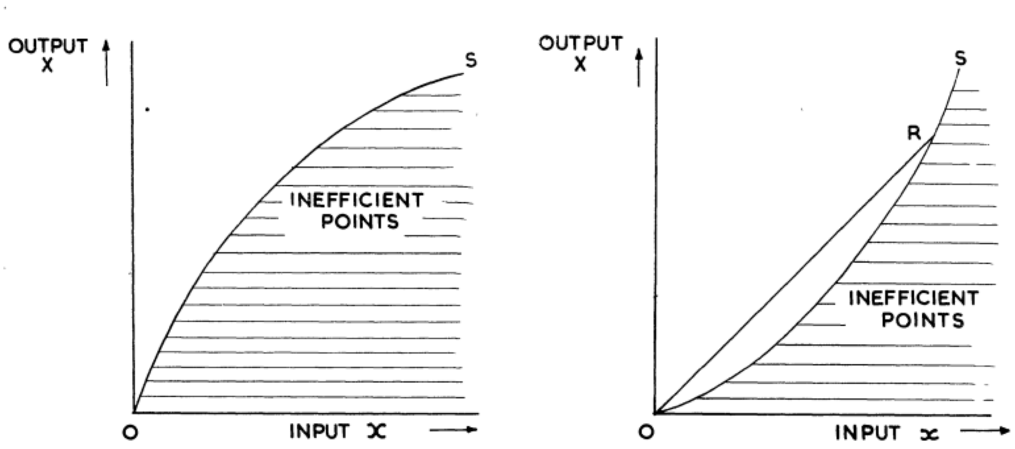


DIAGRAM 3*a*.—Diseconomies of scale. DIAGRAM 3*b*.—Economies of scale.

**Figure 3.9: Efficient frontiers**

Stochastic Frontier Analysis – a statistical approach to frontier analysis – was mainly introduced by a paper from Meeusen and van den Broeck (the 'MB' paper 1977) and another from Aigner Lovell and Schmidt (1977) (the 'ALS' paper, cited in (Kumbhakar and Knox Lovell, 2000, p. 8) and in (Coelli et al., 2005, p. 242)). These two papers introduced a Stochastic Frontier Analysis model which can be expressed as in Formula 3.4 (Kumbhakar and Knox Lovell, 2000, p. 8).

$$
\begin{aligned}
&y = f(x;\beta).\exp\{v-u\}\\
&\boldsymbol{where}:\\
&y\ is\ a\ scalar\ output,\\
&x\ is\ a\ vector\ of\ inputs,\\
&\beta\ is\ a\ vector\ of\ technology\ parameter\ and\\
&v\ is\ intended\ to\ capture\ the\ effect\ of\ statistical\ noise\\
&while\ u \geq 0\ is\ intended\ to\ capture\ the\ effect\ of\ technical\ inefficiency.
\end{aligned}
$$

**Formula 3.4 An expression of the ALS Stochastic Frontier model**

This model also implies that producers operate on (*u* = 0) or below (*u* > 0) the production frontier $[f(x;\beta).\exp\{v\}]$. Kumbhakar mentions that different papers assign different distributions (e.g. MB assigned an exponential distribution to *u* whilst ALS assigned both an exponential and half normal distribution). Parameters to be estimated are β, *v*, and *u*'s variance. Jondrow et al. (1982) described a method to estimate technical inefficiencies which 'greatly enhanced the appeal of SFA' (Kumbhakar and Knox Lovell, 2000, p. 9). Furthermore, the introduction of the *v* component allows SFA to account for statistical noise or measurement error which traditional production functions such as Cobb-Douglas do not (Coelli et al., 2005, p. 242).

Farrell's paper (1957) – which did not correctly address mix inefficiencies (Cooper et al., 2007, pp. 46-47) – prompted Charnes, Cooper and Rhode (CCR, 1978) to develop another frontier analysis method called Data Envelopment Analysis (DEA). DEA is a non-parametric benchmarking method (i.e. which does not use statistical distribution) which akin to SFA measures productivity by considering a system of inputs and outputs. The performance of entity is measured as shown in Formula 3.5 (Cooper et al., 2007, p. 21).

$$Performance = \frac{virtual\ output}{virtual\ input} = \frac{u_1 y_{1o} + \ \dots + \ u_s y_{so}}{v_1 x_{1o} + \ \dots + \ v_m x_{mo}}$$

$$where:$$

$$x\ and\ y\ are\ resp.\ the\ input\ and\ output\ vectors\ and$$

$$u_s\ output'_s s\ weight, v_m\ output'_m s\ weight.$$

**Formula 3.5: The DEA virtual performance ratio**

The fractional performance ratio above corresponds to the objective function of the model first introduced by CCR in 1978. Since then, many different DEA models have been introduced (e.g. BCC, SBM, ADD or FDH (Cooper et al., 2007)), some being drastically different from this original model. The ratio above accounts for all outputs and inputs. This type of measure is called Total Productivity Factor see (Cooper et al., 2007, p. 1) and (Grosskopf, 1993, p. 162)).

A linear mathematical optimisation process is then carried out for each entity participating in the benchmarking study. This process optimises the performance ratio illustrated above by finding an optimal set of weights whilst being constrained by all of the other entities' input and output values. Following this linear optimisation process, DEA determines the following:

- ⇨ Whether the unit is efficient, i.e. a best in class. The group of efficient units determining the efficient frontier (known as the production frontier in SFA).
- ⇨ If the unit is not efficient, how much input reduction (whilst keeping output levels constant) is necessary in order to reach efficiency (or vice versa. This is called the 'technical or radial inefficiency').
- ⇨ Whether the unit has any potential slack for all inputs or outputs and if so, it quantifies these (this is called the 'mix inefficiency'). This is a major improvement since Farrell's (1957) paper.
- ⇨ For inefficient units, the list of all the efficient units that represent the local best practices is provided (this is called the reference set).

Although the optimisation process provides information on the inputs reduction (or outputs increase) that is required to reach efficiency, some models allow for exogenously fixed factors to be unaffected during the optimisation process. This ensures that the optimisation results will not advise an impossible change for these factors (e.g. the optimisation results will not advise an input reduction if it is not possible to reduce this particular input). These variables are called non-controllable or non-discretionary variables (the latter being a variant relaxing the equality constraint).

DEA is also sensitive to the degree of freedom issues. Dyson (2001, p. 248) advises that the number of entities, called Decision Making Units (DMU), should be as shown in Formula 3.6 (Dyson et al., 2001).

$$Number\ of\ entities\ \geq 2\ \times m\ \times s$$
$$\boldsymbol{where}:$$
$$m\ is\ the\ number\ of\ inputs\ and\ s\ the\ number\ of\ outputs.$$

**Formula 3.6: Advised minimum number of DMUs in DEA (1)**

Cooper et al (2007, p. 284) suggest a different formula for the minimum number of entities. This is shown in Formula 3.7.

$$Number\ of\ entities\ \geq MAX(m\ \times s;\ 3\ \times (m+s))$$

**Formula 3.7: Advised minimum number of DMUs in DEA (2)**

The advantage of this formulation is that for a small number of inputs and outputs, $3\ \times (m+s)$ better limits the risk of degree of freedom issues.

As explained by Cullinane *et al* (2006, pp. 355-356), SFA accounts for both random shocks and measurement errors and this statistical approach has more solid ground in economic theory. However, it can be considered risky to make assumptions on the production technology by choosing a functional form (e.g. Cobb Douglas or more generally a translog - (Cullinane et al., 2006, p. 356)). The authors also warn against the difficulties in specifying the error structure ($u$) and the potential error this specification might create.

On the other hand, DEA does not make any assumption on the distribution of the error terms (Cullinane et al., 2006, p. 356) and its data oriented approach allows inferences to be drawn directly from the observed data' (Cooper and Tone, 1997, p. 72). In effect, DEA inferences are drawn from solutions optimal for each observation whilst inferences with SFA are drawn from optimisation over all observations (Cooper and Tone, 1997, p. 73). However, because DEA does not allow for measurement error, any potential existing error would inevitably be attributed to efficiencies (or inefficiencies). Furthermore, the choice of inputs and outputs can drastically affect the results. Nonetheless, this non-parametric linear approach allows appropriately cleansed data to speak for itself without relying on assumptions on the specific structure of the error term or the production technology.

Tingley *et al* (2005, p. 366) note that many papers have extensively discussed the differences in efficiency scores obtained with SFA and DEA methodologies. The authors mention that more recently interest has moved to the factors affecting efficiency rather than efficiency *per se*. Cooper and Tone (1997, p. 81) additionally state that the different specific attributes of SFA and DEA can lead to both techniques to be sometimes used in combination.

#### 3.3.3.1. DEA applied to transport operations

DEA has been applied in many different industries and has been widely used in transport although mainly to measure big structures' efficiency, mainly ports and airports. This section will discuss the application of DEA to transport operations.

Cullinane *et al* (2006) applied DEA to the measurement of container port efficiency. Their studies compared DEA and SFA in the measurement of efficiency. The authors measured container efficiency by considering 'Container Throughput' as an output and 'Terminal Length', 'Terminal Area', 'Quayside Gantry', 'Yard Gantry' and 'Straddle Carrier' as inputs and using the CCR and BCC models. They concluded that such approaches 'can be employed for the purpose of informing government ports policy [...] or management decision making'

(Cullinane et al., 2006, p. 370). Furthermore, they observed that the Spearman's rank order correlation coefficient of the technical efficiency measure by DEA and SFA indicates that these two approaches 'yield similar efficiency rankings'. Koster (2009) also applied DEA to measure container port efficiency. The authors state that one major problem in measuring container port efficiency resides in the fact that some information is considered highly confidential by the different port operators and hence not disclosed to the public domain. They also acknowledge some striking differences between sources of data in the public domain (Koster et al., 2009, p. 1145). The authors state that most publications use both CRTS and VRTS (usually CCR and BCC) and that their study will follow a similar approach. The authors conclude that DEA is a powerful tool but that it is sensitive to data errors. The authors state that relying on public data (prone to error) can lead to major errors. Interestingly Koster *et al* acknowledged efficiency differences between small and big terminals and therefore decided to include only big terminals with an annual throughput greater than 500,000 TEU (twenty-foot equivalent unit – a 20 foot long container used to measure vessel capacity and terminal throughput). To address this issue, SangHyun (2009) applied a tiered DEA approach which considered different categories based on terminal operations.

DEA studies of airports and container ports are generally quite similar as both ports and airports are large structures that move freight and in the case of airports, passengers. Yoshida and Fujimoto (2004) applied DEA to the measurement of airport efficiency. The authors used 'Runway Length', 'Terminal

Size', 'Access Cost' and 'Labour' as inputs and 'Passenger Volume', 'Cargo Handling' and 'Aircraft Movements as outputs. Both CRTS and VRTS were used. The authors compared the results of DEA with another approach that of Total Factor Productivity. The authors state that the strong correlation observed between the results demonstrate the robustness of their conclusion. In a similar fashion, Jessop (2009) uses both CRTS and VRTS approaches to measure the efficiency of some Italian airports. The author concludes that the study successfully highlighted the few inefficient airports.

DEA has also been applied to road transport.  Several studies in relation to the efficiency of road transport organisations have been published. For example, Husain et al (2000) applied DEA to multiple service units of the Malaysian Road Transport Department at Selangor using 'Estimated Labour' and 'Labour Costs' as inputs and 'Revenue' as output. DEA was also applied to measure the efficiency of companies running transport operations. For instance Kertens (1996) compared the efficiency of French urban transit companies using a Free Disposal Hull (a DEA model which does not make a piecewise-linear assumption and which requires a mix-integer algorithm to be solved). The author used 'Number of Vehicles', 'Number of Employees' and 'Fuel' as inputs and 'Vehicle Kms' and 'Seat Kms' as outputs. Finally, DEA has also been applied to benchmark the energy efficiencies of different transport modes in India (Ramanathan, 2000). The author uses 'Energy Consumption' as input and 'Passengers Km' and 'Tonnes Km' as output to measure energy efficiency. The author runs the models for different period of time to measure efficiency changes. He concludes that the efficiency of rail transport has increased over the measurement period and suggests that huge 'savings in energy consumption' could be made if rail transport is 'made to capture future requirements'.

The measurement of road transport operations' efficiency seems to have received less attention from research. The few papers that can be found appear to focus solely on bus operations (Cowie and Asenova, 1999) and truck efficiency in the construction and maintenance business (Odeck and Hjalmarsson, 1996). The application of DEA to truck efficiency was apparently solely conducted by Odeck and Hjalmarsson. This research, undertaken in the mid nineties, led to the publication of two papers (Hjalmarsson and Odeck, 1996, Odeck and Hjalmarsson, 1996). The authors measured truck productivity in the Norwegian construction and maintenance operations by applying the CCR model. They used 'Wage', 'Fuel', 'Rubber' and 'Maintenance' as inputs and 'Annual Distance Travelled' as output. The authors observed significant differences in efficiency levels amongst trucks. They were also able to correlate these scores with geographical characteristics. Odeck also applied his knowledge of DEA to the bus industry along with Alkadi (2001). Their study highlighted some significant differences between bus companies.

Following an on-going review of the literature, it seems that DEA applications to transport have been concentrated on large structures such as ports or airports and that very little research could be found on day to day transport operations.

The applicability of these different performance measurement methods (traditional benchmarking, ranking approaches and efficiency frontier methods) to this research will be discussed in section 4.1.

# 4. Methodology

The literature review showed this study's potential based on creating an improved fuel efficiency measurement based on fuel card data for van operations (chapter 2). It also listed several candidate methods which could potentially improve fuel efficiency measurement in this industry (section 3.3). This chapter will explain which method is to be retained for this study and why it has been selected. The main characteristics of the chosen method will be subsequently discussed although the technical details will be confined to several appendices. Finally, this case study's protocol will be introduced at the end of this methodology chapter.

## 4.1. Reasons for this study to use Data Envelopment Analysis

KPIs are essential to operators in day to day business operations as they can easily reflect the different aspects of performance and be used as descriptive, diagnostic and predictive measures. Whilst KPIs alone cannot indicate the possible magnitude of improvements, they can be used in benchmarking studies so as to seek this externally oriented information. However, the external information – necessary to the interpretation of the measure itself – cannot generally be easily included in the KPI measure itself. For example, vehicle weight cannot be easily incorporated in the mpg measure despite the fact that knowing the vehicle weight is essential to interpret mpg. Furthermore, both methods struggle to reflect all aspects of performance in a single measure. Instead several KPIs have to be used to reflect all aspects of performance. In this case, this means mile per gallon (mpg) and pence per mile (ppm) have to be both used to reflect these two aspects of fuel efficiency.

Although weighted averages are often used to address this issue, Laise (2004, p. 624) warns on the risk associated with using simple weighted averages to find best in class performers. Similarly, Cooper *et al* (2007) explain how problems can arise when arbitrarily choosing weights. This makes finding best in class performers using KPIs or traditional benchmarking hard and potentially impractical task.

On the other hand, literature has shown that outranking methods are better suited to ranking different entities. AHP, a method first introduced by Saaty (1980), calculates each criterion's weight through matrix calculations based on dominance values given by managers generally on a 1-9 scale although some are known to use different scales (Government of Canada, 2002). This process has the advantages of appraising each criterion's weight by translating human opinions of dominance to actual weights (a more robust process than arbitrarily choosing the weights). The process also checks on the consistency of the manager's perception of criteria dominance. Nevertheless, the ELECTRE methods do not only weight each criteria individually but works on a dominance basis instead (although weights can be used to relax the notion of strict dominance), see (Buchanan and Vanderpooten, 2007)). These two different methods address traditional benchmarking limitations in regards to finding best in class performers for multi-criteria situations. They could consequently both be used to find best performers in terms of fuel cost and fuel use (multi-criteria benchmarking). However, they do not offer a satisfactory method to include the factors that are necessary to the interpretation of the mpg measure (e.g. vehicle weight) and thus, do not answer all its limitations. Consequently, only vehicles with the same weight, type, age and operating in similar conditions could be compared without bias using these outranking methods (which in turn limits the usefulness of these methods).

In contrast, frontier methods – and as explained in the previous chapter – provide suitable mechanisms to measure performance against several different criteria (by considering them as either inputs or outputs – in our case, this is to simultaneously incorporate 'fuel used' and 'fuel cost' in a single measure). Moreover these methods offer mechanisms to incorporate the variables necessary to the interpretation of the mpg measure (e.g. weight, vehicle age...) which cannot be satisfactorily included within traditional or outranking approaches. SFA, the production frontier method which looks at efficiency from a statistical perspective, can incorporate these kind of variables as exogenous variables although the existing literature on the subject is rather slim (see Kumbhakar and Knox Lovell, 2000, p. 261). Similarly, DEA offers adequate mechanisms to take into account in the calculations exogenous or undesirable factors. The literature on the inclusion of exogenous and undesirable factors seems more extensive in DEA than in SFA. Additionally, because SFA relies on a statistical approach, the confidence in the inferences drawn from datasets in which producers are only observed once (these datasets are called single cross section) is severely limited (see Kumbhakar and Knox Lovell, 2000, p. 95 and p. 166). However, in normal operations, nothing guarantees that several observations will be available for each vehicle (in our case this mainly concerns fuel card transactions). This is especially true for fuel performance measurement as some managers are eager to see fuel efficiency figures on a weekly basis (and some vehicles might not refill many times within such a short operational period). Although DEA requires adequate and intelligent data cleansing to ensure that no measurement error is assigned to a Decision Making Unit's (DMU) efficiency or inefficiency, it performs well with single observation datasets. Additionally, DEA provides very efficient and relatively easy ways to analyse the factors affecting efficiency (Cooper et al., 2007, see meaning of optimal weights and slack analysis) – a feature which importance was highlighted by Tingley *et al* (2005).

Due to its statistical approach, SFA is less robust than DEA at measuring performance when dealing with datasets having a limited number of observations. As some vehicles are expected to only have a small number of refills for short measurement periods, DEA seems to be a more robust choice in this respect. In light of the previous theory, DEA can also be used as descriptive, diagnostic, and predictive performance measures. Effectively, DEA scores can be used to quantify observed performance (descriptive), weights and slacks are a powerful tool to understand the reasons behind performance (Cooper et al., 2007, see meaning of optimal weights and slack analysis), and extensive research was also conducted on measuring performance over time using DEA (Cooper et al., 2007, see malmquist index, p. 328). Similarly, the fuel efficiency DEA model can encompass all the relevant families of measure (productivity/resource utilisation and cost) into a single model and thus, into a single measure. Issues related to data cleansing will be discussed in section '5.3 Data Cleansing'. Using a benchmarking approach is also not an issue as advanced performance measurement methods are likely to only be of interest to fleet managers running fleets of more than 15 or 20 vehicles (this is the expected minimum number of vehicles required to avoid issues with degree of freedom).

Finally, as introduced earlier in section 3.3.3.1, DEA ‘has been widely applied in [...] the transport industry’ (Cullinane et al., 2006, p. 356) although the literature concentrates mainly on ports (Cullinane et al., 2006, SangHyun, 2009), airports (Yoshida and Fujimoto, 2004, Yu, 2004, Pestana Barros and Dieke, 2007), or other important structures rather than directly on road transport.

Only a limited number of papers have been found dealing with the use of DEA to measure road operations (Hjalmarsson and Odeck, 1996, Odeck and Hjalmarsson, 1996, Kerstens, 1996, Cowie and Asenova, 1999), and **– despite the potential interest highlighted in the aforementioned literature – none could be found on van operations or fuel efficiency measurement. This lack of research brings originality to this study**.

For all the above reasons, this study will use Data Envelopment Analysis as a means of measuring van fuel efficiency.

Table 4.1 lists the performance measures introduced so far and compares their different characteristics.

| PM \ Characteristics | Measure can be used as a Descriptive measure | Measure can be used as Diagnostic measure | Measure can be used as Predictive measure | Can appropriately include other factors in the measure | Benchmarking | Compare against best performance | Can easily draw inferences from limited observations |
|---|---|---|---|---|---|---|---|
| KPI | ✓ | ✓ | ✓ | | | | ✓ |
| Traditional benchmarking | ✓ | | | | ✓ | | ✓ |
| AHP | ✓ | ✓ | | | ✓ | | ✓ |
| ECOGRAI/ELECTRE | ✓ | | | | ✓ | | ✓ |
| Stochastic Frontier Analysis (SFA) | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | |
| Data Envelopment Analysis (DEA) | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |

Table 4.1: Performance Measures comparison table

## 4.2. Introduction to DEA

The previous section discussed different performance measurement methods and explained, between those identified, why DEA appears to best address the mpg measure's limitations. This section will take a closer look at DEA and discuss most of its core concepts through a series of small examples. A summary of all DEA's key characteristics will be made at the end of this section as these are essential to understand the 'Case Study and Results' and 'Summary of Results and Discussion' chapters. All the other technical aspects of DEA such as model types (CCR, BCC, SBM – these will be introduced in section 4.2.5), model orientation, returns to scale, non-discretionary, non-controllable and undesirable (anti-isotonic) variables are discussed in the appendices (sections 8.2 to 8.4). Although such information is necessary to comprehend the remainder of the thesis, the current section is intended to provide enough information to understand the next sections.

### 4.2.1. Performance Ratio

As seen earlier in the section 3.3 'Performance Measurement Methods', efficiency is commonly measured through the mean of a performance ratio which takes the form illustrated in Formula 4.1 (Cooper et al., 2007):

$$Efficiency = \frac{Output}{Input}$$

**Formula 4.1: Efficiency ratio**

For example, a common efficiency ratio in road transport operations is miles per gallon (the number of miles is the output while gallon the input). More generally, efficiency ratios can also be used to reflect productivity such as with the number of jobs per day/vehicle (where the number of jobs is the output and day/vehicle the input). These measures are called '*partial productivity measures*' in an effort to differentiate them from '*total productivity measures*' (Hayes et al., 1988); the latter attempting to take into account all outputs and all inputs under the same efficiency ratio (Cooper et al., 2007, p.1). A total productivity efficiency ratio can be illustrated as in Formula 4.2.

$$Total\ Productivity\ Ratio = \frac{\sum_{j=1}^{s} output_j \times weight_j}{\sum_{i=1}^{m} input_i \times weight_i}$$

$$where\ s\ is\ the\ number\ of\ outputs\ and\ m\ the\ number\ of\ inputs,$$
$$weight_j\ is\ the\ weight\ of\ output_j\ and$$
$$weight_i\ is\ the\ weight\ of\ input_i.$$

Formula 4.2: Total factor productivity ratio

The choice of weights in DEA is not arbitrary but is rather the result of an optimisation process completed for each entity. One interesting feature of total productivity measures is that they reduce the risk the chances of attributing gains to one factor which are in fact caused by another factor (or other factors). For instance, if a supermarket's sales increase following an advertising campaign, the ratio 'sales / labour' would also be likely to improve. However labour's performance could have potentially decreased during that same period and this could go unnoticed (or worse, the sales increase could be attributed to labour). The total productivity approach used by DEA avoids this problem by directly including all parameters under the same ratio and simultaneously measure the impact of all factors.

### 4.2.2. Single Input, Single Output

The performance ratio concept can be easily illustrated with a single input / single output example. The exercise is to measure the efficiency of 8 depots based on their sales performance and their number of employees (supposing the sales unit is £100,000.00, and the employee unit 1,000 employees). The data are illustrated in Table 4.2 (Cooper et al., 2007, p. 3).

| Depot | A | B | C | D | E | F | G | H |
|---|---|---|---|---|---|---|---|---|
| Employee | 2 | 3 | 3 | 4 | 5 | 5 | 6 | 8 |
| Sale | 1 | 3 | 2 | 3 | 4 | 2 | 3 | 5 |
| Sale/Employee | 0.5 | 1 | 0.667 | 0.75 | 0.8 | 0.4 | 0.5 | 0.625 |

Table 4.2: Single Input, Single Output example data

The last row is calculated using the efficiency ratio formula introduced in Formula 4.1 above. This formula reflects the productivity of each depot but can also be used to treat more generic cases of efficiency. This ratio helps identifying store B (with an efficiency ratio of 1) as the best efficiency score, and F the worst depot (with an efficiency ratio of 0.4).

The depots' performance can be plotted on a graph with the number of employees on the 'x' axis and the sales on the 'y' axis (Cooper et al., 2007, p. 4). This is illustrated in Figure 4.1.

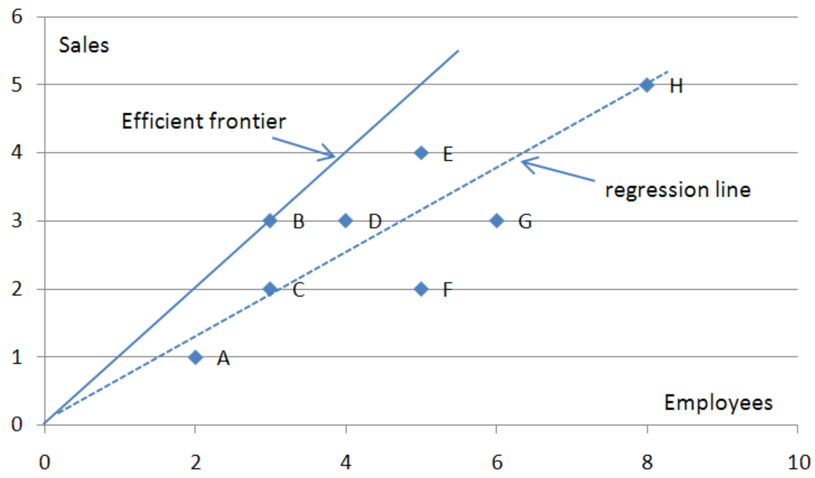


Figure 4.1: Graph comparison of depot's efficiency

The slope of the line starting from the origin O and passing through B corresponds to B's sales per employee ratio. As B is an efficient depot (demonstrating the best sales per employee ratio), this line is called the efficient frontier. This frontier touches at least one point (depot) and all the other depots will therefore be on (should their sales/employee ratio also be equal – in this case – to 1) or under this frontier (should their sales/employee ratio be lower than 1). As mentioned earlier DEA compares against the efficient frontier and this frontier is said in mathematics to 'envelop' the data – hence DEA's name.

Statistical approaches such as regression analysis provide an estimate function which describes the relation between the different variables (if it exists – Curwin and Slater, 2002, p. 390). DEA differs from these statistical approaches as it defines the efficient frontier from the observed data and compares all different entities against this efficient frontier. It is not however always realistic to believe this frontier stretches to infinity and this assumption will be relaxed later. For the moment however, the frontier is assumed to be constant for the range of operations; this is called *constant returns to scale*.

As just mentioned, DEA evaluates entities' performance against the best observed performance. In this particular example, it is possible to rewrite each depot's performance in respect to B's performance as in Formula 4.3 (Cooper et al., 2007, p. 4).

$$0 \leq \frac{Sales\ per\ employee\ of\ others}{Sales\ per\ employee\ of\ B} \leq 1$$

**Formula 4.3: Relative efficiency ratio**

Following this formula, each depot's efficiency would consequently be as in Table 4.3: Relative efficiency (Cooper et al., 2007, p. 5). Because B's 'sales per employee' ratio is equal to 1, the depots' efficiency is equal to the row Sales / Employee in

‘Table 4.2: Single Input, Single Output example data’ although this is just a coincidence.

| Depot | A | B | C | D | E | F | G | H |
|---|---|---|---|---|---|---|---|---|
| Efficiency | 0.5 | 1 | 0.667 | 0.75 | 0.8 | 0.4 | 0.5 | 0.625 |

**Table 4.3: Relative efficiency**

Let’s suppose Depot B still be the unique efficient Depot with a ‘sales per employee’ ratio of 1.5, the efficiency results illustrated in Table 4.3 would have changed in comparison with Table 4.2: Single Input, Single Output example data (e.g. if its sales output was 9 instead of 6 in this previous table). The ratio of ratios illustrated in Formula 4.3 makes the new relative efficiency measure independent from the unit used. This property, called unit invariance, is an essential characteristic in performance measurement as it ensures the efficiency measured does not vary depending on the unit choice.

The inefficient units, i.e. all the units strictly below the efficient frontier, would need to reach the efficient frontier in order to become efficient. This can be done in mainly two different ways:

- ⇨ Either through increasing the outputs whilst keeping the inputs constant, or
- ⇨ Through reducing the inputs whilst keeping the output levels constant, or
- ⇨ Simultaneously reducing inputs and increasing outputs (ADD and SBM models).

The improvement process can be illustrated as in Figure 4.2 for Depot A (Cooper et al., 2007, p. 5).

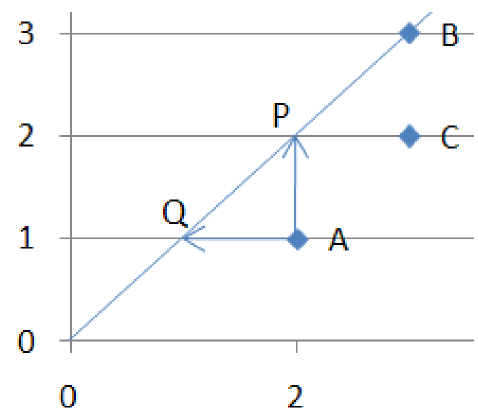


Figure 4.2: Improvement of Depot A

Depot A can become efficient by either reaching P (keeping input levels constant while increasing output levels), or by reaching Q (i.e. keeping output levels constant but reducing input levels). This two-way approach to improve efficiency is relaxed by the Slack Based Models the latter being described in the appendix 8.3.2 SBM Model.

Interestingly, multiplying A's employees input by its efficiency score matches Q's X coordinates (2 * 0.5 = 1). Similarly, multiplying A's sales output by the inverse of the efficiency score would in a similar fashion give P's coordinates (1 * (1 / 0.5) = 1 * 2 = 2).

### 4.2.3. Single Input – Two Output Case

A similar approach can be taken to measure a depot's efficiency in a one input, two outputs scenario. Here, depot's efficiency is measured by observing for each depot, the relation between the number of vehicles (input) and the volume of sales (in £10,000s) and utilisation (in units of 15% of total capacity) as outputs. In order to have a unitised frontier, all the variables are normalised by the number of vehicles (e.g. both inputs and outputs are divided by the number of vehicles). Data are as in Table 4.4.

| Depot | A | B | C | D | E | F | G |
|---|---|---|---|---|---|---|---|
| Vehicles | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| Sales / Vehicles | 1 | 2 | 4 | 2.5 | 4 | 5 | 5.5 |
| Utilisation / Vehicles | 4 | 5 | 1.5 | 3 | 4 | 3 | 1 |

Table 4.4: One input – two outputs example

These data can be illustrated graphically as in Figure 4.3.

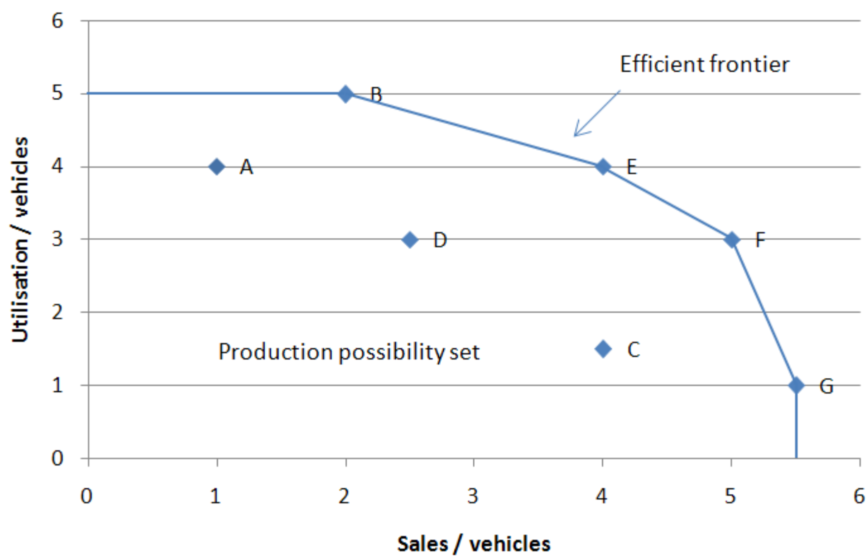


Figure 4.3: One input – two outputs example

In this particular scenario, it is logical that the efficient frontier is defined by the performers demonstrating maximum output levels in terms of both utilisation and sales. Thus the frontier is defined by the depots B, E, F and G. Entities in DEA are generally called a Decision Making Unit (DMU) as each can make decisions which can affect its efficiency. The depots and other entities will from now be referred to as DMUs.

The region bounded by the frontier is called the production possibility set; i.e. the region of possible production levels as defined by the best observed performance. The production possibility set should more accurately be called the piecewise 'linear production possibility set assumption' (Cooper et al., 2007, p. 7) as it is not sure whether the actual production frontier behaves in a linear manner between the performance points observed (i.e. in this case between the points B, E, F, G). The frontier is projected vertically and horizontally on the border (at point G and B). Note that there is no discrimination on how a DMU can be efficient. This is easily demonstrated by considering DMU B's operations (high utilisation but low sales) which are radically different from DMU G's operations (very low utilisation rate but high volume of sales). In both cases, the DMUs are evaluated efficient.

As explained earlier, inefficient DMUs (i.e. the depots not on the efficient frontier: A, D and C) can become efficient by reaching the efficient frontier. Thus DMU's efficiency can be calculated by their relative 'distance' to the frontier. For instance, C's efficiency can be calculated with the 'radial measure' as in Formula 4.4.

$$Efficiency_C = \frac{d(O,C)}{d(O,Q)} = 0.76$$

$$where\ d(O,X)\ is\ the\ distance\ between\ the\ origin\ and\ point\ X.$$

Formula 4.4: Efficiency in a single input – two output case

This is illustrated graphically in Figure 4.4.

Because DMUs in this particular example aim at maximising their outputs, the precedent ratio could be rather re-written as in Formula 4.5.

$$Efficiency_C = \frac{d(O,Q)}{d(O,C)} = \frac{1}{0.76} = 1.31$$

Formula 4.5: Output efficiency in a single input – two output case

Formula 4.5 shows that DMU C would have to proportionally increase its sales and utilisation by 31% to reach the efficient frontier and become efficient. In this case DMU C would be at coordinates of Q. Note that because this measure of efficiency is unit invariant, the results would be the same if the utilisation or sales units chosen were different.

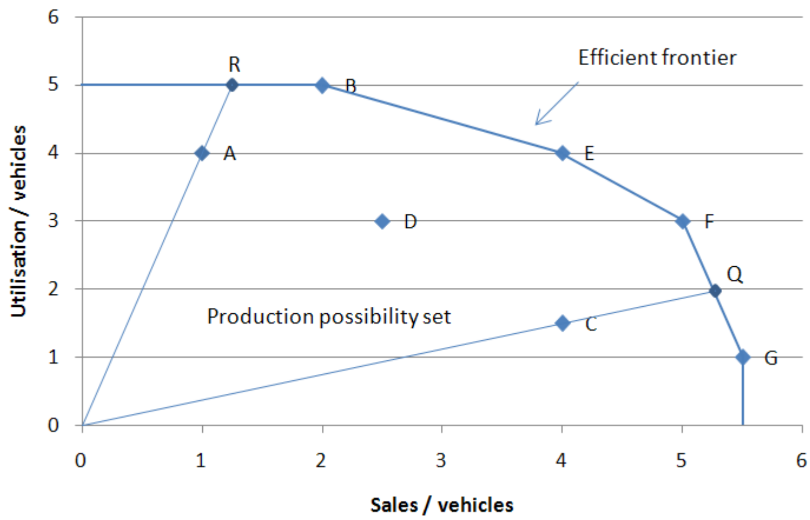


Figure 4.4: Graph of improvement in a single input – two output case

Point Q – DMU C's projection on the efficient frontier – is on the line connecting points F and G. The set made of DMU F and G is DMU C's reference set. The reference set of an inefficient DMU consists of the efficient DMUs which were significant in the evaluation of the DMU's efficiency.

The radial inefficiencies that can be addressed by proportionally increasing all outputs (and decreasing all inputs) are called technical inefficiencies. Technical inefficiencies can consequently be removed without changing outputs (or inputs) proportions. Conversely, another type of inefficiency also exists where proportions in which inputs are used (or output produced) have to be changed in order to attain efficiency. These are called mix inefficiencies (Cooper et al., 2007) and can be illustrated by considering DMU A's efficiency.

In order to eliminate its technical inefficiency, DMU A needs first to reach point R on the efficient frontier (see Figure 4.4: Graph of improvement in a single input – two output case). However, DMU B (2, 5) demonstrates a greater Sales output than point R (1.25, 5) (R's coordinates are calculated by solving intersection of $y = 4x$ $and$ $y = 5$). Even if A eliminates all its technical inefficiencies by reaching R, another DMU demonstrates a better output mix (DMU B). Thus in order for A to become fully efficient (and not just technically efficient), it needs to change its output mix and increase its Sales by 0.75 in order to move from R and reach B thus, becoming fully efficient.

#### 4.2.4. A brief introduction on DEA computational process

Although DEA's key concepts have been introduced in sections 4.2.1 to 4.2.3, little has been said about DEA's process itself. DEA works by adapting the fractional ratio introduced earlier (the performance ratio outputs/inputs) and transforming it into a linear equivalent. This transformation is important as it enables the use of linear optimisation methods.

An optimisation process is carried out for each DMU in order to calculate the DMU's efficiency. This optimisation process aims at maximising the value of a performance ratio as introduced earlier: $\frac{\sum_{j=1}^{s} output_j \times weight_j}{\sum_{i=1}^{m} input_i \times weight_i}$ by finding the optimal weights for the DMU. The weights are constrained by the existing data so that when used with other DMUs' values, they do not provide results above the maximum performance levels observed (this is how the data are enveloped by the efficient frontier).

The optimisation result provides amongst other things: the DMU's score, and for inefficient DMUs quantification of the potential slacks and indications on how to improve performance with the weights and the reference set.

### 4.2.5. A brief introduction to DEA's main models

Many different DEA models exist which all have their own specific characteristics. It is generally useful to use different models to better understand the nature of the data and of the frontier (i.e. whether variable or constant returns to scale apply). In effect, it is often crucial to test the performance against different models to understand the relation between the performance measured and the model choice.

This study will use the three following models:

- ⇨ ***Charnes, Cooper & Rhode - Input oriented*** model (CCR-I) (Charnes et al., 1978). This model assumes constant returns to scale (as in Figure 4.1: Graph comparison of depot's efficiency). Constant returns to scale means that any DMU is supposed capable of reaching the best efficiency ratio measured in the data set, regardless of scale effect.
- ⇨ ***Banker, Charnes & Cooper – Input oriented*** model (BCC-I) (Banker et al., 1984). This model assumes full returns to scale (i.e. the efficient frontier assumes a convex shape around the data as in Figure 4.5: below). In opposition to the CCR model, scale effects are considered in this model. This model calculates a radial measure of efficiency.
- ⇨ ***Slack Based Model Constant RTS Input oriented*** model (SBM-CI) (Tone, 2001). This model assumes constant returns to scales as with the CCR

model (a convexity constraint can be added as with the BCC model). However, instead of calculating efficiency through a radial measure as with the CCR or BCC model, the SBM model tries to maximise slacks for all concerned variables. This means that both technical and scale inefficiencies are directly and simultaneously taken into account in the objective function, thus giving new insights on the reasons behind performance. This is illustrated with DMU F in Figure 4.5.

A comparison between the frontier under VRTS and CRTS can be found in Figure 4.5.

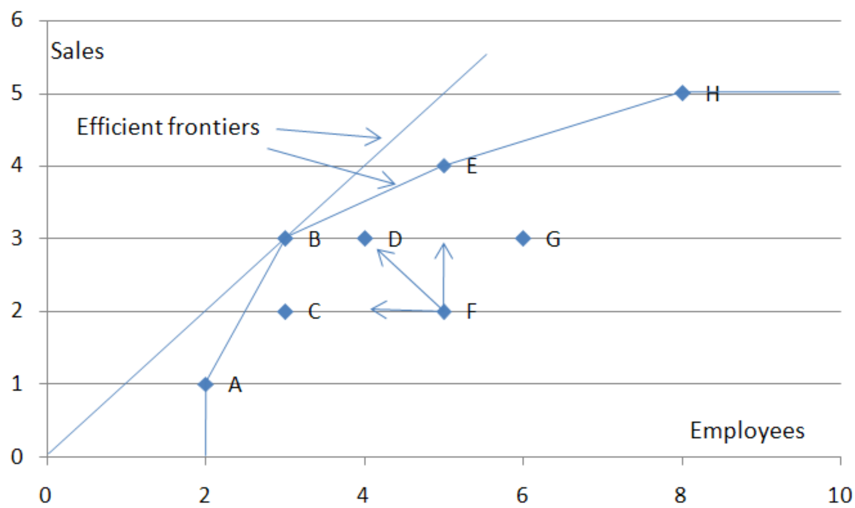


Figure 4.5: Illustrating different DEA models

These three models have been retained as they help in identifying the nature of the efficient frontier and - through some efficiency ratios – also help appraising whether inefficiencies are due to scale or to inefficient operations. In Figure 4.5 the frontier assuming Constant Return to Scale (CRTS) is the straight line passing through the graph origin and by B. Conversely, the frontier assuming Variable Return to Scale (VRTS) is illustrated by the piecewise linear frontier joining A, B, E and H. Both the CCR and BCC models project horizontally or vertically on the frontier while the SBM model – minimising all slacks – allows projections in any direction. These concepts are further detailed in Appendix 3: Other DEA models.

### 4.2.6. Summary of Introduction to DEA

This brief section explained how DEA uses a *total factor productivity* approach to measure performance and illustrates this principle with a series of examples. Further non-ratio DEA models (known as models in their envelopment form) will be introduced in the next sections although their relation to this ratio form (known as multiplier models) will be explained (see section 8.2.1 'Transforming the Fractional Problem').

The notion of *efficient frontier,* which is determined by the observed best performance, was reviewed using a simple example. Differences between DEA and traditional statistical 'trend' approaches were discussed. The notion of *production possibility set* (the area bounded by the efficient frontier) was also mentioned along with the notion of *reference set* (the list of efficient DMUs against which an inefficient DMU is evaluated). Finally, DMU B with DMU G demonstrated in the single input two outputs (depot operation) example how some very *different DMUs* can be *efficient* nonetheless.

The concept of *relative efficiency*, i.e. the new ratio constructed by dividing the performance of an entity by a best entity's performance was studied. This brought to light the important notion of *unit invariance* which ensures results will not be dependent on the units chosen.

The DMU's performance in the examples was measured under *constant returns to scale*, i.e. it was assumed that production levels would be constant regardless of the DMU size. This assumption, although possibly true on small scales, is however sometimes inadequate on larger scales. This will be relaxed later on with a discussion on variable returns to scale with the use of the BCC model to calculate fuel efficiency (see section 5.4.2.2 'Taking a closer look at Variable Returns To Scale').

Finally, the concepts of *technical inefficiency* (the inefficiency which can be eliminated by proportionally increasing all outputs – or decreasing all inputs) and of *mix inefficiencies* (inefficiency which can only be removed by changing the mix proportion of inputs or outputs) was introduced. It was then explained that a DMU can only be efficient if it has no technical and mix inefficiencies (some DEA models do not clearly distinguish these two types of inefficiencies but this will be discussed later; see 8.3.2 SBM Model).

As mentioned earlier, the concepts introduced in this section should be sufficient to understand most of the Case Study and Discussion chapters. However, some aspects of the case study and the discussion will go beyond the basic knowledge introduced here. Consequently, further information on some of DEA's technical aspects will be found in the appendices (resp. 'Appendix 2: From Econometrics to the Charnes Cooper and Rhodes model' and 'Appendix 3: Other DEA models'. These sections will discuss the different key aspects of the model).

## 4.3. Case Study Protocol

In order to evaluate whether DEA can improve fuel efficiency measurement, a series of case studies will be conducted within different companies running van operations. It is advantageous to use a case study approach for this particular research, as a single successful case study would be enough to demonstrate DEA's potential in measuring fuel efficiency.

**Population selection**

In order to conduct this case study, it is necessary to recruit some participants. Participants must have an interest in participating to the study as they will have to share data as well as comment on the model results. Obtaining this professional opinion on the models results is crucial in order to validate the study's findings.

Participants should all use vans as a necessary part of their daily business lives (they can also use other types of vehicles although these will not be included in the study). Potential suitable business operations could for example be: gas engineers, electricians, network or TV providers (e.g. SKY, BT...) or small building repair operations. It is important to observe that although each company use vans of potential different sizes, their operations are also likely to be different. Consequently, each of their operational environments will slightly differ from one another and the benchmarking exercises should only be conducted internally for each company (this will ensure the impact of environmental factors is kept to a minimum). A cross-company benchmarking should be conducted afterwards in order to appraise the importance and sensitivity of the environmental factors. This will be discussed in section 5.5.2 Multi-companies benchmark.

**Model description**

This step involves analysing the different components which should be included in a more comprehensive fuel efficiency measure. This should include both factors which impede the reading of the measure (e.g. vehicle type or weight) and other factors which use resources and have a potential impact on the efficiency measure itself (servicing). Other factors which add another dimension to fuel efficiency such as fuel cost should also be used in the model.

Both the resulting list and the feasibility of each variable will be assessed in section 5.2.2.

**Model conception**

Once the list of parameters to be included in the model has been decided, it will be necessary to conceptualise the DEA model. All variables need first to be organised in a series of inputs and outputs. The variable type needs also to be specified so that it is understood whether a unit is treated normally or as a discretionary, non-controllable, or anti-isotonic/undesirable variable.

The previous step 'Model description' has listed all the variables which are relevant to fuel efficiency; i.e.: the factors which are necessary to the interpretation of the measure, other factors which have an impact on fuel efficiency or others which add an extra dimension to fuel efficiency. These variables were split into inputs and outputs by considering how they are related to the production process. In this instance, both volume used and cost were obvious inputs of the production process while miles travelled related more to the outputs. Despite not being obvious inputs, vehicle weight and age nonetheless participate to the production process as anti-isotonic inputs thus are to be incorporated to the model as such.

One of DEA's main criticisms is that similar models can provide very different results as soon as one variable is different; the literature shows that researchers tackled similar problems with models which had different inputs and outputs. There is no very clear rule that can address this problem except that all variables which have an impact on the production process should be included in the model and that the model should be built in a step by step approach to ensure the impact of each variable is clearly appraised and understood. This is the approach that this study takes. Section 5.2.2 Fuel Efficiency Model General Considerations will review in more details the potential variables to include in the model while Table 5.4: Factors to include in the fuel efficiency model will list all variables and summarise whether they are to be included in the model or not.

**Data collection**

Adequate and appropriate data will have to be collected. This can ideally be done via Microsoft Excel spreadsheet as a support for data collection. As the vans correspond to the DMUs in this study, data will have to be collected for each van.

**Data cleansing**

As highlighted earlier in the literature review, DEA is highly sensitive to measurement error and data noise. Because the study is based on fuel card data which has a relatively high rate of error (mainly because part of this data are generally keyed in manually at the petrol station), the data will have to be thoroughly cleansed prior to being used.

Errors to be cleansed include:

- ⇨ Registration misspellings,
- ⇨ Missing / additional fuel transaction (from incorrectly cleansed registration misspelling),
- ⇨ Fuel jerry cans (to be used for plant engine for example) filled with the same transaction as a vehicle refill,
- ⇨ Fuel theft.

The error type which will have the most impact on the model results is missed transactions (caused by misspelt registration which could not be allocated to any vehicle or was assigned to an incorrect vehicle). These could cause a vehicle's fuel performance to be artificially inflated and, should that vehicle be evaluated efficient, artificially move the production frontier further away from the real production area. This could consequently potentially cause many DMUs' performance to be under-evaluated.

It is important to observe that cleansing data from companies which have a telematics solution is more effective as the telematics data help the cleansing process with cross-comparisons.

**Data smoothing**

The literature review chapters pointed out that fuel data were sometimes misused in fuel trials; some operators wrongly assumed that the fuel drawn at the pump during measurement period was the fuel used. Although this assumption is acceptable for long periods and high mileage, the accuracy of this assumption is often unacceptable for shorter periods. A possible alternative is to have all vehicles refuelling at the exact beginning and exact end of a fuel trial; which is often unrealistic or impossible and could not be easily done for this specific study. More practically, a small data smoothing algorithm will be designed so that the fuel used is estimated based on actual observed performance. This approach is deemed more consistent and reliable that relying on an inaccurate assumption.

**Model development process**

A step by step approach should be taken for developing the models. A very basic one input – one output model (fuel ⇒ miles) should be developed first. The fuel efficiency calculated with this DEA model should then be compared against the corresponding mpg measure. Different types of DEA models should also be tested (e.g. CCR, BCC and SBM) and the differences between the models appraised. This should help appraise whether the performance levels observed are due to the variable (data) itself or rather to a specific model. Each variable should then be added to the model, one variable at a time.

This step by step approach should allow for variables to be added safely to the fuel efficiency model. This method also allows a safe evaluation of the impact each variable has on the performance score.

**Analysis**

The model results should be verified against other available DEA solvers (when possible). The results should also be validated by the participants (van operation experts). Comparison against traditional methods should also be appropriately made.

Analysis on some of DEA technical aspects should also be carried out. This could include aspects such as data sensitivity or returns to scale. This aspect of the analysis is essential as it will provide an indication on the strengths and weaknesses of DEA used in this specific context of van operation fuel efficiency measurement.

Finally, the same models should also be run across all companies in order to assess whether environmental factors have a non-negligible impact on the measurement. If environmental factors are predominant on the measurement, each company's frontier should envelop each other in a more or less neat fashion (this is similar to Russian dolls nested within each other; just extended to an *n* dimensions space concept).

# 5. Case Study and Results

The Introduction chapter explained this study aimed at improving van fuel efficiency measurement through the use modern performance measurement methods. The Methodology chapter which followed justified – based on the information presented in the literature review – that the most appropriate method for such an investigation was DEA. It is important to observe that this case study will solely focus on the actual measurement of van fuel efficiency and not the actions taken based on this measurement. Thus, the study focuses on performance measurement and not on performance management.

The Case Study Theoretical Background section will first consider the best way to approach the application of DEA to van fuel efficiency measurement and justify why the case study approach was retained. Some background case study theory will also be discussed.

This will be followed by a detailed description of the research process in terms of population selection and model details. Once the case study details are explained, this Case Study and Results chapter will concentrate exclusively on more technical details such as data collection & cleansing to finish with the results themselves and their analysis. This follows the protocol described in the Methodology chapter.

The entire case study – which is a technical chapter – will be briefly summarised in Chapter 6 ‘Summary of Results and Discussion’ before the results are discussed. Although reading the case study should give an in depth understanding of this study, Chapter 6 should be understandable without reading this current chapter ‘Case Study and Results’.

## 5.1. Case Study Theoretical Background

Many different classifications of the different research types exist and the boundaries between each type can sometimes be a little fuzzy. Kontio (2005) distinguishes the following three however:

- ⇨ **Exploratory research** which structures and identifies new issues and problems.
- ⇨ **Constructive research** which identifies and develops methods to solve issues or problems.
- ⇨ **Empirical research** which tests a solution’s feasibility using empirical data.

This particular research is exploratory due to the way in which the literature review investigated the current state of research and identified gaps within it. However, due to the experimental aspects this research demonstrates (i.e. to test the feasibility of measuring van fuel efficiency using DEA), the study can also be classified as a quantitative empirical research. This research is consequently both exploratory and empirical.

In regards to applying DEA to van fuel efficiency measurement, the following research question has been formulated:

- ⇨ How can DEA be applied to van fuel efficiency measurement?

This prompts a series of other more specific research questions:

- ⇨ What are the factors affecting fuel efficiency?
- ⇨ What is each factor's exact effect on fuel efficiency?
- ⇨ How easy is applying DEA to van fuel efficiency measurement?
- ⇨ How useful is applying DEA to van fuel efficiency?

Yin (1994, p. 1) explains that case studies are only one of several ways to do research. Possible alternatives can be: experiments, surveys, the analysis of past research (history or archival analysis); each of these options having their own specific pros and cons. Yin explains that the decisions of which approach to retain depends on three conditions: 'the type of research question, the control an investigator [or researcher] has over actual behavioural events, and the focus on contemporary as opposed to historical phenomena'. As Yin comes from a more theoretical and social sciences background, only the first criterion (the type of research question) is relevant to this discussion.

The application of DEA to fuel efficiency measurement is believed to be entirely new, it is consequently not possible to look at past research (although there is extensive literature on DEA's application to transport; see section 4.1). Furthermore, experts' opinion gathered from methods such as the Delphi technique or conventional surveys can only probe people's opinion on this specialist subject which would unfortunately not really answer the different research questions listed above (in this case experts would be the people using the measure (i.e. fleet managers) and academics specialised in the transport industry). In effect, the only way to answer the research questions is through an experiment. Yin (1994, p. 15) describes case studies as ‘a way of investigating an empirical topic by following a set of pre-specified procedures’. Although Yin (1994) explains that experimental case studies are generally better suited to answer the ‘what’ and ‘why’ research questions, it is clear that in order to conduct this quantitative empirical research, conducting experimental case studies is the most appropriate method.

Yin mentions that most texts about case study methodology tend to focus chiefly on data collection. He argues that the design and analysis steps are as important as the data collection step despite being often neglected. This section will consequently briefly discuss all of these important steps.

Yin lists five components of importance for case study research design. These are:

> The **study’s questions**. These are questions generally written in the form of ‘who’, ‘what’, ‘where’, ‘how’ and ‘why’ questions. Writing these questions helps deciding which research method should be used. This study’s questions have already been listed above.
>
> The **study’s proposition**. This is essential as it helps the researcher understanding what needs to be researched and answered. The proposition helps the researcher to move in the right direction and to look at the right place to find evidences. Yin notes that some studies do not have a research proposition. This can be the case for some experiments or surveys. The study’s proposition corresponds to the hypothesis which was introduced at the very beginning of this thesis (see the Hypothesis section).

The **unit of analysis**. This relates to ‘what the case is’ (or cases are); in many social sciences studies the unit of analysis is an individual. In this particular study however, fuel efficiency performance is measured for each van. Yet, as DEA is an efficient frontier benchmarking technique, fuel efficiency can only be calculated for a group of vans or more precisely companies’ van fleets. Furthermore, and although analysis can be made individually for each van, fleet operators tend to consider fuel trials as a fleet wide exercise. Consequently, and although individual van performance analysis will be conducted for some vans, this study’s real units of analysis is a whole van fleet.

**Linking data to proposition**. This step needs to be done in order to connect the data, or data results, to the hypothesis. There is no clearly defined method to link data to the research proposition although the thorough observation of the DEA results along with traditional mpg benchmarking analysis should provide a robust link to the proposition.

**Criterion for interpreting the study’s finding**. These criterion are essential to test the results’ validity and analyse the results. This study will use fleet managers’ opinion on the DEA and traditional mpg benchmarking results to evaluate the validity and usefulness of this study’s approach against those of others.

Criteria to interpret study’s findings include:

- ⇨ The measure is coherent with fuel efficiency operator’s understanding.
- ⇨ The measure can be easily understood.

- ⇨ The measure includes factors impacting fuel efficiency other than miles travelled and fuel used which is all the information mpg captures – and is an essential point in justifying an improvement on the mpg measure.
- ⇨ The measure can help fleet operators to make better informed decisions, which could in turn lead to better fuel efficiency (this point is also essential in justifying an improvement on the mpg measure).
- ⇨ The measure's calculations are reproducible (this refers to the method reliability).

When it comes to testing the research design, four tests have been commonly used (Ellinger et al., 2005, p. 12). Yin (1994, p. 32) summarises them as in Table 5.1 (Yin, 1994, p. 33).

The construct validity will be tested by gathering data from several different companies (this study will take a multi-case studies approach). Internal validity is only relevant to explanatory and causal studies so it does not apply to this study's exploratory approach. External validity is obtained through the replication of the protocol in each different company. Finally, the study's reliability can be tested through the evaluation of the protocol's robustness and accuracy. This protocol will be further detailed in the following section 'Research Process'.

| Test | Case Study Tactic | Phase of research in which tactic occurs |
|---|---|---|
| Construct Validity | Use multiple sources of evidence | Data collection |
| | Establish chain of evidence | Data collection |
| | Have key information review draft case study report | Composition |
| Internal Validity | Do pattern-matching | Data analysis |
| | Do explanation-building | Data analysis |
| | Do time series analysis | Data analysis |
| External Validity | Use replication logic in multiple-case studies | Research design |
| Reliability | Use case study protocol | Data collection |
| | Develop case study data base | Data collection |

Table 5.1: Case study design tests

The data to be collected consist of fuel and vehicle information. The information is purely quantitative and will have to be provided by the participating companies. It is essential to conceptualise the DEA model first as this will tell what data need to be gathered. Because the data are of a quantitative nature, the data collection steps do not demonstrate the traditional caveats of qualitative data analysis in social sciences. This is further reinforced as all subjective data (e.g. data depending on driver's memories such as tyre pressure checks) are discarded from this study. Nonetheless, data cleansing can prove quite challenging and will be extensively discussed in the following 'Data Cleansing' section. A data cleansing protocol will also be detailed in this next section.

As introduced in the Case Study Protocol section, this study's analysis should be done by comparing individual van's DEA performance score with their corresponding mpg and against fleet managers' perception of the measure. Similarly, the ranking provided by the DEA models should be compared with a corresponding mpg benchmarking analysis. This theoretical triangulation (Bryman, 2001) should hopefully assist in appraising the differences between DEA and traditional mpg analysis results.

## 5.2. Research Process

This section will detail all the case study research process steps. This includes population selection, data collection, data cleansing, model development and the following analysis.

### 5.2.1. Population Selection

To ensure some of the case study theoretical limitations are correctly addressed, a multi-case study needs to be conducted (Yin, 1994, p. 33). As explained in the previous section, data from several companies is to be collected. This is necessary for the case study approach to be valid. This section will discuss the requirements in terms of participating companies.

Three companies provided their data. They are:

- ⇨ **FSH Maintenance** specialised in building maintenance.
- ⇨ **Carillion** which is a corporate group working in many different sectors. This specific depot is specialised in the property maintenance operations.
- ⇨ **Avonline**, one of the leading resources and solutions providers to the UK telecoms, media and technology.

Each of these companies runs vans to conduct their daily operations. Typical operations consist of one to several jobs to carry out during the day and generally at different sites. Each company runs vans of different size. All vans weigh between 1,500 kg and 3,500 kg gross vehicle weight. Each company generally operates at regional level and their operations always stay within UK borders. The scope limitation to operation within the UK should not impede one from drawing inferences on this study's applicability to other countries. Table 5.2 summarises the fleet size and area of operations of each company.

| Company Name | Fleet Size (only vans) | Area of operation |
|---|---|---|
| **Avonline** | 287 | Bristol |
| **Carillion** | 142 | Manchester |
| **FSH Maintenance** | 69 | West Yorkshire |

**Table 5.2: Participating companies details table**

The consequent number of vehicles involved helps ensuring the study is externally valid (see (Yin, 1994) and Table 5.1: Case study design tests) but also reduces the risk of non-sampling error as mentioned earlier in the Data Gathering section.

Although these companies' operations are similar in terms of business model (all engineering services jobs), there are some non-negligible operational differences in terms of equipment weight and area of operations (hill, wind...). This is not of concern to this study as a separate case study will be conducted for each of these companies. A simultaneous measurement of all three companies vehicles' performance will nonetheless be conducted in order to appraise whether these environmental factors are significant or not.

All vans are tracked by the same telematics company Masternaut Three X hardware. Telematics information can help cleansing the data and can provide consistent mileage information. Restricting the study to companies which are equipped with telematics devices can theoretically create a small bias although this will be considered appropriately in the Summary of Results and Discussion chapter.

All the companies aforementioned do not retrieve CANbus information and use fuel card data to calculate their fuel efficiency. The selected companies use a variety of fuel card types (e.g. driver card or vehicle card; see section 2.2.3 Fuel Card Management above for more explanations on this). The limitations of mpg measurement based on fuel card data have been addressed in 'Literature Review – Van Fuel Efficiency Measurement'. Furthermore, one of the three companies mentioned that fuel performance was measured fortnightly assuming the volume refilled was the volume used – and despite the known caveats of this approach (see Smoothing Algorithm below for more information on this issue).

### 5.2.2. Fuel Efficiency Model General Considerations

One of DEA's major strength is that no assumption has to be made on the production technology and that any input or output of the production process can be included in the model (see (Cullinane et al., 2006) and section 3.3.3 about this). However, this characteristic introduces an element of appreciation on which variable should be actually included in a DEA model – which is a frequent criticism against DEA's robustness. Cooper *et al* (2007) recommend a careful selection of the model variables to ensure the model is robust and correctly reflects the performance process.

The literature review chapters explained how external information – such as vehicle weight – needs to be known when interpreting mpg. In order to create a DEA model which would improve fuel efficiency measurement, it is important to include all the variables which can impede the interpretation of mpg. This section will list all the variables of interest and will justify why some should be included whilst others should not.

The first variables to include are the 'fuel used' and the number 'miles travelled' (during the measurement period) so that the model could be illustrated as in Figure 5.1.

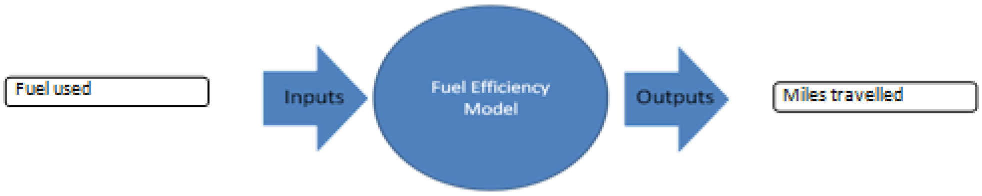


**Figure 5.1: Fuel Efficiency model – fuel used**

This model uses '*Fuel Used*' as an isotonic input and '*Miles Travelled*' as an isotonic output. Isotonic inputs are inputs which have a beneficial impact on the outputs production; i.e. an increase in isotonic input levels should translate to greater output levels (e.g. a vehicle should do more miles with two tanks worth of fuel in comparison to a single one).

The fuel information is to be collected from the fleets' fuel card records (see the Data Cleansing section for more information). Mileage, on the other hand, is obtained from the telematics unit. Although this can be slightly less accurate than the distance provided by the odometer, this data source is more consistent than drivers writing down the odometer readings (and rounding it up should they forget to write it down at the time of the refill). Furthermore using telematics information should avoid any misreading of poorly written odometer reading.

Another aspect of fuel efficiency is fuel cost. Indeed, it is conceivable that a vehicle can be mpg efficient (i.e. in respects to the litres of fuel used to cover a distance), but pence per mile (ppm) inefficient (i.e. in this case the cost of fuel per mile). To reflect the possible fact that a vehicle might be mpg efficient but ppm inefficient the '*Fuel cost*' is added to the previous model. This is illustrated as in Figure 5.2.

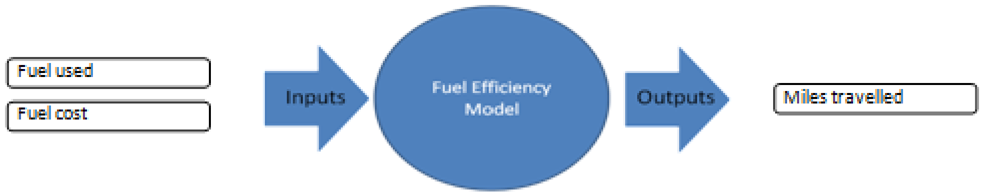


**Figure 5.2: Fuel Efficiency model – cost spent on fuel**

In addition, the fuel cost information is to be collected from the fleets' fuel card records. The *total cost spent on fuel* is used instead of the average pence per litre (ppl) value. The main reason for avoiding the use of ppl is that averages could cause issues (for possible issues in regards to using averages, see section 3.3.2 'Pair-wise and outranking methods' and (Laise, 2004)) but also because it is important to reflect the total contribution toward the mileage achieved. Furthermore, Cooper *et al* (2007, see 'Problem 1.4' p. 19) warn against using processed measures in DEA models as this could potentially put (undesired) emphasis on some variables (in this case the impact of 'fuel used' as an input could be reduced).

Vehicle weight is another factor which has a direct impact on fuel efficiency. In effect, heavier vans demonstrate worse mpg performance than lighter vans. This is illustrated in Table 5.3 which shows a list of vans sorted by weight (lightest on top, heaviest at the bottom).

| 1 | DMUName | Veh Type | Volume | Weight | Miles | MPG |
|---|---|---|---|---|---|---|
| 2 | 8 | Medium Van | 304.35 | 2185 | 3943.3 | 58.9012785 |
| 3 | 20 | Medium Van | 236.52 | 2185 | 2839.5 | 54.5773207 |
| 4 | 3 | Medium Van | 429.19 | 2185 | 4905.7 | 51.9624477 |
| 5 | 14 | Medium Van | 213.32 | 2185 | 2425 | 51.6795088 |
| 6 | 2 | Medium Van | 257.57 | 2185 | 2901.5 | 51.2112652 |
| 7 | 5 | Medium Van | 250.88 | 2185 | 2766.1 | 50.1233448 |
| 8 | 21 | Medium Van | 229.04 | 2185 | 2507.2 | 49.7640655 |
| 9 | 1 | Medium Van | 309.36 | 2185 | 3367.7 | 49.4888597 |
| 10 | 10 | Medium Van | 218.97 | 2185 | 2377.2 | 49.353654 |
| 11 | 17 | Medium Van | 470.22 | 2185 | 5011.4 | 48.4502676 |
| 12 | 18 | Medium Van | 425.04 | 2185 | 4366.2 | 46.6994786 |
| 13 | 9 | Medium Van | 318.34 | 2185 | 3233 | 46.1692374 |
| 14 | 19 | Medium Van | 258.92 | 2185 | 2604.4 | 45.7277989 |
| 15 | 15 | Medium Van | 155.01 | 2185 | 1553 | 45.5459692 |
| 20 | 16 | Medium Van | 467.7 | 2185 | 4155.1 | 40.3879974 |
| 21 | 12 | Medium Van | 262.1 | 2185 | 2262.8 | 39.2479844 |
| 22 | 7 | Medium Van | 242.13 | 2185 | 1641.8 | 30.8254807 |
| 23 | 28 | Large Van | 291.32 | 2861 | 3314.3 | 51.7201439 |
| 24 | 33 | Large Van | 305.68 | 2861 | 3468.9 | 51.5896955 |
| 25 | 31 | Large Van | 297.49 | 2861 | 3284.2 | 50.1874851 |
| 26 | 23 | Large Van | 544.92 | 2861 | 5374.7 | 44.8393893 |
| 27 | 42 | Large Van | 339.14 | 2861 | 3117.1 | 41.7839919 |
| 28 | 26 | Large Van | 327.1 | 2861 | 2841.9 | 39.4972134 |
| 29 | 30 | Large Van | 486.6 | 2861 | 3966.3 | 37.0554135 |
| 30 | 29 | Large Van | 123.29 | 2861 | 992.2 | 36.5855492 |
| 42 | 34 | Large Van | 542 | 2861 | 3781.6 | 31.7186366 |
| 43 | 39 | Large Van | 706.37 | 2861 | 4818.8 | 31.0130775 |
| 44 | 35 | Large Van | 616.3 | 2861 | 4163.8 | 30.7139662 |
| 45 | 68 | Large Van | 580.06 | 2900 | 4229.9 | 33.1509052 |
| 46 | 69 | Large Van | 459.74 | 2900 | 3320.5 | 32.834424 |
| 47 | 66 | Large Van | 447.92 | 3000 | 3266.7 | 33.1548454 |
| 48 | 61 | Large Van | 576.51 | 3000 | 4066.6 | 32.0673314 |
| 49 | 63 | Large Van | 683.56 | 3000 | 4686.3 | 31.1667599 |
| 50 | 62 | Large Van | 625.5 | 3000 | 4000.2 | 29.0731842 |
| 51 | 65 | Large Van | 447.7 | 3000 | 2862.8 | 29.0698053 |
| 52 | 64 | Large Van | 211.39 | 3000 | 1224.6 | 26.3358916 |
| 53 | 57 | Large Van | 273.94 | 3300 | 2070.4 | 34.3587232 |
| 56 | 46 | Large Van | 407.95 | 3300 | 2953 | 32.907487 |
| 57 | 58 | Large Van | 431.75 | 3300 | 3096 | 32.5991904 |
| 58 | 51 | Large Van | 411.18 | 3300 | 2914.7 | 32.2255314 |
| 59 | 44 | Large Van | 326.1 | 3300 | 2231.2 | 31.1046924 |
| 60 | 56 | Large Van | 487.77 | 3300 | 3281.1 | 30.5803597 |
| 63 | 53 | Large Van | 829.69 | 3300 | 5388.8 | 29.5266665 |
| 64 | 48 | Large Van | 624.4 | 3300 | 3858.6 | 28.0934499 |
| 65 | 50 | Large Van | 792.75 | 3300 | 4827.2 | 27.6819864 |
| 66 | 52 | Large Van | 483.45 | 3300 | 2872 | 27.0066726 |
| 67 | 45 | Large Van | 467.76 | 3300 | 2738.6 | 26.6160579 |
| 68 | 54 | Large Van | 756.11 | 3300 | 4294.4 | 25.8199693 |
| 69 | 60 | Large Van | 621.2 | 3300 | 3498.3 | 25.6014057 |
| 70 | 67 | Large Van | 519.09 | 3500 | 3123.9 | 27.3585244 |

Table 5.3: Impact of weight on the traditional mpg measure

Table 5.3 clearly shows that 'vehicle weight' needs to be taken into account to appraise fuel efficiency when looking at vans' mpg (none of the heavier vans are green). In a similar manner as with 'fuel cost', 'vehicle weight' can be added to the fuel efficiency model as illustrated by Figure 5.3.

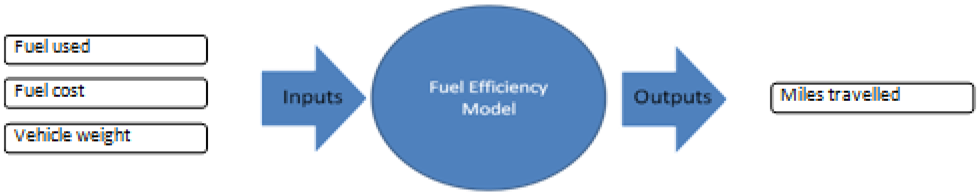


Figure 5.3: Fuel Efficiency model – vehicle weight

Given that heavier vans burn up more fuel than lighter ones, vehicle weight is an anti-isotonic input; i.e. the heavier the vehicle is, the fewer miles it would travel for a given amount of fuel. Furthermore, whilst a vehicle's weight can generally – to some extent – be reduced, this study will consider 'vehicle weight' fixed (this is because most fleet managers would not be able to reduce vehicle weight in order to improve fuel efficiency). As 'fixed' variables are not the standard behaviour DEA expects from inputs, appropriate calculations should be made to correctly include weight in the fuel efficiency model. The different possible methods to reflect this 'fixed' characteristic will be discussed appropriately in section 5.4.4 'Adding the Weight'. The vehicle weight information is obtained from the Vehicle Certification Agency (VCA) based on the vehicle make, model and description given by the fleet managers. This database only informs of the vehicle gross weight as the vehicle net weight generally depends on the type of equipment fitted inside the van. Because the exact vehicle net weight was not known exactly (no weight bridge like with bulk LGVs vehicles), using vehicle gross weight was considered a less biased option and was retained for this study. The limitations attached to using vehicle gross weight instead of net weight will be discussed in section 7.4 Potential for Further Research.

Finally, the vehicle age should also be included in the model to reflect the fact older vehicles are less efficient that new ones. This can be caused by engine ageing, but also because newer vehicles often offers better performance in terms of combustion power for fuel used. This can be illustrated as illustrated by Figure 5.4.

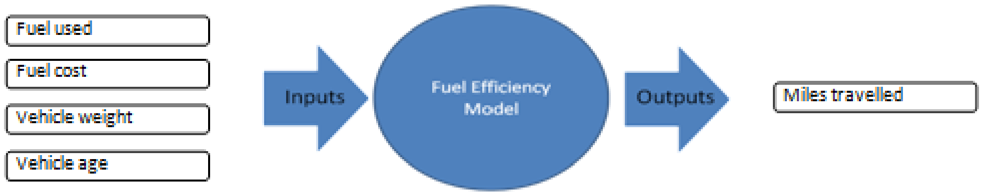


**Figure 5.4: Fuel Efficiency model – age**

This variable is also anti-isotonic and 'fixed' and this should equally be reflected in the calculations. Each vehicle age is obtained from the vehicle registrations (all registrations are standard registrations). There could be a slight rounding error in the vehicle age as a vehicle bought at the end of a registration year would have done less mileage than another bought at the beginning of the same year. This approach also ignores the cases where a vehicle was not registered within 6 months after being manufactured.

Although servicing could have a potential impact on fuel efficiency it will not be included in the study. This is because many big fleets have short term lease contracts which include servicing (as this is also the case for these companies). This generally hides the real cost of servicing and is the main reason for not including servicing cost in the fuel efficiency model.

Tyre pressure also impacts fuel efficiency (FBP, 2005) and should ideally be included in the study. However, this information can generally only be provided by drivers who might not be able to remember it correctly or who could provide misleading information (e.g. from fearing the management's reaction). Furthermore, it would be difficult to provide fleet managers results suggesting their drivers should check their vehicle's tyre pressure less often. Moreover, there is a strong health and safety issue in doing so as small tyre punctures are still possible and checking tyre pressure regularly can potentially prevent this type of accident. Because of the risks associated with including the number of tyre pressure checks are too serious in light of the potential benefits gained from including this variable, tyre checks will not be included in the model.

The types of operations also affect the fuel efficiency (FBP, 2005); a vehicle travelling a constant speed of 60 mph on the motorway will demonstrate a better mpg performance than the same vehicle driving in a busy town centre. However, no common agreement on vehicles operational environment metrics (e.g. such as a 'difficulty score' linking type of operations to fuel efficiency) exist. Therefore, including these would make it hard to demonstrate that the efficiency levels measured are actually caused by the vehicles' performance and not by the specific method used to measure the environmental factors. Furthermore, vehicles within a sample (in this case a company) are assumed to operate under similar conditions so it is acceptable to ignore environmental factors as long as the benchmarking studies are conducted within each company (although a test will be conducted with all the vehicles from all the companies in order to appraise whether there truly exist environmental factors).

Engine size also has an impact on fuel efficiency and could potentially be included in the fuel efficiency model. However, discussions with fleet managers uncovered that the choice of engine size is generally decided based on the type of operations (e.g. carrying boilers or other heavy products) which required more or less powerful engines. Due to the relative small sample sizes it was estimated unwise to include engine size in the model (not enough vehicles in each sub-'operation types' for the model to provide meaningful answer).

Finally, driver behaviour information should also not be included in the model for the reason that it is already reflected in the 'fuel used', 'fuel cost' and the 'mileage' variables. Furthermore, the study's objective is not to find the regression parameters which should predict what fuel performances levels will be based on driver behaviour information, but instead to find the best possible way to measure fuel efficiency.

Because of the many variables included in the fuel efficiency model, several different aspects of performance should be reflected in this measure. In light of the families of measure introduced earlier in section 3.2.3, the fuel efficiency model exhibits the following characteristics:

- ⇨ It is a productivity measure.
- ⇨ It should help appraising resources utilisation and allocation.
- ⇨ It should reflect the operational cost.

The fuel efficiency model does not cover the quality, timeliness, cycle-time and safety performance measurement families. These should probably not be included as not really related to the notion of fuel efficiency.

Table 5.4 summarises the factors included or discarded from the model.

| Factor | Included | Explanation |
|---|---|---|
| **Fuel used** | YES | Intrinsic to the notion of fuel efficiency. |
| **Fuel Cost** | YES | Important aspect of fuel efficiency to include in the measure. |
| **Vehicle weight** | YES | Important factor impacting fuel efficiency. |
| **Vehicle age** | YES | Potential factor impacting fuel efficiency. |
| **Tyre pressure** | NO | Information difficult to obtain, bias in drivers' answer. |
| **Driver behaviour** | NO | Measure should reflect driver behaviour, not include it as an input to 'predict' fuel efficiency levels. |
| **Maintenance** | NO | Potentially interesting although the company selected are under fix maintenance contract with their manufacturers; hence maintenance price difficult to obtain. |
| **Topography** | NO | Potentially interesting but difficult to measure. Assumed similar for a long enough measurement period. |
| **Weather** | NO | Assumed similar for all vehicles across the measurement period. |

| | | |
|---|---|---|
| **Traffic** | NO | Difficult to measure. Assumed similar for a long enough period of measurement. |
| **Vehicle speed** | NO | Hard to differentiate necessary speed from unnecessary speed. Similarly to driver behaviour, the measure should reflect driver behaviour, not include it as an input to 'predict' fuel efficiency levels. |

**Table 5.4: Factors to include in the fuel efficiency model**

## 5.3. Data Cleansing

As with any modelling, it is crucial to cleanse the collected data to ensure the inferences drawn from the results are correct. In this study, the fuel used and fuel cost information is obtained from fuel card records. However – and as previously introduced in the Case Study Theoretical Background section above – fuel card data tend to generally not be very accurate. This is mainly caused by the human process of drivers spelling their vehicle registration out to the person at the till, who then types it in (for driver fuel card, see section 2.2.3 Fuel Card Management). Because DEA is sensitive to measurement error, it is consequently crucial to adequately cleanse the fuel card data before using it in the model (Avkiran and Thoraneenitiyan, 2009). This section will explain how the fuel card data are cleansed.

Fuel card records generally give at least the following information (the exact list of information available depends on the fuel card provider):

- ⇨ Vehicle registration

- ⇨ Transaction date (and sometimes time for some fuel card provider like ReD)
- ⇨ Card number
- ⇨ Volume of fuel drawn
- ⇨ Net cost
- ⇨ VAT

Because the only common identifier between the fuel card dataset and other vehicle information (i.e. the vehicle weight and mileage) is the vehicle registration, it is necessary to ensure the registrations in the fuel card data are correct. In order to appropriately cleanse these data, a small cleansing algorithm was designed. Part of this algorithm has been coded using Visual Basic for Application (the programming language for the Microsoft Office suite; in this specific case used on Microsoft Excel); the remaining part of the algorithm was executed manually (this includes manually querying some of the telematics databases). This section will describe the different steps of this algorithm.

### 5.3.1. Cleansing Algorithm

The algorithm works in a series of steps described below. The algorithm starts at step one and automatically carries on to the next step unless a match is found.

**Step 1**

Registration spaces are taken out. This means that 'W135 OFD' becomes 'W135OFD' for any following comparison. The fuel cards files are cleansed so as to only keep actual fuel transactions (oil and other products are discarded). Registrations from vehicle fuel cards should match with fleet details at that point (as the fuel card registration should be the registration embossed on the card).

**Step 2**

If an exact match exists on the left 7 digits, then the process stops and is started again for the next fuel card record (if any).

This means that registrations like 'W135 OFD', 'W135OFD', 'W135 OFD – SOME ADDED TEXT' and 'W135OFD – ADDED TEXT' would all match with each other (this is important as some people add extra information after the registration both on the fuel card and sometimes on the telematics systems). In some rarer cases, information is added at the beginning of the registration. This will cause registration to mismatch but is corrected manually.

**Step 3**

A list of similar registrations is built based on phonetic mistakes.

For example some zeros in the following registration 'VO05 IFD' can be sometimes pronounced as 'o' (the letter 'o'). This causes many records to be typed as 'VOO5 IFD' or 'V005 IFD'. By default, if the algorithm finds a misspelt registration (as in 'VOO5 IFD') it will try finding an exact match on the corrected registration 'VO05 IFD'. If the match is found, the process stops and is started again for the next fuel card record (if any). Variant of the 'o' '0' error are also calculated (e.g. 'W135 F0D' (with a zero) is in fact 'W135 FOD').

For other phonetic errors, (e.g. ‘W135 OFD’ where the D is mistaken for an ‘E’ in the fuel card file (‘W135 OFE’)), potential phonetic matches are computed. Two registrations are matched only if there is no other similar registration which could match. For example, there is an unmatched registration ‘W135 OFD’ on the fuel card file which can only be matched phonetically with ‘W135 OFE’ (i.e. there does not exist any registration OFB, OFC, OFP or OFV). The registration ‘W135 OFD’ is then matched with the existing vehicle ‘W135 OFE’. Here again, if a match can be found, the process stops for this registration and starts again for the next fuel card record (if any).

**Step 4**

If the fuel card file has the date and time of transaction, the telematics records of the whole fleet are queried at the time of the fuel transaction in order to obtain the list of all the vehicles which stopped at the time of the fuel transaction. If a single vehicle stopped at that time (i.e. engine stopped within 10 minutes before of the fuel transaction and started again within 5 minutes following the time of the transaction), then the registration is matched with the registration of the unique vehicle which was stopped at the time of refill (and the process stops and starts again for the next fuel card record if any).

**Step 5**

If the cleansing process reaches step 5, then:

⇨ There is either no phonetic match or several possible matches.

- ⇨ There are several vehicles that stopped at the time of the fuel transaction.

The telematics records are consequently queried for the list of vehicles which stopped at the time of the fuel transaction and which are potential phonetic matches. If a unique vehicle matches these two criteria, the registration is matched against this vehicle.

**Step 6**

The mpg performance is calculated for all the vehicles around the fuel transaction (or all the vehicles which were stopped at the time of the transaction if telematics information is available). If a single vehicle has an unrealistically good mpg and if the registrations are relatively similar (human appreciation), the registration is matched against this vehicle. High transactional mpg are checked because any vehicle missing a fuel transaction would 'look' like it has travelled many miles without using a lot of fuel (because one refill was not accounted for). This concept is illustrated below in 'Notes on cleansing algorithm'.

If step 6 fails to find a match, the vehicle is flashed as unmatched.

**Notes on cleansing algorithm**

A consequent number of registrations (as high as 30%) were matched in step 4 and 5 which explicitly rely on telematics information to find potential registration matches. This illustrates the importance of telematics information when cleansing fuel card data. Step 4 and 5 also rely on the fuel card files to provide the time of refill. This suggests that telematics equipped vehicles and fuel card files showing the transaction time would help improving the quality of the fuel cards data cleansing. Keyboard misspellings (i.e. mistaking a ‘q’ for a ‘w’ – the two keys being next to each other on a keyboard) were not taken into account in this algorithm as too many possible matches made the process ineffective.

The mpg in step 6 can be either calculated using telematics mileage (and the time of the transaction) or using the odometer reading mentioned at the petrol station. However this last option relies on the driver and the person at the petrol station’s till to provide the correct information.

Failing to correctly attribute a fuel transaction to a vehicle would artificially increase its fuel performance (the vehicle would seem to have travelled the same amount of miles using less fuel). This dramatically increases the risk of having a DMU artificially becoming efficient – which could affect many other inefficient DMU’s scores (i.e. all the DMUs which would have the artificially efficient DMU in their reference set). To limit this, the mpg figure can be calculated between each fuel transaction and any vehicle having a mpg greater than usual can be discarded from the study. This principle is illustrated in Figure 5.5.

| Transaction ID | Miles travelled since last refill | Gallons | Mpg | Litres |
|---|---|---|---|---|
| 1 | 500 | 13 | 38.46 | 59.09 |
| 2 | 550 | 13.5 | 40.74 | 61.36 |
| 3 | 350 | 9.4 | 37.23 | 42.72 |
| 4 | 300 | 8 | 37.50 | 36.36 |
| | | | | |
| | | | | |
| MPG over period (missing greyed transaction) | | 55.92 | | |
| Third transaction's mpg (missing greyed transaction) | | 95.74 | | |
| Ratio | | 1.71 | | |
| Actual MPG | | 38.72 | | |

Figure 5.5: Spot missing transactions

In this example, the greyed transaction could not be matched so that the apparent vehicle’s mpg over the period is 55.92 (the real mpg which takes all transactions into account is 38.72). The mpg calculated for the third transaction is calculated using 900 miles and only 9.4 gallons (missing the 13.5 gallons of the previous refill in grey), hence the mpg for the third transaction shows an unrealistic mpg performance of 95.74. This extremely high mpg is significantly greater than the (apparent) average mpg of 55.92 (the exact threshold is up for debate but it is reasonable to consider suspect anything above a ratio of 1.3). This shows it is likely there is a missing fuel transaction and that the vehicle should consequently be discarded from the study. Not refilling up to the top of the tank can also have a similar affect although this can be spotted as the following transactions would compensate for this partial refill. Although a missing transaction can go unnoticed when considering average mpg over a period of time, considering the individual mpg between fuel refills can help spotting vehicles which have missing fuel transactions. It is important to note that not filling up to the top of the tank can also cause artificially higher mpg figures.

Commercial Motor regularly publishes the results of its mpg marathons (Tonkin, 2009b). This gives mpg figures as high as 55 for medium to ‘small heavy’ vans (Vauxhall Vivaro 2.0 CDTI) and a mpg figure of 40 for heavy vans (Vivaro 2700 SWB CTDI) (MPG marathon, 2009). Assuming the Commercial Motor mpg marathon is an industry standard, any medium (or small heavy) van with a mpg performance higher than 55 would be deemed suspect and could be discarded from the analysis. Similarly any heavy van demonstrating a performance greater than 40 could also be deemed suspect and could be discarded from the study. The exclusion process considers different variables such as engine size or type of operations but also considers historical fuel information to appraise the reliability of the data. For example, having consistent transactional mpg figures (e.g. no theft, or missing transaction) would bring confidence in the data and, should the engine size be small for the vehicle category, a slightly higher mpg would not be deemed suspect thus, the vehicle would not be discarded.

Several final points are also worth noting:

- ⇨ Any vehicles with inconsistent fuel card records (e.g. fuel card records for only half of the measurement period) were discarded from the study.
- ⇨ The mileage and vehicle weight were deemed accurate thus no data cleansing was conducted on these variables.
- ⇨ Finally, no telematics unit was reported faulty during the measurement period (as otherwise this would have had an impact on mileage).

### 5.3.2. Theft detection

Discussions with operating managers and telematics experts indicated that theft can be detected in three different ways:

- ⇨ The vehicle mpg between two refills is unrealistically low (and previous not unrealistically high).
- ⇨ A fuel transaction for a volume greater than the vehicle's fuel tank occurred.

- ⇨ The vehicle was not at the petrol station at the time of the transaction (requires telematics).

There are several problems in trying to detect fuel theft. Defining ‘unrealistically low’ obviously depends on several factors such as vehicle model, variable load and type of operations. This can make individuals disagree on what mpg limits or what criteria to use. A potential leak of the fuel system is also still possible thus it would be hard to incriminate a driver based on this information alone.

When refilling a vehicle, there is almost always some fuel left in the tank. Thus, the volume refilled should nearly always be lower than the fuel tank capacity. In some cases however, some drivers need to fill a jerry can for use on a plant or workshop (e.g. for a mini-digger in the case of Carillion). This is however bad practice and a specific fuel card/registration code should be used for jerry can refilling which would allow proper performance monitoring. In this case again, it is not possible to draw conclusions on this information alone.

The last criterion can spot cases where a driver gives its fuel card to somebody to refill a vehicle which does not belong to the company. This method obviously relies on telematics and would not work if both the company’s vehicle and the fraudulent vehicles are at the same petrol station at the time of the transaction (as the company vehicle will be at the petrol station at the time of refill).

Although theft detection is crucial for business operations, it is less so for this particular study as any vehicle from which fuel has been stolen (either via siphoning or via the fraudulent use of a fuel card) would demonstrate a poor fuel efficiency performance. Poor performance should be an incentive for the management to closely monitor the drivers' performance and challenge them to improve their performance. For this reason, theft detection was not used to discard any vehicle from this measurement study.

### 5.3.3. Smoothing Algorithm

As previously introduced at the end of section 2.2.3 Fuel Card Management, fleet managers often need to measure fuel efficiency across the whole fleet over a period of time (e.g. a month). This measurement period has also to be the same for all vehicles in order to avoid any bias. However, mpg can only be accurately calculated between refills – and only if the refills are always made up to the top of the tank. It is also generally not possible for fleet managers to have all their vehicles refilled at the exact beginning and end of fuel trials. This makes measuring fuel efficiency during fuel trials a tricky problem. Because the study will have to use the volume of fuel used between the start and end of the measurement period, this problem will have to be properly addressed. This section will explain in more details why refills always need to be made up to the top of the tank. A small algorithm which can address the inaccuracy caused by measuring mpg between two dates (on which refill did not occur) will also be introduced.

#### 5.3.3.1. On the Necessity to Refill Up To the Top of the Tank

The tank level varies as the vehicle is being used. Typical tank level variations are illustrated in Figure 5.6.

This graph shows that each time the vehicle is being driven the tank fuel level goes down. Similarly, every time the vehicle is being refilled, the tank level goes up. Note that the third refill has not been made up to the top of the tank (which is 70 litres in this example). With proportional scales, each refill should be represented by a vertical line; however, there is a slight slope for the refills on the graph above due to the way the graphic generator handles the data (especially the third small refill); however, this should not prevent from explaining the concept.

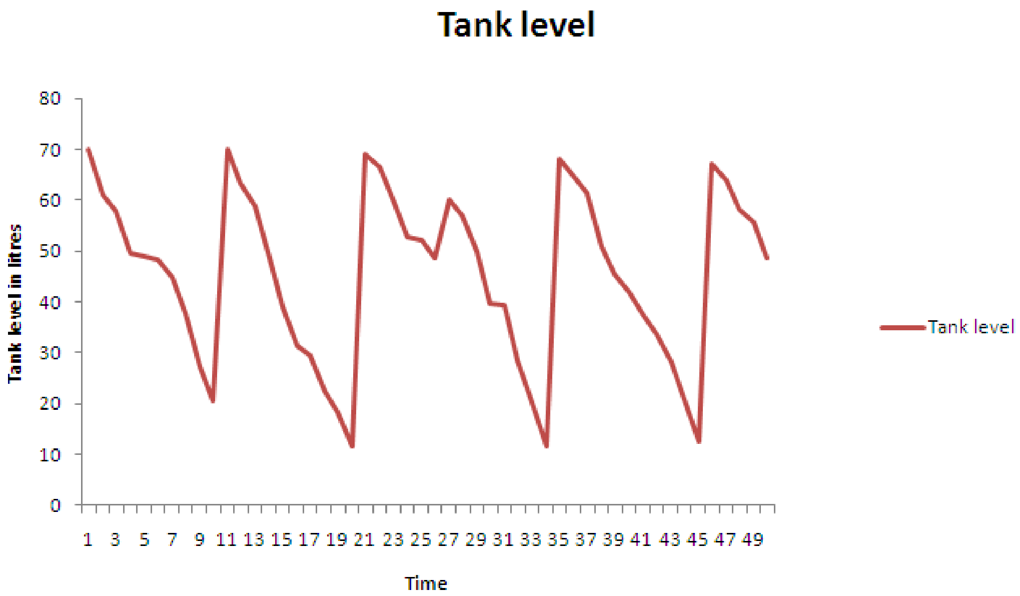


Figure 5.6: Tank level

Figure 5.7 illustrates how mpg is calculated between transactions using fuel card information.

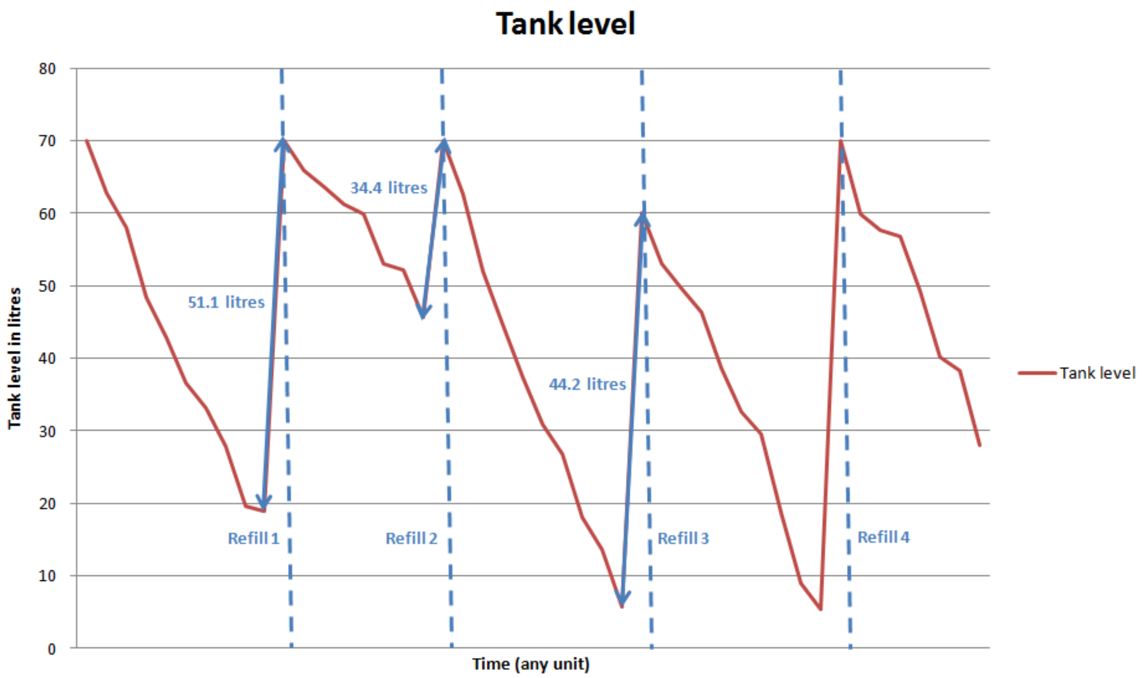


Figure 5.7: How not refilling to the top of tank affects measurement

In Figure 5.7, the vehicle's mpg between Refill 1 and Refill 2 is calculated by taking the distance travelled between the two refills and dividing it by the number of litres of Refill 2.

On the other hand, Refill 3 was not made up to the top of the tank and the quantity of fuel used to cover the distance between the two refills, is not 44.2 litres (Refill 3's volume), but 54.2 litres. Because the Refill 3 was not made up to the top of the tank, the mpg calculated between Refill 2 and Refill 3 is incorrect.

Although a fuel transaction's mpg is always erroneous when the refill is not made up to the top of the tank, it is important to observe that the error will be smoothed out on longer periods providing the first and last refills are up to the top of the tank. Because partial refills would cause the mpg between two transactions to be artificially high, they could also cause false positive in Step 6 of the previous algorithm (which is meant to spot missing registration; see Cleansing Algorithm above).

#### 5.3.3.2. On the Issue of Measuring mpg Over a Period of Time

As explained before, it is possible to make an accurate estimation of the quantity of fuel used in between two refills as long as both refills are made up to the top of the tank. When measuring fuel performance across a whole fleet, it is hard (or impossible) to have all the vehicles refill at the exact beginning and exact end of the period. Consequently, it is not possible to know the exact volume each vehicle used over the period. Thus, it is not possible to calculate fuel consumption or mpg accurately over such period (at least not with fuel cards; in opposition CANbus could provide this information). This concept is illustrated in Figure 5.8.

The beginning and end of period are indicated with vertical dotted lines, the refills are indicated with small triangles on the timeline.

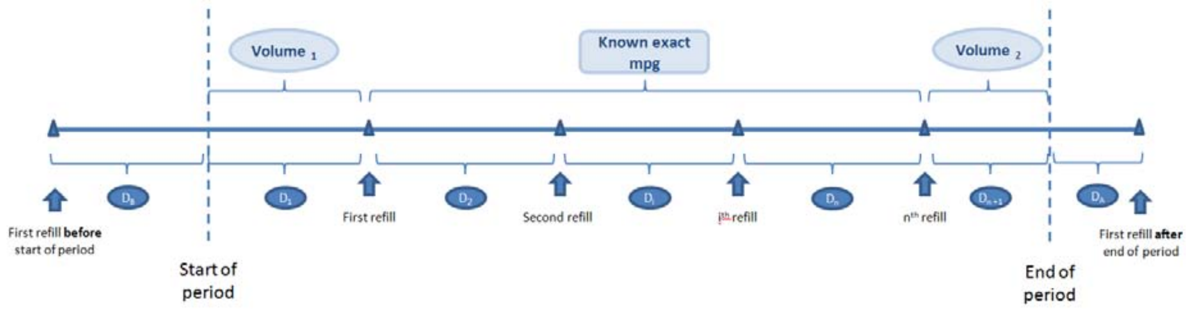


Figure 5.8: Measuring fuel efficiency over a period of time

Considering the beginning of the measurement period, the fuel put in the tank at the first refill before the start of the period was used to cover both distance $\mathrm{D}_1$ and distance $\mathrm{D}_\mathrm{B}$ (Distance Before). This is illustrated in Figure 5.9.

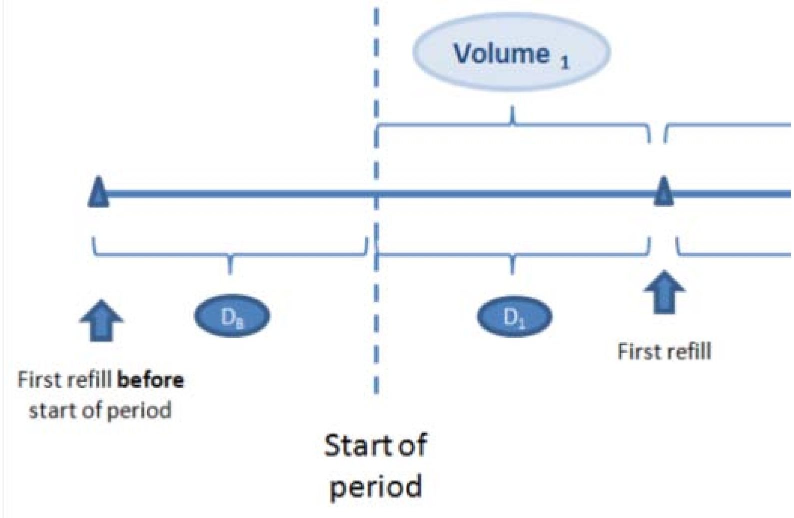


Figure 5.9: Inaccuracy at the period start

Similarly, the fuel put in the tank during the last refill '*n*' was used to cover both distance $\mathrm{D}_{\mathrm{n}+1}$ and distance $\mathrm{D}_\mathrm{A}$ (Distance After). This is illustrated in Figure 5.10.

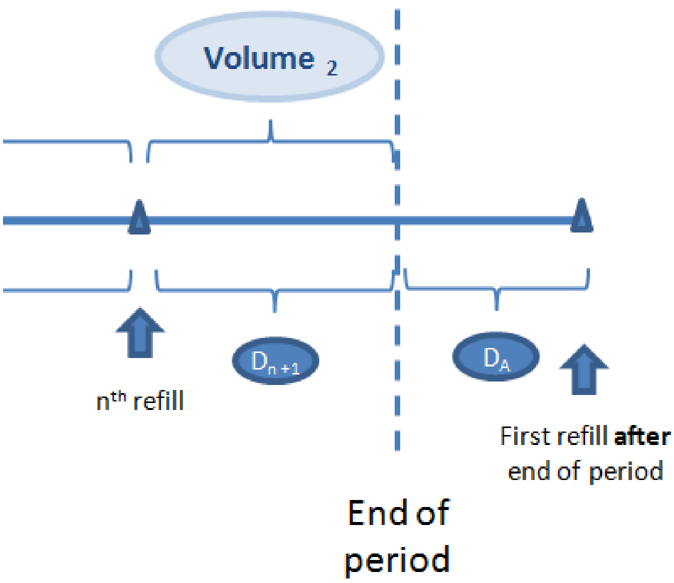


Figure 5.10: Inaccuracy at the period end

Because of this, calculating the mpg using the total volume of fuel drawn during the period is inaccurate. Formula 5.1 shows the calculations necessary to calculate this (incorrect) mpg.

$$Incorrect\ mpg = \frac{\sum_{i=1}^{n+1} D_i}{\sum_{j=1}^{n} Volume\ (Gallons)\ of\ refill_j}$$

Formula 5.1: Incorrect mpg formula

#### 5.3.3.3. Explaining the algorithm

The smoothing algorithm works by evaluating the amount of fuel used at the beginning and end of the period instead of relying on any assumption that these volumes would even out (it is also incorrect to think mpg over a period would compensate the mpg of the next period as mpg is not a linear measure). This section will explain the smoothing algorithm principles.

Supposing that the measurement period is long enough and that the vans operations are intense enough, most vans would then refill several times during the measurement period. Following, section 5.3.3.1, the exact mpg between two refills can be calculated as long as the vehicle was refilled up to the top of the tank. Thus, in most cases, an accurate mpg can be calculated between the first and the last refills of the measurement period (there would be a small error if the first and last refills are not made up to the top of the tank). This concept is illustrated in Figure 5.11.

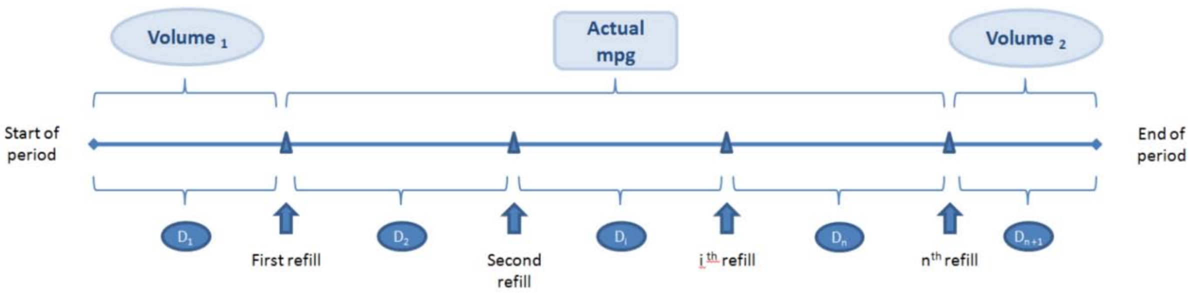


Figure 5.11: An accurate mpg between first and last refills

The mpg calculations for the known mpg are illustrated in Formula 5.2.

$$Actual\ mpg = \sum_{i=2}^{n} \frac{D_i}{Volume\ (Gallons)\ of\ refill_i}$$

Formula 5.2: Accurate mpg formula

It is then possible to estimate the volume used to cover known distance $\mathrm{D}_1$ and distance $\mathrm{D}_{n+1}$ using the previously calculated mpg (the distances can be gathered either using telematics devices or odometer readings). These calculations are illustrated in Formula 5.3.

$$Smoothed\ Volume = \sum_{i=2}^{n} Volume\ of\ refill_i + \frac{D_1 + D_{n+1}}{mpg_a}$$

$$where\ mpg_a\ is\ the\ 'actual\ mpg'\ calculated\ above$$

Formula 5.3: Smoothed volume formula

Formula 5.4 illustrates the full factorised calculations.

$$Smoothed\ Volume = \sum_{i=2}^{n} Volume\ of\ refill_i \times \left(1 + \frac{D_1 + D_{n+1}}{\sum_{i=2}^{n} D_i}\right)$$

Formula 5.4: Factorised smoothed volume formula

The Smoothed mpg can then be calculated using this Smoothed Volume and the total mileage travelled during the measurement period. This is illustrated in Formula 5.5.

$$Smoothed\ mpg = \frac{Total\ Distance\ Travelled}{Smoothed\ Volume}$$

Formula 5.5: Smoothed mpg formula

A detailed example of the calculations can be found in section 8.5 Appendix 5: Smoothing Algorithm Calculations Example.

This formula smoothes the mpg and is particularly effective when a refill has occurred soon before the start of the measurement period, or straight after the end. The Smoothing Algorithm results are illustrated in Table 5.5.

| | VehCode * | Bad MPG * | Smoothed MPG * | Fuel Drawn * | Fuel Used * |
|---|---|---|---|---|---|
| 1 | 1 | 53.671 | 49.49 | 285.26 | 309.356 |
| 2 | 2 | 53.025 | 51.211 | 248.76 | 257.569 |
| 3 | 3 | 51.452 | 51.963 | 433.45 | 429.187 |
| 4 | 4 | 43.252 | 42.999 | 121.39 | 122.104 |
| 5 | 5 | 50.554 | 50.124 | 248.74 | 250.876 |
| 6 | 6 | 42.361 | 43.726 | 170.45 | 165.129 |
| 7 | 7 | 29.653 | 30.826 | 251.71 | 242.128 |
| 8 | 8 | 63.313 | 58.901 | 283.14 | 304.351 |
| 9 | 9 | 44.121 | 46.169 | 333.12 | 318.34 |
| 10 | 10 | 59.164 | 49.354 | 182.66 | 218.967 |

Table 5.5: Smoothed mpg results

It should be noted that the smoothing process can either reduce or increase the ‘bad’ mpg (i.e. the mpg calculated with the incorrect formula as introduced earlier). Reduction would generally occur when the mpg has been artificially increased by a refill soon before the start of the period or soon after the end of the period (or both). An increase on the other hand is likely to happen when the vehicle has refilled straight after the beginning of the period or just before the end (or both).

Although mpg performance calculated from fuel card data is generally consistent as long as the fuel cards are correctly used, there is no guarantee that the difference between the smoothed mpg and the ‘bad’ mpg is greater than the variability of the actual transactional mpg around the smoothed mpg (and this could have an impact on the Smoothed mpg should there be only a few refills). It would be interesting to compare this smoothing approach with CANbus fuel consumption and appraise whether the benefits truly are significant. This cannot be done in this study however as none of the companies has CANbus enabled vehicles. Nonetheless, by discarding the randomness around the refills, the smoothing algorithm offers a logical and robust approach to measure fuel efficiency based on observed mpg performance.

An alternative could also appraise fuel used in the first segment by considering the mpg between the first refill before the start of the period and the first refill after the start of the period (and vice versa for the end of the period). However, because this approach uses only two refills inaccuracies caused by not refilling to the top of the tank could not be compensated. Similarly, there is no guarantee the first transaction after the end of the period will be available. Consequently, this approach should be avoided.

Because the smoothing algorithm has been proven robust, the DEA models developed in this study will use the *smoothed volume* calculated for the measurement period. The smoothed cost was also calculated based on the same principles. It is important to observe that without telematics distance information, this algorithm also requires the odometer reading to be taken at the beginning and end of the period (not just at the refills).

## 5.4. The fuel efficiency model

This section will detail how the fuel efficiency model was built, both in regards to the computational aspects and to the modelling approach. As previously mentioned, a step by step approach will be used, starting from a basic model which should demonstrate a high correlation with the traditional mpg measure. All the variables discussed in the previous sections (e.g. vehicle gross weight, vehicle age) should then be progressively added to this original model. This approach is essential to ensure each variable added to the model truly impacts performance. Each of these steps will be detailed in the following sections. Analysis of the model results will be conducted at each step of the model building process. This approach is specific to this study as each step is a logical continuation depending on the results of the previous one. The models' results will be summarised the section 'Summary of Results'.

### 5.4.1. Computational aspects and research process

In order to test the fuel efficiency model, it is necessary to run some DEA models with the collected data. There are two possible alternatives here: either use some already available DEA software or write the code to solve these DEA problems.

Most existing DEA software available at the moment are either restricted in their options, or do not describe accurately which DEA models were used to generate the results (e.g. Banxia's software, see (Virtos, 2010)). Two excellent Excel based DEA programs; one developed by Zhu (2010) and another by Cooper *et al* (2007) are available but limited in the number of DMUs they can handle (at least in their student version) and for the software written by Zhu, with limited result information (e.g. missing weights). Professor Emmanual Thannassoulis and Dr Ali Emrouznejad have also recently developed a DEA program called Performance Improvement Management (PIM) which is probably one of the most advanced DEA software available at the moment. However, at the time this study was conducted, PIM was not yet available.

Because none of the options available at the time this study was conducted was satisfactory, some specific code was developed to solve the appropriate DEA problems. The code was developed in C#, a programming language created by Microsoft™. The reason for using C# and not a more open language such as C++ is that C# runs in a managed environment which makes the process or writing and debugging code a lot easier and can eventually result in a greater productivity as long as portability to a non-Windows platform is not an issue. The JAVA language has also a managed environment and could have been equally used.

The optimisation process used the LINDO library (LINDO, 2010) which can solve linear problems on a large scale. The code developed reads data from MySQL server databases using different queries and the optimisation results are stored both into a comma separated file and written back into a database table. Code was developed for the (input oriented) CCR, BCC, SBM, NCN, SBM, SBM-NC (non-controllable) and SBM-ND (non-discretionary) models. The data manipulation is mainly made through arrays as these are faster and can closely relate to matrix calculations (used in the DEA algorithms). The code was tested against manual calculations and, for the existing DEA models, against the Cooper *et al* free DEA solver.

The CCR, BCC, and SBM models will be first used to measure the fuel efficiency of the basic fuel and miles model. This should allow a better understanding of how performance is measured by the different models as well as providing scale and mix efficiency ratios for each DMU (see Cooper et al., 2007, p. 153). This information should help deciding which model should be used to measure van fuel efficiency.

Although the CCR, BCC and SBM models could suit the basic fuel efficiency model requirements, they do not offer the mechanism to appropriately incorporate variables such as vehicle weight (which is mainly due to the fact the vehicle weight is an anti-isotonic variable). Consequently, several other models were written which ultimately led to the writing of an extension to the SBM model.

### 5.4.2. Model with Fuel & Miles Alone

The first DEA model to be tested only included 'fuel used' as input and 'miles' as output. The fuel used is evaluated using the Smoothing Algorithm. Because this input is isotonic, it does not require any special calculation. Miles travelled are gathered from the telematics system for the relevant measurement period (from 01/04/2009 to 30/06/2009).

This first model is fundamental as it enables the evaluation of DEA as an appropriate alternative to the traditional mpg measure. Because this first DEA model only uses ‘fuel used’ and ‘miles’, DEA scores should be perfectly correlated to the mpg measure. The scores might not however be exactly proportional as the computer calculations use decimals and not exact values which can generally create some small rounding errors.

As discussed earlier, this model results are calculated using the CCR, BCC and SBM models as these models should give sufficient information to understand efficiency and to calculate mix and scale efficiency ratios.

#### 5.4.2.1. Overall results

Results for FSH Maintenance are illustrated in Table 5.6. Results obtained with the other companies’ datasets were similar to those obtained with FSH maintenance data, thus only FSH maintenance results will be illustrated in this section. Because the results obtained are substantial, only a relevant selection of the results will be shown. For instance, most tables will show part of FSH maintenance’s vehicles in order to demonstrate the concepts discussed.

In the table below, the SBM, CCR, BBC scores and the Smoothed MPG columns have been highlighted with a gradient of Red Amber Green (RAG) colouring. The table was also sorted descending on the Smoothed MPG column so that the vehicles demonstrating the best mpg performance are on top and the vehicles demonstrating a worse performance are at the bottom of the list. The list shows the results for both medium and heavy vans. Vehicles 1 to 21 are ‘medium’ vans (all weighing 2185 kg) while vehicles 22 to 69 are ‘heavy’ vans (with their vehicle weight varying from 2861 kg to 3500 kg).

It can be observed in Table 5.6 that DMU 28, 31 and 33 – which all belong to the ‘heavy’ van category (and weighing 2861 kg) – demonstrated a mpg performance as high as 51.72. Although this is higher than the set 44 mpg limit for heavy vans, these vans are relatively small (engine size 1560cc which is the same as some ‘medium’ vans) which could explain their higher performance. These three vans are consequently not discarded as previously explained in the section 5.3.1 Cleansing Algorithm. Similarly, DMU 8 demonstrates a very high mpg of 58. However, fleet managers confirmed the driver driving the vehicle corresponding to DMU 8 was one of their best, thus this DMU was also not discarded.

Only DMU 8 was evaluated efficient by the CCR and SBM models. This was to be expected as DMU 8 is the vehicle which demonstrates the best mpg performance. It is interesting to observe that the SBM and the CCR model provide absolutely identical performance scores. This is due to the fact the CCR model only accounts for technical inefficiencies while SBM accounts for all inefficiencies (both technical and mix inefficiencies). However, because there is only one input and one output there cannot be any mix inefficiencies thus the SBM scores can only reflect technical inefficiencies and – in this one input one output example – are consequently equivalent to the CCR scores.

| DMUName | SBM Score | CCR Score | BCC Score | RTS | MIX | SE | Fuel Used (litres) | Distance Travelled During period (miles) | Smoothed MPG |
|---|---|---|---|---|---|---|---|---|---|
| 8 | 1 | 1 | 1 | CRS | 1 | 1 | 304.351 | 3943.267 | 58.901 |
| 20 | 0.927 | 0.927 | 0.982 | IRS | 1 | 0.944 | 236.523 | 2839.498 | 54.577 |
| 3 | 0.882 | 0.882 | 1 | DRS | 1 | 0.882 | 429.187 | 4905.681 | 51.963 |
| 28 | 0.878 | 0.878 | 0.904 | IRS | 1 | 0.972 | 291.318 | 3314.263 | 51.72 |
| 14 | 0.877 | 0.877 | 0.962 | IRS | 1 | 0.912 | 213.315 | 2425.048 | 51.682 |
| 33 | 0.876 | 0.876 | 0.894 | IRS | 1 | 0.979 | 305.685 | 3468.879 | 51.588 |
| 2 | 0.869 | 0.869 | 0.917 | IRS | 1 | 0.948 | 257.569 | 2901.498 | 51.211 |
| 31 | 0.852 | 0.852 | 0.878 | IRS | 1 | 0.97 | 297.491 | 3284.21 | 50.187 |
| 5 | 0.851 | 0.851 | 0.906 | IRS | 1 | 0.939 | 250.876 | 2766.071 | 50.124 |
| 21 | 0.845 | 0.845 | 0.919 | IRS | 1 | 0.919 | 229.042 | 2507.23 | 49.764 |
| 1 | 0.84 | 0.84 | 0.862 | IRS | 1 | 0.974 | 309.356 | 3367.747 | 49.49 |
| 10 | 0.838 | 0.838 | 0.922 | IRS | 1 | 0.908 | 218.967 | 2377.18 | 49.354 |
| 17 | 0.823 | 0.823 | 0.967 | DRS | 1 | 0.85 | 470.216 | 5011.415 | 48.451 |
| 18 | 0.793 | 0.793 | 0.845 | DRS | 1 | 0.938 | 425.042 | 4366.23 | 46.7 |
| 9 | 0.784 | 0.784 | 0.81 | IRS | 1 | 0.967 | 318.34 | 3233.01 | 46.169 |
| 19 | 0.776 | 0.776 | 0.837 | IRS | 1 | 0.927 | 258.917 | 2604.371 | 45.728 |
| 15 | 0.773 | 0.773 | 0.956 | IRS | 1 | 0.809 | 155.006 | 1553.005 | 45.547 |
| 11 | 0.772 | 0.772 | 0.911 | IRS | 1 | 0.847 | 181.067 | 1810.955 | 45.468 |
| 13 | 0.767 | 0.767 | 0.833 | DRS | 1 | 0.921 | 454.395 | 4516.327 | 45.185 |
| 23 | 0.761 | 0.761 | 0.997 | DRS | 1 | 0.764 | 544.924 | 5374.719 | 44.839 |
| 6 | 0.742 | 0.742 | 0.911 | IRS | 1 | 0.815 | 165.129 | 1588.274 | 43.726 |
| 4 | 0.73 | 0.73 | 1 | IRS | 1 | 0.73 | 122.104 | 1154.916 | 42.999 |
| 42 | 0.709 | 0.709 | 0.738 | IRS | 1 | 0.961 | 339.135 | 3117.124 | 41.785 |
| 16 | 0.686 | 0.686 | 0.709 | DRS | 1 | 0.966 | 467.697 | 4155.121 | 40.388 |
| 26 | 0.671 | 0.671 | 0.71 | IRS | 1 | 0.944 | 327.101 | 2841.89 | 39.497 |
| 12 | 0.666 | 0.666 | 0.742 | IRS | 1 | 0.898 | 262.095 | 2262.818 | 39.249 |
| 30 | 0.629 | 0.629 | 0.632 | DRS | 1 | 0.996 | 486.605 | 3966.302 | 37.055 |
| 29 | 0.621 | 0.621 | 0.99 | IRS | 1 | 0.627 | 123.29 | 992.182 | 36.585 |
| 27 | 0.61 | 0.61 | 0.638 | IRS | 1 | 0.956 | 384.567 | 3040.251 | 35.94 |
| 25 | 0.608 | 0.608 | 1 | DRS | 1 | 0.608 | 832.821 | 6565.123 | 35.837 |
| 38 | 0.608 | 0.608 | 0.8 | DRS | 1 | 0.76 | 684.363 | 5393.127 | 35.826 |
| 24 | 0.59 | 0.59 | 0.628 | IRS | 1 | 0.94 | 363.729 | 2781.251 | 34.762 |
| 32 | 0.584 | 0.584 | 0.77 | DRS | 1 | 0.758 | 714.878 | 5404.969 | 34.372 |
| 57 | 0.583 | 0.583 | 0.664 | IRS | 1 | 0.878 | 273.936 | 2070.355 | 34.358 |
| 43 | 0.574 | 0.574 | 0.756 | DRS | 1 | 0.758 | 726.743 | 5400.928 | 33.785 |
| 41 | 0.572 | 0.572 | 0.614 | IRS | 1 | 0.932 | 360.33 | 2671.507 | 33.705 |
| 36 | 0.572 | 0.572 | 0.628 | IRS | 1 | 0.91 | 323.272 | 2395.183 | 33.683 |
| 40 | 0.571 | 0.571 | 0.606 | IRS | 1 | 0.941 | 379.248 | 2803.591 | 33.607 |
| 47 | 0.569 | 0.569 | 0.613 | IRS | 1 | 0.929 | 356.866 | 2631.856 | 33.527 |
| 22 | 0.569 | 0.569 | 0.603 | IRS | 1 | 0.943 | 384.165 | 2831.464 | 33.507 |
| 66 | 0.563 | 0.563 | 0.581 | IRS | 1 | 0.969 | 447.916 | 3266.705 | 33.155 |
| 37 | 0.563 | 0.563 | 0.694 | IRS | 1 | 0.812 | 214.902 | 1567.138 | 33.152 |
| 68 | 0.563 | 0.563 | 0.589 | DRS | 1 | 0.956 | 580.064 | 4229.928 | 33.151 |
| 49 | 0.559 | 0.559 | 0.564 | IRS | 1 | 0.99 | 512.14 | 3709.077 | 32.924 |
| 46 | 0.559 | 0.559 | 0.587 | IRS | 1 | 0.951 | 407.947 | 2952.951 | 32.907 |
| 69 | 0.557 | 0.557 | 0.573 | IRS | 1 | 0.972 | 459.738 | 3320.496 | 32.835 |
| 58 | 0.553 | 0.553 | 0.577 | IRS | 1 | 0.96 | 431.749 | 3095.99 | 32.599 |
| 51 | 0.547 | 0.547 | 0.577 | IRS | 1 | 0.949 | 411.176 | 2914.674 | 32.226 |
| 61 | 0.544 | 0.544 | 0.556 | DRS | 1 | 0.98 | 576.513 | 4066.571 | 32.067 |
| 34 | 0.539 | 0.539 | 0.542 | IRS | 1 | 0.993 | 541.996 | 3781.575 | 31.719 |
| 63 | 0.529 | 0.529 | 0.586 | DRS | 1 | 0.903 | 683.563 | 4686.315 | 31.167 |
| 44 | 0.528 | 0.528 | 0.59 | IRS | 1 | 0.895 | 326.103 | 2231.239 | 31.105 |
| 39 | 0.527 | 0.527 | 0.592 | DRS | 1 | 0.89 | 706.367 | 4818.786 | 31.013 |
| 7 | 0.523 | 0.523 | 0.636 | IRS | 1 | 0.823 | 242.128 | 1641.824 | 30.826 |
| 35 | 0.521 | 0.521 | 0.54 | DRS | 1 | 0.965 | 616.304 | 4163.77 | 30.714 |
| 56 | 0.519 | 0.519 | 0.535 | IRS | 1 | 0.97 | 487.772 | 3281.102 | 30.58 |
| 59 | 0.506 | 0.506 | 0.547 | DRS | 1 | 0.925 | 682.854 | 4476.99 | 29.806 |
| 55 | 0.505 | 0.505 | 0.515 | IRS | 1 | 0.981 | 535.342 | 3505.82 | 29.771 |
| 53 | 0.501 | 0.501 | 0.659 | DRS | 1 | 0.761 | 829.691 | 5388.759 | 29.526 |
| 62 | 0.494 | 0.494 | 0.498 | DRS | 1 | 0.99 | 625.503 | 4000.187 | 29.073 |
| 65 | 0.494 | 0.494 | 0.522 | IRS | 1 | 0.945 | 447.696 | 2862.814 | 29.07 |
| 48 | 0.477 | 0.477 | 0.479 | IRS | 1 | 0.997 | 624.405 | 3858.563 | 28.093 |

Table 5.6: Fuel and Miles model results

It is also interesting to observe that the SBM and CCR scores are highly correlated to the mpg performance (if the optimisation results were not rounded up they should in fact theoretically be perfectly correlated). This is illustrated in the Figure 5.12.

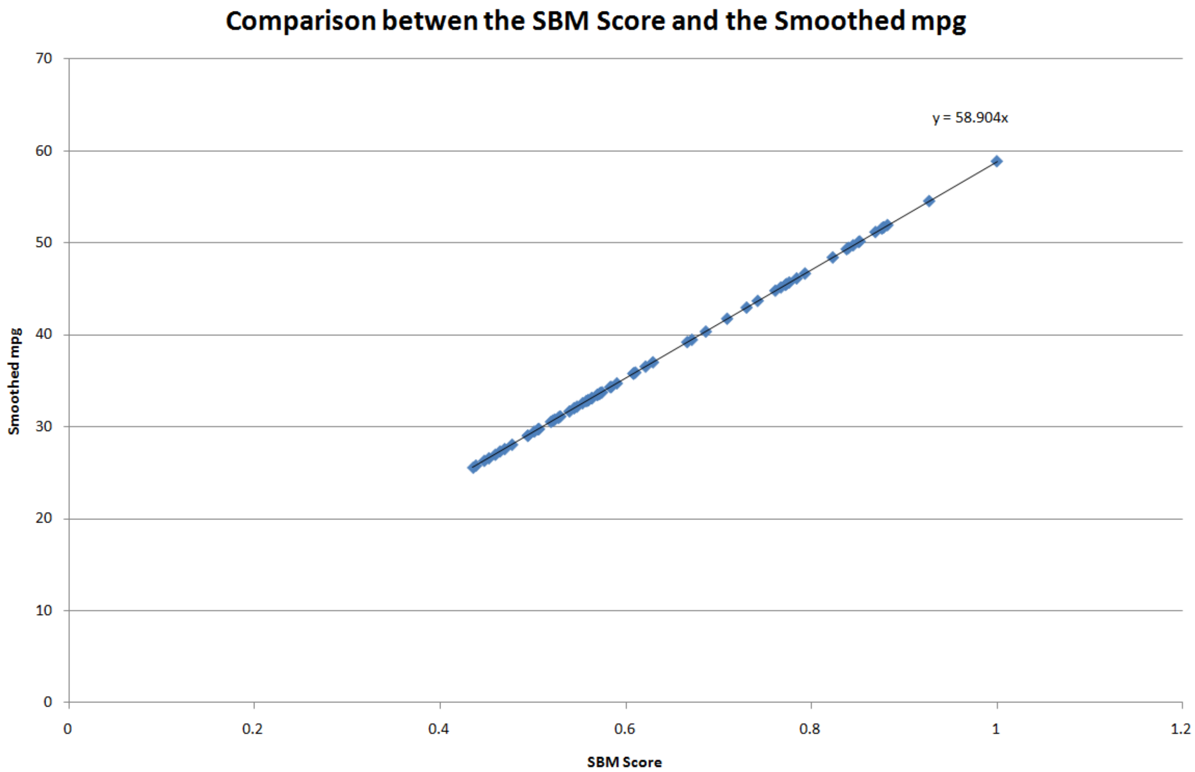


Figure 5.12: Regression between the SBM scores and mpg

Regression analysis gives an R square of 0.999995 (extremely close to 1) and a Pearson coefficient equal to 0.9999979. As mentioned at the end of section 5.4.2 Model with Fuel & Miles Alone, the DEA scores are highly correlated to the mpg figures which can be explained as the DEA model uses the same variables as the mpg measure (i.e. fuel used and distance travelled).

The model results can be illustrated graphically as in Figure 5.13. Because there are only one input and one output, the graph displays the fuel used (input) on the x axis and the miles travelled (output) on the y axis. As explained in section 4.2.2 ‘Single Input, Single Output’, efficient DMUs under constant Return To Scale will display the best output on input ratios. The frontier line passing through the origin and the efficient DMUs will consequently exhibit the highest slope. This is illustrated in the same figure in which the frontier is represented with a blue line.

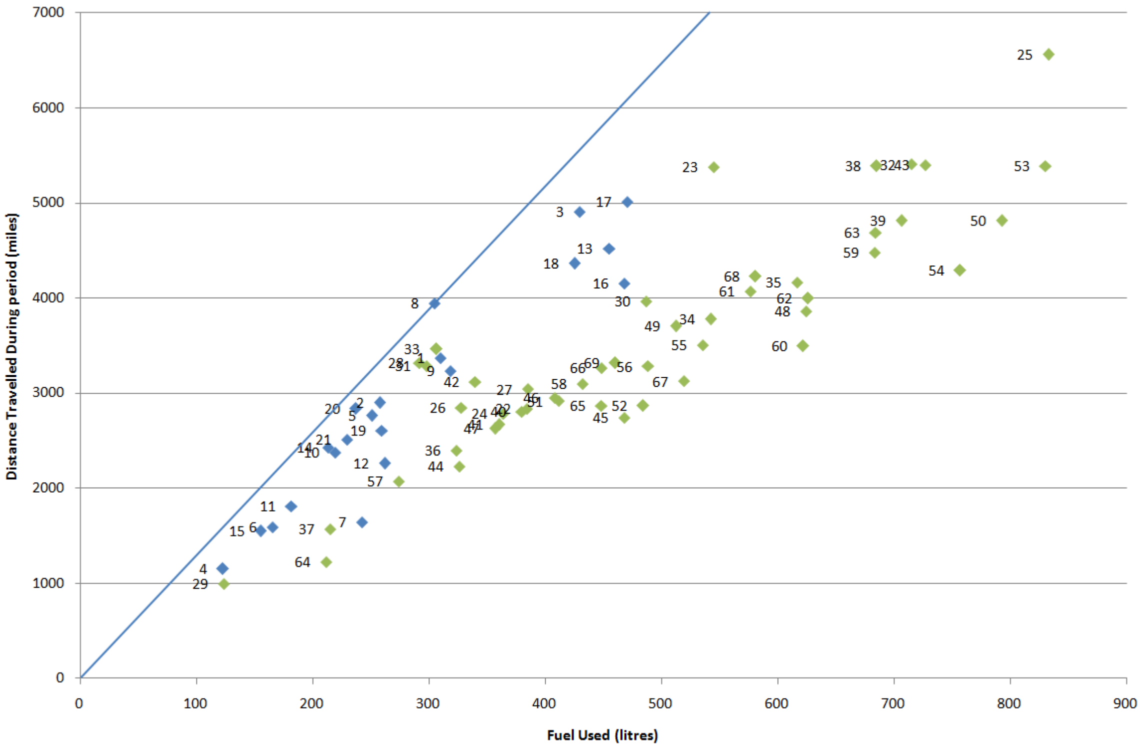


Figure 5.13: Graphical illustration of fuel -> miles data

It is interesting to observe that the 'heavy' vans (displayed in green above) are all inefficient and demonstrate a generally worse performance than 'medium' vans (in blue above). Effectively 'heavy' vans are generally further away from the efficient frontier, resp. On average 217 litres of potential improvements for heavy vans versus 56 litres for medium vans (improvements in fuel used represent the horizontal distance to the frontier). This is illustrated in Table 5.7.

| VehicleType | Average Distance to the Frontier | Average Improvement in Proportion of Fuel Used |
|---|---|---|
| Large Van | 217.72 | 43.08 |
| Medium Van | 56.73 | 20.16 |

Table 5.7: Comparison of average distance to frontier based on vehicle type

This figure shows the absolute 'distance' to the efficiency frontier for medium and large vans, but also the distance percentage in regards to the volume of fuel used.

#### 5.4.2.2. Taking a closer look at Variable Returns To Scale

In order to appraise whether fuel efficiency should be measured under a variable RTS assumption, the data was run against the BCC model. This revealed that efficiency scores calculated by the BCC model differed rather significantly from the CCR and SBM models. In effect, the BCC model identifies several other DMUs as efficient (DMUs 3, 4, 8 and 25). The BCC scores were not proportional to the mpg measure and several DMUs evaluated as efficient illustrated a rather poor or average mpg performance. This is illustrated by DMU 25 for example which the BCC model evaluates as efficient (i.e. with a score of 1 and no slack) while demonstrating a mpg of 35.8. However, this DMU is only evaluated efficient because it is the DMU which travelled the most miles which – under the VRTS assumption – makes this DMU efficient. This is illustrated in Figure 5.14.

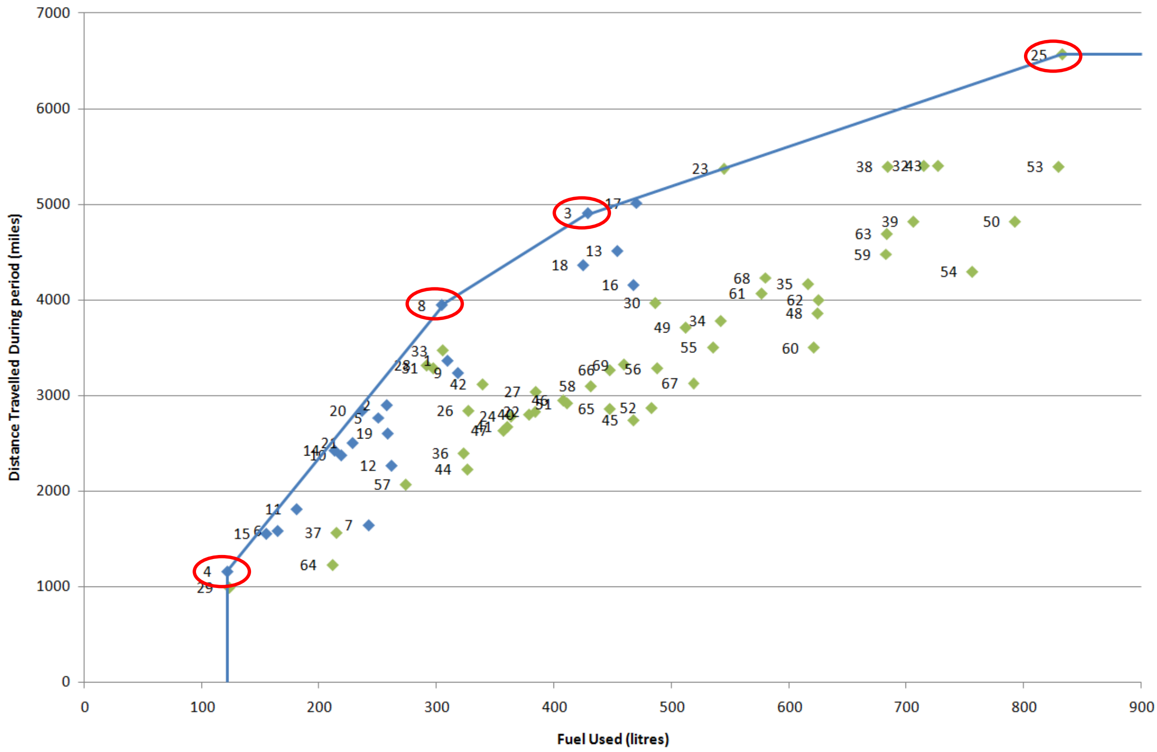


Figure 5.14: Graphical illustration of BCC scores

Evaluating a vehicle as fuel efficient on the sole reason it has travelled more miles than any other vehicle is not appropriate to the notion of fuel efficiency. This was also further confirmed by discussions with the different fleet managers. This consequently implies the Variable RTS assumption is not appropriate to measure vehicle fuel efficiency (or at least not in the range of operations observed above, it might be possible that VRTS could be appropriate for very small operations below as a few tens of miles but the interest in using DEA for such small operations might then be questionable). In plain terms this means that no vehicle should demonstrate a greater or lower fuel efficiency performance in regards to the volume of fuel they have used (or miles they travelled). This suggests that scale efficiency analysis is likely to be inappropriate to fuel efficiency measurement. Consequently, only Constant Returns To Scale approaches will be used to measure van fuel efficiency.

#### 5.4.2.3. Conclusion on the ‘Fuel Used / Miles’ model

The strong correlation between the SBM (or CCR) score with the mpg measure demonstrates it is possible to use DEA to measure fuel efficiency (at least using fuel and distance alone). The results obtained from the BCC model show that the Variable Returns To Scale assumption is not appropriate to the measurement of fuel efficiency. These results were also observed with the other companies’ datasets which confirms the inadequacy of VRTS for fuel efficiency measurement. Consequently only models under a CRTS assumption will be used to measure fuel efficiency.

The next section (5.4.3 Adding the Cost) will evaluate whether the cost has a significant impact on fuel efficiency. In effect, it is important to see first whether cost adds a new dimension to fuel efficiency before adding the weight to the model (as both the ‘fuel used’ and ‘fuel cost’ are radial measures, thus requiring a similar treatment, while weight will be more a categorical variable).

### 5.4.3. Adding the Cost

Fuel costs vary depending on the petrol station brand, the location and also over time. This suggests vehicles could be mpg efficient but pence per mile inefficient (and vice versa). This section will consequently test the impact cost has on fuel efficiency.

As Cooper *et al* (2007, see 'Problem 1.4' p. 19) pointed out, using processed measure such as pence per mile is dangerous as some variables can be over or under evaluated as a result. Using this knowledge, the total amount paid for the fuel used over the period will be used instead of the average ppm over the same period.

A strong correlation was observed between the ‘amount paid’ at petrol stations and the corresponding ‘volume drawn’. A strong correlation was also observed between the smoothed ‘volume used’ and smoothed ‘amount paid’ (i.e. the amount of money paid for the ‘fuel used’). Both these correlations demonstrated an $R^2$ of 0.9999 and a Pearson coefficient of 0.999 (across all vehicle types). This would suggest the fuel cost variability between vehicles is compensated over a period of time (i.e. a vehicle might buy cheaper fuel at some point, but it will buy more expensive fuel later on and the overall average cost will eventually even out amongst the vehicles). However, Dyson *et al* (2001, see Pitfall 4.2 (Correlated factors) p. 249) advise testing the effect of a correlated variable on the DEA model results before discarding it; indeed, in some cases discarding a strongly correlated variable can lead to significant changes in efficiency (this is also true for perfectly correlated variables with an $R^2$ equal to 1). Because it is not advised to discard a strongly correlated variable on the sole ground of correlation, the impact which cost has on fuel efficiency will be tested in this section.

For clarity purposes the previous model (i.e. ‘Fuel used’ -> ‘Miles’) will be referred to as Scenario 1. Conversely, the model studied in this section (i.e. ‘Fuel used’, ‘Fuel cost’ -> ‘Miles’) will be referred to as Scenario 2.

In order to test for the impact of fuel cost, the smoothed cost will be added to the model discussed in the previous section (Scenario 1). ‘Fuel cost’ is an isotonic variable thus it will not require any special treatment (the more money spent on fuel, the more miles the vehicle should travel). The resulting model is illustrated in Figure 5.15.

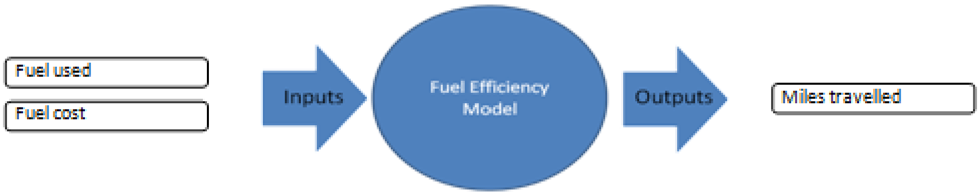


Figure 5.15: Fuel Efficiency model – cost spent on fuel

This model will be tested with CCR and SBM models under the CRTS assumption. First step is to compare the results obtained with the CCR model for Scenario 1 and 2. Then the results obtained with the SBM-C model for Scenario 1 and 2 should in a similar manner be compared against each other. A last section will discuss the differences between CCR and SBM-C results for Scenario 2. This last section will also look at how the 'fuel cost' could potentially be used instead of volume of 'fuel used'.

#### 5.4.3.1. CCR model results comparison between Scenario 1 and 2

The results from the CCR model are illustrated in Table 5.8.

DMU 8 remains the only efficient DMU in this Scenario 2. The score difference is rather small with an average of 0.01013 and a standard deviation equal to 0.0099147. The maximum score difference is observed for DMU 29 with a 5.27% score increase from Scenario 1.

| DMUName | Fuel -> Miles | Fuel, Cost -> Miles | Score difference | Rank Fuel -> Miles | Rank Fuel, Cost -> Miles | Absolute Rank Difference |
|---|---|---|---|---|---|---|
| 1 | 0.840232475 | 0.852178288 | 0.0119458 | 11 | 11 | 0 |
| 2 | 0.869455247 | 0.881923421 | 0.0124682 | 7 | 6 | 1 |
| 3 | 0.882208614 | 0.886968243 | 0.0047596 | 3 | 4 | 1 |
| 4 | 0.730027223 | 0.748739643 | 0.0187124 | 22 | 22 | 0 |
| 5 | 0.850986667 | 0.865416802 | 0.0144301 | 9 | 10 | 1 |
| 6 | 0.742370361 | 0.758728204 | 0.0163578 | 21 | 21 | 0 |
| 7 | 0.523359527 | 0.548223782 | 0.0248643 | 54 | 49 | 5 |
| 8 | 1 | 1 | 0.0000000 | 1 | 1 | 0 |
| 9 | 0.783852567 | 0.797278933 | 0.0134264 | 15 | 15 | 0 |
| 10 | 0.837918826 | 0.888541536 | 0.0506227 | 12 | 3 | 9 |
| 11 | 0.771945965 | 0.771945965 | 0.0000000 | 18 | 20 | 2 |
| 12 | 0.666360795 | 0.679043231 | 0.0126824 | 26 | 25 | 1 |
| 13 | 0.767132503 | 0.776315051 | 0.0091825 | 19 | 18 | 1 |
| 14 | 0.877440129 | 0.88335454 | 0.0059144 | 5 | 5 | 0 |
| 15 | 0.773290889 | 0.779348629 | 0.0060577 | 17 | 17 | 0 |
| 16 | 0.685705511 | 0.691370879 | 0.0056654 | 24 | 24 | 0 |
| 17 | 0.82258638 | 0.823871976 | 0.0012856 | 13 | 13 | 0 |
| 18 | 0.792854164 | 0.793306927 | 0.0004528 | 14 | 16 | 2 |
| 19 | 0.776355856 | 0.803124795 | 0.0267689 | 16 | 14 | 2 |
| 20 | 0.92658814 | 0.941542965 | 0.0149548 | 2 | 2 | 0 |
| 21 | 0.844884995 | 0.844884995 | 0.0000000 | 10 | 12 | 2 |
| 22 | 0.568868384 | 0.581080525 | 0.0122121 | 40 | 40 | 0 |
| 23 | 0.761269405 | 0.772281825 | 0.0110124 | 20 | 19 | 1 |
| 24 | 0.590174995 | 0.599408898 | 0.0092339 | 32 | 32 | 0 |
| 25 | 0.608428779 | 0.620355847 | 0.0119271 | 30 | 29 | 1 |
| 26 | 0.670569723 | 0.670569723 | 0.0000000 | 25 | 27 | 2 |
| 27 | 0.610177198 | 0.61307324 | 0.0028960 | 29 | 30 | 1 |
| 28 | 0.878088314 | 0.880029517 | 0.0019412 | 4 | 7 | 3 |
| 29 | 0.621129334 | 0.673889098 | 0.0527598 | 28 | 26 | 2 |
| 30 | 0.629111698 | 0.637215304 | 0.0081036 | 27 | 28 | 1 |
| 31 | 0.852070709 | 0.879987366 | 0.0279167 | 8 | 8 | 0 |
| 32 | 0.583552352 | 0.595026621 | 0.0114743 | 33 | 33 | 0 |
| 33 | 0.875857739 | 0.877522889 | 0.0016652 | 6 | 9 | 3 |
| 34 | 0.538511752 | 0.547224393 | 0.0087126 | 50 | 50 | 0 |
| 35 | 0.521447145 | 0.533196611 | 0.0117495 | 55 | 53 | 2 |
| 36 | 0.571859262 | 0.587995395 | 0.0161361 | 37 | 36 | 1 |
| 37 | 0.562840491 | 0.590657669 | 0.0278172 | 42 | 35 | 7 |
| 38 | 0.608236775 | 0.608935961 | 0.0006992 | 31 | 31 | 0 |
| 39 | 0.526533232 | 0.5328994 | 0.0063662 | 53 | 55 | 2 |
| 40 | 0.570571279 | 0.584818573 | 0.0142473 | 38 | 39 | 1 |
| 41 | 0.572235035 | 0.57243107 | 0.0001960 | 36 | 42 | 6 |
| 42 | 0.709414425 | 0.726322668 | 0.0169082 | 23 | 23 | 0 |
| 43 | 0.573595954 | 0.58498312 | 0.0113872 | 35 | 38 | 3 |
| 44 | 0.528092314 | 0.539196658 | 0.0111043 | 52 | 52 | 0 |
| 45 | 0.451878821 | 0.455370917 | 0.0034921 | 66 | 67 | 1 |
| 46 | 0.55869019 | 0.565436653 | 0.0067465 | 45 | 45 | 0 |
| 47 | 0.569213903 | 0.585167419 | 0.0159535 | 39 | 37 | 2 |

**Table 5.8: CCR results for the Fuel, Cost -> Miles model**

Taking a closer look at the DMUs showing the greatest score difference (DMU 10 and DMU 29) indicated that these two DMUs have the greatest 'Slack Fuel used' on 'fuel used' ratio (resp. 0.0506227 and 0.052759). Interestingly, these values are also equal to the two DMUs' score difference listed in Table 5.8: CCR results for the Fuel, Cost -> Miles model. This can be explained by the fact these two

DMUs have no 'fuel cost' slack so that the only difference between the CCR score in Scenario 1 and 2 is the 'fuel used' slack. Graphically, these two DMUs also are the furthest away from the regression line than any other DMU. These two DMUs are outlined and circled in red in Figure 5.16.

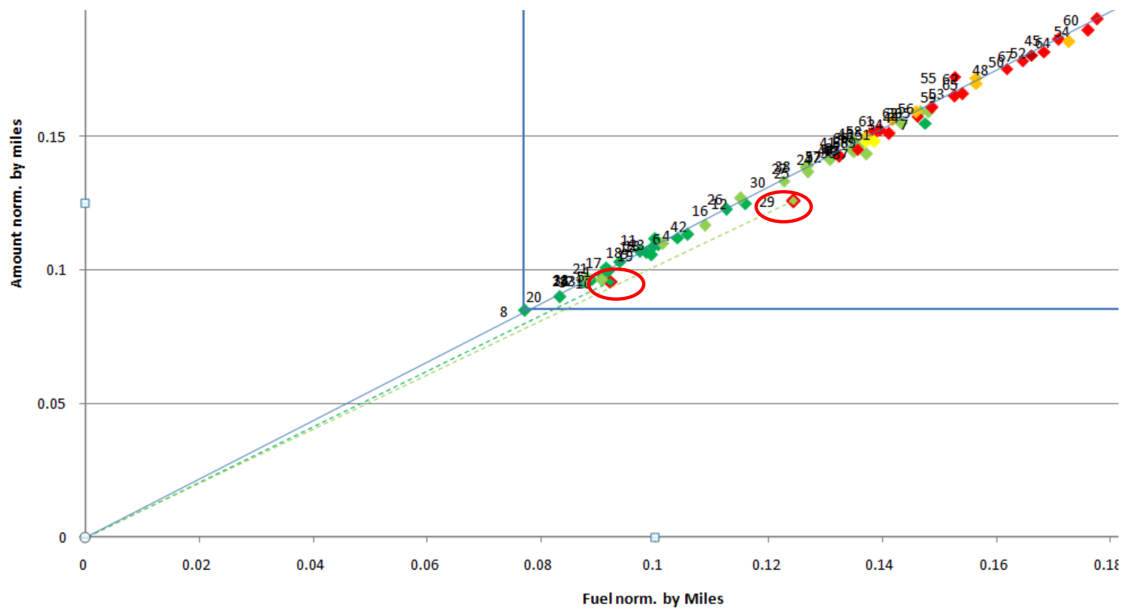


Figure 5.16: Graphical illustration of DMU 10 and 29

The score difference spread is illustrated in Figure 5.17.

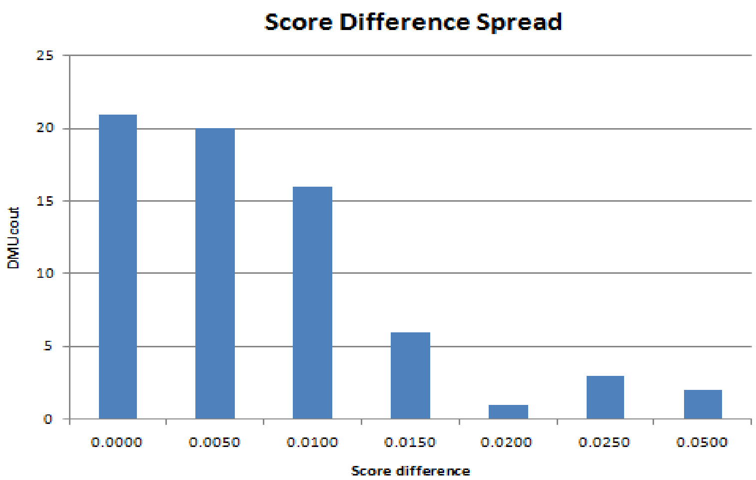


Figure 5.17: Score difference spread when adding cost – CCR

This highlights the generally low impact the cost has on the CCR efficiency scores obtained in Scenario 1. It can be easily observed that – except for some noise around 0.0250 and 0.0500 – the score differences tends to be relatively low which gives a half Poisson shape to this distribution. This is mainly caused by the strong correlation observed earlier (especially the very high Pearson coefficient value). This is illustrated in Figure 5.18.

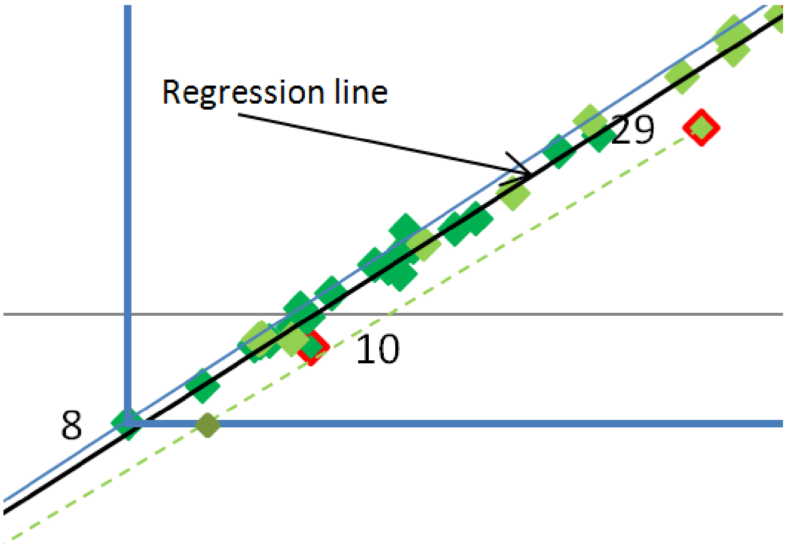


Figure 5.18: Explaining why score difference is low

The graph above is a magnified area of Figure 5.17 to which the regression line has been added (the line in black). The strong correlation, and especially the high Pearson coefficient value (0.999), implies that all points will tend to be relatively close to the black regression line. Due to its specific input and output values, DMU 8 is also very close to the regression line. As a consequence the line starting from the origin and passing through DMU 8 is very close to this regression line. This blue line is not the efficient frontier but a line on which DMUs can only demonstrate technical efficiency and no mix inefficiencies. Thus any DMU on this blue line would demonstrate identical scores between Scenario 1 and 2 (both for the CCR and the SBM models). Subsequently, the greatest score difference would be observed for the DMU the farthest away below this blue line (in this case DMU 29 – and only for fuel used slacks). However, because of this strong Pearson coefficient and of the closeness of the blue line to the regression line, nearly all the DMUs are really close to this blue line. Therefore, the score differences between Scenario 1 and 2 tend to be rather small.

Considering DMU ranking, the maximum difference is observed for DMU 10 with a change in rank of 9 positions. As discussed above, this DMU is demonstrating the greatest score difference as well. DMU 10's rank difference can be explained by the combination of a high score difference in a score range regrouping a consequent number of DMUs. This is however not a serious issue as the rank difference tends to be relatively low (92.75% of DMUs will have experienced a rank change lower or equal to 3) as illustrated in Figure 5.19.

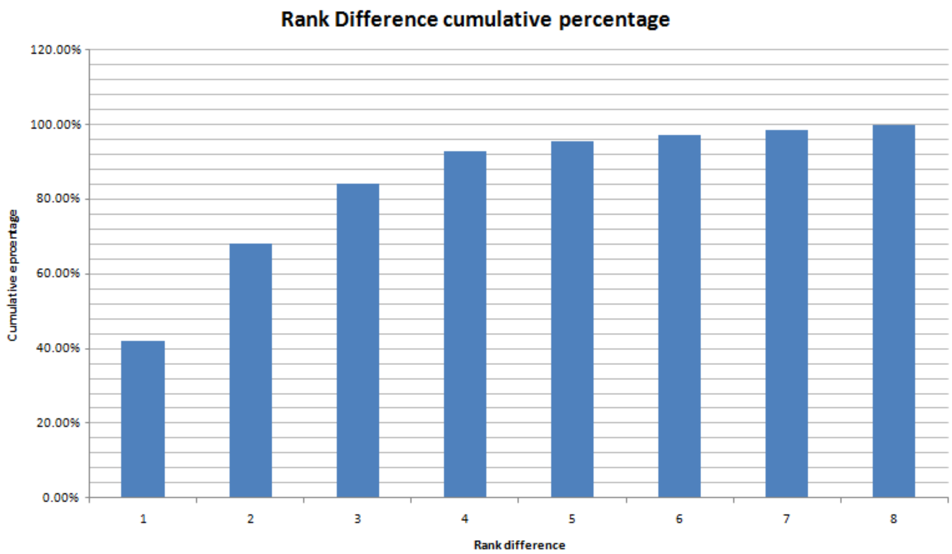


Figure 5.19: Rank difference spread when adding cost – CCR

Furthermore the rank difference generally consists of position swaps between two or three DMUs and there is no DMU that 'jumps' position in a drastic manner. The low rank difference observed in the figure above is a logical consequence of the low score difference observed earlier.

The relatively low score difference associated with a relative slight impact on the DMU rank suggest adding the cost to the ‘Fuel used’ -> ‘Miles’ model does not significantly affect the CCR model results.

##### 5.4.3.2. SBM model results comparison between Scenario 1 and 2

Results from the SBM model are illustrated in Table 5.9.

As with the CCR model, DMU 8 remains the only efficient DMU while DMUs 10 and 29 show the greatest score difference between Scenario 1 and 2 with respectively 0.0253114 for DMU10 and 0.0263799 for DMU29. This coincides with the previous results obtained from the CCR model. Average score difference is 0.0052882 with a standard deviation of 0.0048382.

| DMUName | Fuel -> Miles | Fuel, Cost -> Miles | Score difference | Rank Fuel -> Miles | Rank Fuel, Cost -> Miles | Absolute Rank Difference |
|---|---|---|---|---|---|---|
| 1 | 0.840232475 | 0.846205382 | 0.0059729 | 11 | 11 | 0 |
| 2 | 0.869455247 | 0.875689334 | 0.0062341 | 7 | 7 | 0 |
| 3 | 0.882208614 | 0.884588429 | 0.0023798 | 3 | 3 | 0 |
| 4 | 0.730027223 | 0.739383433 | 0.0093562 | 22 | 22 | 0 |
| 5 | 0.850986667 | 0.858201734 | 0.0072151 | 9 | 10 | 1 |
| 6 | 0.742370361 | 0.750549282 | 0.0081789 | 21 | 21 | 0 |
| 7 | 0.523359527 | 0.535791654 | 0.0124321 | 54 | 51 | 3 |
| 8 | 1 | 1 | 0.0000000 | 1 | 1 | 0 |
| 9 | 0.783852567 | 0.79056575 | 0.0067132 | 15 | 15 | 0 |
| 10 | 0.837918826 | 0.863230181 | 0.0253114 | 12 | 9 | 3 |
| 11 | 0.771945965 | 0.766071738 | 0.0058742 | 18 | 20 | 2 |
| 12 | 0.666360795 | 0.672702013 | 0.0063412 | 26 | 25 | 1 |
| 13 | 0.767132503 | 0.771723777 | 0.0045913 | 19 | 18 | 1 |
| 14 | 0.877440129 | 0.880397334 | 0.0029572 | 5 | 4 | 1 |
| 15 | 0.773290889 | 0.776319759 | 0.0030289 | 17 | 17 | 0 |
| 16 | 0.685705511 | 0.688538195 | 0.0028327 | 24 | 24 | 0 |
| 17 | 0.82258638 | 0.823229178 | 0.0006428 | 13 | 13 | 0 |
| 18 | 0.792854164 | 0.793080546 | 0.0002264 | 14 | 14 | 0 |
| 19 | 0.776355856 | 0.789740326 | 0.0133845 | 16 | 16 | 0 |
| 20 | 0.92658814 | 0.934065552 | 0.0074774 | 2 | 2 | 0 |
| 21 | 0.844884995 | 0.843182649 | 0.0017023 | 10 | 12 | 2 |
| 22 | 0.568868384 | 0.574974454 | 0.0061061 | 40 | 40 | 0 |
| 23 | 0.761269405 | 0.766775615 | 0.0055062 | 20 | 19 | 1 |
| 24 | 0.590174995 | 0.594791946 | 0.0046170 | 32 | 32 | 0 |
| 25 | 0.608428779 | 0.614392313 | 0.0059635 | 30 | 29 | 1 |
| 26 | 0.670569723 | 0.669505492 | 0.0010642 | 25 | 26 | 1 |
| 27 | 0.610177198 | 0.611625219 | 0.0014480 | 29 | 30 | 1 |
| 28 | 0.878088314 | 0.879058916 | 0.0009706 | 4 | 5 | 1 |
| 29 | 0.621129334 | 0.647509216 | 0.0263799 | 28 | 27 | 1 |
| 30 | 0.629111698 | 0.633163501 | 0.0040518 | 27 | 28 | 1 |
| 31 | 0.852070709 | 0.866029038 | 0.0139583 | 8 | 8 | 0 |
| 32 | 0.583552352 | 0.589289487 | 0.0057371 | 33 | 33 | 0 |
| 33 | 0.875857739 | 0.876690314 | 0.0008326 | 6 | 6 | 0 |
| 34 | 0.538511752 | 0.542868072 | 0.0043563 | 50 | 50 | 0 |
| 35 | 0.521447145 | 0.527321878 | 0.0058747 | 55 | 55 | 0 |
| 36 | 0.571859262 | 0.579927328 | 0.0080681 | 37 | 35 | 2 |
| 37 | 0.562840491 | 0.57674908 | 0.0139086 | 42 | 39 | 3 |
| 38 | 0.608236775 | 0.608586368 | 0.0003496 | 31 | 31 | 0 |
| 39 | 0.526533232 | 0.529716316 | 0.0031831 | 53 | 54 | 1 |
| 40 | 0.570571279 | 0.577694926 | 0.0071236 | 38 | 37 | 1 |
| 41 | 0.572235035 | 0.572333052 | 0.0000980 | 36 | 41 | 5 |

Table 5.9: SBM results for the Fuel, Cost -> Miles model

The score difference obtained with SBM is half that of the CCR model (for the DMUs having some non-zero slacks on ‘fuel used’). In this case it is because the SBM – I objective function divides the input slacks by the number of inputs (in this case 2, see Formula 5.6).

$$SBM - I \ \ \rho_I^* = min_{\lambda,s^-} \quad 1 - \frac{1}{m}\sum_{i=1}^{m} s_i^- / x_{io}$$

Formula 5.6: The SBM – I objective function

Because of this model behaviour, the score difference spread is much lower than with the CCR model (the reason for this low score difference is similar to the explanation given for the CCR model). This is illustrated in Figure 5.20: Score difference spread when adding cost – SBM.

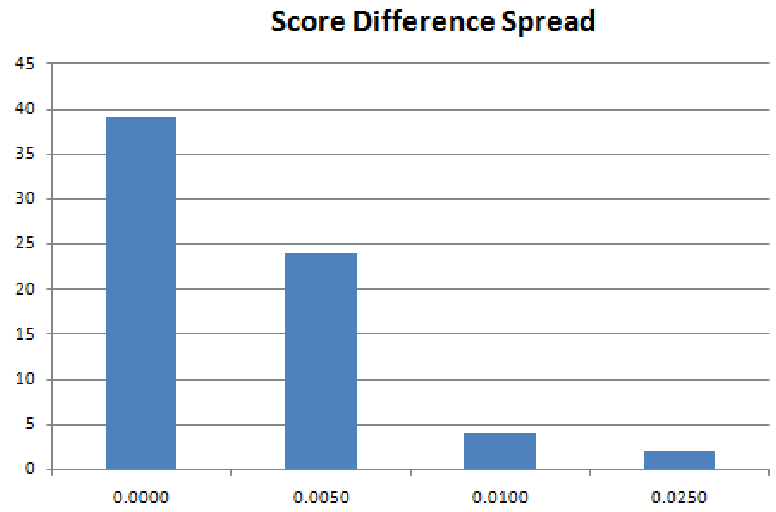


Figure 5.20: Score difference spread when adding cost – SBM

Conversely, rank differences are relatively low with a maximum of 5 while 98.55% of the DMUs have a rank difference lower or equal to 3. This is illustrated in Figure 5.21.

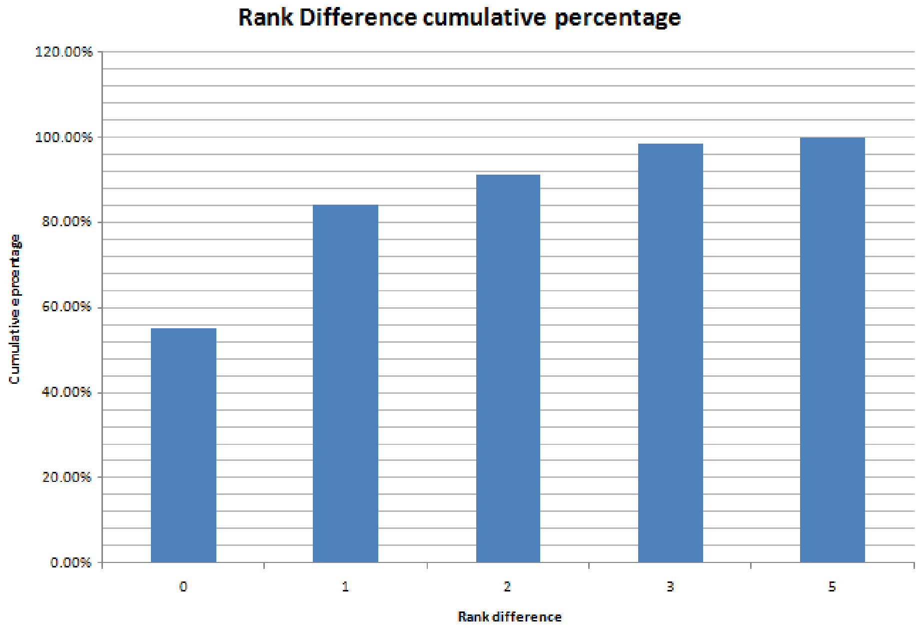


Figure 5.21: Rank difference spread when adding cost – SBM

As with the CCR model, adding the cost to the Scenario 1 model does not significantly affect the results of the Slack Based Model. This is again likely caused by the high correlation between fuel cost and fuel volume as explained earlier. These preliminary results suggest it is not relevant to simultaneously use the 'volume of fuel used' and the 'fuel cost' as inputs and that instead 'volume fuel used' could be used as the sole input.

#### 5.4.3.3. Further analysis

Although the previous section suggested that adding the 'fuel cost' variable to the 'Fuel used' -> 'Miles' model did not have any significant impact on the results obtained from either the CCR or the SBM model, further analysis should be conducted to better understand the relation between cost and fuel efficiency.

This analysis should evaluate the differences between the CCR and the SBM results in Scenario 2 and explore the relation between mpg and the 'Fuel cost' -> 'Miles' model. For convenience, this last model will be referred to as Scenario 3.

**Comparison between the CCR and SBM model results for Scenario 2**

Comparison of results obtained from the CCR and SBM models in Scenario 2 are summarised in Table 5.10. As demonstrated by Tone (2001, p. 502 Theorem 2.), the CCR scores are all greater or equal to the SBM scores. This is because the SBM model accounts for all inefficiencies whereas the CCR model only accounts for purely technical inefficiencies (This obviously has had an impact on the slacks found by each model) (Cooper et al., 2007, p. 103 Theorem 4.8). The score differences between CCR and SBM in Scenario 2 are logically equal to the score differences between SBM Scenario 1 and SBM Scenario 2. This is because the score differences between the CCR and SBM models in Scenario 2 and between SBM Scenario 1 and SBM Scenario 2 are the mix inefficiencies created when adding the cost to the fuel efficiency model.

Conversely the rank difference between CCR and SBM in Scenario 2 is rather small which is mainly due to the low score difference observed earlier. There is consequently no significant difference between CCR and SBM in Scenario 3 except the CCR model provides higher scores than SBM.

| DMUName | Score CCR | Score SBM | Score difference | Rank CCR | Rank SBM | Rank difference |
|---|---|---|---|---|---|---|
| 1 | 0.852178288 | 0.846205382 | 0.005972907 | 11 | 11 | 0 |
| 2 | 0.881923421 | 0.875689334 | 0.006234087 | 6 | 7 | 1 |
| 3 | 0.886968243 | 0.884588429 | 0.002379815 | 4 | 3 | 1 |
| 4 | 0.748739643 | 0.739383433 | 0.00935621 | 19 | 19 | 0 |
| 5 | 0.865416802 | 0.858201734 | 0.007215068 | 8 | 8 | 0 |
| 6 | 0.758728204 | 0.750549282 | 0.008178921 | 17 | 17 | 0 |
| 7 | 0.548223782 | 0.535791654 | 0.012432128 | 43 | 45 | 2 |
| 8 | 1 | 1 | 0 | 1 | 1 | 0 |
| 9 | 0.797278933 | 0.79056575 | 0.006713183 | 10 | 10 | 0 |
| 10 | 0.888541536 | 0.863230181 | 0.025311355 | 2 | 6 | 4 |
| 11 | 0.771945965 | 0.766071738 | 0.005874227 | 13 | 13 | 0 |
| 12 | 0.679043231 | 0.672702013 | 0.006341218 | 15 | 15 | 0 |
| 13 | 0.776315051 | 0.771723777 | 0.004591274 | 11 | 11 | 0 |
| 14 | 0.88335454 | 0.880397334 | 0.002957206 | 2 | 2 | 0 |
| 15 | 0.779348629 | 0.776319759 | 0.00302887 | 9 | 9 | 0 |
| 16 | 0.691370879 | 0.688538195 | 0.002832684 | 11 | 11 | 0 |
| 17 | 0.823871976 | 0.823229178 | 0.000642798 | 6 | 6 | 0 |
| 18 | 0.793306927 | 0.793080546 | 0.000226381 | 7 | 6 | 1 |
| 19 | 0.803124795 | 0.789740326 | 0.01338447 | 6 | 6 | 0 |
| 20 | 0.941542965 | 0.934065552 | 0.007477412 | 1 | 1 | 0 |
| 21 | 0.844884995 | 0.843182649 | 0.001702346 | 4 | 4 | 0 |
| 22 | 0.581080525 | 0.574974454 | 0.00610607 | 20 | 20 | 0 |
| 23 | 0.772281825 | 0.766775615 | 0.00550621 | 4 | 4 | 0 |
| 24 | 0.599408898 | 0.594791946 | 0.004616952 | 11 | 11 | 0 |
| 25 | 0.620355847 | 0.614392313 | 0.005963534 | 8 | 8 | 0 |
| 26 | 0.670569723 | 0.669505492 | 0.001064232 | 6 | 5 | 1 |
| 27 | 0.61307324 | 0.611625219 | 0.001448021 | 7 | 7 | 0 |
| 28 | 0.880029517 | 0.879058916 | 0.000970601 | 1 | 1 | 0 |
| 29 | 0.673889098 | 0.647509216 | 0.026379882 | 4 | 4 | 0 |
| 30 | 0.637215304 | 0.633163501 | 0.004051803 | 4 | 4 | 0 |
| 31 | 0.879987366 | 0.866029038 | 0.013958328 | 1 | 2 | 1 |
| 32 | 0.595026621 | 0.589289487 | 0.005737134 | 4 | 4 | 0 |
| 33 | 0.877522889 | 0.876690314 | 0.000832575 | 1 | 1 | 0 |
| 34 | 0.547224393 | 0.542868072 | 0.00435632 | 17 | 18 | 1 |
| 35 | 0.533196611 | 0.527321878 | 0.005874733 | 19 | 21 | 2 |
| 36 | 0.587995395 | 0.579927328 | 0.008068066 | 5 | 4 | 1 |
| 37 | 0.590657669 | 0.57674908 | 0.013908589 | 4 | 7 | 3 |
| 38 | 0.608935961 | 0.608586368 | 0.000349593 | 2 | 2 | 0 |
| 39 | 0.5328994 | 0.529716316 | 0.003183084 | 17 | 17 | 0 |
| 40 | 0.584818573 | 0.577694926 | 0.007123647 | 5 | 4 | 1 |
| 41 | 0.57243107 | 0.572333052 | 9.80177E-05 | 6 | 5 | 1 |

Table 5.10: SBM – CCR model results comparison for the Fuel, Cost -> Miles model

**Using fuel cost as sole input**

As explained earlier, the model ‘fuel cost’ -> ‘Miles’ will be referred to as Scenario 3.

As expected, the CCR and SBM models provided identical scores when ‘fuel cost’ was used as sole input (this is because there cannot be any slack so that all inefficiencies are thus purely technical). The score differences between Scenario 1 and Scenario 3 were comparable to those observed between Scenario 1 and 2 for the CCR model. This suggests ‘fuel cost’ can be used instead of ‘fuel used’ and that the results produced should be similar. Table 5.11 illustrates the differences between the CCR score for Scenario 1 and 3.

| DMUName | CCR Scenario 1 | CCR Scenario 3 | Score difference | Rank Scenario 1 | Rank Scenario 3 | Absolute Rank Difference |
|---|---|---|---|---|---|---|
| 1 | 0.840232475 | 0.852178288 | 0.0119458 | 11 | 11 | 0 |
| 2 | 0.869455247 | 0.881923421 | 0.0124682 | 7 | 6 | 1 |
| 3 | 0.882208614 | 0.886968243 | 0.0047596 | 3 | 4 | 1 |
| 4 | 0.730027223 | 0.748739643 | 0.0187124 | 22 | 22 | 0 |
| 5 | 0.850986667 | 0.865416802 | 0.0144301 | 9 | 10 | 1 |
| 6 | 0.742370361 | 0.758728204 | 0.0163578 | 21 | 21 | 0 |
| 7 | 0.523359527 | 0.548223782 | 0.0248643 | 54 | 49 | 5 |
| 8 | 1 | 1 | 0.0000000 | 1 | 1 | 0 |
| 9 | 0.783852567 | 0.797278933 | 0.0134264 | 15 | 15 | 0 |
| 10 | 0.837918826 | 0.888541536 | 0.0506227 | 12 | 3 | 9 |
| 11 | 0.771945965 | 0.760197511 | 0.0117485 | 18 | 20 | 2 |
| 12 | 0.666360795 | 0.679043231 | 0.0126824 | 26 | 25 | 1 |
| 13 | 0.767132503 | 0.776315051 | 0.0091825 | 19 | 18 | 1 |
| 14 | 0.877440129 | 0.88335454 | 0.0059144 | 5 | 5 | 0 |
| 15 | 0.773290889 | 0.779348629 | 0.0060577 | 17 | 17 | 0 |
| 16 | 0.685705511 | 0.691370879 | 0.0056654 | 24 | 24 | 0 |
| 17 | 0.82258638 | 0.823871976 | 0.0012856 | 13 | 13 | 0 |
| 18 | 0.792854164 | 0.793306927 | 0.0004528 | 14 | 16 | 2 |
| 19 | 0.776355856 | 0.803124795 | 0.0267689 | 16 | 14 | 2 |
| 20 | 0.92658814 | 0.941542965 | 0.0149548 | 2 | 2 | 0 |
| 21 | 0.844884995 | 0.841480302 | 0.0034047 | 10 | 12 | 2 |
| 22 | 0.568868384 | 0.581080525 | 0.0122121 | 40 | 40 | 0 |
| 23 | 0.761269405 | 0.772281825 | 0.0110124 | 20 | 19 | 1 |
| 24 | 0.590174995 | 0.599408898 | 0.0092339 | 32 | 32 | 0 |
| 25 | 0.608428779 | 0.620355847 | 0.0119271 | 30 | 29 | 1 |
| 26 | 0.670569723 | 0.66844126 | 0.0021285 | 25 | 27 | 2 |
| 27 | 0.610177198 | 0.61307324 | 0.0028960 | 29 | 30 | 1 |
| 28 | 0.878088314 | 0.880029517 | 0.0019412 | 4 | 7 | 3 |
| 29 | 0.621129334 | 0.673889098 | 0.0527598 | 28 | 26 | 2 |
| 30 | 0.629111698 | 0.637215304 | 0.0081036 | 27 | 28 | 1 |
| 31 | 0.852070709 | 0.879987366 | 0.0279167 | 8 | 8 | 0 |
| 32 | 0.583552352 | 0.595026621 | 0.0114743 | 33 | 33 | 0 |
| 33 | 0.875857739 | 0.877522889 | 0.0016652 | 6 | 9 | 3 |
| 34 | 0.538511752 | 0.547224393 | 0.0087126 | 50 | 50 | 0 |
| 35 | 0.521447145 | 0.533196611 | 0.0117495 | 55 | 53 | 2 |
| 36 | 0.571859262 | 0.587995395 | 0.0161361 | 37 | 36 | 1 |
| 37 | 0.562840491 | 0.590657669 | 0.0278172 | 42 | 35 | 7 |
| 38 | 0.608236775 | 0.608935961 | 0.0006992 | 31 | 31 | 0 |
| 39 | 0.526533232 | 0.5328994 | 0.0063662 | 53 | 55 | 2 |
| 40 | 0.570571279 | 0.584818573 | 0.0142473 | 38 | 39 | 1 |
| 41 | 0.572235035 | 0.57243107 | 0.0001960 | 36 | 42 | 6 |
| 42 | 0.709414425 | 0.726322668 | 0.0169082 | 23 | 23 | 0 |
| 43 | 0.573595954 | 0.58498312 | 0.0113872 | 35 | 38 | 3 |
| 44 | 0.528092314 | 0.539196658 | 0.0111043 | 52 | 52 | 0 |
| 45 | 0.451878821 | 0.455370917 | 0.0034921 | 66 | 67 | 1 |
| 46 | 0.55869019 | 0.565436653 | 0.0067465 | 45 | 45 | 0 |
| 47 | 0.569213903 | 0.585167419 | 0.0159535 | 39 | 37 | 2 |

Table 5.11: CCR model results comparison for Scenario 1 and 3

The scores were again strongly correlated to mpg ($R^2$ = 0.9949, Pearson = 0.997). However, there were a few occasions where the score rank shows discrepancies with the mpg rank. This is illustrated in Table 5.12 where the data is sorted by mpg descending and inhomogeneous scores are displayed in red.

| Score Scenario | mpg |
|---|---|
| 1 | 58.9006 |
| 0.941542965 | 54.5766 |
| 0.886968243 | 51.9626 |
| 0.880029517 | 51.7199 |
| 0.88335454 | 51.6817 |
| 0.877522889 | 51.5885 |
| 0.881923421 | 51.2114 |
| 0.879987366 | 50.1875 |
| 0.865416802 | 50.1236 |
| 0.841480302 | 49.7642 |
| 0.852178288 | 49.4902 |
| 0.888541536 | 49.3539 |
| 0.823871976 | 48.4508 |
| 0.793306927 | 46.6996 |

Table 5.12 Ranking discrepancies between score and mpg in Scenario 3

Such discrepancies were however logically not observed with ‘fuel used’.

#### 5.4.3.4. Conclusion on fuel cost

This section demonstrated adding the ‘fuel cost’ to the ‘fuel used’ -> ‘miles’ model did not have any significant impact so it was not relevant to simultaneously use ‘fuel cost’ and ‘fuel used’ in the same model. It was also demonstrated that although scores are relatively similar regardless of whether ‘fuel used’ or ‘fuel cost’ were used as input, of whether ‘fuel used’ tended to provide more consistent results in regards to mpg. Consequently, only ‘fuel used’ will be retained in further models.

Because the volume of ‘fuel used’ and the corresponding ‘fuel cost’ are radial measures, using the CCR model results in this specific instance would seem more appropriate than non radial models such as the Slack Based Model (although the SBM model would shows greater discrimination amongst DMUs as incorporating mix inefficiencies in the measurement of efficiency). However, the SBM might be more appropriate when incorporating variables such as vehicle weight or age as these variables are unlikely to behave in a radial manner. The incorporation of vehicle weight to the fuel efficiency model will be discussed in the following section.

### 5.4.4. Adding the Weight

This section will investigate how the vehicle gross weight can be incorporated into the fuel efficiency model. This new fuel efficiency model is illustrated in Figure 5.22.

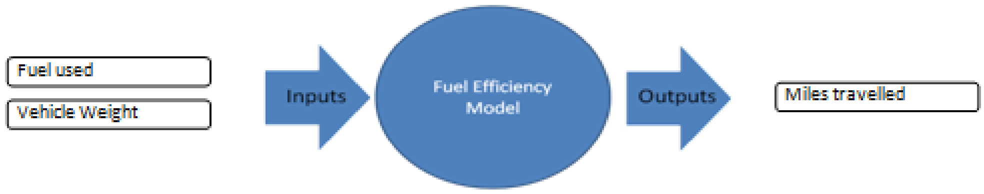


Figure 5.22: Fuel Efficiency model – vehicle weight

Vehicle weight is well known for having a strong impact on fuel efficiency. It is possible to illustrate the impact of weight is to plot the different frontiers per van category. The figure below illustrates the different frontiers for each van weight category starting with green (medium van 2185kg) and finishing with burgundy (heavy van 3,500kg). The colour code is illustrated in Table 5.13.

| Weight Category | Colour Name | Associated colour |
|---|---|---|
| 2185 kg | Green | |
| 2861 kg | Light Green | |
| 2900 kg | Yellow | |
| 3000 kg | Orange | |
| 3300 kg | Red | |
| 3500 kg | Burgundy | |

Table 5.13: Van gross weight category colouring

And the different frontiers for each van weight category are illustrated in Figure 5.23.

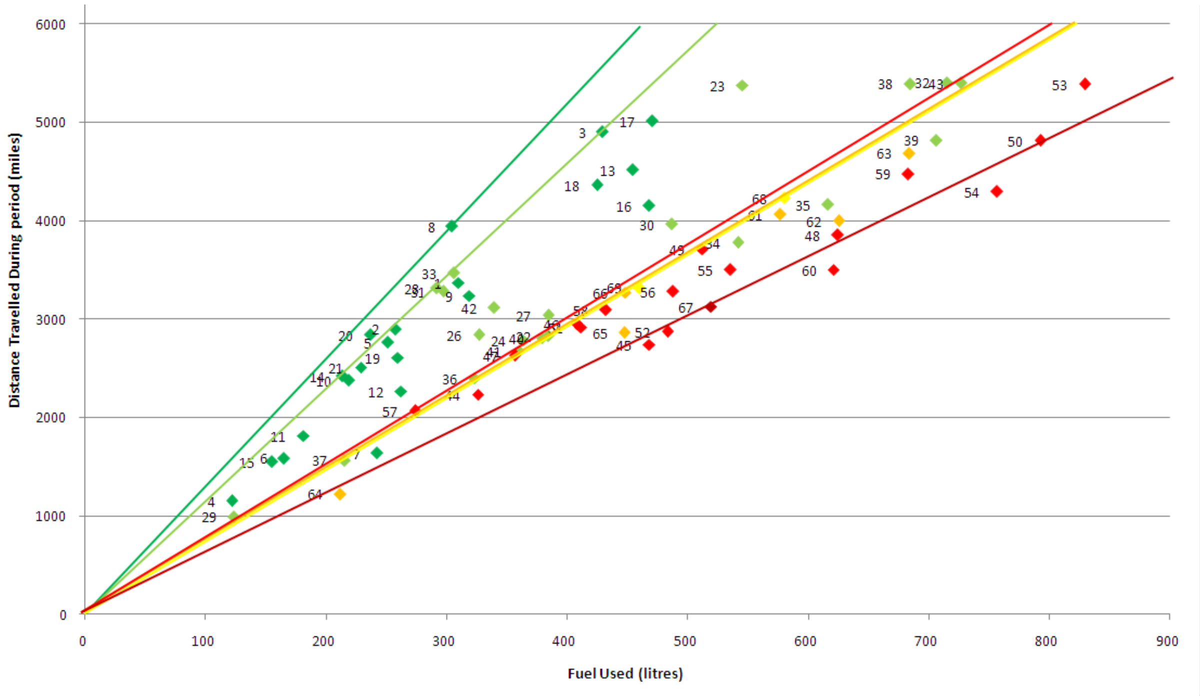


Figure 5.23: Frontier for each weight category

It is interesting to observe that the yellow (2,900 kg) and orange (3,000 kg) frontiers are nearly superimposed (they are on the figure above although their equations differ slightly) and very close to the red (3,300 kg) frontier. In effect, the maximum mpg difference between the best DMUs in the yellow category and the best DMU in the red category is 1.207 which explains why the three frontiers look really close to each other. Not surprisingly, the frontiers for the medium vans show the best mpg (green lines) while the heaviest van frontier shows a worse mpg performance (burgundy frontier). This further illustrates that it is rather unfair to compare van mpg across different weight categories (unless the weight difference is really small) and prompts the need to incorporate vehicle weight in the model especially for fleets having vehicles in several weight categories.

This effect of weight was also blatant when ordering a list of vehicles by their respective weight and mpg earlier. This is illustrated again in Table 5.14.

| 1 | DMUName | Veh Type | Volume | Weight | Miles | MPG |
|---|---|---|---|---|---|---|
| 2 | 8 | Medium Van | 304.35 | 2185 | 3943.3 | 58.9012785 |
| 3 | 20 | Medium Van | 236.52 | 2185 | 2839.5 | 54.5773207 |
| 4 | 3 | Medium Van | 429.19 | 2185 | 4905.7 | 51.9624477 |
| 5 | 14 | Medium Van | 213.32 | 2185 | 2425 | 51.6795088 |
| 6 | 2 | Medium Van | 257.57 | 2185 | 2901.5 | 51.2112652 |
| 7 | 5 | Medium Van | 250.88 | 2185 | 2766.1 | 50.1233448 |
| 8 | 21 | Medium Van | 229.04 | 2185 | 2507.2 | 49.7640655 |
| 9 | 1 | Medium Van | 309.36 | 2185 | 3367.7 | 49.4888597 |
| 10 | 10 | Medium Van | 218.97 | 2185 | 2377.2 | 49.353654 |
| 11 | 17 | Medium Van | 470.22 | 2185 | 5011.4 | 48.4502676 |
| 12 | 18 | Medium Van | 425.04 | 2185 | 4366.2 | 46.6994786 |
| 13 | 9 | Medium Van | 318.34 | 2185 | 3233 | 46.1692374 |
| 14 | 19 | Medium Van | 258.92 | 2185 | 2604.4 | 45.7277989 |
| 15 | 15 | Medium Van | 155.01 | 2185 | 1553 | 45.5459692 |
| 20 | 16 | Medium Van | 467.7 | 2185 | 4155.1 | 40.3879974 |
| 21 | 12 | Medium Van | 262.1 | 2185 | 2262.8 | 39.2479844 |
| 22 | 7 | Medium Van | 242.13 | 2185 | 1641.8 | 30.8254807 |
| 23 | 28 | Large Van | 291.32 | 2861 | 3314.3 | 51.7201439 |
| 24 | 33 | Large Van | 305.68 | 2861 | 3468.9 | 51.5896955 |
| 25 | 31 | Large Van | 297.49 | 2861 | 3284.2 | 50.1874851 |
| 26 | 23 | Large Van | 544.92 | 2861 | 5374.7 | 44.8393893 |
| 27 | 42 | Large Van | 339.14 | 2861 | 3117.1 | 41.7839919 |
| 28 | 26 | Large Van | 327.1 | 2861 | 2841.9 | 39.4972134 |
| 29 | 30 | Large Van | 486.6 | 2861 | 3966.3 | 37.0554135 |
| 30 | 29 | Large Van | 123.29 | 2861 | 992.2 | 36.5855492 |
| 42 | 34 | Large Van | 542 | 2861 | 3781.6 | 31.7186366 |
| 43 | 39 | Large Van | 706.37 | 2861 | 4818.8 | 31.0130775 |
| 44 | 35 | Large Van | 616.3 | 2861 | 4163.8 | 30.7139662 |
| 45 | 68 | Large Van | 580.06 | 2900 | 4229.9 | 33.1509052 |
| 46 | 69 | Large Van | 459.74 | 2900 | 3320.5 | 32.834424 |
| 47 | 66 | Large Van | 447.92 | 3000 | 3266.7 | 33.1548454 |
| 48 | 61 | Large Van | 576.51 | 3000 | 4066.6 | 32.0673314 |
| 49 | 63 | Large Van | 683.56 | 3000 | 4686.3 | 31.1667599 |
| 50 | 62 | Large Van | 625.5 | 3000 | 4000.2 | 29.0731842 |
| 51 | 65 | Large Van | 447.7 | 3000 | 2862.8 | 29.0698053 |
| 52 | 64 | Large Van | 211.39 | 3000 | 1224.6 | 26.3358916 |
| 53 | 57 | Large Van | 273.94 | 3300 | 2070.4 | 34.3587232 |
| 56 | 46 | Large Van | 407.95 | 3300 | 2953 | 32.907487 |
| 57 | 58 | Large Van | 431.75 | 3300 | 3096 | 32.5991904 |
| 58 | 51 | Large Van | 411.18 | 3300 | 2914.7 | 32.2255314 |
| 59 | 44 | Large Van | 326.1 | 3300 | 2231.2 | 31.1046924 |
| 60 | 56 | Large Van | 487.77 | 3300 | 3281.1 | 30.5803597 |
| 63 | 53 | Large Van | 829.69 | 3300 | 5388.8 | 29.5266665 |
| 64 | 48 | Large Van | 624.4 | 3300 | 3858.6 | 28.0934499 |
| 65 | 50 | Large Van | 792.75 | 3300 | 4827.2 | 27.6819864 |
| 66 | 52 | Large Van | 483.45 | 3300 | 2872 | 27.0066726 |
| 67 | 45 | Large Van | 467.76 | 3300 | 2738.6 | 26.6160579 |
| 68 | 54 | Large Van | 756.11 | 3300 | 4294.4 | 25.8199693 |
| 69 | 60 | Large Van | 621.2 | 3300 | 3498.3 | 25.6014057 |
| 70 | 67 | Large Van | 519.09 | 3500 | 3123.9 | 27.3585244 |

Table 5.14: Impact of weight on the traditional mpg measure

Although the impact of 'vehicle weight' is blatant, it is unclear however, whether the relation between 'vehicle weight' and mpg is linear (which is different from piece-wise linear). Indeed, external factors such as driver behaviour, vehicle aerodynamics or other factors impacting on fuel efficiency might impact on this relation. Regression analysis gives a $R^2$ of 0.6283 and a Pearson coefficient of -0.7926 which suggests the relation is not really (or perfectly) linear. This is illustrated in Figure 5.24.

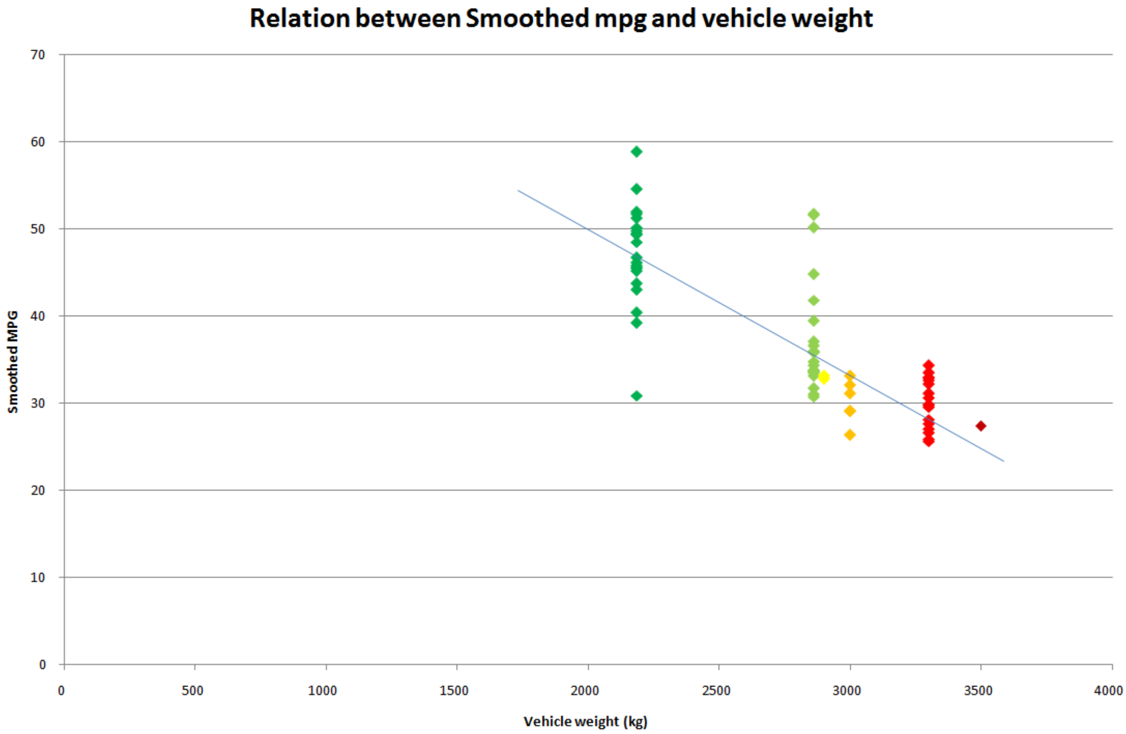


Figure 5.24: Relation between Smoothed mpg and vehicle weight

The linear regression analysis based on averages for each vehicle weight category gives better results with a $R^2$ of 0.9089 and a Pearson coefficient value of -0.9533. However, averages hide the variability highlighted in the graph above which is not a desired characteristic in DEA studies. Consequently, it cannot be concluded the 'vehicle weight' behaves in a radial manner in regards to 'fuel used'. This implies that slack based approaches (e.g. the SBM model) are likely to be better suited than radial ones (CCR model) when incorporating the 'vehicle weight' variable.

In many situations, a vehicle's weight is fixed and cannot or should not be changed (at least to some extent). It is consequently important to incorporate the 'vehicle weight' as a 'categorisation' variable essential to produce a fair fuel efficiency measure (the categorical model is not used here; this will be explained in the Conclusion). However, because the vehicle weight cannot be changed (or at least this assumption is made), it is crucial the fuel efficiency model does not allow for improvements on 'vehicle weight'.

Because 'vehicle weight' is both anti-isotonic and 'non-improvable' specific processing will be required. This will be discussed in the following two sections.

#### 5.4.4.1. Anti-isotonic variable treatment

DEA models work on the assumption that more input should generate more output, thus efficiency could be gained by either:

- ⇨ Reducing the inputs while keeping the outputs constant or
- ⇨ Increasing the outputs while keeping the inputs constant or
- ⇨ Reducing the inputs while increasing the outputs.

An anti-isotonic variable works in an opposite manner; i.e. more of an anti-isotonic input would worsen the output and vice versa. Thus, anti-isotonic variables – such as 'vehicle weight' for the fuel efficiency model – need to be processed specifically in DEA.

Koopmans (1951) highlighted in his conference work that the production process might generate undesirable variables. Literature has subsequently shown an interest in incorporating undesirable variables in DEA models. Many different models were subsequently developed (Seiford and Zhu, 2002, Färe and Grosskopf, 2004, Seiford and Zhu, 2005, Jahanshahloo et al., 2005, Hadi Vencheh et al., 2005) with interest for both undesirable inputs and outputs. Most of the models developed allow for the undesirable variables to be treated accordingly; i.e. increase an undesirable input (input oriented model) or decrease an undesirable output (output oriented model) which is the same behaviour an anti-isotonic variable should illustrate. However, in this particular instance 'vehicle weight' is 'non-improvable' (i.e. 'non-changeable'). Specific models allowing for appropriate undesirable variables processing are consequently not necessary as 'vehicle weight' should not be reduced. It is consequently not necessary to use a model which allows for the 'vehicle weight' input to reduce but simply to transform 'vehicle weight' into an isotonic variable.

Dyson (2001, Pitfalls 5.3 p.251) listed three different approaches to turn an anti-isotonic variable isotonic. These are listed below:

- ⇨ Invert the variable,
- ⇨ Move the variable from the input to the output side (and vice versa),
- ⇨ Subtract the value from a large number.

Inverting the variable breaks the ratio or interval scale (presumed to exist, Dyson et al., 2001) and the data would require further treatment. Similarly, although moving the variable to its opposite side is logical from a mathematical or theoretical point of view, the resulting model loses contact with the reality and might become difficult to explain to fleet managers. Finally the results obtained by subtracting the value from a large number (K) are generally sensitive to the value of K and should K be too large, it could dominate the data. This last approach is nonetheless retained as there is no need for further data processing and this also seems the easiest approach to communicate. Appropriate testing for different values of K should ensure K does not dominate the data.

This section will only test the following fuel efficiency model: 'Fuel used', 'Fuel cost' -> 'Miles' against several different DEA models (these will be detailed later on). Because this section only discusses the same fuel efficiency model (i.e. same inputs and outputs), each section will not be referred to as Scenario x as before but by the DEA model name (e.g. SBM-ND-I for Slack Based Model – Non-Discretionary – Input Oriented).

#### 5.4.4.2. Modelling non-controllable and non-discretionary variables

The literature distinguishes two main different approaches to incorporate exogenously fixed variables (or at least variables which cannot or should not be changed):

- ⇨ Non-controllable variables which simply do not allow any non-zero slack on the non-controllable variables.
- ⇨ The non-discretionary approach which allows non-zero slacks to appear on the non-discretionary constraints but prevent these slacks to impact the scores.

The non-controllable approach can be modelled as in Figure 5.25.

$$
\begin{aligned}
&(NCN)\ min_{\theta,\lambda}\ \theta \\
&subject\ to: \\
&\theta x_O^C \geq X^C\lambda \\
&y_O^C \leq Y^C\lambda \\
&x_O^N = X^N\lambda \\
&y_O^N = Y^N\lambda \\
&L \leq e\lambda \leq U \\
&\lambda \geq 0 \\
&where: \\
&C\ refers\ to\ controllable\ variables\ and \\
&N\ refers\ to\ non\ controllable\ variables.
\end{aligned}
$$

**Figure 5.25: The NCN (non-controllable) model**

As with standard DEA terminology, x refers to the input vector while y refers to the output vector. λ are the variables of the linear optimisation problem and $e$ the unity vector.

The model calculates efficiency through a radial measure $\theta$. In the figure above, the letter 'C' refers to the controllable variables while 'N' refers to the non-controllable variables. The equality constraint on the non-controllable variables prevents any slack from being assigned to these variables. The last constraint details the upper and lower bound condition for eλ (this enables accurate control of the model's scale assumption but is not relevant to this discussion).

The non-discretionary approach can be modelled as in Figure 5.26.

$$(NDSC)\, min\, \theta - \epsilon \left( \sum_{i \in D} s_i^- + \sum_{r=1}^{s} s_r^+ \right)$$

$$subject\ to:$$

$$\theta x_{iO} = \sum_{j=1}^{n} x_{ij}\lambda_j + s_i^-, i \in D$$

$$x_{iO} = \sum_{j=1}^{n} x_{ij}\lambda_j + s_i^-, i \in ND$$

$$y_{rO} = \sum_{j=1}^{n} y_{rj}\lambda_j - s_i^+, r = 1, \dots, s$$

$$where:$$

$$D\ refers\ to\ discretionary\ variables\ and$$

$$ND\ refers\ to\ non\ discretionary\ variables.$$

**Figure 5.26: The NDSC (non-discretionary) model**

This ND model's formulation is similar to the NCN model introduced above except that slacks are allowed even for non-discretionary variables. However, the slacks assigned to non-discretionary variables do not enter in the evaluation of the score (this is illustrated by the second constraint where $x_{iO}$ is *not* multiplied by $\theta$ (unlike for discretionary constraints) so cannot impact on the score. This model also calculates efficiency through a radial measure $\theta$.

The formulation and terminology of these two approaches are related (Cooper et al., 2007, p. 219) and the choice between them relies mainly on management understanding and model behaviour testing. This implies both models will have to be tested to determine which one is best suited for fuel efficiency analysis.

Because previous analysis showed radial measures were probably not appropriate when incorporating vehicle weight into the fuel efficiency model, the non-controllable and non-discretionary approaches will be implemented into the SBM model. The resulting SBM models will assume variables cannot be improved in a radial manner, and ensure specific variables are treated as non-controllable (SBM-NC) or non-discretionary (SBM-ND).

The formulation of the SBM-NC model is illustrated in Figure 5.27.

$$(LSBM - NC) \min \tau = t - \frac{1}{m}\sum_{i=1}^{m} S_i^{C-}/x_{iO} \quad i \in C$$

$$where\ m\ only\ accounts\ for\ controllable\ inputs$$

$$subject\ to:$$

$$1 = t + \frac{1}{s}\sum_{r=1}^{s} S_r^{C+}/y_{rO} \quad r, s \in C$$

$$tx_O^C - X^C\Lambda - S^{C-} = 0$$

$$tx_O^N - X^N\Lambda \qquad = 0$$

$$ty_O^C - Y^C\Lambda + S^{C\mp} = 0$$

$$ty_O^N - Y^N\Lambda \qquad = 0$$

$$t \geq 0;\ \Lambda \geq 0;\ S^{C-} \geq 0;\ S^{C+} \geq 0$$

$$where:$$

$$C\ refers\ to\ controllable\ variables\ and$$

$$N\ refers\ to\ non\ controllable\ variables.$$

Figure 5.27: The SBM-NC model in its linear form

This formulation is **similar to the Slack Based Model** except that non-controllable variables were discarded from the objective and that the constraint matrix was partitioned to prevent any slacks on non-controllable variables.

The first constraint is a consequence of the transformation of the SBM-NC model from its fractional to its linear form, process described in 8.4 Appendix 4: The Charnes Cooper transformation. This model fractional form can be found in 8.6

Appendix 6: Details on non-controllable and non-discretionary models. This model can be transformed in its input oriented form by ignoring the output denominator of the fractional form (see appendix 8.6). This is illustrated in Figure 5.28.

$$(SBM-NC-I)\ min_{\lambda,s^-}\ \ \rho_i\ =\ 1-\frac{1}{m}\sum_{i=1}^{m} s_i^- / x_{iO}\ \ i\in C$$

$$where\ m\ only\ accounts\ for\ controllable\ inputs$$

$$subject\ to:$$

$$x_O^C = X^C\lambda +\ s^{C-}$$

$$x_O^N = X^N\lambda$$

$$y_O^C \leq Y^C\lambda$$

$$y_O^N = Y^N\lambda$$

$$\lambda\ \geq 0;\ s^{C-}\geq 0$$

Figure 5.28: The SBM-NC-I model

Similarly, the formulation of the SBM-ND is illustrated in Figure 5.29.

$$(LSBM-ND)\ \min\tau = t-\frac{1}{m}\sum_{i=1}^{m}\mathrm{S}_\mathrm{i}^{D-}/x_{iO}\ \ i\ \in D$$

$$where\ m\ only\ accounts\ for\ discretionary\ inputs$$

$$subject\ to:$$

$$1 = t+\frac{1}{s}\sum_{r=1}^{s}\mathrm{S}_\mathrm{r}^{D+}/y_{rO}\ \ r\,,s\in D$$

$$tx_O -\ X\Lambda -\ S^- = 0$$

$$ty_O -\ Y\Lambda +\ S^+ = 0$$

$$t\ \geq 0;\ \Lambda\ \geq 0;\ S^-\geq\ 0;\ S^+\geq\ 0$$

Figure 5.29: The SBM-ND model in its linear form

This formulation is very close to the SBM-NC-except that non-zero slacks are allowed for the non-discretionary variables but they do not enter into the objective function. Here again, the first constraint is a result of the fractional to linear transformation. Because the fuel efficiency model is input oriented it is important to formulate the model in its input oriented form. This can be done by ignoring the output denominator and is illustrated in Figure 5.30.

$$(SBM - ND - I)\ min_{\lambda,s^-}\ \rho_i \ = \ 1 - \frac{1}{m}\sum_{i=1}^{m} s_i^- / x_{iO} \ \ i \in D$$

$$where\ m\ only\ accounts\ for\ disretionary\ inputs$$

$$subject\ to:$$

$$x_O = X\lambda + \ s^-$$

$$y_O = Y\lambda - \ s^+$$

$$\lambda \ \geq 0;\ s^- \geq 0;\ s^+ \geq 0$$

**Figure 5.30: The SBM-ND-I model**

These two models are logical non-controllable and non-discretionary extensions of the SBM model. No records of such models could be found in the literature at the time this thesis was written. However, Saen (2005) who has developed a SBM model addressing non-discretionary variable using extra parameters on slacks which allowed defining the extent to which some specific inputs can be nondiscretionary. Similarly (Hahn, 2007) considered non-discretionary variables along with SBM but used a Tobit regression to reflect the impact of these external non-discretionary factors on the efficiency measurement. None of these approaches treated non-discretionary factors in an appropriate manner for this study which prompted the development of the two models introduced above.

#### 5.4.4.3. Results of the SBM-NC-I model

The previous two sections explained how 'vehicle weight' should be made isotonic and which methods were appropriate to ensure 'vehicle weight' was not considered as a standard input by the fuel efficiency model. The next two sections will introduce the results for the SBM-NC-I and the SBM-ND-I models. In all the subsequent models, vehicle weights are made isotonic by subtracting the vehicle gross weight from 3501 (all vehicles are lighter than 3501 kg so all isotonic weights are positive).

The results for SBM-NC-I model are illustrated in the table further down. The results are relatively consistent in regards to mpg as shown by the RAG colouring. The best in class vehicle for each weight category shows a better efficiency than the other vehicle in the class and is also generally displayed in 'green' even for heavy vehicles. However, some DMUs' scores are not consistent with their mpg value and this would require further investigation. This is the case for DMU 4 which is evaluated efficient while demonstrating a mpg performance of 42.99. This is inconsistent with other DMU's performance such as DMU 8 which is also efficient but demonstrates a mpg performance of 58.9. Similarly the efficient DMU 23 demonstrates a mpg performance of 44.8 while another efficient vehicle in the same weight category (DMU 33) demonstrates a mpg performance of 51.58. These inconsistencies are illustrated in the Table 5.15 (the list is ordered by weight and SBM-NC-I Score).

| VehCode | Fuel Used | Isotonic Weight (3501) | Distance Travelled During Period | mpg | SBM-NC-I Score |
|---|---|---|---|---|---|
| 8 | 304.351 | 1316 | 3943.267 | 58.90059 | 1 |
| 4 | 122.104 | 1316 | 1154.916 | 42.99904 | 1 |
| 20 | 236.523 | 1316 | 2839.498 | 54.57659 | 0.98175868 |
| 14 | 213.315 | 1316 | 2425.048 | 51.68174 | 0.961582636 |
| 15 | 155.006 | 1316 | 1553.005 | 45.54729 | 0.955596242 |
| 3 | 429.187 | 1316 | 4905.681 | 51.96261 | 0.93706747 |
| 10 | 218.967 | 1316 | 2377.18 | 49.35391 | 0.922473893 |
| 21 | 229.042 | 1316 | 2507.23 | 49.76423 | 0.919007984 |
| 2 | 257.569 | 1316 | 2901.498 | 51.21143 | 0.917272019 |
| 11 | 181.067 | 1316 | 1810.955 | 45.46807 | 0.911169888 |
| 6 | 165.129 | 1316 | 1588.274 | 43.72605 | 0.910974679 |
| 5 | 250.876 | 1316 | 2766.071 | 50.12362 | 0.906460999 |
| 17 | 470.216 | 1316 | 5011.415 | 48.45082 | 0.87815958 |
| 1 | 309.356 | 1316 | 3367.747 | 49.49019 | 0.862226479 |
| 19 | 258.917 | 1316 | 2604.371 | 45.72782 | 0.837490685 |
| 18 | 425.042 | 1316 | 4366.23 | 46.69958 | 0.817198723 |
| 9 | 318.34 | 1316 | 3233.01 | 46.16938 | 0.810229648 |
| 13 | 454.395 | 1316 | 4516.327 | 45.18456 | 0.79798554 |
| 12 | 262.095 | 1316 | 2262.818 | 39.24905 | 0.742160763 |
| 16 | 467.697 | 1316 | 4155.121 | 40.38846 | 0.696787136 |
| 7 | 242.128 | 1316 | 1641.824 | 30.82619 | 0.635731402 |
| 33 | 305.685 | 640 | 3468.879 | 51.58854 | 1 |
| 23 | 544.924 | 640 | 5374.719 | 44.83922 | 1 |
| 28 | 291.318 | 640 | 3314.263 | 51.71992 | 0.995368631 |
| 31 | 297.491 | 640 | 3284.21 | 50.18747 | 0.964446022 |
| 25 | 832.821 | 640 | 6565.123 | 35.83682 | 0.892070321 |
| 38 | 684.363 | 640 | 5393.127 | 35.82551 | 0.800724137 |
| 42 | 339.135 | 640 | 3117.124 | 41.78493 | 0.795937536 |
| 32 | 714.878 | 640 | 5404.969 | 34.37158 | 0.769300163 |
| 30 | 486.605 | 640 | 3966.302 | 37.05505 | 0.756519607 |
| 43 | 726.743 | 640 | 5400.928 | 33.78514 | 0.755815447 |
| 26 | 327.101 | 640 | 2841.89 | 39.49695 | 0.739690975 |
| 29 | 123.29 | 640 | 992.182 | 36.58489 | 0.709877342 |
| 27 | 384.567 | 640 | 3040.251 | 35.9398 | 0.681588363 |
| 39 | 706.367 | 640 | 4818.786 | 31.01312 | 0.672650407 |
| 24 | 363.729 | 640 | 2781.251 | 34.76166 | 0.648257096 |
| 35 | 616.304 | 640 | 4163.77 | 30.71355 | 0.637533211 |
| 34 | 541.996 | 640 | 3781.575 | 31.71866 | 0.636420839 |
| 40 | 379.248 | 640 | 2803.591 | 33.60699 | 0.627717725 |
| 22 | 384.165 | 640 | 2831.464 | 33.50668 | 0.627058398 |
| 41 | 360.33 | 640 | 2671.507 | 33.70498 | 0.623414074 |
| 36 | 323.272 | 640 | 2395.183 | 33.68285 | 0.607993824 |
| 37 | 214.902 | 640 | 1567.138 | 33.15164 | 0.582125712 |
| 68 | 580.064 | 601 | 4229.928 | 33.1509 | 0.70531248 |

Table 5.15: SBM-NC-I model results

In order to explain this inconsistent behaviour, it is best to plot the data as in Figure 5.31.

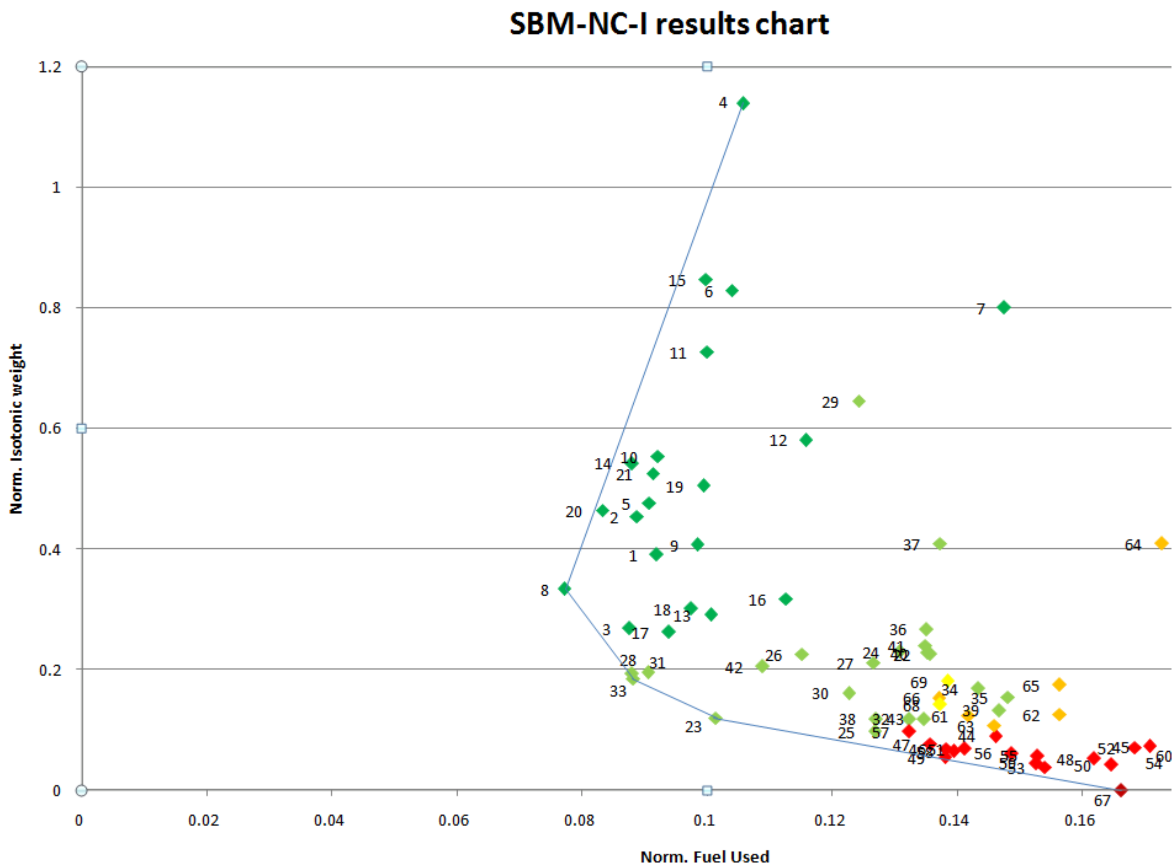


Figure 5.31: Graphical results of the SBM-NC-I model

In order to display data with 2 inputs and 1 output, both inputs are normalised against the output. This means that both inputs are divided by the output value. This result can then be multiplied by the same coefficient for all DMUs (in this case 1000). This allows both the fuel used and the isotonic weight to be proportional to the number of miles travelled and thus makes the resulting data comparable and displayable on a graph. All the subsequent diagrams will use a similar approach to display the data.

Figure 5.31 shows normalised 'fuel used' on the x axis and normalised 'vehicle weight' on the y axis. Because the weight has been made isotonic, light vehicles are at the top and heavy vehicles at the bottom. Conversely, an efficient vehicle will tend to be on the left hand side of the graph while inefficient vehicles will tend to be on the right hand side. Looking at the SBM-NC-I results it is possible to draw the efficient frontier. This frontier connects DMU 4, 8, 33, 23 and 67. Two obvious problems stand out from looking at this graph.

The first problem is caused by the model preventing any slacks on non-controllable variables (in this case weight) which precludes any 'vertical projection' of the frontier in the graph above. DMUs such as DMU 4 cannot then project on this non-existent vertical frontier and there consequently is a risk that they will be deemed efficient (which is not a desired characteristic in view of their poor mpg performance). This limitation impacts on any DMU above DMU 8. All these DMUs would thus have an artificially increased fuel performance as a consequence of this (e.g. DMU 15, DMU 6, etc). This concept is illustrated in Figure 5.32 below. This figure shows the frontier which would exist should slack be allowed on vehicle weight (which is similar to the previous frontier except for the line linking DMU 8 to DMU 4 which is replaced by a vertical plain line). With such a frontier, DMU 4 could then be projected and would consequently be deemed inefficient (which would be logical in view of its mpg performance).

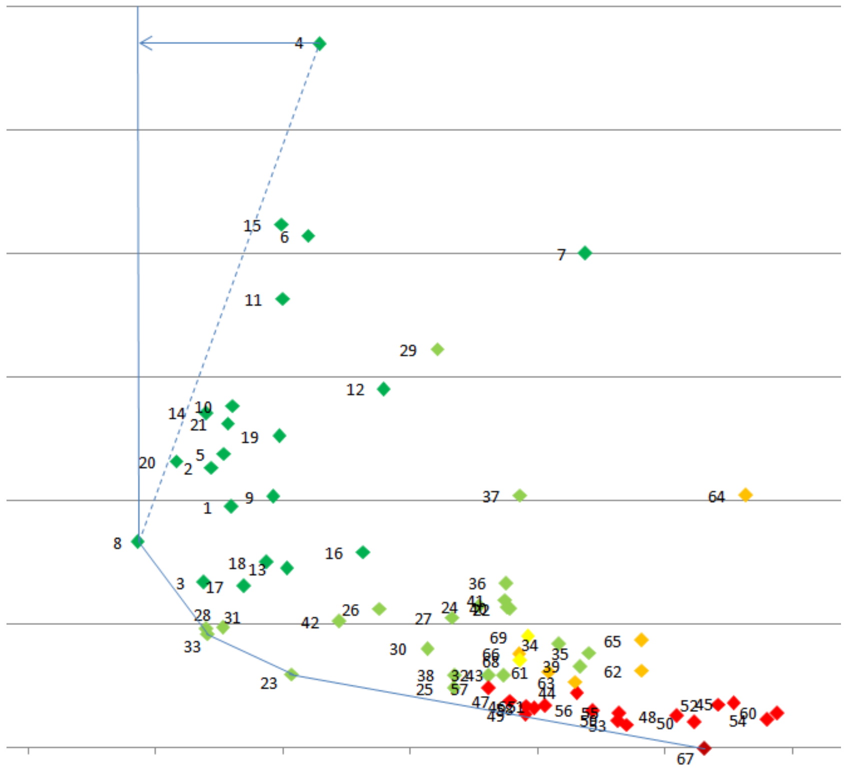

Figure 5.32: SBM-NC and the need to allow for slacks

The second problem is caused by the inconsistent impact weight has on efficiency measurement in relation to the number of miles the vehicle has made. For example DMU 4 and 15 (visible on the top of the figure above next to the efficient frontier) have the same weight (2185 kg) but because the vehicle weights are normalised based on the number of miles travelled, DMU 4 appears to ‘use’ more of the input ‘vehicle weight’. This is also the reason for DMU 23 to be evaluated as efficient within its weight category while its mpg performance is worse than efficient DMU 33 by 6 mpg (the latter demonstrating a mpg performance of 51.58). Conversely, DMU 28, the most mpg efficient vehicle in this weight category (mpg of 51.7) is nonetheless deemed inefficient due to a disadvantageous weight to miles ratio.

Because slacks are allowed in the SBM-ND model (and its variants) and because they do not enter into the efficiency calculations, the SBM-ND model could potentially address the first problem (non-vertical frontier projection). The second problem relating to the weight to miles ratio could be addressed by further data processing prior to the DEA calculations. To ensure the effects of each solution are independently measurable, it is necessary to first test how the SBM-ND-I model addresses the first issue, to then evaluate how the second issue can be addressed by further data processing.

#### 5.4.4.4. Results of the SBM-ND-I model

Results of the SBM-ND-I model, which can be found in Table 5.16, demonstrated less inconsistency with the mpg measure than the SBM-NC-I model. DMU 4 was also no longer evaluated as efficient which suggests it can now project on the efficient frontier (a vertical projection) while the resulting non-zero slacks on weight will not enter into the measurement of the efficiency itself.

| VehCode | Fuel Used | Isotonic Weight (3501) | Distance Travelled During Period | mpg | SBM-ND-I Score |
|---|---|---|---|---|---|
| 8 | 304.351 | 1316 | 3943.267 | 58.90059 | 1 |
| 3 | 429.187 | 1316 | 4905.681 | 51.96261 | 0.93706747 |
| 20 | 236.523 | 1316 | 2839.498 | 54.57659 | 0.92658814 |
| 17 | 470.216 | 1316 | 5011.415 | 48.45082 | 0.87815958 |
| 14 | 213.315 | 1316 | 2425.048 | 51.68174 | 0.877440129 |
| 2 | 257.569 | 1316 | 2901.498 | 51.21143 | 0.869455247 |
| 5 | 250.876 | 1316 | 2766.071 | 50.12362 | 0.850986667 |
| 21 | 229.042 | 1316 | 2507.23 | 49.76423 | 0.844884995 |
| 1 | 309.356 | 1316 | 3367.747 | 49.49019 | 0.840232475 |
| 10 | 218.967 | 1316 | 2377.18 | 49.35391 | 0.837918826 |
| 18 | 425.042 | 1316 | 4366.23 | 46.69958 | 0.817198723 |
| 13 | 454.395 | 1316 | 4516.327 | 45.18456 | 0.79798554 |
| 9 | 318.34 | 1316 | 3233.01 | 46.16938 | 0.783852567 |
| 19 | 258.917 | 1316 | 2604.371 | 45.72782 | 0.776355856 |
| 15 | 155.006 | 1316 | 1553.005 | 45.54729 | 0.773290889 |
| 11 | 181.067 | 1316 | 1810.955 | 45.46807 | 0.771945965 |
| 6 | 165.129 | 1316 | 1588.274 | 43.72605 | 0.742370361 |
| 4 | 122.104 | 1316 | 1154.916 | 42.99904 | 0.730027223 |
| 16 | 467.697 | 1316 | 4155.121 | 40.38846 | 0.696787136 |
| 12 | 262.095 | 1316 | 2262.818 | 39.24905 | 0.666360795 |
| 7 | 242.128 | 1316 | 1641.824 | 30.82619 | 0.523359527 |
| 33 | 305.685 | 640 | 3468.879 | 51.58854 | 1 |
| 23 | 544.924 | 640 | 5374.719 | 44.83922 | 1 |
| 28 | 291.318 | 640 | 3314.263 | 51.71992 | 0.995368631 |
| 31 | 297.491 | 640 | 3284.21 | 50.18747 | 0.964446022 |
| 25 | 832.821 | 640 | 6565.123 | 35.83682 | 0.892070321 |
| 38 | 684.363 | 640 | 5393.127 | 35.82551 | 0.800724137 |
| 42 | 339.135 | 640 | 3117.124 | 41.78493 | 0.795937536 |
| 32 | 714.878 | 640 | 5404.969 | 34.37158 | 0.769300163 |
| 30 | 486.605 | 640 | 3966.302 | 37.05505 | 0.756519607 |
| 43 | 726.743 | 640 | 5400.928 | 33.78514 | 0.755815447 |
| 26 | 327.101 | 640 | 2841.89 | 39.49695 | 0.739690975 |
| 27 | 384.567 | 640 | 3040.251 | 35.9398 | 0.681588363 |
| 39 | 706.367 | 640 | 4818.786 | 31.01312 | 0.672650407 |
| 24 | 363.729 | 640 | 2781.251 | 34.76166 | 0.648257096 |
| 35 | 616.304 | 640 | 4163.77 | 30.71355 | 0.637533211 |
| 34 | 541.996 | 640 | 3781.575 | 31.71866 | 0.636420839 |
| 40 | 379.248 | 640 | 2803.591 | 33.60699 | 0.627717725 |
| 22 | 384.165 | 640 | 2831.464 | 33.50668 | 0.627058398 |
| 41 | 360.33 | 640 | 2671.507 | 33.70498 | 0.623414074 |
| 29 | 123.29 | 640 | 992.182 | 36.58489 | 0.621129334 |
| 36 | 323.272 | 640 | 2395.183 | 33.68285 | 0.607993824 |
| 37 | 214.902 | 640 | 1567.138 | 33.15164 | 0.562840491 |
| 68 | 580.064 | 601 | 4229.928 | 33.1509 | 0.70531248 |
| 69 | 459.738 | 601 | 3320.496 | 32.83453 | 0.641595652 |
| 63 | 683.563 | 501 | 4686.315 | 31.16672 | 0.740584465 |
| 61 | 576.513 | 501 | 4066.571 | 32.06694 | 0.709256313 |
| 66 | 447.916 | 501 | 3266.705 | 33.15519 | 0.688720129 |
| 62 | 625.503 | 501 | 4000.187 | 29.07295 | 0.640384362 |
| 65 | 447.696 | 501 | 2862.814 | 29.07021 | 0.575811623 |
| 64 | 211.388 | 501 | 1224.639 | 26.33698 | 0.447142865 |

Table 5.16: SBM-ND-I model results

This can be readily verified by looking at DMU 4 which has some non-zero slack allocated to its 'Isotonic Vehicle Weight' as slacks are now allowed on the 'vehicle weight' variable. As specified earlier, 'vehicle weight' slacks do not however enter into the measurement of efficiency. This is illustrated in the following calculations in which the score calculations details are shown. It is apparent that the 'vehicle weight' slack does not enter into the measurement of efficiency by considering Formula 5.7.

$$1 - \left({}^{Slack\ Fuel\ Used}/_{Fuel\ Used}\right) = 1 - \left({}^{32.9647}/_{122.104}\right) = 0.730$$

Formula 5.7: Manual SBM-ND score calculations

This value can be further verified by considering Table 5.17 and column 'Manual Check of Score Calculation'. This column shows that only the 'fuel used' variables enters into the evaluation of efficiency.

| DMUName | Score | Slack Fuel Used | Slack Isotonic Weight (3501) | Fuel Used | Manual Check of Score Calculation |
|---|---|---|---|---|---|
| 4 | 0.7300272232 | 32.9647559377 | 930.5659282012 | 122.104 | 0.730027223205628 |

Table 5.17: Manual verification of SBM-ND-I score calculation

Figure 5.33 illustrates the frontier change.

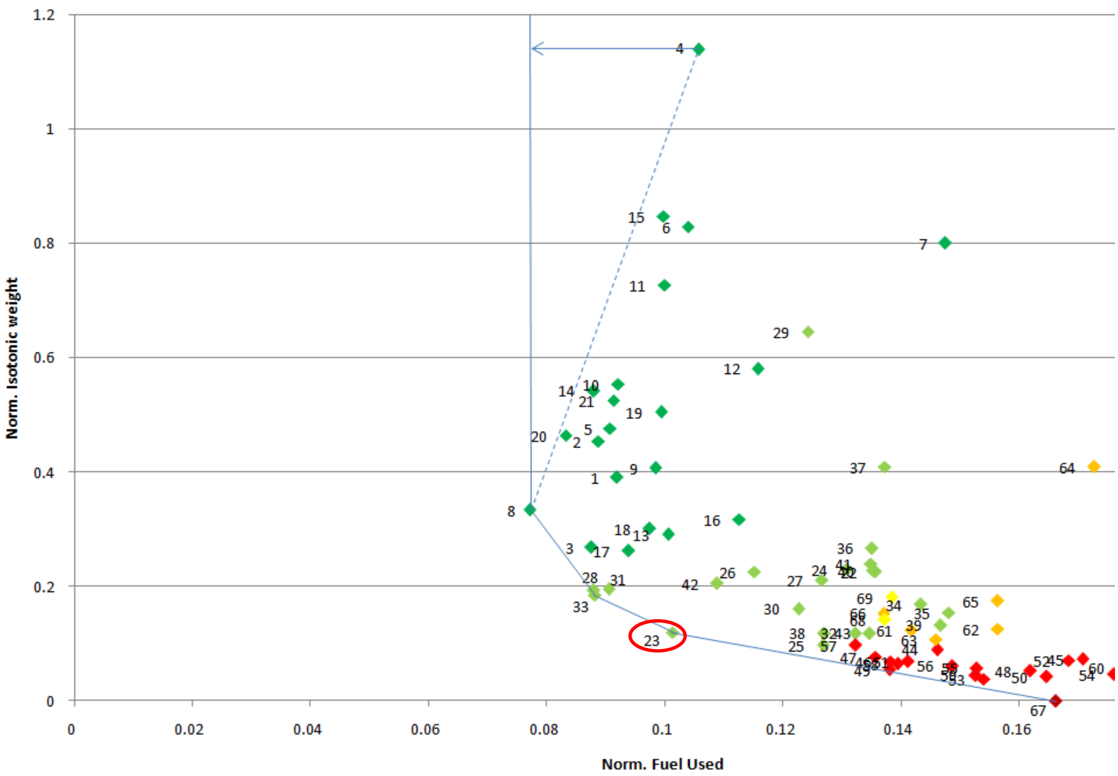


Figure 5.33: The SBM-ND-I model

Because the SBM-ND-I model allows for slacks on the non-discretionary variables DMUs can be projected on this vertical extension of the efficient frontier. Although this addresses the first issue, vehicles' efficiency measurement is still inconsistent in relation to the 'vehicle weight' to 'miles travelled' ratio. This is illustrated by DMU 23 which is still evaluated as efficient by the SBM-ND-I model. This issue should be addressed and will consequently be discussed in the next section.

#### 5.4.4.5. Addressing the 'vehicle weight' on 'miles' issue

The weight problem of the previous model was caused by the fact the impact of each DMU's gross weight depends on the mileage travelled. This is obviously unfair and such models would provide inconsistent results. In effect, the vehicle weight should be used as a *constant* which indicates the DMU's weight category to the model. Practically, this means the relation between vehicle gross weight and miles should be constant for a given vehicle weight category while 'fuel used' should be kept proportional to the distance travelled (so that the vehicle mpg performance is not destroyed by the data transformation). This can be done by normalising miles to 1,000 for all the vehicles whilst regressing 'fuel used' proportionally and leaving vehicle gross weight untouched (the value of 1,000 is arbitrary and has no effect on the end results). The result is a preserved ratio between 'fuel used' and 'distance travelled' and a constant 'vehicle gross weight' to 'distance travelled' ratio within each weight category. This is illustrated in Table 5.18.

| VehCode | Normalised Fuel Used | Isotonic Weight (3501) | Distance Travelled |
|---|---|---|---|
| 8 | 77.18 | 1316 | 1000 |
| 20 | 83.30 | 1316 | 1000 |
| 3 | 87.49 | 1316 | 1000 |
| 14 | 87.96 | 1316 | 1000 |
| 2 | 88.77 | 1316 | 1000 |
| 5 | 90.70 | 1316 | 1000 |

**Table 5.18: Processed data for weight treatment**

In the figure above, 'Normalised fuel used' and Distance travelled has been calculated using Formula 5.8.

$$Normalised\ Variable = \frac{Variable}{Distance\ Travelled/_{1000}}$$

Formula 5.8: Normalised variable formula

The ratio between 'Fuel used' and 'Distance' is preserved as the two variables are divided by the same coefficient. This ratio also normalises 'Distance Travelled' to 1,000 which gives a unique 'weight to distance' ratio for each vehicle weight category. This unique weight can be observed in the column named 'Isotonic Weight (3501)' (in this column, 1316 corresponds to a van gross weight of 2185kg. The difference is explained by the isotonic treatment 'vehicle weight' received).

The SBM-ND-I model results with this data are illustrated in Table 5.19.

| VehCode | Normalised Fuel Used | Isotonic Weight (3501) | Distance Travelled | mpg | Score |
|---|---|---|---|---|---|
| 8 | 77.18 | 1316 | 1000 | 58.90246022 | 1 |
| 20 | 83.3 | 1316 | 1000 | 54.57493253 | 0.926530612 |
| 3 | 87.49 | 1316 | 1000 | 51.9612742 | 0.882157961 |
| 14 | 87.96 | 1316 | 1000 | 51.68362756 | 0.877444293 |
| 2 | 88.77 | 1316 | 1000 | 51.21202974 | 0.869437873 |
| 5 | 90.7 | 1316 | 1000 | 50.12229195 | 0.850937156 |
| 21 | 91.35 | 1316 | 1000 | 49.76564729 | 0.844882321 |
| 1 | 91.86 | 1316 | 1000 | 49.48935206 | 0.840191596 |
| 10 | 92.11 | 1316 | 1000 | 49.35503072 | 0.837911193 |
| 17 | 93.83 | 1316 | 1000 | 48.45030246 | 0.822551423 |
| 18 | 97.35 | 1316 | 1000 | 46.69842712 | 0.79280945 |
| 9 | 98.47 | 1316 | 1000 | 46.16727816 | 0.783792018 |
| 19 | 99.42 | 1316 | 1000 | 45.72613036 | 0.776302555 |
| 15 | 99.81 | 1316 | 1000 | 45.54745897 | 0.773269212 |
| 11 | 99.98 | 1316 | 1000 | 45.4700128 | 0.771954391 |
| 13 | 100.61 | 1316 | 1000 | 45.18528854 | 0.767120565 |
| 6 | 103.97 | 1316 | 1000 | 43.72503491 | 0.742329518 |
| 4 | 105.73 | 1316 | 1000 | 42.99718037 | 0.729972572 |
| 16 | 112.56 | 1316 | 1000 | 40.38816525 | 0.685678749 |
| 12 | 115.83 | 1316 | 1000 | 39.24796581 | 0.666321333 |
| 7 | 147.48 | 1316 | 1000 | 30.82514158 | 0.523325197 |
| 28 | 87.9 | 640 | 1000 | 51.71890648 | 1 |
| 33 | 88.12 | 640 | 1000 | 51.58978529 | 0.997503404 |
| 31 | 90.58 | 640 | 1000 | 50.18869375 | 0.970412895 |
| 23 | 101.39 | 640 | 1000 | 44.83767512 | 0.866949403 |
| 42 | 108.8 | 640 | 1000 | 41.78393272 | 0.807904412 |
| 26 | 115.1 | 640 | 1000 | 39.49688862 | 0.763683753 |
| 30 | 122.68 | 640 | 1000 | 37.05650375 | 0.716498207 |
| 29 | 124.26 | 640 | 1000 | 36.58532014 | 0.707387735 |
| 27 | 126.49 | 640 | 1000 | 35.94032635 | 0.694916594 |
| 25 | 126.86 | 640 | 1000 | 35.83550276 | 0.6928898 |
| 38 | 126.9 | 640 | 1000 | 35.82420709 | 0.692671395 |
| 24 | 130.78 | 640 | 1000 | 34.76136932 | 0.672121119 |
| 32 | 132.26 | 640 | 1000 | 34.37238681 | 0.66460003 |
| 43 | 134.56 | 640 | 1000 | 33.78486831 | 0.65324019 |
| 41 | 134.88 | 640 | 1000 | 33.70471441 | 0.651690392 |
| 36 | 134.97 | 640 | 1000 | 33.68223961 | 0.651255835 |
| 40 | 135.27 | 640 | 1000 | 33.60753959 | 0.649811488 |
| 22 | 135.68 | 640 | 1000 | 33.50598379 | 0.647847877 |
| 37 | 137.13 | 640 | 1000 | 33.1516946 | 0.640997594 |
| 34 | 143.33 | 640 | 1000 | 31.71765771 | 0.613270076 |
| 39 | 146.59 | 640 | 1000 | 31.01229197 | 0.599631626 |
| 35 | 148.02 | 640 | 1000 | 30.71268666 | 0.59383867 |
| 68 | 137.13 | 601 | 1000 | 33.1516946 | 0.669768158 |
| 69 | 138.45 | 601 | 1000 | 32.8356221 | 0.663382503 |
| 66 | 137.12 | 501 | 1000 | 33.15411231 | 0.743593062 |
| 61 | 141.77 | 501 | 1000 | 32.06667052 | 0.719203503 |
| 63 | 145.86 | 501 | 1000 | 31.16750226 | 0.699036615 |
| 62 | 156.37 | 501 | 1000 | 29.07266023 | 0.6520527 |
| 65 | 156.38 | 501 | 1000 | 29.07080113 | 0.652011003 |
| 64 | 172.61 | 501 | 1000 | 26.33736099 | 0.590704366 |
| 57 | 132.31 | 201 | 1000 | 34.35939748 | 1 |
| 47 | 135.59 | 201 | 1000 | 33.52822391 | 0.975809426 |
| 49 | 138.08 | 201 | 1000 | 32.92360863 | 0.95821263 |
| 46 | 138.15 | 201 | 1000 | 32.90692638 | 0.957727108 |
| 58 | 139.45 | 201 | 1000 | 32.6001569 | 0.948798853 |
| 51 | 141.07 | 201 | 1000 | 32.22578776 | 0.937903169 |
| 44 | 146.15 | 201 | 1000 | 31.10565775 | 0.905302771 |
| 56 | 148.66 | 201 | 1000 | 30.58046468 | 0.89001749 |
| 59 | 152.53 | 201 | 1000 | 29.80457536 | 0.867435914 |
| 55 | 152.7 | 201 | 1000 | 29.77139411 | 0.866470203 |
| 53 | 153.97 | 201 | 1000 | 29.52582893 | 0.859323245 |
| 48 | 161.82 | 201 | 1000 | 28.09351057 | 0.817636881 |
| 50 | 164.61 | 201 | 1000 | 27.61734937 | 0.803778628 |
| 52 | 168.33 | 201 | 1000 | 27.00702121 | 0.786015565 |
| 45 | 170.8 | 201 | 1000 | 26.616463 | 0.774648712 |
| 54 | 176.07 | 201 | 1000 | 25.81979826 | 0.751462487 |
| 60 | 177.57 | 201 | 1000 | 25.6016888 | 0.745114603 |
| 67 | 166.16 | 1 | 1000 | 27.35972484 | 1 |

Table 5.19: SBM-ND-I results with treated weight

The SBM-ND-I model with treated weight data provides consistent scores in regards to vehicle mpg and weight category; i.e. there is no DMU within a weight category which has a worse mpg but a higher score than another. Only DMU 8, DMU 28, DMU 57 and DMU 67 are now efficient. DMU 23 and 33 are no longer efficient although DMU 33 demonstrates a score of 0.99 (it has 0.13 mpg difference with efficient DMU 28).

The results are plotted in Figure 5.34.


Figure 5.34: Plot of SBM-ND-I results with treated weight

Figure 5.34 above clearly illustrates the unique 'vehicle gross weight' 'distance' ratio per weight category (each weight category is on the same horizontal weight 'line'). The efficient frontier links the efficient DMUs and infinitely spreads vertically from DMU 8 as the model assumes constant RTS and allows for slacks on 'vehicle gross weight' (non-discretionary model). It is important to observe the model does not always assume the best DMU in its weight category is efficient. This is because the model assumes a piecewise linear frontier in between each efficient DMUs (this piecewise linear is a fundamental assumption of Data Envelopment Analysis). This implies that between each weight category the fuel efficiency performance is supposed to be linear. This is exemplified by DMU 66 and DMU 68 which demonstrate the best performance for their respective weight category (resp. 3000 and 2900 kg) but are not efficient (respective scores of 0.74 and 0.67). This is because the efficient DMUs of the weight categories under and above DMU 66 and DMU 68 demonstrate a proportionally better performance than DMU 66 and DMU 68 (these two efficient DMUs are DMU 28 and DMU 57). This behaviour is an essential characteristic of the SBM-ND-I model as it means the performance of vehicles in different weight categories can impact a van's fuel efficiency measurement.

Interestingly, with processed 'vehicle weight' data, the SBM-NC-I model gave identical results to the SBM-ND-I model. This is a consequence of both the treatment made on the vehicle gross weight which forces DMUs of the same weight to be aligned and also the fact the best DMU of any weight category does not demonstrate a worse fuel efficiency performance than the best DMU of a heavier category. This concept is illustrated in Figure 5.35.

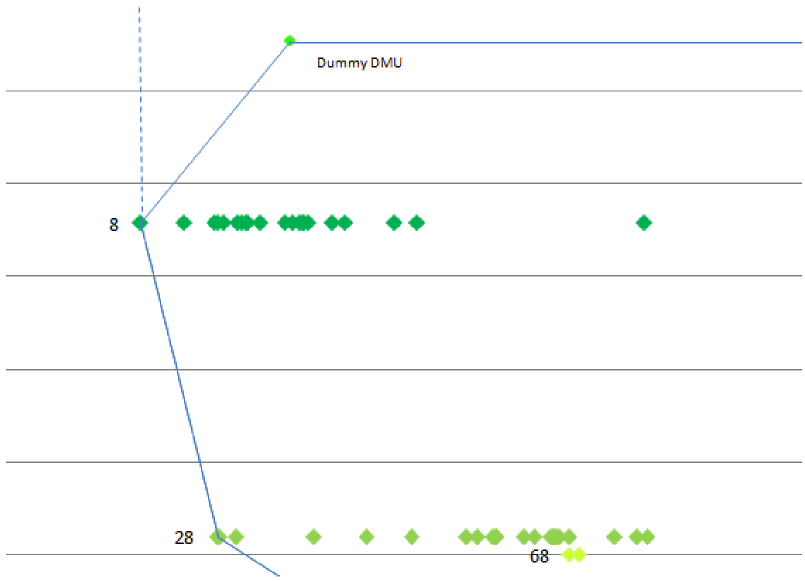


Figure 5.35: Explaining the disparities between SBM-ND and SBM-NC

The figure above is a selection of the previous graph to which was added a dummy DMU indicated in fluorescent green. This DMU has the particularity of being lighter than the DMUs in plain green (DMU 8's category) while surprisingly demonstrating a mpg performance worse than the best DMU of a heavier category (DMU 8). In this scenario, the SBM-ND model frontier would evaluate the dummy DMU as inefficient and the frontier would span from DMU 8 and would vertically stretch to infinity (the dotted line – in practice the production possibility set will stop at the minimum possible vehicle weight so this frontier is not actually infinite). However, the SBM-NC model would not allow any slack on vehicle weight. Consequently the dummy DMU would be evaluated as efficient by the non-discretionary model (in this case the efficient frontier stretches horizontally from the dummy DMU). It is obviously not fair to evaluate a lighter van demonstrating a worse mpg performance than heavier vans as efficient thus, and although this situation should rarely be observed, the SBM-NC model should not be used for fuel efficiency measurement.

In order to ensure the fuel efficiency model developed above is robust, it is necessary to test the results sensitivity to the choice of K. This will be studied in the following section 5.4.4.6.

#### 5.4.4.6. Results sensitivity to changes of K

As Dyson *et al* (2001) warned that DEA results could be sensitive to the choice of K and that the data could be dominated should K be too big. It is consequently essential to ensure that the value used for K is appropriate. This can be done by comparing the results of the SBM-ND-I model with different values of K. Model results will be calculated with the following values of K: 4,000; 6,000; 10,000 and 30,000. Model results with a K equal to 4,000 should be hopefully identical or very similar to those obtained with 3,501. This would indicate the model results are not extremely sensitive to the choice of K. The other values of K will help appraising how quickly K can dominate the data.

The SBM-ND-I results with K equal to 4,000 were identical to those obtained with K equal to 3,501. The resulting data were just translated as illustrated in Figure 5.36.

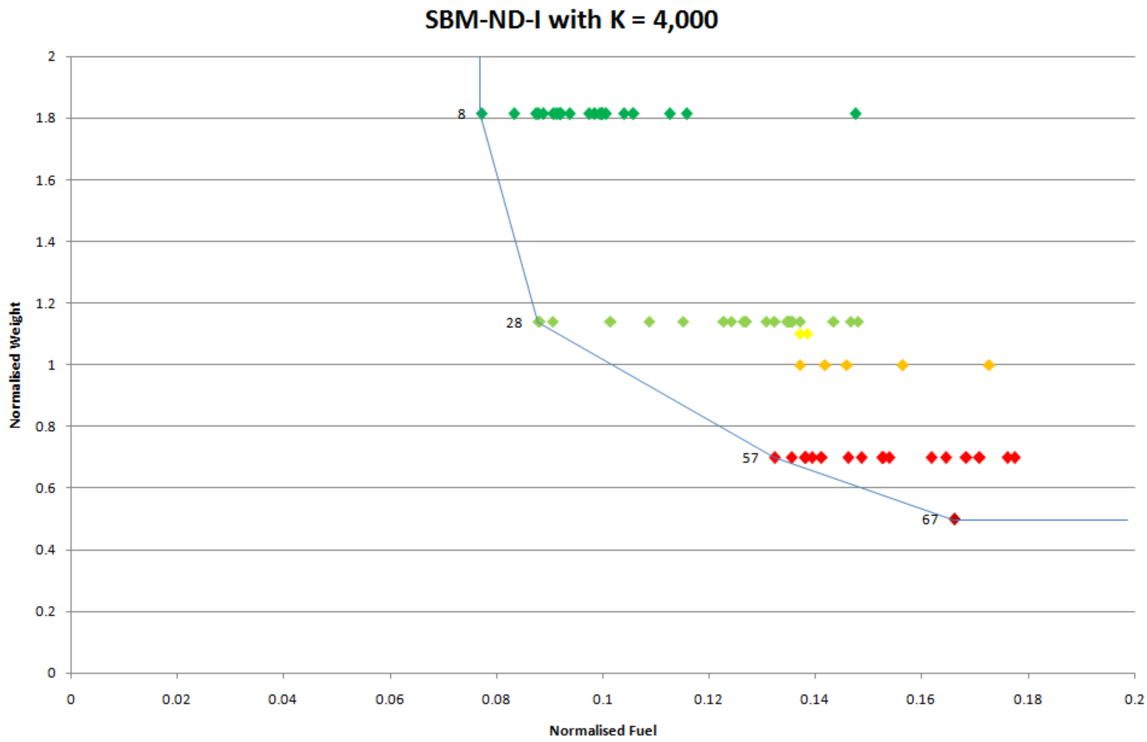


Figure 5.36: SBM-ND-I results with K = 4,000

Further observations logically revealed the SBM-ND-I model provided identical results for all the values of K listed above. It appeared the data was just translated vertically as illustrated in Figure 5.37 with K equal to 30,000.

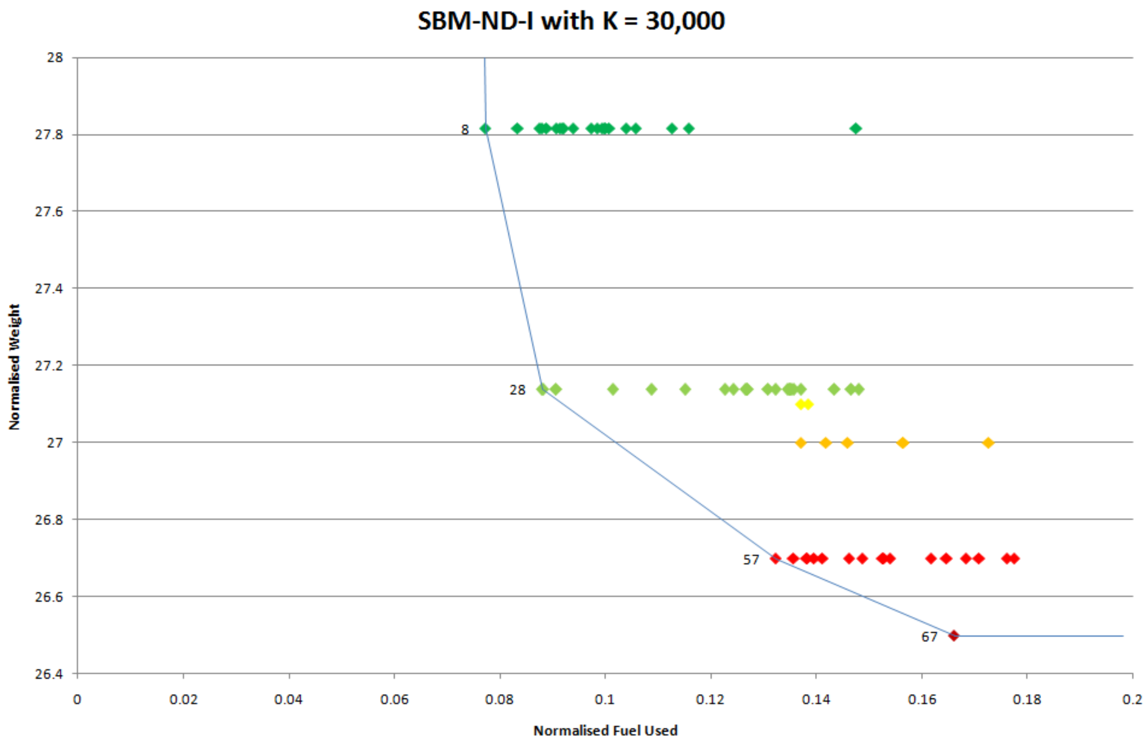


Figure 5.37: SBM-ND-I results with K = 30,000

The figure above shows that the vehicle weight to miles ratios now range from 26 to 28 which is more than 20 times greater than with the first model. Nonetheless, the different DMUs keep their relative position and their efficiency is not altered.

Because the SBM-ND-I model prevents any non-discretionary variable from affecting the score, 'vehicle weight' does not impact the score evaluation in any way and only 'fuel used' enters into the efficiency evaluation. However, because both the 'fuel used' to 'miles' ratio and the 'distance proportions' between the different weight categories are not affected by K, the fuel slacks are identical regardless of the value of K used in the calculations. This is not in contradiction with Dyson *et al* (2001) stating an excessive K could dominate the data and is in fact a consequence of the data processing undertaken on vehicle weight.

The vehicle weight slacks are also identical because they are all equal to zero. This is a consequence of both the data processing performed on vehicle weight which forces all vehicles to be aligned in their respective weight category and also of the fact there is no DMU best in class of a lighter category which demonstrates a mpg performance worse than the best in class of a van in a heavier category (see Figure 5.35: Explaining the disparities between SBM-ND and SBM-NC).

This implies that any K greater than the maximum vehicle weight can be used with the SBM-ND-I fuel efficiency model. For the remainder of this study, 3,501 would be used.

The 'Adding the Weight' section has demonstrated it is possible to incorporate 'vehicle weight' in the fuel efficiency model. The results obtained were effectively taking weight into account (as a categorisation variable) and were consistent with mpg. As a result, heavy vehicles can be evaluated as efficient (unlike with mpg) and most importantly, vehicle's efficiency is compared to the performance of other vehicles in different weight categories (this is for example the case for DMU 66 and DMU 68).

#### 5.4.4.7. Conclusion on vehicle weight

This section demonstrated that 'vehicle gross weight' can be effectively included in the fuel efficiency model provided this anti-isotonic variable is processed appropriately. It was also demonstrated that the characteristics of the SBM-NC-I model made it unsuitable for the measurement of fuel efficiency and that vehicle weight slacks should be allowed even though not accounted for in the calculation of efficiency (in the SBM-ND-I model). The SBM-NI-I model results provided a clearer and fairer measurement of efficiency.

### 5.4.5. Adding the Age

The previous section demonstrated the importance of including ‘vehicle weight’ in the fuel efficiency model. This current section will test the impact of ‘vehicle age’ on fuel efficiency.

Akin to ‘vehicle weight’, ‘vehicle age’ is an anti-isotonic variable – as older vehicles should demonstrate worse fuel efficiency performance than younger ones (because of both newer engine design and of engine wear & tear). ‘Vehicle age’ should consequently be transformed into an isotonic variable to ensure the model allows inefficient vehicles to ‘age’ (and not to ‘rejuvenate’ instead) which would reduce any potential slack found on this variable. However, it is yet unsure whether this would be a desired characteristic for the model. Therefore, age will be tested under two scenarios: the first considering ‘vehicle age’ as an ‘improvable’ variable while the second considering ‘vehicle age’ as ‘non-improvable’. This approach allows re-using the same models developed earlier.

In order to make vehicle age isotonic, ‘vehicle age’ will be subtracted from a bigger number L in a similar manner as with ‘vehicle weight’. Because ‘vehicle age’ is given as a year value and that some vehicles were manufactured in 2006 (i.e. the oldest vehicle will be 3 years old), L is equal to 4 as this will provide a small positive number in each case (although the SBM-ND model allows for semi-positive data).

Because the impact 'vehicle age' has on fuel efficiency should not depend on the number of miles travelled over the period, the isotonic 'vehicle age' will be simply added to the previous dataset. The ratio 'Miles Travelled' to 'Vehicle age' will consequently be unique per age category (as the 'miles' variable is always equal to a 1,000).

The model should look as in Figure 5.38.

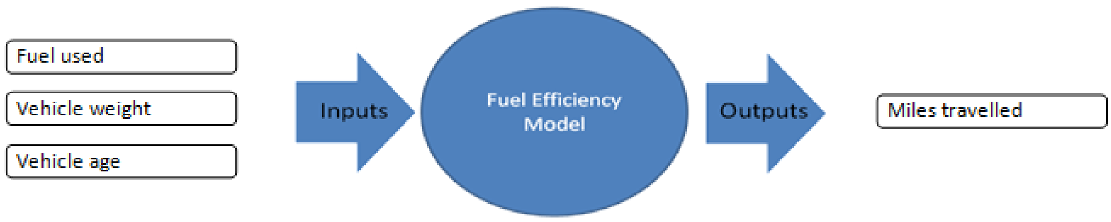


Figure 5.38: Fuel Efficiency model – vehicle weight

#### 5.4.5.1. Results with 'Vehicle Age' as an 'improvable' variable

Adding the 'vehicle age' variable to the previous model resulted in an average score difference of 12.31% in comparison with the previous results. Standard deviation was equal to 0.0698 while the maximum score difference was equal to 33.76% and observed for DMU 58. The score difference had a strong impact on the ranking with a maximum difference of 38 ranks (which account for 55.07% of the total number of vehicles) and an average of 10.8. One more DMU was found efficient by this model as well. This is understandable as the model can find efficient DMUs both in relation to the weight (like in the previous scenario) but also in relation to the vehicle age. This is illustrated in Table 5.20.

| DMUName | Isotonic Weight (3,501) | Vehicle Age | mpg | Score 20_0 |
|---|---|---|---|---|
| 47 | 201 | 1 | 33.52822391 | 1 |
| 49 | 201 | 1 | 32.92360863 | 0.990983488 |
| 46 | 201 | 1 | 32.90692638 | 0.990734709 |
| 51 | 201 | 1 | 32.22578776 | 0.980577019 |
| 56 | 201 | 1 | 30.58046468 | 0.95604063 |
| 48 | 201 | 1 | 28.09351057 | 0.918953158 |
| 50 | 201 | 1 | 27.61734937 | 0.911852257 |
| 44 | 201 | 2 | 31.10565775 | 0.713872734 |
| 55 | 201 | 2 | 29.77139411 | 0.693975115 |
| 53 | 201 | 2 | 29.52582893 | 0.690313048 |
| 52 | 201 | 2 | 27.00702121 | 0.65275055 |
| 45 | 201 | 2 | 26.616463 | 0.64692623 |
| 54 | 201 | 2 | 25.81979826 | 0.635045721 |
| 57 | 201 | 3 | 34.35939748 | 1 |
| 58 | 201 | 4 | 32.6001569 | 0.611159914 |
| 59 | 201 | 4 | 29.80457536 | 0.56946994 |
| 60 | 201 | 4 | 25.6016888 | 0.506793096 |

Table 5.20: The impact of 'Vehicle Age' on the results

In this figure DMU 57 is efficient because it demonstrates the best mpg (fuel to miles ratio) of its weight category. On the other hand, DMU 47 which belongs to the same weight category demonstrates a mpg performance of 33.52 which is 1 mpg lower than DMU 57. However, because it is two years older (Vehicle age has been made isotonic in the figure above), it demonstrates the best 'vehicle age' to 'miles' ratio and is therefore evaluated efficient by the model.

The envelopment map (a 2 dimensions matrix which shows which DMUs are used in the reference set of other DMUs) shows that 95% of the DMUs are compared to DMU 28 or DMU 47 which are both 3 years old. Consequently, the average age target is 3 years old in 95% of the cases and across the whole fleet. In fact, this figure does not mean vehicles become fuel efficient when they reach 3 years old but rather that the two efficient vehicles were 3 years old and that most of the fleet's performance is similar to DMU 28 and 47 so were therefore compared to these two vehicles. In fact, older vehicles could perfectly be fuel efficient.

Although this behaviour was to be expected from this model and correctly reflects potential efficiencies (or inefficiencies) in relation to age, it suffers from one drawback. In some cases a DMU can be evaluated efficient just from being in an advantageous age category. This is illustrated in Table 5.21.

| DMUName | Isotonic Weight (3,501) | Vehicle Age | mpg | Score 20_0 |
|---|---|---|---|---|
| 14 | 1316 | 1 | 51.68362756 | 0.999658936 |
| 18 | 1316 | 1 | 46.69842712 | 0.95146379 |
| 15 | 1316 | 1 | 45.54745897 | 0.94033664 |
| 13 | 1316 | 1 | 45.18528854 | 0.936835305 |
| 16 | 1316 | 1 | 40.38816525 | 0.890458422 |
| 12 | 1316 | 1 | 39.24796581 | 0.879435379 |
| 8 | 1316 | 2 | 58.90246022 | 1 |
| 3 | 1316 | 2 | 51.9612742 | 0.759561567 |
| 2 | 1316 | 2 | 51.21202974 | 0.745099696 |

**Table 5.21: One drawback of incorporating 'Vehicle Age'**

In the figure above, DMU 8 is efficient because it demonstrates the best mpg for its weight category. DMU 14 which belongs to the same weight category demonstrates a fuel performance 7 mpg lower than DMU 8 but is nonetheless evaluated nearly efficient (0.04% close from being efficient) simply because it is a year older (and thus demonstrates an excellent 'vehicle Age' on 'Miles' ratio). Furthermore, DMU 3, only a year older than DMU 14, is evaluated more than 25% less efficient than this DMU while demonstrating a similar mpg performance. Although this was to be expected from the model, fleet managers mistrusted the model because of the impact such small age differences had on the score. This prompted testing the model in which 'Vehicle Age' is 'non-improvable'.

#### 5.4.5.2. Results with ‘Vehicle Age’ as an ‘non-improvable’ variable

This model is similar to the previous one in all aspects except that ‘Vehicle Age’ is now considered a non-discretionary variable. Interestingly, the differences with the model from section ‘Adding the Weight’ appeared only where the envelopment map was changed. Changes in the envelopment map could be caused by either a vehicle becoming more like another after ‘Vehicle age’ was injected in the model or alternatively by a vehicle previously inefficient but becoming efficient after age was incorporated in the model. In both cases, DMUs compared to these new efficient vehicles would see their respective projections, slacks and ultimately scores change.

This is illustrated in Table 5.22.

In Table 5.22, the score differences between the model from the section ‘Adding the Weight’ and the model with ‘Vehicle age’ as a non-discretionary variable appear only for some specific DMUs. Taking a close look reveals that these DMUs previously only had DMU 57 in their reference set but were all compared to efficient DMU 47 after the age was incorporated in the fuel efficiency model (some DMUs were compared both to DMU 47 and 57). This explains the score difference between this model results and the results obtained from the SBM-ND-I model with ‘fuel used’, ‘vehicle weight’ and ‘miles’.

| DMUName | Score 18_0 | Score 20_1 | Score difference | Isotonic Weight (3,501) | Vehicle Age | mpg | Score 20_1 |
|---|---|---|---|---|---|---|---|
| 43 | 0.65324019 | 0.65324019 | 0 | 640 | 1 | 33.78486831 | 0.65324019 |
| 68 | 0.669768158 | 0.669768158 | 0 | 601 | 3 | 33.1516946 | 0.669768158 |
| 69 | 0.663382503 | 0.663382503 | 0 | 601 | 3 | 32.8356221 | 0.663382503 |
| 66 | 0.743593062 | 0.743593062 | 0 | 501 | 3 | 33.15411231 | 0.743593062 |
| 63 | 0.699036615 | 0.699036615 | 0 | 501 | 3 | 31.16750226 | 0.699036615 |
| 62 | 0.6520527 | 0.6520527 | 0 | 501 | 3 | 29.07266023 | 0.6520527 |
| 65 | 0.652011003 | 0.652011003 | 0 | 501 | 3 | 29.07080113 | 0.652011003 |
| 61 | 0.719203503 | 0.719203503 | 0 | 501 | 2 | 32.06667052 | 0.719203503 |
| 64 | 0.590704366 | 0.590704366 | 0 | 501 | 2 | 26.33736099 | 0.590704366 |
| 58 | 0.948798853 | 0.948798853 | 0 | 201 | 4 | 32.6001569 | 0.948798853 |
| 59 | 0.867435914 | 0.867435914 | 0 | 201 | 4 | 29.80457536 | 0.867435914 |
| 60 | 0.745114603 | 0.745114603 | 0 | 201 | 4 | 25.6016888 | 0.745114603 |
| 57 | 1 | 1 | 0 | 201 | 3 | 34.35939748 | 1 |
| 44 | 0.905302771 | 0.916524119 | 0.011221348 | 201 | 2 | 31.10565775 | 0.916524119 |
| 55 | 0.866470203 | 0.877210216 | 0.010740013 | 201 | 2 | 29.77139411 | 0.877210216 |
| 53 | 0.859323245 | 0.86997467 | 0.010651426 | 201 | 2 | 29.52582893 | 0.86997467 |
| 52 | 0.786015565 | 0.795758332 | 0.009742767 | 201 | 2 | 27.00702121 | 0.795758332 |
| 45 | 0.774648712 | 0.784250586 | 0.009601874 | 201 | 2 | 26.616463 | 0.784250586 |
| 54 | 0.751462487 | 0.760776964 | 0.009314477 | 201 | 2 | 25.81979826 | 0.760776964 |
| 47 | 0.975809426 | 1 | 0.024190575 | 201 | 1 | 33.52822391 | 1 |
| 49 | 0.95821263 | 0.981966976 | 0.023754345 | 201 | 1 | 32.92360863 | 0.981966976 |
| 46 | 0.957727108 | 0.981469417 | 0.023742309 | 201 | 1 | 32.90692638 | 0.981469417 |
| 51 | 0.937903169 | 0.961154037 | 0.023250868 | 201 | 1 | 32.22578776 | 0.961154037 |
| 56 | 0.89001749 | 0.912081259 | 0.02206377 | 201 | 1 | 30.58046468 | 0.912081259 |
| 48 | 0.817636881 | 0.837906316 | 0.020269435 | 201 | 1 | 28.09351057 | 0.837906316 |
| 50 | 0.803778628 | 0.823704514 | 0.019925885 | 201 | 1 | 27.61734937 | 0.823704514 |
| 67 | 1 | 1 | 0 | 1 | 1 | 27.35972484 | 1 |

Table 5.22: Score differences when adding 'Vehicle age' as a non-discretionary variable

Note: Scenario 18_0 refers to the model with Weight alone, Scenario 20_1 to the model with 'Vehicle age' as a non-discretionary variable.

Although the model behaviour illustrated in Table 5.22 is perfectly logical from a DEA point of view, the fleet managers to whom the results were presented were rather confused by the results and showed mistrust towards the model incorporating 'Vehicle Age'. This is mainly because some DMUs show a better fuel efficiency performance than others while demonstrating a worse mpg performance.

#### 5.4.5.3. Conclusion on vehicle age

Despite showing a unique 'age' to 'miles travelled' ratio and being tested through two different approaches (as an 'improvable' variable and as a ‘non-improvable’ variable), the impact of ‘vehicle age’ further segmented the efficiency results of the model with ‘fuel used’ and ‘vehicle weight’ in a way fleet managers evaluated incorrect. This behaviour might be caused by the fact there is no difference in fuel efficiency ‘potential’ between a brand new vehicle and a vehicle less than – say – 3 or 4 years old and that the step changes impact between engine generations was not clearly observable. Thus, for similar operations, a brand new vehicle and a 3 years old vehicle would demonstrate a similar mpg performance. However, because the older vehicles would have used less input (isotonic age), it would be evaluated as efficient while the brand new vehicle would be evaluated as inefficient. Because of this behaviour, ‘vehicle age’ should not be incorporated in the fuel efficiency model.

#### 5.4.6. Sensitivity Analysis

Sensitivity analysis is the topic which relates to the robustness and stability of the results to changes in the data or model (Cooper et al., 2007, p. 283). Part of this study’s sensitivity analysis was already conducted when – each time a variable was added to the model – its impact on fuel efficiency was evaluated. However, in order to ensure the results are robust and stable, it is necessary to conduct further sensitivity analysis in relation to changes in the data. This sensitivity analysis will be conducted on the results of the SBM-ND-I model with ‘vehicle weight’. ‘Vehicle age’ is not included in the model as explained in the previous section.

Several different methods to measure DEA results’ sensitivity to variations in the data exist. Attention to this topic of sensitivity was originally brought by the work of

Charnes and Cooper (Charnes and Cooper, 1968, cited in, Cooper et al., 2007, p. 284) who developed algorithms evaluating data variation sensitivity based on the inverse of the simplex matrix (without having to recalculate the inverse each time). This approach seems to have however received less attention from research than other methods.

Another alternative is based on the concept of 'distance' (or vector norm) to determine a DMU's radii of stability. One of the models based on this concept is the Chebychev norm illustrated in Figure 5.39.

$$\begin{aligned}
&\max \delta \\
&subject\ to: \\
&y_{rO} = \sum_{j=1}^{n} y_{rj}\lambda_j - s_r^+ - \delta d_r^+, \quad r = 1, \dots, s \\
&x_{iO} = \sum_{j=1}^{n} x_{ij}\lambda_j + s_i^- + \delta d_i^+, \quad i = 1, \dots, m \\
&1 = \sum_{j=1}^{n} \lambda_j
\end{aligned}$$

**Figure 5.39: The Chebychev norm model**

In the figure above, all variables are constrained to be non-negative while the $d_i^+ \ and \ d_r^-$ constant serve as weights which are generally set to unity. Charnes *et al* (Charnes et al., 1996 cited in, Cooper et al., 2007, p. 287) recommended not solely using these approaches as they do not reflect non-zero slacks.

The two previous approaches only treat one DMU at a time which poses a problem when it is not sure which DMU should receive attention. Thompson *et al* (1994) have initiated a third approach which considers simultaneous changes in all the data. Their analysis is carried out via the multiplier model (the model form with allocates a weight to each input and output) as a pair of optimal vectors *u** and *v** will generally remain valid over some variations in the data. This approach was further developed by Cooper *et al* (Cooper et al., 2007, pp. 287-291) by developing the function for a vector *w* = (*u, v*) illustrated in Formula 5.9.

$$
\begin{aligned}
&h_j(w) = \frac{f_j(w)}{g_j(w)} = \frac{\sum_{r=1}^{s} u_r y_{rj}}{\sum_{i=1}^{m} v_i x_{ij}} \\
&next\ let\ set: \\
&h_O(w) = max_{j=1,\dots,n}\ h_j(w) \\
&so\ that: \\
&h_O(w) \geq h_j(w)\ \forall\ j.
\end{aligned}
$$

**Formula 5.9: Introducing relation around vector *w***

Because the model returns to the ratio form of the CCR model, there is no need for concern in regards to the condition $vx_O = 1$ (Cooper et al., 2007, p. 288). For any efficient point O the relation illustrated in Formula 5.10 holds.

$$h_O(w^*) > h_j(w^*)\ \forall\ j \neq O$$

**Formula 5.10: Relation between efficient and inefficient DMU with *w***

In this relation, $DMU_O$ is said to be top ranked. It is then possible to create variations in the data until this relation does not hold any more and another inefficient DMU outranks $DMU_O$. The sensitivity to data variation can therefore be appraised by the amount of variation necessary to outrank the efficient DMU.

Because this method enables the simultaneous testing of results sensitivity to data variation it will be retained for this study. However, in order to use this method, it would be necessary to adapt it to the SBM-ND-I model previously illustrated in Figure

5.30. This is possible by using the dual form of the SBM-ND-I model as illustrated in Figure 5.40.

$$(DSBM - ND - I)\ max_{\ \xi,v,u}\ \xi$$
$$subject\ to:$$
$$\xi + vx_O - uy_O = 1$$
$$-vX + uY \leq 0$$
$$v \geq \frac{1}{m}[{}^{1}/{x_O}]$$

**Figure 5.40: The SBM-ND-I model in its dual form**

Which in a similar fashion gives the $h_j$ of Formula 5.11.

$$h_j(w) = 1 + uy_j - vx_j$$

**Formula 5.11: Introducing relation around vector *w* for SBM-ND-I**

For which Formula 5.12 holds over small variation of the data.

$$h_O(w^*) > h_j(w^*)\ \forall\ j$$
$$where\ DMU_O\ is\ efficient$$

**Formula 5.12: Relation between efficient and inefficient DMU with *w* for SBM-ND-I**

It is consequently possible to measure results sensitivity using Formula 5.12. This can be done as follows:

- ⇨ The relation $h_O(w^*) > h_j(w^*)\ \forall\ j$ is true when data is not changed.
- ⇨ The data is modified,
- ⇨ Vector $h_j$ is applied on the modified data,
- ⇨ The process is iterated until a $h_j(w^*)$ outranks $h_O(w^*)$.

Thompson *et al* (1994) allow the data to be varied in many different ways (Cooper et al., 2007, p. 289). In this particular instance, 'vehicle weight' will not be modified as this variable is ‘fixed’ and the data is deemed accurate (manufacturing information). Furthermore, there is no assumption made on the equipment or driver own body weight hence no data variation would be made on ‘vehicle weight’. Similarly, ‘Miles Travelled’ will not be altered as this would break the unique ratios ‘Fuel Used’ to ‘Miles’ and ‘Vehicle weight’ to ‘Miles’ which would then distort the weight categorisation carefully designed in the previous section (see section ‘Adding the Weight’). Conversely, ‘Fuel Used’ will be varied proportionally as follows:

- ⇨ The ‘Fuel Used’ value will be increased by a percentage x for all efficient DMU,
- ⇨ The ‘Fuel Used’ value will be decreased by the same percentage x for all inefficient DMU.

The percentage value starts at 1% and will be increased gradually by 1% steps until the efficient DMUs are outranked. This means a 1% difference will be applied on the data as described above and the vector $h_j(w)$ applied to the DMUs.

Using respectively $w = (u_{DMU_8}, v_{DMU_8})$ and $w = (u_{DMU_{28}}, v_{DMU_{28}})$ ($DMU_8$ and $DMU_{28}$ were both deemed efficient in the model ‘fuel used’, ‘vehicle weight’ ⇨ ‘miles travelled’), sensitivity analysis revealed that just a 1% variation in the data was necessary for DMU 33 (inefficient DMU) to outrank DMU 8 and DMU 28 (efficient DMUs). This is illustrated with $w_8$ in Table 5.23 where the fourth column corresponds to the vector ‘hj(w)’ calculated for each DMU.

| MUName | Score | Manual Score | hj(w) | Weight Normali... | Weight Isotonic... | Weight Distanc... |
|---|---|---|---|---|---|---|
| 33 | 0.9975034044 | 0.997503391808 | 1.008567 | 0.0113481616 | 0.0011479997 | 0.0017322232 |
| 8 | 1 | 0.9999999414... | 0.990000 | 0.0129567245 | 0.0002054676 | 0.0012703953 |
| 28 | 1 | 0.99999998203 | 0.988611 | 0.0113765643 | 0.0001804094 | 0.001115462 |
| 31 | 0.9704128947 | 0.970412881474 | 0.977012 | 0.0110399647 | 0.0011168219 | 0.0016851789 |
| 20 | 0.9265306122 | 0.92653064973 | 0.931498 | 0.0120048019 | 0.000190372 | 0.0011770602 |

Table 5.23: Hj(w) at 1%

In the figure above, data was modified as per explained above and $(w_8^*)$ applied to all DMUs until one outranks $h_8(w_8^*)$. As illustrated in the figure above, a 1% modification in the data was enough for DMU 33 to outrank both DMU 8 and DMU 28.

This is further illustrated in Figure 5.41 by the SBM-ND-I model results with the data modified at 1% where DMU 33 has become efficient.

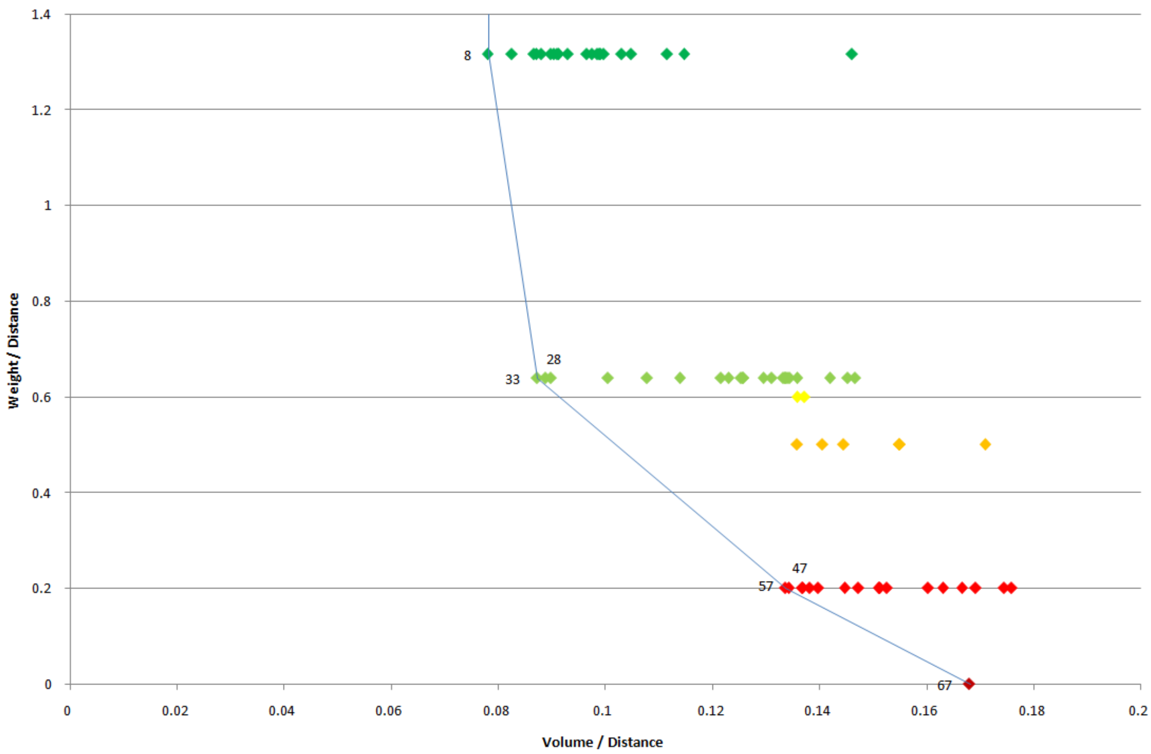


Figure 5.41: SBM-ND-I results with data modified at 1%

Similarly, a 3% change in the data was necessary for DMU 47 to outrank DMU 57 and DMU 67 (with $w_{57}$ *and* $w_{67}$).

This sensitivity analysis demonstrates the fuel efficiency model results are rather sensitive to small modifications in the data. Logically, DMU 33 and DMU 47 – which outranked other efficient DMU with a data variation of respectively 1% and 3% – are both extremely close to the frontier. This is consistent with their mpg difference: 0.12 mpg difference between DMU 28 (inefficient) and DMU 33 (efficient), and 0.83 mpg difference between DMU 57 (inefficient) and DMU 47 (efficient).

The sensitivity results indicate the fuel efficiency SBM-ND-I model results are sensitive to data variations. This is a consequence of the model design and the fuel card data; mpg variations between best in class are generally small and the categorisation designed in the previous section implies that small changes in fuel consumption could logically have a significant impact on the DMU ranking. Due to this high sensitivity, and although the results should reflect a picture close to the actual performance, it would be sensible to avoid challenging drivers on small efficiency differences. This is especially true as error (or discrepancy) created by the 'consistent mpg' assumption of the smoothing algorithm might be greater than the actual mpg differences observed between the different DMUs. The results of the sensitivity analysis further reinforce the importance of the data cleansing and smoothing algorithm.

This section cited different approaches to evaluate these results sensitivity to variation in the data. This study used the original method first initiated by Thompson *et al* (1994) to measure results' sensitivity. It could also have been possible to use statistical approaches such as the 'composed error' approach originally developed by

ALS (Aigner et al., 1977) – which was not used for reason similar to those explained in section ‘Reasons for this study to use Data Envelopment Analysis’. Bootstrapping, a method which draws statistical inferences from the data to measure properties when sampling from an approximated distribution, could also have been used. This method originally received quite extensive criticism (Simar and Wilson, 1999a, Simar and Wilson, 1999b, on the work of, Ferrier and Hirschberg, 1999), but it is now a widely used and recommended technique (Hahn, 2007).

## 5.5. Summary of Results

The Summary of Results section will discuss verification, validation and testing, as well as the multi-companies benchmark (where the data from all companies were used altogether) and results communication.

### 5.5.1. Verification, validation and testing

Balci (1998, p. 336) explains why model verification, validation and testing are fundamental in modelling. He defines these terms as follows:

Model verification ensures that the model has been built correctly. The author explains this as follows: ‘Model verification deals with building the model *right*’.

Model validation on the other hand aims at testing whether the model behaves as it should in regards to the study’s objectives and the desired level of accuracy. Balci (1998) describes model validation as follows: ‘Model validation deals with building the *right* model’.

Finally, model testing is asserting whether there is any discrepancy or inaccuracies within the model. This can be done through tests such as sending some test data and checking the results. If the results are different from what was to be expected the test has failed and the model needs to be tuned or revised.

The verification, validation and testing stages were all conducted in this study as detailed below.

The fuel efficiency model was designed through the careful observation of mpg's limitations. The results of this observation were discussed with fleet managers in order to understand which elements of the model were crucial and how each should be implemented. Although it was not certain how some of these parameters should be exactly modelled (e.g. which model to use, should variables be non-discretionary or non-controllable, etc), each option was carefully investigated and the results discussed with the fleet managers so that the resulting model corresponded to reality. Furthermore, the mathematics behind the model were discussed with academics to ensure the model was designed and built correctly. These steps ensured the model was appropriately verified.

During its design and development, the fuel efficiency model results were constantly checked versus the mpg measure. This was important as although the fuel efficiency model was meant to be a new fuel efficiency measure, mpg captures the core principle of fuel efficiency (i.e. the ratio between the fuel used and the mileage covered). Any fuel efficiency measure should consequently, to some extent, also relate to this 'fuel used' to 'miles' ratio. This was further demonstrated when the fleet managers mistrusted the model which included ‘vehicle age’ as they felt this was somehow incoherent with their perception of the notion of fuel efficiency. Furthermore, the model results were regularly discussed with the fleet managers throughout the model development in order to ensure they were consistent with what was perceived as ‘fuel efficiency’ (although this very notion of fuel efficiency was also debated). Checking the model results against mpg and fleet managers’ perception of the results provided a triangulation necessary to provide some confidence in the results. Finally, external validity (see the Case Study Theoretical Background section) was verified by comparing the results from the different companies’ vehicles. Each company’s results were similar to others which imply the procedure could be reproduced in difference environments.  This demonstrated the models outputs were the result of the data processed through the model and not of the process alone thus validating the model.

Original basic models were all tested against some available code and free DEA software. It was not however possible to check more advanced models such as SBM-ND-I as no other software or code offered this model. Testing for the SBM-ND-I model was thus conducted against some data samples and calculation made manually on Microsoft Excel. Besides, rare remaining errors or anomalies were uncovered during the analysis of the results and corrected afterwards. The results analysis as well as the manual testing ensured the model produced correct results.

### 5.5.2. Multi-companies benchmark

As discussed earlier, external environmental factors were not considered in this study as each case study was conducted internally within each company and environmental factors within each company were deemed homogeneous (perhaps not over a day but over the 3 month measurement period).

In order to appraise whether environmental factors have a significant impact on the results, data from all companies were run through the fuel efficiency model. Some data was discarded as discussed in the Cleansing Algorithm section but also because the student version of the LINDO imposes a limit on the number of variables / constraints which can be processed (the discarded DMUs demonstrated very poor mpg performance so these would have not affected the frontier).

The results were quite interesting as each company's individual frontier did not cross any other company's frontier. This is illustrated in Figure 5.42.

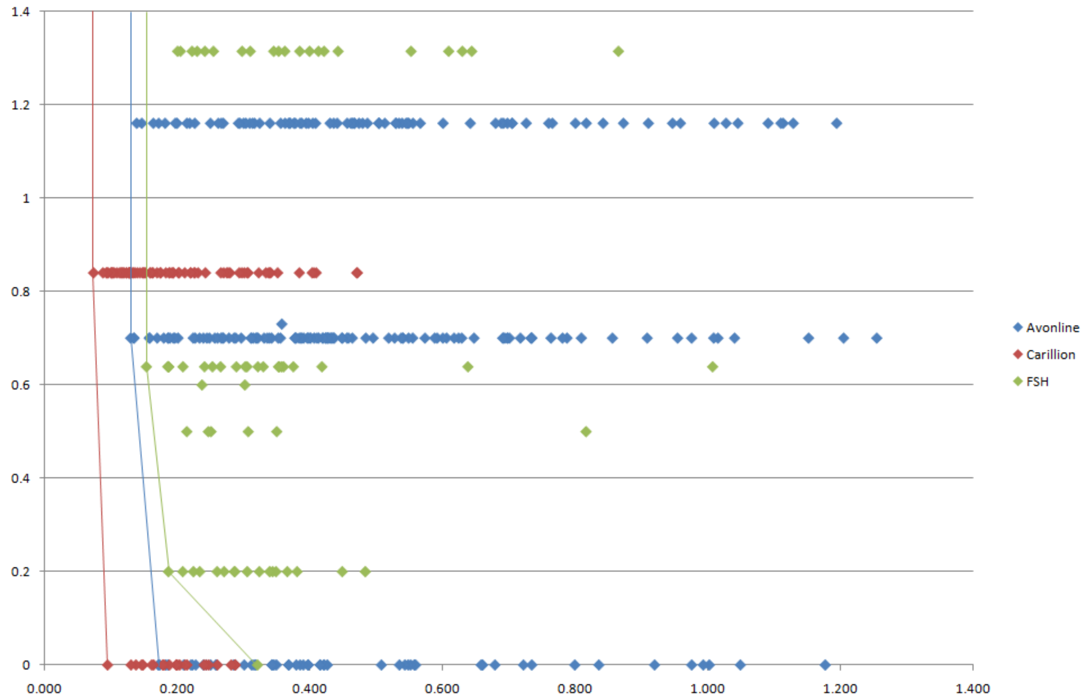


Figure 5.42: Multi-companies benchmark graph

This suggests environmental factors have a great effect on the data as each company's efficient DMUs demonstrate different level of efficiency. This indicates factors such as type of operations, vehicle load weight, or landscape seem to have such an impact on the fuel efficiency model that multi-companies benchmark should not be attempted unless the model accounts for these factors in an appropriate manner.

### 5.5.3. Communicating the results

Communicating the model results to the fleet managers was important in many different respects. As seen above, this was first essential to verify the fuel efficiency model and appraise its validity. Communicating the results was also necessary to evaluate how the results were perceived by the fleet managers. This information was critical to answer some of the research questions on results applicability and usefulness as well as to discuss some criteria necessary to interpret the study's findings.

The Case Study Theoretical Background section listed key criteria to interpret the study's findings. The points relevant to communicating the results are listed and discussed below:

*The measure is coherent with fuel efficiency operator's understanding*. This was confirmed by the fleet managers appreciations of the SBM-ND-I fuel efficiency model with 'fuel used', 'vehicle weight' and 'miles' while the model with 'vehicle age' was mistrusted.

*The measure can be easily understood*. This was directly discussed with the fleet managers. Although the intricacies of the model could be confusing, the benchmarking approach as well as the fact that weight was directly incorporated to the measure were generally easily understood. The RAG table with the score and the improvements (projections) was the most appreciated way of communicating the results. Besides, a target mpg could be provided by using the target litres (fuel used) and the mileage (1,000 miles). This helps fleet managers appraising the performance gap to reach efficiency. In order for fleet managers to read the results more easily, it is possible to transform the previous DEA results graph to a graph where units used are mpg and gross vehicle weight and where all units are made isotonic. This is illustrated in Figure 5.43.

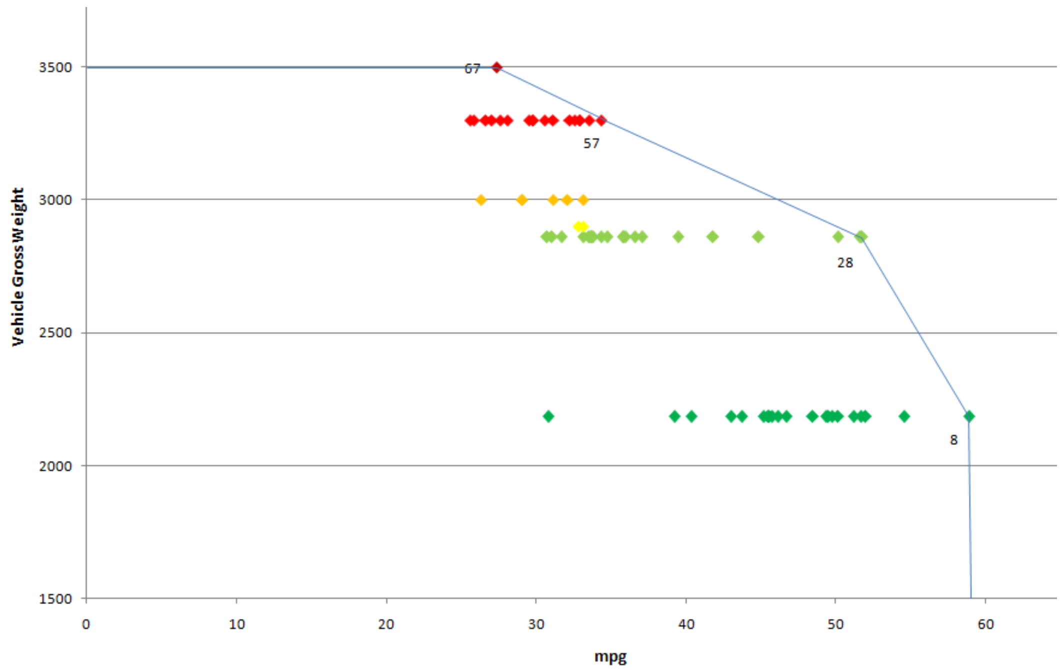


Figure 5.43: Reverted results graph

*The measure can help fleet operators to make better informed decisions*, which could in turn lead to better fuel efficiency (this point is also essential in justifying of an improvement on the mpg measure). Fleet managers were particularly keen on the data cleansing function as they recognised fuel card data cleansing was a major limitation to conducting appropriate and accurate fuel management within their company. Although their experience gave them a pretty good acumen in estimating what a vehicle's mpg should be in respect of their weight and load, they mentioned they could struggle when there were only a few vehicles in a weight category or when the vehicle was loaded with heavy extra equipment. They found the model particularly useful in this respect as it allowed for extra weight to be incorporated in the measure (although this was not done in this study, this would just consist of adding the equipment weight to the vehicle's gross weight – generally not varying for vans). Another appreciated characteristic was that a van's efficiency could be evaluated by comparing it to vans in different weight categories. This last point was on one occasion more difficult to communicate as the fleet manager struggled to comprehend how a van’s mpg (or more generally fuel efficiency) in one weight category could be comparable to the performance of vans in other weight categories.

The feedback from the fleet managers were overall positive as they generally understood the model results relatively easily and found the results useful. The overall study will be discussed in the following Summary of Results and Discussion chapter.

# 6. Summary of Results and Discussion

This chapter will succinctly recapitulate the results as most of the results were already discussed in the previous chapter. The entire research will then be discussed.

## 6.1. On the results and their usefulness

### 6.1.1. Brief summary of the case study chapter

The previous chapter explained the type of research which this study belonged to and justified why the multi-case study approach was best suited for this research. The variables that should potentially be included in the fuel efficiency model were then discussed with the fleet managers and other industry experts. It was demonstrated that the fuel efficiency model should include the following variables:

- ⇨ Fuel Used (input calculated using the smoothing algorithm)
- ⇨ Fuel cost (input calculated using the smoothing algorithm)
- ⇨ Vehicle weight (anti-isotonic, non-improvable input)
- ⇨ Vehicle age (anti-isotonic and potentially a non-improvable input)
- ⇨ Miles travelled (output)

Other interesting variables such as driver behaviour, type of operations, tyre pressure checks, engine size or servicing were not included in the fuel efficiency model for diverse reasons (see end of section 5.2.2 for more information on this).

The data cleansing algorithm was effective at cleansing fuel card data and allowed up to 95% of registrations to be matched with fleet details. All remaining problematic data was discarded from the study as explained earlier. This ensured the data used in the models was correct (although this meant some vehicles could not be included in the study). As stated earlier in section 5.3.1 Cleansing Algorithm, many matches between fuel card records and vehicle registrations were made in steps 4 and 5 which suggest telematics information significantly improve fuel card data cleansing. Theft detection, which is an essential part of fuel performance monitoring, was not used to discard any vehicle from the study as the poor performance these vehicles demonstrated was reflected in the fuel efficiency models.

The Smoothing Algorithm section explained how to use fuel card data to calculate the volume used in between two dates. This section demonstrated it is essential vehicles are refilled up to the top of the tank to allow effective fuel performance measurement. Assuming a driver's mpg performance constant (in these cases drivers tend to drive the same vehicle), the smoothing algorithm appraises the average mpg performance between the first and last refill within the measurement period and used this information along with the distance travelled to calculate an average 'fuel used' (and cost) over this period. The Smoothed mpg accuracy would suffer should the first or last fuel transaction not be made up to the top of the tank (e.g. if the first transaction is not up to the top of the tank, the fuel refilled at Refill 2 will cover some of the mileage before the first refill; conversely, if the last transaction is not up to the top of the tank, some fuel used over the period will not be measured). Similarly it is only sensible to use the smoothing algorithm when vehicles are driven by a single driver (as otherwise the assumption that the 'average mpg' should be consistent over time is erroneous). However, even in situations where vehicles are being driven by several different drivers, the smoothed mpg still tends to discard blatantly incorrect mpg figures. This implies that even though it would not be logical to use the smoothing algorithm when drivers drive different vehicles, the algorithm would still successfully discard blatantly wrong mpg figures.

The SBM and CCR DEA models scores obtained for the 'fuel used' and 'miles' model demonstrated a strong correlation with mpg. This indicated it is possible to use DEA to measure fuel efficiency under CRTS. On the other hand, the BCC DEA model, working under VRTS, evaluated some DMUs as efficient while these were actually demonstrating an average mpg performance. This is inconsistent with the common understanding of fuel efficiency and implies that VRTS models should not be used for fuel efficiency measurement. It was then demonstrated that adding the 'fuel cost' to the previous model had a limited impact on the efficiency scores. Finally, the high Pearson coefficient between the fuel cost and the fuel used (volume) justifies the low impact adding the 'fuel cost' had on fuel efficiency measurement.

Successfully incorporating 'vehicle weight' in the fuel efficiency model proved more complicated than merely adding the variable to the model. Analysis revealed that the SBM-ND-I model was suitable to measure fuel efficiency with 'vehicle weight' although further data processing was required to incorporate 'vehicle weight' in the model (in order for the 'vehicle weight' to 'miles' ratio to be unique for each weight category). The results demonstrated the significant impact of 'vehicle weight' on fuel performance and were also importantly were trusted by the fleet managers. However, adding 'vehicle age' to the model' increased the 'segmentation' within each weight category. This impacted the scores in a way which fleet managers mistrusted which consequently makes the age segmentation at least impractical in this instance if not irrelevant.

The objective of this research is to create an improved fuel efficiency measure which would reflect driver behaviour performance and potentially fuel theft or leaks. It is consequently not relevant to incorporate these driver behaviour parameters in the fuel efficiency model but instead find the other relevant parameters which would accurately reflect driver behaviour performance (as well as potential fuel leaks and theft). The results obtained when incorporating 'vehicle age' in the model suggest that adding any variable having an impact on fuel efficiency lower than driver behaviour would create further results segmentation which could ultimately lead to fleet manager mistrusting the model.

The models were tested in 3 different companies and the results obtained for each company showed similar effectiveness. This was critical in verifying the external validity of the study. Besides, the data for FSH Maintenance were entirely re-processed to ensure the results were reproducible. Except for some minor differences (e.g. different lambdas), efficiency status were identical between the two processes. Finally, the model results were regularly compared against the mpg measure but also reviewed by fleet managers. This triangulation ensured the results obtained by the fuel efficiency model were robust.

#### 6.1.2. Usefulness of the Results

The section ‘Fuel Cards Management and the Limitations of CANbus’ listed the limitations of the mpg measure. These were mainly divided in three distinct categories:

- ⇨ Factors which are necessary for the interpretation of the measure but yet not included in the measure itself (e.g. vehicle weight),
- ⇨ Cost, which is another dimension of fuel efficiency (a vehicle might be mpg efficient but ppm inefficient – even though this assumption was demonstrated incorrect),
- ⇨ The inappropriate way the measure is often used (in regards to the measurement period and time of refills).

As discussed earlier, the type of operations was not included in the fuel efficiency model as no measure could satisfactorily capture the environmental factors in a numerical form. However this was not an issue as each company was supposed to have relatively homogeneous operations (this was confirmed during discussions with the companies). All the other limitations – except ‘fuel cost’ and ‘vehicle age’ which were evaluated inappropriate for measuring fuel efficiency – were addressed by the Smoothing Algorithm and the fuel efficiency model. This is a clear improvement in comparison with the mpg measure.

DEA’s characteristics also helped providing a detailed feedback on the reasons behind performance. Each unit is assigned an efficiency score which can be compared across the whole fleet without the need to use the vehicle weight to interpret the result. This is particularly useful for fleets that have vans of many different weights or when extra equipment is towed or fitted on the van (e.g. when towing a small trailer or mini-digger or when some vans are equipped with extra racking). Another key attribute of DEA is that it lists – for each inefficient unit – the similar efficient units against which the inefficient unit's performance was measured (the reference set). This can help fleet managers in finding champions for each weight category if any (some categories will not have any champion but will have champions in different weight categories; see Figure 5.34: Plot of SBM-ND-I results with treated weight and explanatory paragraph for further information). Finally, the results are easy to communicate in a table with RAG colouring which simply highlights good and bad performers. It is also possible to communicate the results visually with the graph introduced in the section 'Communicating the results'.

The data processing undertaken on 'fuel used' and on mileage made the resulting model look as if it was merely a traditional benchmarking analysis conducted within each weight category. Although this is partially true in most cases, it fails to take into account cases where the efficiency of vehicles is measured by comparing their performance to the performance of vans in different weight categories. This was exemplified by the vehicles 66 and 68 in Figure 5.34: Plot of SBM-ND-I results with treated weight. This model characteristic highlights its usefulness for fleets running vans of many different weight categories. Although it is possible to reproduce the calculations using some linear algebra, DEA offers a more robust approach to measure efficiency.

As seen above, results from the DEA model are easy to communicate, universal (for van fleet), fair in measuring performance and useful. The Smoothing algorithm and the fuel efficiency model require similar information to that which a standard mpg analysis requires (with the addition of the vehicle, and eventual racking or other accessories' weight). Yet, they can address most of traditional mpg analysis limitations which demonstrates the improvements made on the mpg measure. There are however several limitations to this new measure which will be discussed in the following section.

## 6.2. Limitations of the Results

This section will discuss the limitations for the cleansing algorithm, the Smoothing algorithm and of the fuel efficiency model itself. Limitations in relation to DEA will be discussed first, followed by a more general section on limitation of the study itself.

### 6.2.1. Limitations in relation to data cleansing and fuel cards

One of DEA's major issues is that it is sensitive to measurement error (i.e. small changes in the data can significantly affect the output). Data sensitivity analysis highlighted that small changes in the data (below 5% when changes are non compensatory) could affect DMU's efficiency status. This was mainly caused by the fact some inefficient DMUs were really close to the efficient frontier (hence a small change in their data could easily turn them efficient).

Moreover, in this study the 'fuel used' variable is not exactly the fuel which a driver has been using but rather an approximation based on the average mpg the vehicle demonstrated during the measurement period. Although the model results should reflect a picture relatively close to reality, this limitation implies it is impossible to know whether small performance gaps are due to driver's performance or measurement error. It is consequently not advisable to challenge drivers over small performance gaps. Finally, it is important to note that this measurement accuracy issue is not caused by DEA but rather by the approximation that has to be done when using fuel card data. In fact, when using fuel card data, both traditional mpg analysis and DEA would suffer from inaccuracies in evaluating the fuel used in between two non-refill dates. That is, DEA is not a disadvantage in comparison with traditional mpg analysis.

Similar to the issue above, the cleansing algorithm cannot always cleanse data to obtain 100% match between the fuel card file and the fleet details. Vehicles with missing transactions would consequently show a better fuel consumption and thus demonstrate higher level of efficiency. This becomes problematic when a vehicle which misses one of its fuel transactions is evaluated efficient as this will affect the performance of all the inefficient DMU having this efficient DMU in their reference set. Although this problem is again caused by fuel card data, its importance is exacerbated by DEA's characteristics. The results of the cleansing algorithm also highlighted the importance of telematics in cleansing the fuel card data. Due to DEA's high sensitivity to data measurement error, it seems essential to use telematics in order to measure fuel effectively using fuel card data. Finally, cleansing fuel card data can sometimes be a tedious and resource intensive task which can be a serious limitation in busy operational environments.

### 6.2.2. Limitations in relation to DEA

One key limitation of DEA is that it can only measure performance in comparison to the best observed performance. While this is an excellent way to uncover inefficiencies against observed best performance, it does not provide efficient DMUs any indication on how to improve their performance and will consequently not give specific ideas for innovation although it should stimulate management to look for possible ways to improve performance. Although this study does not address this issue, it should be possible – for fleets of a consequent size – to record each vehicle's characteristics (in relation to fuel use, e.g. vehicle weight, engine size and model or any fuel intervention). It could then be potentially possible to correlate fuel efficiency with some of these vehicle interventions and, if deemed appropriate, use some on the efficient trucks. This of course would only work if the efficient vehicles do not all use the same interventions. Finally, it could also be possible to train the efficient driver further although this would probably not be the best allocation of potentially limited training resources (as training bad drivers would result in greater return on investment). Furthermore, there is no guarantee the efficient driver has not reached a peak of efficiency (this could potentially be confirmed by mpg averages of the industry sector). In this case, instead of wasting resources trying to further improve the driver's fuel efficiency, it would be more interesting to control fuel efficiency by regular measurement.

Another issue in relation to the model is caused by the need for a unique ratio 'vehicle weight' to 'outputs'. This study addressed this need by normalising 'fuel used' by the mileage while leaving 'vehicle weight' untouched. However, this approach only worked because there was only a single output. In effect, with the exception of some very specific cases, it might be impossible to normalise the inputs should the model have several outputs (unless the outputs are multiples of one another but in this case they should be discarded from the DEA model). While this is not an issue for this specific study, it might be more problematic should this approach be employed in a situation where several outputs need to be taken into account.

Finally, in the 'Efficient Frontier Analysis' section, this benchmarking approach would require a minimum of 9 $\left(3 \times (m + s)\right)$ vehicles in order to avoid degree of freedom issues. It would consequently not be advisable to use the fuel efficiency model for fleets of fewer than 9 vehicles.

#### 6.2.3. Limitation of the study itself

This section will discuss a list of limitations in relation to the overall approach taken by this study.

This research preferred using fuel cards over CANbus to obtain the necessary fuel information despite CANbus capability to provide accurate per driver information. This decision was made based on the fact CANbus still remains an expensive elite technology while fuel cards are omnipresent in the industry. As a consequence, and because it is generally infeasible to have all vehicles refilling at the exact beginning and end of a measurement period, some inaccuracies are created when appraising the volume used over the measurement period. This problem can be totally discarded if accurate CANbus data is available.

The model was also only tested with vans (with weights ranging from about 2000 kg to 3500 kg) and there is no guarantee it would work for rigid, HGVs or a mix of these. In effect, although it is reasonable to expect differences in the fuel efficiency of vans with different weights, their operations tend to remain relatively similar within a company (at least this was the case for the three companies involved in this study). These operational similarities are essential for the model to work and without incorporating the environmental factors (or type of operations) it might be impractical to simultaneously evaluate the efficiency for vans and HGV in the same fuel efficiency model. However, it might still be possible to use the fuel efficiency model to measure the performance of HGV and rigids. This will be discussed in the Conclusion chapter.

There are a number of limitations related to the fact the ’fuel used’ information is derived from fuel card data. The main constraints relate to the correct use of fuel cards (e.g. fill up to the top of the tank). Another important constraint is that the fuel efficiency measure relates to the vehicle and not a driver. This makes this approach inappropriate if several different drivers drive the same vehicle over the measurement period. However, CANbus – technology which associated to telematics can provide driver fuel consumption regardless whether the driver drives different vehicles or not – can potentially be used to solve this problem. The situation is summarised in Table 6.1.

| Situation | Vehicles are equipped with CANbus | Only fuel card data is available |
|---|---|---|
| **Drivers only drive their own respective vehicle** | ✔ | ✔ |
| **Drivers drive different vehicles** | Possible but might involve further data processing | ✘ |

**Table 6.1: Table summarising model applicability in regards to fuel information**

As illustrated in the table above, when drivers drive their own vehicle, it is possible to use either CANbus or fuel cards to obtain the fuel used information (top row cases). A notable difference between the two is that fuel cards reflect cost while CANbus can only show driver driving behaviour performance. This implies that using fuel cards would highlight both drivers with a poor driving behaviour and potential fuel thieves (whilst CANbus would miss the last aspect).

If the drivers drive several different vehicles and the company only uses fuel card (bottom right case), it is impossible to know which amount of fuel each driver has used; thus the approach taken in this study is not recommended (as the assumption of constant mpg is likely to be incorrect). In order to use the model one of the following would have to be done:

- ⇨ The drivers only drive a single vehicle during some time (in between two fuel transactions, or an arbitrary period of time). The only potential issue with this is that the measurement periods might not coincide between drivers which could create a small bias.
- ⇨ The vehicles are equipped with CANbus telematics technology.

Finally, if the drivers drive different vehicles equipped with CANbus telematics technology (bottom left case), it is again possible to know the amount of fuel each driver has used on each vehicle. In this case, there are two possible scenarios:

- ⇨ The driver only drives vehicles with the same gross weight, the total distance and total 'fuel used' can be used in the model or,
- ⇨ The driver drives vehicles of different weights.

In the second case, the 'fuel used' information is available for each vehicle thus a fuel efficiency score can be calculated for each driver / vehicle weight. The fuel efficiency model can thus be used to calculate a score for each driver / vehicle weight. This will give each driver an idea on how well they are performing within each weight category. This could be really useful to help them understanding where they should concentrate on improving their driving behaviour. If the management requires a unique score per driver, it is also possible to aggregate the score using a weighted average based on the fuel used (the scores could also be aggregated using the mileage as a weight although because the model is input oriented it seems more logical to use 'fuel used' as the weight in the weighted average).

When the fuel efficiency model was originally designed, many different variables were considered. After testing, some variables were discarded and only 'fuel used', 'vehicle weight' and 'miles travelled' were retained. However, because weight is merely used as a categorical variable, it can be argued that the fuel efficiency results could potentially be calculated without having to resort to DEA. While it is potentially possible to calculate the DEA scores for each vehicle without using DEA, DEA remains useful as it provides a robust method to accurately map the frontier and spot cases where DMU's efficiency is calculated from using the performance of vans in different weight categories.

Another small limitation concerns the fact all the data used in this study is obtained from companies using telematics. This creates a small bias as telematics information was used to cleanse the fuel card data whilst its importance was highlighted in the Data Cleansing section. However it is possible that cleansing fuel card information without telematics data to support the process would not be effective enough. This could have a dramatic impact on the model results (as the fuel efficiency model is very sensitive to data measurement).

Finally, the study made the assumption that operations were similar within a company. Although this was the case for the three companies there is no guarantee this should always be the case for companies running van operations. It is possible that during the whole measurement period and within one company some vans do a single job a day a hundred miles from the depot and others do several jobs a day locally. These two populations should not be compared together and the model could be used independently within each population.

### 6.2.4. Aspects not studied

Some aspects of DEA were not used in this study. This section briefly lists most of these aspects and discuss the reason why there were not included in this study.

In the models discussed in this study, the production possibility sets emerged from technical aspects of the companies' operations. These operations set the constraints defining the production possibility set (along with the data collected) and there was no need for further constraint on the data. In some situations however, it is possible to make assumptions outside the data (which are thus decided by the management). These further constraints are imposed on the multiplier vectors *v* and *u* and by two different approaches: the 'assurance region' approach and the 'cone-ratio' approach.

The assurance region approach formulates constraints on the weights in Formula 6.1.

$$L_{1,2} \le \frac{v_2}{v_1} \le U_{1,2}$$
$$\textit{which is equivalent to}:$$
$$L_{1,2}\, v_1 \le v_2 \le U_{1,2}\, v_1$$

**Formula 6.1: Assurance region weights constraints**

This formulation offers greater control over the values weights can take and limits extreme weighting divergence (e.g. zero weights).

A similar method to weight restriction is the cone ratio approach. This method defines a polyhedral convex cone in the space defined by *v* (for the input constraints and *u* for the output constraints). The input vectors (or output) are then constrained to be within this cone.

These two approaches are not relevant to this study as there was no need for external constraints on the data. This can be further corroborated by the fact there was no zero weight in the resulting data.

Similarly, allocation models, which deal with situations where DEA can be used to identify inefficiencies in relation to information on price and cost, are not of interest since fuel cost has been discarded from the fuel efficiency model. Likewise, scale elasticity and congestion are not relevant to this study since return to scale discussions were deemed inappropriate in regards to fuel efficiency.

Super efficiency corresponds to the efficiency results obtained when data from $DMU_O$ is removed from the dataset during $DMU_O$'s efficiency evaluation. Super-efficiency is useful for ranking efficient DMUs or comparing performance of two groups. The aim of the fuel efficiency model consists more of measuring vehicles' fuel efficiency rather than to obtain a detailed ranking of efficient drivers or vehicles. Nonetheless, this could potentially be applied should the management decide to reward a best driver.

Finally, the dynamic nature of companies operations can lead to efficiency changes over time. Efficiency changes can be measured using two different techniques: the window analysis and the Malmquist index. Window analysis can be done by dividing the measurement period in small segments (e.g. quarters) and measuring progressively the efficiency of segments Q1 to Q4, then of segment Q2 to Q5, etc and observing changes in efficiency. The Malmquist index is based on ratio variations in relation the efficient frontier changes over time. Because the fuel efficiency is likely to change over time, studying the changes in efficiency might have been appropriate for this study. No such study was undertaken as no particular emphasis was put by the management on fuel efficiency during the measurement period (which could have led to better fuel efficiency levels and a change in the efficient frontier). Furthermore, the short period hindered an adequate measurement of efficiency change overtime.

## 6.3. Contribution and Applicability

The previous section focused entirely on summarising the results and discussing their usefulness and related limitations. This section will discuss the contribution of this study to the body of research as well as the applicability of the findings.

### 6.3.1. Contributions

This study provides theoretical contributions to the research but also a practical contribution in terms of improved companies' operations measurement. This section will detail these contributions.

The first contribution relates to the application of DEA, a well utilised and researched method, to a field where it has never been applied before. Effectively, the section 'Reasons for this study to use Data Envelopment Analysis' explained that although DEA was extensively applied to the transport industry, it was - as far as this research could tell – never applied to the measurement of fuel efficiency. Testing the applicability of DEA to a field where it was never applied before is an original contribution to the body of research.

The second theoretical contribution relates to the SBM-ND and SBM-NC model. Although the two DEA models are logical extensions of the SBM model, no paper could be found discussing these model developments. This is again, a theoretical contribution to the body of research.

The third and last theoretical contribution to research concerns the development of the algorithms to cleanse and smooth fuel card data. Here again, no paper or research could be found relating to the cleansing or the smoothing of fuel card data using a rigorous, algorithmic approach.

These theoretical contributions both led to the publication of two research conference papers (VIRTOS et al., 2009, VIRTOS et al., 2010). The research outcomes were also presented at the Operational Research 51 conference and at an international European Logistics Association conference in 2009 for which a grant was awarded.

This research's practical contribution relates to the ability to provide an improved fuel efficiency measurement with fuel card data to companies. In effect, the contribution is threefold. Firstly, fuel card data cleansing has been improved which enables fleet managers to better measure drivers' fuel efficiency, but also allow them to better control fuel costs and potentially spot theft via poor mpg performance. Secondly, the smoothing algorithm enables the legitimate use of fuel cards to measure fuel efficiency across a whole fleet of vehicles and between two arbitrary dates (when not all vehicles can refill at the exact beginning and end of the measurement period). Finally, the use of DEA to incorporate 'vehicle weight' directly in the fuel efficiency measure improved the readability of the fuel efficiency measure, but also enabled effective comparison between vans of different gross weights.

### 6.3.2. Applicability

While the previous section discussed this study's contributions both theoretically but also in terms or practical contribution to companies' operations, this section will discuss the applicability of this study's findings.

The data cleansing and smoothing algorithms are applicable to any company using fuel card data (i.e. a majority). As introduced earlier, these two algorithms can benefit from telematics data but could also be used without. These algorithms provide companies with robust methods to measure fuel efficiency based on fuel card data, but also to control fuel cost and conduct fuel theft analysis. The Smoothing algorithm can also be practically applied in the industry which was demonstrated by the company Masternaut Three X applying the algorithm to its mpg calculations.

The fuel efficiency model developed is only applicable to companies using fuel cards, running vans and in which drivers do not share their vehicle. Although these conditions significantly reduce the applicability scope, section 'Limitation of the study itself' explained how this could be extended to different scenarios if CANbus information is available. With this technology, the fuel efficiency model can then be potentially applied to any company running vans.

Although the fuel efficiency model is only applicable to van fuel efficiency measurement, the same approach could potentially be used to measure HGV and rigids' fuel efficiency. This would however require further research as the load weight – which could safely be ignored with vans (unless the internal racking / equipment weight is taken into account) – will need to be addressed appropriately with HGV.

The data cleansing is very specific to the road industry and fuel card data thus it is unlikely this could be re-used in any other area of research but transport. However, the concept of smoothing volume can be potentially re-used in other industry where measuring a production unit consumption is important (and where the consumption can be assumed constant).

Finally, the SBM-NC and SBM-ND algorithms can potentially be re-used in many different situations where some of the inputs or outputs are non-controllable or non-discretionary and the radial assumption irrelevant.

# 7. Conclusion

## 7.1. Research Summary

This research has been undertaken using a structured approach to improve fuel efficiency measurement in the van operations sector. This work has led to the development of two new DEA models and of the fuel efficiency DEA model – the latter addressing the limitations of traditional mpg analysis.

The work undertaken stems out from the hypothesis introduced in section 1.2. This hypothesis is follows:

***It is possible to develop a form of vehicle fuel efficiency measurement that gives a fleet manager more relevant information than currently available***

The subsequent aims and objectives have provided more detail on how this study aimed to test this hypothesis but also helped in defining the study's scope and provided indications on how practical solutions could be found. The research aims were as follows:

1. To analyse the main fuel performance measurement methods used in the transport industry.
2. To evaluate the limitations of these measures and discuss the consequent impact on fuel efficiency measurement in transport businesses.

3. To develop an advanced performance measurement method in order to produce a more constructive measure and assess the extent to which it is a better measure.

4. To apply this advanced fuel efficiency performance measurement method to selected companies which operate vans.

5. To evaluate the extent to which this methodology is of operational value to transport businesses.

**Aim 1** was addressed in chapter 2 which listed most of the factors impacting fuel efficiency but also most of the interventions which can positively impact fuel efficiency. Section 2.3 evaluated the different fuel savings interventions as specified in **Aim 2**. **Aims 3 and 4** were addressed in chapter 5 where the DEA models were both designed and applied to real operational data. Finally, chapter 6 discussed and evaluated the research applicability to businesses running road transport operations, hence addressing **Aim 5**.

The study's objectives were as follows:

1. To demonstrate the relevance of fuel efficiency to transport operations
2. To critically review the factors and techniques which can have a positive impact on fuel efficiency

3. To develop an advanced performance measurement method in order to produce a more effective measure and to assess its usefulness as a better measure.

4. To review the existing literature on performance measurement & performance measurement methods
5. To evaluate the applicability of some appropriate performance measurement methods
6. To demonstrate the relevance of DEA as a suitable performance measurement method
7. To identify the companies relevant to the study and collect the appropriate information
8. To develop a new fuel efficiency measure and appropriate (DEA) performance measurement models
9. To apply the developed model to this selection of companies
10. To evaluate the model results in collaboration with the participants
11. To iteratively improve these models with the participants feedback
12. To analyse the results
13. To critically analyse the results in comparison with traditional measurement methods
14. To appraise the applicability, usefulness and limitations of the new fuel efficiency performance measure

It is believed all these different objectives were achieved throughout the study's different chapters.

**Objective 1** was addressed in section 2.1 where this study's interest in fuel efficiency was explained partially because previous research demonstrated it is generally the budget showing the most variability but also because fuel spending is where most savings can generally be made (Wilson, 1987). Section 2.2 reviewed the key fuel saving interventions available to fleet managers as specified in **Objective 2**. The pros and cons of each were summarised in section 2.3 while section 2.3.4 more specifically demonstrated the relevance of improving fuel efficiency measurement based on fuel card information as specified in **Objective 3**.

**Objective 4** was addressed in both chapter 3 – which briefly reviewed the fundamentals of performance measurement theory – and in section 3.3 – which reviewed some relevant performance measurement methods. These two sections, together with as chapter 2 constitute a comprehensive review of the existing literature on transport operations in relation to fuel efficiency, on performance measurement and Data Envelopment Analysis. The evaluation of the applicability of the relevant performance measurement methods (which corresponds to **Objective 5**), as well as selection of DEA as the most suitable performance measure (which corresponds to **Objective 6**) were discussed in section 4.1. The end of this section highlighted the gap in DEA research on micro level transport operations and fuel efficiency measurement.

The selection of relevant companies operating vans was undertaken and the details of the three companies selected are laid out in section 5.2.1 as specified by **Objective 7**. **Objectives 8 to 12** – which encompass developing the DEA models, applying these models to the selected companies, iteratively review and analyse the model results with the companies and improve the models where necessary – were all addressed throughout the remaining parts of chapter 5.

The critical analysis of the fuel efficiency DEA model results in comparison with more traditional approaches such as mpg analysis was undertaken in chapter 6 and particularly in section 6.1 (this corresponds to **Objective 13**). This section explained that although the fuel efficiency DEA model suffers from data measurement error (e.g. missing fuel transactions) in a similar manner to traditional mpg analysis, it nonetheless addressed several limitations of this traditional method – thus, representing an improvement in comparison with such fuel measurement methods.

Finally, the applicability, usefulness and limitations of this research was discussed both in the remaining of chapter 6 and in the conclusion (section 7.2 and following) as specified by **Objective 14**.

By carrying out the research aims and objectives and by answering each research questions listed in section 5.1 'Case Study Theoretical Background', **this study fully tested the hypothesis and demonstrated that it was possible to improve fuel efficiency measurement based on fuel card information using Data Envelopment Analysis.**

The research findings from this project repeatedly have been promulgated via several research papers and were also presented at several research conferences (see section 6.3.1 Contributions).

The following sections discuss the study's findings, limitations and potential for further research.

## 7.2. Findings

The research has led to a number of key findings including:

- ⇨ Not all performance measurement methods are likely to suit fuel efficiency measurement.
  - o The traditional mpg measure has several limitations in terms of parameters necessary to the interpretation of the measure which are not directly incorporated in the measure itself. Furthermore, some other aspects of fuel efficiency are not reflected in the mpg measure (e.g. ppm).
  - o Methods such as ELECTRE or AHP – although addressing some limitations of traditional mpg analysis / pence per mile analysis, mainly in relation to averages – relate more to ranking methods and do not provide a performance score. Further they do not provide a satisfactory method to include the factors necessary to the interpretations of the measure.
  - o Finally, efficient frontier analysis approaches, which are advanced benchmarking methods based on the concept of efficient frontier, were evaluated as the ideal approaches for this research. Due to some technicalities (single cross section dataset and the weak inferences that could be drawn from such datasets – see 'Reasons

for this study to use Data Envelopment Analysis' for more information on this), Data Envelopment Analysis (DEA) was finally applied as the chosen tool for this research.

- ⇨ Fuel card data can be appropriately cleansed in order to measure fuel efficiency.
  - o Algorithms can help cleansing fuel card data.
  - o Telematics information can help improving the data cleansing quality.
  - o Several utilisation rules should be observed in order to effectively measure fuel efficiency using fuel card data. This includes amongst others refilling to the top of the tank or not refilling both a vehicle and jerry can in the same transaction.

- ⇨ It is possible to appropriately measure fuel efficiency during an arbitrary period using fuel card data only. This requires using the smoothing algorithm to smooth the volume used at the 'edges' of the period.

- ⇨ It is possible to use DEA to improve fuel efficiency measurement. This was demonstrated by both appropriate comparisons between the fuel efficiency DEA model and traditional mpg analysis but also by the fleet managers trusting the newly created measure.
  - o Under the test conditions there is no significant difference between fuel cost efficiency and fuel used efficiency (this was demonstrated in section 5.4.3 Adding the Cost).

- The vehicle weight (in this study vehicle gross weight) has a significant impact on the fuel efficiency DEA measure.
- Age – although a logical input variable of the model – adds complexity to the fuel efficiency model. More importantly, vehicles demonstrating a roughly similar mpg performance could sometimes have drastically different fuel efficiency scores explained only by minimal age difference (e.g. one year). This behaviour is not coherent with the notion of fuel efficiency shared between industry experts and fleet managers thus 'vehicle age' was not included in the final fuel efficiency DEA model.

## 7.3. Limitations of the Research

Although the research demonstrated that it is feasible to measure vehicles' fuel efficiency with the SBM-ND-I model, a series of related limitations have recognised. The first limitation was in relation to the fuel data obtained and the inaccuracy in appraising the volume used in between two non-refill dates. This can potentially cause issues to businesses as fleet managers often want to see an average fuel efficiency score (traditionally mpg) over a period of time and for all their vehicles. This is understandable as it allows relatively unbiased comparison between all vehicles. Besides, simply relying on fuel efficiency measurement between refills is risky as there is no guarantee the refills are made in each case up to the top of the tank. Furthermore, refills are generally not made to the exact same level which adds further inaccuracy when only measuring fuel efficiency between refills. The Smoothing algorithm addressed, to some extent, this first limitation.

There were also a few limitations in relation to the use of DEA. These were threefold. The first limitation was linked to the fact DEA does not have any mechanism that can suggest what could be done in order to improve efficient DMUs' performance. This limitation is shared amongst all frontier analysis performance measurement methods. Nonetheless, appropriate performance measurement methods such as the one developed in this study can support innovation in providing accurate feedback on the performance levels achieved. Another limitation related to the data processing applied to vehicle weight. In effect, this data processing technique would not be possible if there were several different outputs. Finally, due to the degree of freedom issue, this benchmarking technique can only safely be applied to fleets of more than 9 vehicles.

Similarly, there were also some limitations in relation to the study itself. The first is that the model was only used to measure van fuel efficiency and applying this approach to artics or rigids would require further research. There is also a small bias in the study as all the companies were using telematics – this was shown to help the fuel card data cleansing. Conversely, an assumption of homogeneous operations – essential for the model to be valid – was made for the three companies. Although each company was relatively homogeneous, this might not always be the case for other companies. Finally, another major limitation occurs when drivers use different vehicles during the measurement period which, as seen before, prohibits the use of the fuel efficiency model. This last limitation can however be addressed if the fuel information is obtained from CANbus instead of fuel cards (as long as CANbus gives fuel information by driver).

## 7.4. Potential for Further Research

This research has already suggested some areas of potential further research which could potentially address limitations of this particular study. This section will summarise some key further research opportunities.

Biases were listed in this study; further research could evaluate their respective impact to uncover whether these are significant or not. Conducting the same study with a few companies not using telematics would potentially help quantifying the importance of telematics data in fuel card data cleansing. This is especially important as the odometer reading on fuel card data files can be inaccurate thus the cleansing algorithm would have to be further developed to allow for this potential inaccuracy.

Similarly, further research could look at capturing and quantifying the environmental factors associated with transport operations which could in turn allow the use of the fuel efficiency model across several companies and would provide valuable external information to these companies. This could potentially be done by looking at the number of miles travelled on urban roads, rural roads or motorways. These values could then be used as output variables in the model. Alternatively, they could be used to evaluate the operations category the vehicle belongs to (e.g. determine whether the vehicle should be categorised as operating within the building operations or engineer servicing operations) where each category could be ranked using ranking techniques mentioned in section 3.3 so as to be used in a categorical model. Equally, these values could also be used to calculate a difficulty score which could be used as an input in the fuel efficiency DEA model. In a similar manner, the same model could be run again with vehicle net weight instead of vehicle gross weight (which was the only weight available at the time). Although no significant difference is expected, vehicle net weight might provide a finer discrimination between vehicles as accurately reflecting the actual weight impacting fuel efficiency. Finally, more research could go into making this model compatible for any vehicle type as this could greatly benefit mixed fleet operations with rigid and HGV vehicles.

From a technical perspective, there is room for improvement with the smoothing algorithm, as one of its key weaknesses is the assumption that all refills are made up to the top of the tank. If the last refill is not made up to the top of the tank, the overall mpg would be incorrect which would also cause the volumes for the two segments (beginning and end of the period) to be inaccurate. This could however be addressed by assessing how likely it is that the first and last refill were not made up to the top of the tank although developing these rules would require further research.

More work could be undertaken to evaluate how the smoothing algorithm would behave under the categorical model (where the vehicle weight is used to categorise the different DMUs). This model looks at categorised DMUs based on their operation difficulties. A DMU can only be compared to DMUs within their own category or to DMUs from categories which operate under more difficult conditions.

More extensive research could also be carried out in applying similar DEA models to different areas of efficiency in relation to transport operations as this could potentially improve overall companies' efficiency. DEA could for example be applied to measure depot efficiency (by considering the 'number of vehicles' and 'number of drivers' as model inputs and 'income' and 'number of jobs' as model outputs) but also to the measurement of vehicle utilisation ('number of vehicles' and 'number of drivers' as model inputs and 'vehicle utilisation' as output).

The human behaviour in relation to fuel efficiency measurement was not included in this study. Human behaviour can nonetheless impact this study on two main levels. These are described below:

⇨ The study concentrates on performance measurement. It has been explained in section 2.3.4 that improving performance measurement might not always lead to performance improvements. In effect it is the informed decision based on the measurement results which can lead to improvement. This implies that it is generally crucial for managers to act on the measure in order to see improvements in performance levels. Further research could go into appraising how fuel performance management in van operations and more generally transport is conducted. This could potentially lead to the creation of a performance management framework specific to transport operations and fuel efficiency. The performance management framework have been well researched (Smith and Goddard, 2002) and Freight Best Practice (FBP,

2009) has also listed recommendations which could make a good basis to create such a framework.

- ⇨ Similarly, drivers' reaction to performance measurement was not discussed in this study and could receive attention from further research. Drivers can effectively feel potentially threatened from managers starting to actively manage their driver behaviour and fuel efficiency performance. In effect, this might lead drivers to misuse fuel cards on purpose so that management would not be in a position to use fuel card data to conduct measurement. In effect, the industry is advising fleet managers to positively challenge their drivers through incentive schemes such as Drive for Life (Masternaut Three X, 2010a). Greenroad Inc. also highlights the importance of feedback to the driver with its Greenroad Live solution which give drivers real time feedback on their driving behaviour (Greenroad, 2010). Several different telematics companies provide similar products to answer this need (Masternaut Three X, 2010b, MiX Telematics, 2010, Journey Dynamics, 2010). Further research could go into understanding the variables at play when it comes to driver acceptance of monitoring technology and performance measurement. This could help in ensuring that performance measurement is accepted by everybody as a necessary tool for the whole company to improve its performance and thus ensure that performance measurement can actually lead to performance improvements.

## 7.5. Contribution to Knowledge and Concluding Remarks

As explained in section 4.1, DEA research in transport is extensive although – as far as this research could tell – the application of DEA to fuel efficiency has never been documented. This research contributes to theoretical research as it extends the application of a well known method to an area where it is has apparently never been applied before.

This research also contributes to knowledge through the development of two new Data Envelopment Analysis models (SBM-ND and SBM-NC) which were produced to allow vehicle weight to be correctly incorporated into the fuel efficiency model. This last theoretical contribution relates to the development of both the smoothing and cleansing algorithms which allow for fuel card data to be appropriately cleansed and used to measure fuel efficiency. The SBM-ND and SBM-NC models seem logical extensions of the SBM model and were consequently probably used already although no publication could be found introducing these two models. Similarly, aspects of the cleansing algorithm are likely to have been already carried out in transport operations as such cleansing work is essential to measure fuel efficiency performance using fuel card records. Here again, no publication was found detailing this work, hence the cleansing algorithm can again be considered a theoretical contribution.

This study's practical contributions are twofold. The first main practical contribution relates to the data cleansing and smoothing algorithms. These two algorithms provide fleet managers with the possibility to use fuel cards – virtually omnipresent in van and more generally transport operations – to accurately measure fuel efficiency over a period of time (instead of just between refills). Besides, the underlying concept of the smoothing algorithm implies that it can be used to measure fuel efficiency with any type of card or fuel card records. This means that data from an on-site fuel bunker can also benefit from the smoothing algorithm. As stated in section 6.3.1, these theoretical contributions have led to the publication of two research papers which were introduced in section 6.3.1 Contributions. The work was presented at two international conferences.

The last practical contribution relates to the enhanced fuel efficiency measure provided by the DEA model. This measure enables the comparison of all vehicles against a single common measure (the DEA score). Furthermore, because all vehicles are compared against each other regardless of their weight (although the weight enters the evaluation of efficiency), this performance measurement method is particularly useful to benchmark vehicles when the fleet has a limited number of vehicles in each weight category. Conversely, traditional benchmarking analysis would only compare vehicles within the same weight category – which is of limited use when there are only a few vehicles in each weight category.

This study demonstrated that it is possible to improve van fuel efficiency measurement based on fuel card information through the use a modern performance measurement method, in this case the benchmarking method called DEA, and by directly incorporating 'vehicle weight' in the measure itself. The fuel efficiency model can only be applied within the same company to measure van performance unless environmental parameters are taken into account (see section 5.5.2 Multi-companies benchmark). When only fuel card information is available, drivers need to drive the same vehicle during the measurement period to satisfy the study's assumptions. However, when CANbus information is available, the drivers are free to drive any vehicle during the measurement period. In this case, the fuel efficiency score would be given for each driver-vehicle or each driver-'vehicle weight'. This research also highlighted key limitations of this study. Most of these were addressed and some potential further research was suggested to deal with the others. This research finally suggested that the model was sensitive to data measurement so that accurate data was needed for the model results to be valid.

Fleet managers appreciated the proposed measure and the incorporation of the vehicle weight in the fuel efficiency scores. However, it seems their first interest did not reside in an accurate measurement of fuel efficiency per se, but instead in the data cleansing, the smoothing and the model capability to fairly uncover badly inefficient drivers. This indicates the industry puts an emphasis on uncovering bad drivers and potential fuel theft as addressing these bad practices can quickly lead to improvement and savings. However, many companies are interested in challenging their driver to improve their fuel consumption as many recent competitions demonstrate (Masternaut Three X, 2010a, Low Carbon Vehicle Partnership, 2009). This indicates fleets' latent need to fairly and accurately measure fuel efficiency and suggests that methods such as the one developed in this study could be used more often in a near future.

# 8. Appendices

## 8.1. Appendix 1: The CANbus technology

### 8.1.1. A bit of history

In the early 1980's, the proliferation of electronic devices and wiring looms in cars was causing serious problems all over manufacturing processes as well as adding weight to the vehicle. Besides, excessive wiring was costly and did not provide good control over the vehicle's electronic systems. Acknowledging this problem, R. Bosch started working on an in-vehicle network project as early as 1983. This project led to the development of Controlled Area Network (CAN) technology.

Controlled Area Networks enable different Electronic Control Units (ECU) to exchange information over the same network. Using a centralised network dramatically reduced the need for excessive wiring loom and offered a better control over the vehicle's electronic architecture.

Bosch introduced the CAN protocol to the Society of Automotive Engineers (SAE) in 1986 and licensed the protocol to different electronic manufacturers soon after this. The first licence was given to Intel in 1987 which developed the first CAN controller called the 82526. This controller was released the same year.

The use of the CAN technology is now widespread and CAN controllers can be found amid most industries as well as within home electronic equipments (kitchen...). Although most vehicles manufactured in Europe use the CAN technology (probably more than 95%), some vehicles use other technologies (such as LIN or Flexray) or a mix of several different technologies. This brief text will however only discuss the CAN technology.

CAN is sometimes referred to as CANbus (or CAN bus). This terminology refers to the fact CAN shares information via a centralised digital bus. Both terms can be used to designate the same CAN technology.

### 8.1.2. Available information on vehicles

Due to the increased use of electronics devices and components on modern vehicles, most cars, vans and HGV (i.e. a good proportion of vehicles manufactured after 1996 and most vehicles manufactured after 2000) are equipped with CAN (or similar) networks. There exists however no standard relating to what information should be made available on vehicles and in what format this information should be made available. Nonetheless, the following information is generally available from most vehicles' CAN:

- ⇨ Fuel used
- ⇨ Distance
- ⇨ Engine information (throttle opening, engine rpm, gear, clutch...)
- ⇨ Speed
- ⇨ Braking information

Depending on the vehicle's electronic configuration further information such as the following might be available:

- ⇨ Airbags

- ⇨ Indicators
- ⇨ Warning lights
- ⇨ Hazard lights
- ⇨ Ceiling light
- ⇨ Door opening
- ⇨ Seat belt (not fasten)

This list is not of course exhaustive but should hopefully give a lay person a good overview of the type of information available from CAN networks on vehicles. This wealth of accurate vehicle information available on Controlled Area Networks is a strong incentive for telematics companies to connect to the vehicles' CAN and provide this information to transport companies.

As some messages sent over the CAN are more critical than others (e.g. brake (for ABS) or engine information (for EPS)), critical CAN networks such as powertrain (chassis) are separated from body or convenience (radio) networks. Powertrain networks are also generally faster (>= 1 Mbps) than body or convenient networks.

Although most of the CAN data tend to be accurate, there are some instances where these data were reported inaccurate. It is known for example that the fuel used value, which comes from the Engine Management Electronic Control Unit (ECU), can sometimes be wrong if the injector sensor is badly configured. In this case the fuel used measured by the CAN is consistently different from the actual fuel used by a coefficient. It is also known that the fuel injector accuracy tends to worsen with vehicle ageing. Finally, the accuracy of fuel used also varies depending on the RPM (it is more accurate in normal operating range than at low RPM).

In some cases, the fuel used information measured from the CANbus is inaccurate. In those cases, it is necessary to compare CANbus information with fuel card information as this enables the calculation of the ‘CANbus to reality’ ratio necessary to calibrate the telematics unit. While this process can be time consuming, there is unfortunately no method to know in advance whether the information reported by the vehicle’s CAN is accurate.

CANbus tank level accuracy is again not 100% accurate. This is caused by several reasons:

- ⇨ one cm difference in a HGV tank can represent a volume difference of about 5 litres,
- ⇨ diesel density changes in function of temperature,
- ⇨ 10 degrees increase in temperature will result in a 1% volume increase.

The information described above is conveyed by the CAN network but is defined within Higher Level Protocols which use the CAN as physical and data support layers. The next section (Technical description) will introduce technical details about CAN (the physical and data layers). Details on Higher Level Protocol will be introduced in a section 8.1.4.1 Higher Level Protocols.

### 8.1.3. Technical description

Controlled Area Networks (CAN) consist of a single pair of twisted wires to which Electronic Control Units (ECUs) are connected. The wires are twisted with each other in order to both keep them together but also because external perturbations caught by the wires would more likely cancel each other out if the wires are twisted (this is due to the voltage difference between the two wires). The pair of wires is terminated on both ends with a 120 Ohms resistor.

CAN networks can be of variable speeds and length with respect to the characteristics illustrated in Table 8.1 (Voss, 2005, p. 76):

| Speed | Max permitted length |
|---|---|
| **1 Mpbs** | 40 m |
| **500 Kbps** | 110 m |
| **250 Kps** | 280 m |
| **125 Kps** | 620 m |

**Table 8.1: Maximum CAN length at different speeds**

The reason for shorter maximal permitted network lengths at higher speeds is due the time it will take two ECUs at both extremities of the network to communicate with each other (speed in this sense does not relate to velocity of information on the network but 'reply' speed between ECUs).

The information on a CAN network is encoded with a voltage difference between the two wires called respectively CAN high and CAN low. When CAN is idle, both CAN wires are at 2.5 Volts; however, when information is transmitted, the voltage of CAN high wire jumps up to 3.5 Volts while the voltage of the CAN low wire drops down to 1.5 Volt. Each ECU connected to the CAN network is logically connected to both CAN high and CAN low.

This is illustrated in Figure 8.1 (Voss, 2005).

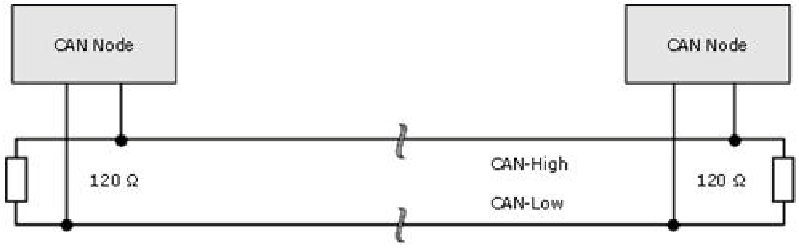


Figure 8.1: Diagram of a CAN network

CAN data is sent thought series of data frames. Bosch's CAN specification distinguishes the following different messages (Bosch, 1991):

- ⇨ A DATA FRAME which carries data from a transmitter to the receivers.
- ⇨ A REMOTE FRAME which is transmitted by a bus unit to request the transmission of the DATA FRAME with the same IDENTIFIER.
- ⇨ An ERROR FRAME which is transmitted by any unit on detecting a bus error.
- ⇨ An OVERLOAD FRAME which is used to provide for an extra delay between the preceding and the succeeding DATA or REMOTE FRAMEs.

This section will only briefly describe the Data frame which is probably best for introducing the basic concepts around CAN.

A data frame is composed of 7 different bit fields: the start of frame, the arbitration field, the control field, the data field, the CRC field, the ACK field and the end of frame field. This can be illustrated by Figure 8.2 (Bosch, 1991).

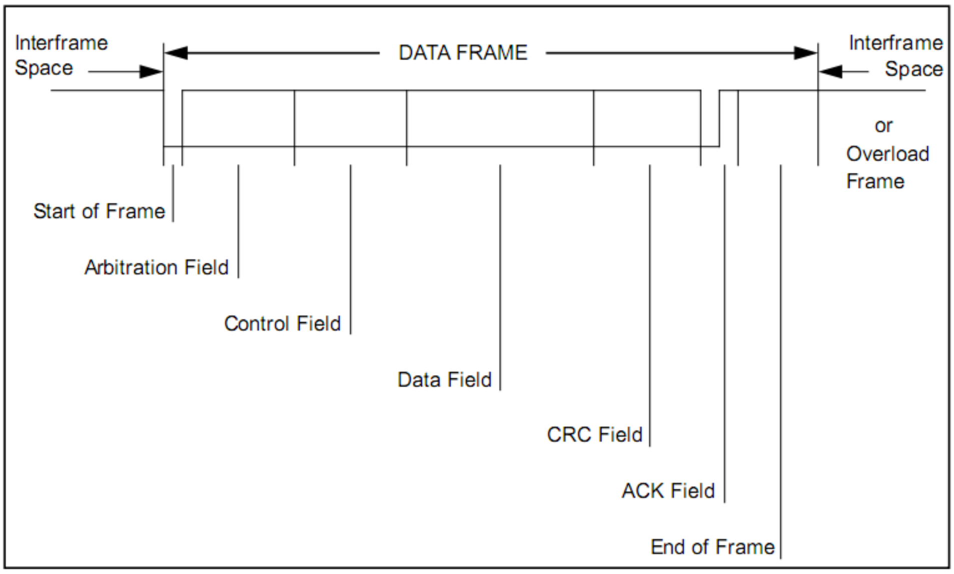


Figure 8.2: Diagram of a CAN data frame

- ⇨ The Start Of Frame (SOF) marks the beginning of Data Frames and Remote Frames. It consists of a single dominant bit. ECUs are only allowed to start transmission when the bus is idle.
- ⇨ The Arbitration Field consists of the identifier and the RTR bit. The identifier field is used by different protocol to identify the type of message encapsulated in the Data Frame. A revision of the CAN provides an extended Identifier field of 29 bits (this is referred to as the 'Extended CAN' and this is the CAN version the J1939 protocol uses). The RTR bit is used to indicate the frame is a remote frame (RTR bit dominant).

- ⇨ The Control Field consists of 6 bits which include the data length code and two other bits reserved for future expansion (used in Higher Level Protocols, see further down).
- ⇨ The Data Length code specifies the number of bytes in the Data Field.
- ⇨ The Data Field consists of the data to be transferred within the Data Frame. It can contain from 0 to 8 bytes (64 bits).
- ⇨ The CRC fields which is a check sum and a recessive bit (the delimiter DEL).
- ⇨ The ACK field which is two bits long and contains both the ACK slot and a (ACK) delimiter bit. These are used to control the quality of the data frame.
- ⇨ Each data frame is finished by 7 recessive bits which form the End Of Frame field.

Controlled Area Networks are also extremely robust. The robustness of CAN has been measured using the Residual Error Probability method. With parameters as follow (1 bit error every 0.7 second, Baud rate of 500 kBits per second, and CAN operating 8 hours a day, 365 days a year), the probability of undetected error was of one every 1,000 years (Voss, 2005).

All CAN networks conform to the physical and data technical characteristics detailed above. However due to the limited functionality the CAN alone provides Higher Level Protocols have been developed. These protocols encapsulate specific data in the different fields described above in order to provide added functionality (relevant data is generally stored in the data field). The J1939 protocol is an example of a higher level protocol based on the Controlled Area Network technology.

### 8.1.4. CAN-bus on vehicles

This section will discuss the information that is generally transferred across CAN networks, to then briefly discuss the notion of Higher Level Protocol into which this information is encapsulated. Further discussion on Higher Level Protocols will show that the automotive industry is not well standardised in this respect. This means each vehicle make or model tends to have a specific CAN protocol (although there tend to be commonalities amongst vehicles of the same make) and this causes quite a challenge for third party to connect to CAN networks and retrieve the relevant information.

#### 8.1.4.1. Higher Level Protocols

CAN has been described in the previous section as a very powerful and versatile type of network offering all the reliability, speed and robustness necessary in harsh environments such as the automotive one. Despite being one of the best choices for small applications, CAN is however not sufficient on its own for more complex operations and applications. This is mainly due to the limited data field size of 8 Bytes but also because some specific networks need a master-slave configuration or enhanced network management (network start-up, node monitoring…) which CAN itself does not provide (Voss, 2005). Higher level protocols (HLP), embedded in CAN, help delivering more functionalities from Controlled Area Networks.

The differences between the CAN technology and HLPs are best explained with the 7 layers reference model from the OSI (Open System Interconnection) illustrated in Figure 8.3 (Voss, 2005).

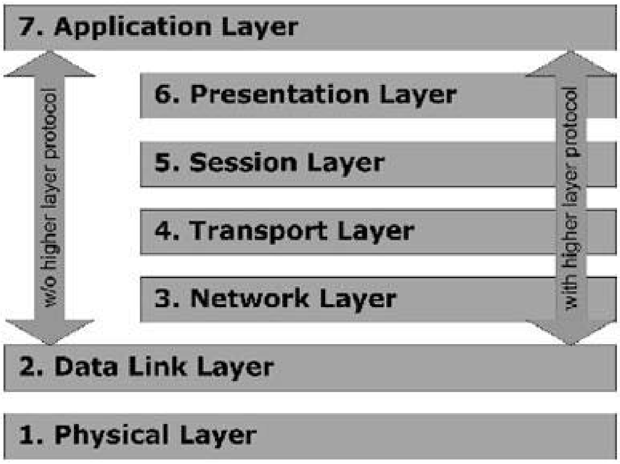


Figure 8.3: ISO/OSI 7 Layer Reference Model

The CAN specification from Bosch only specifies the first two layers, i.e. the Physical layers and the Data Link layer. These two layers respectively describe the physical characteristics of CAN networks (network length, structure, impedance...) and its communication principles (i.e. the data frame, the remote frame...). Higher Level Protocols – which sit at the application level –specify which information is to be held within all the different fields, which enables them to provide added functionalities.

As an example, the J1939 protocol specifies the engine speed as in Figure 8.4 (FMS-Standard, 2002).

1.5 Electronic Engine Controller #1: EEC1

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 00F004 | | | | | | | | PGN Hex |
| 61,444 | | | | | | | | PGN |
| 20 ms | | | | | | | | Rep. Rate |
| Data Byte 1 | Data Byte 2 | Data Byte 3 | Data Byte 4 | Data Byte 5 | Data Byte 6 | Data Byte 7 | Data Byte 8 | Byte No |
| | | | 8 7 6 5 4 3 2 1 | 8 7 6 5 4 3 2 1 | | | | Bit No |
| Not used for FMS-Standard | Not used for FMS-Standard | Not used for FMS-Standard | Engine speed<br>0.125 rpm / Bit gain<br>0 rpm offset<br>5.2.1.9<br>SPN 190 | Engine speed<br>0.125 rpm / Bit gain<br>0 rpm offset<br>5.2.1.9<br>SPN 190 | Not used for FMS-Standard | Not used for FMS-Standard | Not used for FMS-Standard | Name<br>values<br>values<br>values<br>SAE ref<br>SPN |

Figure 8.4: J1939 FMS definition of engine speed

J1939 FMS specifies that engine speed is encoded by the 4th and 5th bytes in the data frames with a PGN value of 00F004 (in this protocol, the PGN (Parameter Group Number) is the data frame id). Furthermore, it specifies that the engine speed is to be encoded in a 0.125 rpm unit.

Higher Level Protocols encodes data in a similar logic. Finally, they will thereafter simply be referred to as protocols.

#### 8.1.4.2. Higher Level Protocols in the automotive industry

The automotive industry offers a very inhomogeneous picture in term of protocols as there is no compulsory norm specifying which protocol should be used. As a result most manufacturers use their own protocol on their vehicles which is generally proprietary. Protocols are generally ‘make’ dependent and also often ‘model’ dependent. As the development of a new vehicle model often results in a modification of its electronic structure, protocols used on a new model (or year of manufacture) can also vary from previous ones.

Most HGV manufacturers have however agreed to provide critical vehicle’s information in a standard CAN format – the J1939 FMS. This protocol, adopted by most of the major manufacturers, specifies the format of the information to be made available from a J1939 FMS compliant gateway (generally just a simple plug with CAN high and CAN low) as well as the frequency at which each parameter is broadcasted through this gateway. The J1939 is a protocol designed for Heavy Duty Vehicles and is recommended by the Society of Automotive Engineers. It is important to note that a truck could provide a J1939 FMS output through a gateway but use a different (generally proprietary) protocol on the truck itself. This is for example the case of Mercedes trucks. Rigid vehicles are also generally equipped with a J1939 FMS gateway. The manufacturers that have agreed to provide a J1939 FMS outputs are: Daimler, MAN, Scania, DAF Trucks, IVECO, Volvo, Renault, and Mercedes. Note that Mercedes did not officially join the fms-standard group although they provide J1939 FMS compliant gateways on their vehicles. Further information on the Fleet Management Standard can be found on their web site (http://www.fms-standard.com) (FMS-Standard, 2002).

There is, on the other hand, no common open protocol on cars and vans (most if not all of them are proprietary), nor is there generally any gateway that broadcasts crucial vehicle information for third parties to collect. This implies that in order to retrieve the information on a car or a van, it is – in most cases – necessary to reverse engineer the vehicle's protocol first.

The European Union has however made the EOBD plug (European On-Board Diagnostic) mandatory for all petrol vehicles starting MY2001 and all diesel vehicles starting MY2003. This connector is a standardised plug giving access to technicians to a wealth of information on the vehicle sub-systems. The EOBD standard – which was originally designed to facilitate car's diagnostic and repair – specifies the format of the EOBD plug. EOBD also specifies the pin out of the connector which lists 5 potential signalling protocols. Some pins can be used with two different protocols. A given vehicle is however likely to only implement one of the protocols. The possible protocols are: J1850 PWM (Pulse-Width Modulation), J1850 VPW (Variable Pulse Width), ISO 9141-2, ISO 14230 KWP (KeyWord Protocol), and ISO 15765 CAN. Diagnostic information is made available from the EOBD connector in order to help quick identification of possible issues on the car. The various parameters available are addressed (as in location) as Parameter Identification Number (PID) and are defined in the J1979 protocol. However, manufacturers are not required to implement all the PID listed in the J1979 and can also implement their own proprietary PID. It is important to note that unlike the J1939 FMS protocol, EOBD is not meant to serve as an information gateway to which telematics units could connect, but merely is a diagnostic plug with listed compatible protocols. Some telematics companies however connect to or behind this plug in order to retrieve some of the key information.

#### 8.1.4.3. From a telematics unit point of view

There are several options for a telematics company to retrieve CANbus information:

If the telematics unit can connect to CAN (although this might expose the company's liability), the unit can be directly connected to the vehicle's CAN and retrieve the relevant information. This implies the vehicle's protocol is already known or has been reverse engineered previously. Most companies solder the unit directly to the CAN (which exposes the company's liability) although Masternaut has developed Electro Magnetic clamps that can retrieve CAN information without direct soldering. Squarell has also recently developed a similar contactless product.

If the telematics unit cannot directly understand CAN signals, the unit will have to rely on a third party interface to interpret the CAN information. The major CAN interface provider in this respect is Squarell (Holland). This company provides a series of products which can connect to the vehicle CANbus (with or without contact depending on which option is chosen). The Squarell interface can then translate this information to either J1939 FMS (in this case the Squarell interface serves as a third party J1939 FMS gateway which can virtually be installed on any Squarell compatible vehicle including cars and vans) or RS232 with a protocol defined by Squarell. Accutest is another company that provides such interfaces. Accutest products connect to the EOBD connector and provide an output in the Accutest RS232 protocol.

Alternatively, some companies (e.g. SPAL, bridge-water electronics) provide small interfaces that read the CAN signals and provide voltage based outputs for each parameter. These interfaces are however more dedicated to taxi companies (that would need speed or distance feed for their meters for example) rather than telematics companies which need more complex and detailed information.

It is important to observe that whilst the J1939 FMS compliant gateways on trucks provide third party with a safe way to connect to the vehicle's electronics in order to retrieve some of the available vehicle's information, connecting directly to the vehicle's CANbus is likely to void the vehicle warranty. Manufacturers are quite blurring on the subject and it is very difficult to get an official and clear answer from them. Legal departments will generally claim that this is a technical question, and technical departments that this is a legal question. Masternaut and Squarell contactless products are, in this respect, critical CANbus telematics developments.

#### 8.1.4.4. Final word on CANbus

The market as evolved quite rapidly in regards to CANbus and whilst most small companies (i.e. lower than 20 vehicles) would not generally spend vast amount of money on CANbus technology, most big companies require this as a minimal requirement in their bids to telematics companies. These companies not only require the basic set of information such as the one found in the J1939 FMS protocol (i.e. fuel, distance, engine and brake information) but also more advanced information such as warning lights (to warn of potential failure or breakdown), airbags, door opening, or indicators. Successful telematics companies that would want to tackle the big fleet market will have to provide this information and be versatile enough to add new information should this be required by customers.

Although using a third party interface to retrieve CAN information tends to be more expensive than connecting directly to the vehicle's CAN (more costly hardware), it gives telematics companies without CAN capabilities the possibility to nonetheless provide CAN information to their customers (and thus focus on other part of their business). Finally, while there exist little evidence which would suggest how liability would be appraised in case laws of accidents where telematics units would be directly connected to the vehicle's CAN, the telematics users' concerns over liability suggest Masternaut and Squarell contactless developments should find their market.

## 8.2. Appendix 2: From Econometrics to the Charnes Cooper and Rhodes model

This appendix introduces the Charnes Cooper and Rhodes (CCR) model and its main characteristics. Although the CCR model was used in this study, not all concepts mentioned in this appendix were specifically used. These were nonetheless discussed as they should provide answers to most of the questions which might arise from this research.

Koopmans (1951) defined efficiency as the point where 'it is impossible to produce more of any output without producing less of some other output or using more of some input'. This definition is often referred to as the Pareto Koopmans definition of efficiency (Cooper et al., 2007, p. 46) as Koopmans work is closely related to the Pareto's definition of efficiency in economics. Although Farrell applied these efficiency concepts to observed data, his application was only able to measure technical efficiency (i.e. it could not correctly estimate slacks). Although Farrell was aware of this shortcoming, he was unable to address it with a mathematical formulation.

Charnes and Cooper were also aware of this limitation and were working on implementing this mathematical programming theory through the introduction of goal programming (Charnes and Cooper, 1961, cited in , Doumpos and Zopounidis, 2002, p. 40). Charnes and Copper worked further on this with Rhodes which ultimately led to the development of the first Data Envelopment model called the Charnes Cooper and Rhodes (CCR) model which will be introduced in this section.

### 8.2.1. Transforming the Fractional Problem

As seen earlier, a DMU's efficiency is determined by the relation between its inputs and outputs. When considering a series of *n* DMUs with m inputs and s outputs, their input and output data can be placed in two matrices, one matrix X for the inputs and a matrix Y for the outputs. This is illustrated in Figure 8.5.

$$X = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ . & . & \cdots & . \\ . & . & \cdots & . \\ x_{m1} & x_{m2} & \cdots & x_{mn} \end{pmatrix}$$

$$Y = \begin{pmatrix} y_{11} & y_{12} & \cdots & y_{1n} \\ y_{21} & y_{22} & \cdots & y_{2n} \\ . & . & \cdots & . \\ . & . & \cdots & . \\ y_{s1} & y_{s2} & \cdots & y_{sn} \end{pmatrix}$$

$where:$

$x_{m1}\ is\ the\ m^{th}\ input\ of\ DMU_1,$

$x_{2n}\ is\ the\ 2^{nd}\ input\ of\ DMU_n,$

$y_{2n}\ is\ the\ 2^{nd}\ output\ of\ DMU_n\ and$

$y_{sn}\ is\ the\ s^{th}\ output\ of\ DMU_n.$

Figure 8.5: Inputs and outputs data matrices

Following the total factor productivity ratio introduced in 4.2.1 Performance Ratio and the above notation, it is possible to formulate a virtual performance ratio as in Figure 8.6.

$$Virtual\ ratio_{DMU_O} = \frac{u_1 \times y_{1_O} + \ldots + u_s \times y_{s_O}}{v_1 \times x_{1_O} + \ldots + v_m \times x_{m_O}}$$

$where:$

$y_s is\ DMU_O's\ s^{th} output\ and$

$x_m is\ DMU_O's\ m^{th} input.$

Figure 8.6: The virtual ratio

Further expanding this approach, Charnes Cooper and Rhodes (1978, p. 430) defined the CCR fractional problem (FP) as in Figure 8.7.

$$CCR_{FP_O} \quad max_{v,u}\, \theta = \frac{u_1 y_{1_O} + u_2 y_{2_O} + \cdots + u_s y_{s_O}}{v_1 x_{1_O} + v_2 x_{2_O} + \cdots + v_m x_{m_O}}$$

$$subject\ to\ (s.t.):$$

$$\frac{u_1 y_{1_O} + \cdots + u_s y_{s_O}}{v_1 x_{1_O} + \cdots + v_m x_{m_O}} \leq 1\ (j = 1, \dots, n)$$

$$v_1, v_2, \dots, v_m\ \geq 0$$

$$u_1, u_2, \dots, u_s\ \geq 0$$

$$where\ u\ and\ v\ are\ resp.\ DMU_O'\!s\ output\ and\ input\ weight\ vectors.$$

**Figure 8.7: The CCR model in its fractional form**

The variables of this fractional problem are '*u*' and '*v*' which are respectively the output and input weight vectors ($u = (u_1, u_2, \dots, u_s)$ $while$ $v = (v_1, v_2, \dots, v_m)$). The model was expressed in a sum (Σ) notation in the CCR paper although for clarity purposes the formulation has here been expanded. The notation in the CCR paper ('*s*' as number of outputs, '*m*' as number of inputs, '*n*' DMUs and $DMU_O$ as the DMU under examination) is generally accepted and used in DEA's literature. Because the above model formulation only measures $DMU_O$'s efficiency, the problem will have to be solved *n* times to measure all the DMUs' efficiency. Linear programming optimisation techniques are used to find the optimal solutions (e.g. the simplex algorithm).

The first constraint ensures that the maximum possible value for the ratio is 1. The last two constraints restrict all the inputs and outputs to be non-negative. Some inputs are allowed to be equal to 0 although at least one input (or output) will need to have a positive value per input vector (or output vector). The input and output vectors are said to be semi-positive.

The formulation above measures the efficiency of $DMU_O$. As illustrated by the first constraint, the set of weights used in the objective function (the '*max*' line) will be used with each of the other DMU's values as per illustrated by the first constraint. This constraint specifies that the set of weights assigned to $DMU_O$ and used with any other DMU's values, must be lower or equal to 1. This particular constraint is responsible for enveloping the data and identifying the efficient frontier.

Because the above formulation is a fractional problem, solving it can be quite difficult. In order to resolve this issue and make the most of the advances in linear computations, the model can be transformed to its linear problem form (LP) as in Figure 8.8.

$$\begin{aligned}
&CCR_{LP_O} \quad max_{v,\mu}\, \theta = \mu_1 y_{1_O} + \mu_2 y_{2_O} + \cdots + \mu_s y_{s_O} \\
&s.t. \\
&\nu_1 x_{1_O} + \nu_2 x_{2_O} + \cdots + \nu_m x_{m_O} = 1 \\
&\mu_1 y_{1_O} + \cdots + \mu_s y_{s_O} \le \nu_1 x_{1_O} + \cdots + \nu_m x_{m_O} \ (j = 1, \dots, n) \\
&\nu_1, \nu_2, \dots, \nu_m \ \ge 0 \\
&\mu_1, \mu_2, \dots, \mu_s \ \ge 0
\end{aligned}$$

**Figure 8.8: The CCR model in its linear form**

Note that the 'weight' variables are now $\nu$ (nu), $\mu$ (mu) instead of $v$ and $u$ as with the fractional problem ($\nu = v \times t \ and \ \mu = u \times t \ where \ t \ \sum_{i=1}^{m} v x_{i_O} = 1$). The two models ($CCR_{FP}$ and $CCR_{LP}$) are equivalent (i.e. they will have the same optimal solution). The transformation from fractional to linear problem was first introduced by Charnes and Cooper (1962) and is called the Charnes Cooper transformation (see Appendix 4: The Charnes Cooper transformation). Most fractional models are transformed with similar technique in order to express them in their linear form.

Because the linear form is easier to solve, most DEA problems are generally expressed in their linear form.

***Definition of CCR efficiency***: A DMU will be efficient if and only if its optimal value $\vartheta$ is 1 and there exist at least one optimal (v*, u*) with v* > 0 and u* > 0 (Cooper et al., 2007, p. 24).

### 8.2.2. Optimal Weights

The optimisation calculates the best possible set of weights for $DMU_O$'s performance ratio. The $v^*$ and $u^*$ obtained for the $LP_O$ are a set of optimal weight found for this $DMU_O$. $v_i^*$ is the optimal weight for input $x_i$ . This vector illustrates how important the input $x_i$ is relatively to other inputs. Similarly $u_r^*$ is the optimal weight for output $y_r$ and reflects how important output $y_r$ was evaluated by the optimisation process. By examining each input and output optimal weight, it is possible to know which input or output contributed to the performance level, but also to see the extent to which they contributed. It is however important to keep in mind that optimal weights are not always unique.

The meaning of weight can be easily illustrated with a small 2 inputs, 1 output example. Data is as in Table 8.2.

| DMU | A | B | C | D | E | F | G |
|---|---|---|---|---|---|---|---|
| Input 1 | 6 | 5 | 1.5 | 2.5 | 3 | 1.5 | 1 |
| Input 2 | 1.5 | 1 | 2 | 1 | 3 | 6.5 | 3 |
| Output | 1 | 1 | 1 | 1 | 1 | 1 | 1 |

Table 8.2: Two inputs – one output weight example

Which can be plotted as illustrated in Figure 8.9.

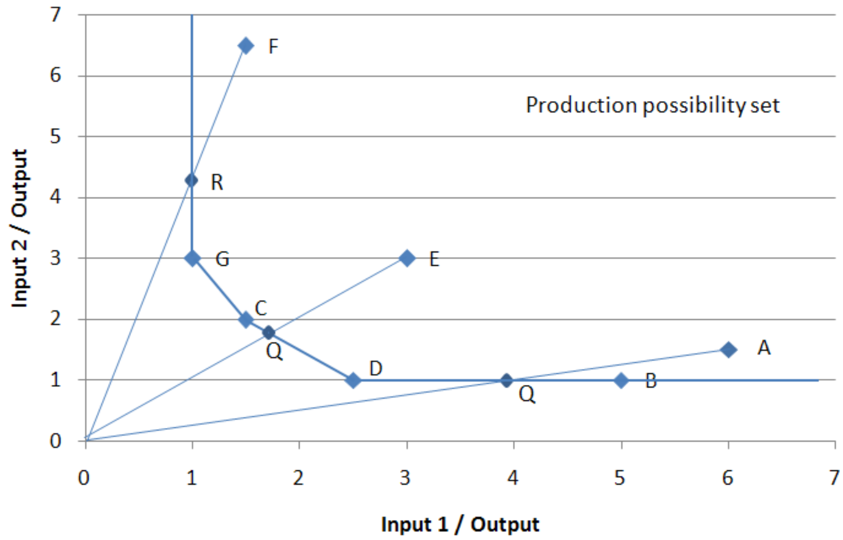


Figure 8.9: Two inputs – one output weight example graph

Because there are two inputs and one output, the data has been normalised by the output values and the efficient DMUs will be those which use the less input. Appraising the DMUs' performance with a CCR model (input oriented, i.e. the optimisation process will try reducing inputs while keeping output level constant) gives the performance levels illustrated in Table 8.3.

| DMU | Score | Weight Input 1 | Weight Input 2 | Weight Output |
|---|---|---|---|---|
| A | 0.666667 | 0 | 0.666666667 | 0.666666667 |
| B | 1 | 0 | 1 | 1 |
| C | 1 | 0.4 | 0.2 | 1 |
| D | 1 | 0.285714286 | 0.285714286 | 1 |
| E | 0.583333 | 0.166666667 | 0.166666667 | 0.583333333 |
| F | 0.666667 | 0.666666667 | 0 | 0.666666667 |
| G | 1 | 1 | 0 | 1 |

Table 8.3: Two inputs – one output optimisation results

$DMU_C$ has two different weights $v_1^* = 0.4\ and\ v_2^* = 0.2$. The ratio $v_1^* \,/\, v_2^* = 2$ seems to indicate that it is more advantageous for $DMU_C$ to weight $Input_1$ twice as much as $Input_2$ when maximising the virtual performance ratio. These values are important to measure the sensitivity of efficiency scores as for example with the Thompson approach (see section 5.4.6 Sensitivity Analysis for more details on this).

$DMU_B$ has an efficiency value of 1. However, one of its weights is equal to 0 (($v_1^*$).) which in respect to the CCR definition above (see ***Definition of CCR efficiency*** above) indicates that unless there is another set of non-zero weights, this DMU is not efficient. In fact, $DMU_B$'s inefficiency can be easily spotted by considering its relative performance to $DMU_D$ which produces the same level of output with half of $DMU_B$'s $Input_1$.

$DMU_G$ on the other hand has also an efficiency score of 1 and a weight equals to zero ($v_2^*$). However, this DMU is efficient as there is another set of optimal non-zero weights. This is exemplified by the set of optimal non-zero weights: $v_1^* = \frac{2}{3} and\ v_2^* = \frac{2}{18}$.

These specificities are not always easy to identify although the dual of the CCR model helps recognising them (during the second computational phase which optimises the slacks). This will be discussed further down (see section 8.2.4 Dual Problem and computational aspects).

### 8.2.3. Production Possibility Set

As explained earlier, the CCR model makes a semi-positive condition on the data. That is the input and output vector values are greater or equal to zero as long as there is at least one positive value in each vector. This can be written as follows for the input vector ($x_j \geq 0, x_j \neq 0\ for\ j = 1, \dots, n$). A pair of input and output vectors is called an *activity* and can be written as follows: $(x, y)\ with\ x \in R^m\ and\ y \in R^s$.

Given this notation, the production possibility set is defined by Formula 8.1 (Yun et al., 2004, p. 90).

$$P = \{(x, y) \mid Y\lambda \geq y, X\lambda \leq x, \lambda \geq 0\}$$
$$where\ is\ a\ semipositive\ vector\ in\ R^n = R^m + R^s.$$

**Formula 8.1: Production possibility set**

Cooper *et al* (2007, p. 42) also list the following properties of the reference set P:

'The observed activities $(x_j, y_j)$ $(j = 1, \dots, n)$ belong to *P*.'

'If an activity (*x*, *y*) belongs to *P*, any semi-positive activity then the activity (t*x*, t*y*) belongs to *P*.' This property is called *constant return to scale assumption* and will consequently only be true for models assuming constant returns to scale.

'For an activity (*x*, *y*) in *P*, any semi-positive activity $(\bar{x}, \bar{y})$ with $\bar{x} \geq x\ and\ \bar{y} \leq y$ is included in *P*.' This means that any activity that uses the same of input or more while producing the same of less output than any other activity, will be included in the production possibility set and is thus considered feasible.

'Any semi-positive linear combination of activities in *P* belongs to *P*' (this is the generalisation of the second property).

This definition of the production possibility set is only applicable to the CCR model described in this section. Other DEA models will have different definition from their respective production possibility set. The BCC model for example – which works under variable returns to scale – has a specific convexity constraint added to the definition of its production possibility set.

### 8.2.4. Dual Problem and computational aspects

The models illustrated in this section are taken from the Cooper, Seiford and Tone book (Cooper et al., 2007).

The $CCR_{LP}$ model introduced in 'Figure 8.8: The CCR model in its linear form' can be expressed in a more concise vector (envelopment) form. This is illustrated by Figure 8.10

$$
\begin{aligned}
&(LP_O)\ max_{v,u}\, uy_O \\
&subject\ to \\
&vx_O = 1 \\
&-vX + uY \ \le \ 0 \\
&v \ge 0, u \ge 0
\end{aligned}
$$

**Figure 8.10: CCR model in the multiplier vector form**

A linear problem has an associated problem with it called the dual problem. The dual problem is really useful for knowing advanced characteristics of the problem. The original linear problem (from which the dual is worked out) is referred to as the primal. The primal and the dual problems have also the same optimal objective function values. Because the number of constraint and the number of variables are swapped between the primal and the dual problems, it is sometimes interesting to solve the dual problem to obtain a solution to a primal. More information on the relation between the primal and the dual problems can be found in 'Operations Research – Application and Algorithms' from Winston (2004, p. 295 onwards) or in other linear algebra books.

For computational reasons (these will be detailed further down below), it is best to solve the $CCR_{DLP}$ than the primal $CCR_{LP}$ problem. The $CCR_{DLP}$, also called the model in its envelopment form, can be written as in Figure 8.11.

$$
\begin{aligned}
&(DLP_O)\ \ min_{\theta,\lambda}\, \theta \\
&subject\ to \\
&\theta x_O - X\lambda \geq 0 \\
&Y\lambda \geq y_O \\
&\lambda \geq 0
\end{aligned}
$$

Figure 8.11: CCR model in dual form

$DLP_O$ has a feasible solution $\theta = 1, \lambda_O = 1, \lambda_j = 0\ (j \neq O)$. This implies that $\theta^*$ is lower or equal to 1. Similarly, because $\mathrm{Y} \geq 0, \mathrm{Y} \neq 0$, the second constraint implies that λ is positive. Because X is also positive, the first constraint forces $\theta$ to be strictly positive. This consequently implies $0 < \theta^* \leq 1$. The optimisation process tries to reduce the inputs in a radial manner while staying in the production possibility set. Thus, it is possible to say that some activities in (*X*λ, *Y*λ) outperform $(\theta x_O, y_O)$ when $\theta^* < 1$ (Cooper et al., 2007, p. 44). In light of this property, it is possible to define the input excesses ($s^-$) and output shortfalls ($s^+$) – or slacks – as in Figure 8.12.

$$
\begin{aligned}
&s^- = \theta x_O - X\lambda, \\
&s^+ = Y\lambda - y_O \\
&with\ s^- \geq 0\ and\ s^+ \geq 0
\end{aligned}
$$

Figure 8.12: CCR model and slacks

It is consequently possible to solve the LP problem in two phases, the first phase aiming at minimising $x_O$ by maximising $\theta$, and a second phase trying to maximise $s^-$ and $s^+$.

Figure 8.13 illustrates the LP to obtain $\theta^*, s^{-*}, s^{+*}$.

$$
\begin{aligned}
&(DLP_O)\\
&Phase\ I\ objective: \min\theta\\
&Phase\ II\ objective: \max\ es^- +\ es^+\\
&subject\ to:\\
&\theta x_O = X\lambda + s^-\\
&\ y_O = Y\lambda -\ s^+\\
&where\ \theta \geq 0, \lambda \geq 0, s^- \geq 0, s^+ \geq 0
\end{aligned}
$$

**Figure 8.13: Computation phases of $CCR_{DLPO}$**

An optimal solution $\theta^*, s^{-*}, s^{+*}$ obtained after solving Phase II is called the max slack solution. The solution obtained is not systematically unique (the score obtained is an optimum so is unique; however, the lambdas and weights might not be unique).

The weights variables (*v*, *u*) from the $LP_O$ problem in its multiplier form can be derived from the columns corresponding to the inputs and outputs slacks in the identity matrix computed while solving the $DLP_O$'s. If applying the simplex criteria to these two columns gives vectors $p^{s-}\ and\ p^{s+}$, the *u* and *v* weight vectors are then given by the relation expressed in Formula 8.2.

$$v^* =\ -p^{s-}\ and\ u^* =\ -p^{s+}$$

**Formula 8.2: Relation between *v*, *u* and the $DLP_O$**

There are several reasons to solve the $DLP_O$ instead of the $LP_O$. In a DEA problem, the number of DMUs is generally greater than the number of constraints. Because the computational effort to solve a LP is likely to increase with the number of constraints, it will require less computer effort to solve the DLP (which number of variables is equal to the number of DMUs) rather than the LP (which number of constraint is equal to the number of DMUs). Cooper *et al* (2007, p. 52) also point out that the memory used to store the basis or its inverse is the square of the number of constraints (the basis is a specific matrix in the simplex linear solving algorithm). The $DLP_O$ will consequently use less memory to solve than the LP. The LP problem does not also allow finding the max slack solution. Finally, the relation between the DLP and the original data is more straightforward than with LP where the solution is a (multiplied) evaluation of the data.

The reference set $E_O$ of $DMU_O$ is defined as by Formula 8.3 (based on the max slack solution described above).

$$E_O = \{j \mid \lambda_j^* > 0\} \ (j \in \{1, \dots, n\})$$

**Formula 8.3: Reference set**

$DMU_O$'s potential improvement (indicated with the symbol Δ) can be calculated with Formula 8.4.

$$\Delta x_O = x_O - (\theta^* x_O - s^{-*}) = (1 - \theta^*)\, x_O + s^{-*}$$
$$\Delta y_O = s^{+*}$$

**Formula 8.4: Possible improvement formulae**

This means that potential improvements in input ($\Delta x_O$) can be made by reducing the technical inefficiency quantified by $\theta^*$ and by reducing any input slack $s^{-*}$. Improvement in output can be made by reducing the output slack $s^{+*}$.

The projections on the efficient frontier (i.e. the point on the frontier which an inefficient DMU needs to reach in order to become efficient – indicated by the symbol '$\widehat{\ }$') are calculated as in Formula 8.5.

$$\hat{x}_O = x_O - \Delta x_O = \theta x_o - s^-$$
$$\hat{y}_O = y_O + \Delta y_O = y_O + s^{+*}$$

**Formula 8.5: Projection formulae**

### 8.2.5. Model Orientation

The CCR models introduced earlier (both multiplier and envelopment forms) were all optimising by reducing the inputs while keeping the outputs constant; this is called input orientation. There is also the possibility however, to try maximising the outputs while keeping the inputs constant; this is called output orientation. The CCR model in its output orientation form (CCR-O instead of CCR-I) is formulated as in Figure 8.14: CCR output oriented model (Cooper et al., 2007).

$$(DLPO_O)\ \max_{\eta,\mu} \eta$$
$$subject\ to$$
$$x_O - X\mu \geq 0$$
$$\eta y_O - Y\mu \leq 0$$
$$\mu \geq 0$$

**Figure 8.14: CCR output oriented model**

An optimal solution to the $DLPO_O$ can be calculated from an optimal solution of the $DLP_O$ (the CCR input oriented model) as in Formula 8.6.

$$\lambda = \mu/\eta$$
$$\theta = 1/\eta$$

**Formula 8.6: Relation between $DLPO_O$ and $DLP_O$**

Replacing $\eta$ and $\mu$ in the $DLPO_O$ model transforms it into the $DLP_O$ model, thus it is possible to write relation in Formula 8.7.

$$\eta^* = 1/\theta^*$$
$$\mu^* = \lambda^*/\theta^*$$

**Formula 8.7: Relation between $DLPO_O$'s optimal solution and $DLP_O$'s**

The optimal slacks of the output oriented model ($t^{-*}, t^{+*}$) are defined as per in the relation expressed in Formula 8.8.

$$t^{-*} = s^{-*}/\theta^*$$
$$t^{+*} = s^{+*}/\theta^*$$

**Formula 8.8: Relation between $DLPO_O$'s slacks and $DLP_O$'s**

Finally, the weights can be obtained by the relation expressed in Formula 8.9 (where q and p are resp. the input and output vectors of the $DLPO_O$ dual problem).

$$p^* = v^*/\theta^*$$
$$q^* = u^*/\theta^*$$

**Formula 8.9: Relation between $DLPO_O$'s weights and $DLP_O$'s**

The very close and simple relation between the CCR-I and the CCR-O are specific to the CCR model and other DEA models will not demonstrate such characteristics.

Model orientation can have impacts on the results and the decision of which orientation to retain should always deserve consideration. Although the CCR model only gives two possibilities for model orientation (input or output oriented), section 8.3 Appendix 3: Other DEA models will introduce other DEA models which will offer more choices regarding the model orientation.

## 8.3. Appendix 3: Other DEA models

This section will introduce the BCC and SBM model which were used in this study.

The CCR model introduced in section 8.2 was built under the assumption of Constant Returns To Scale. This meant that for any observed activity (x, y), it was possible to take a scalar t (where t is a positive scalar) and assume that production (tx, ty) was also possible (i.e. within the production possibility set). Many extensions of the CCR model have since be proposed (Cooper et al., 2007, p. 87) the representative being the model introduced by Banker, Charnes and Cooper (the BCC model). This model is built under a Variable Returns To Scale assumption which allows the production possibility set to span a convex hull around the data as illustrated in Figure 8.15.

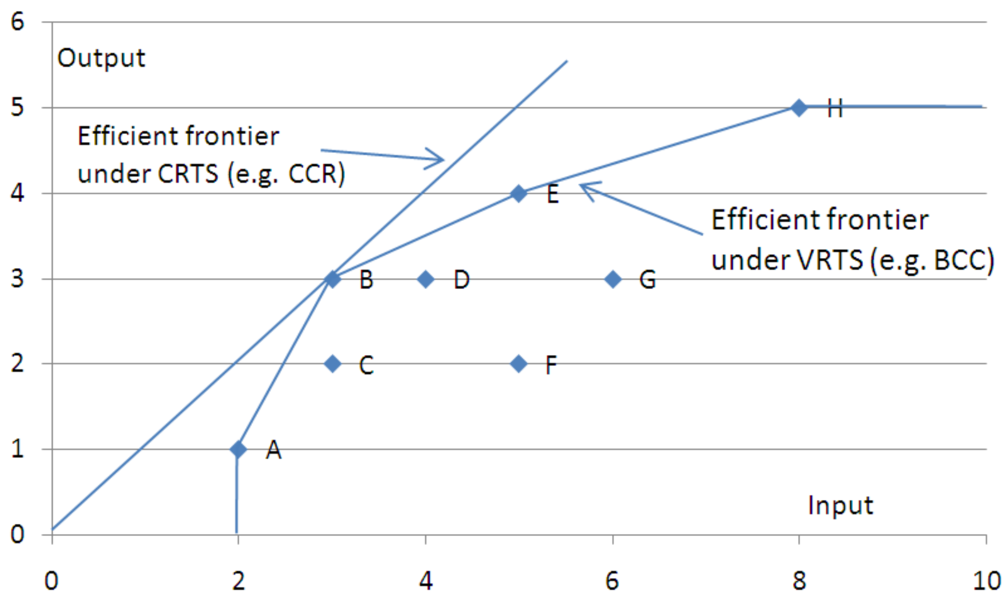


Figure 8.15: Illustration of the CCR and BCC frontiers

This section will introduce two models, the BCC and another called the Slack Based Model (SBM). The SBM model can have the same production possibility set as the CCR or the BCC model (depending whether it has a 'convexity' constraint) but treats slacks in a different manner. There are many more DEA models than the three CCR, BCC and SBM models although these were not used in this study so will not be introduced here.

### 8.3.1. BCC Model

The difference between the CCR and the BCC model can be illustrated by the following small example.

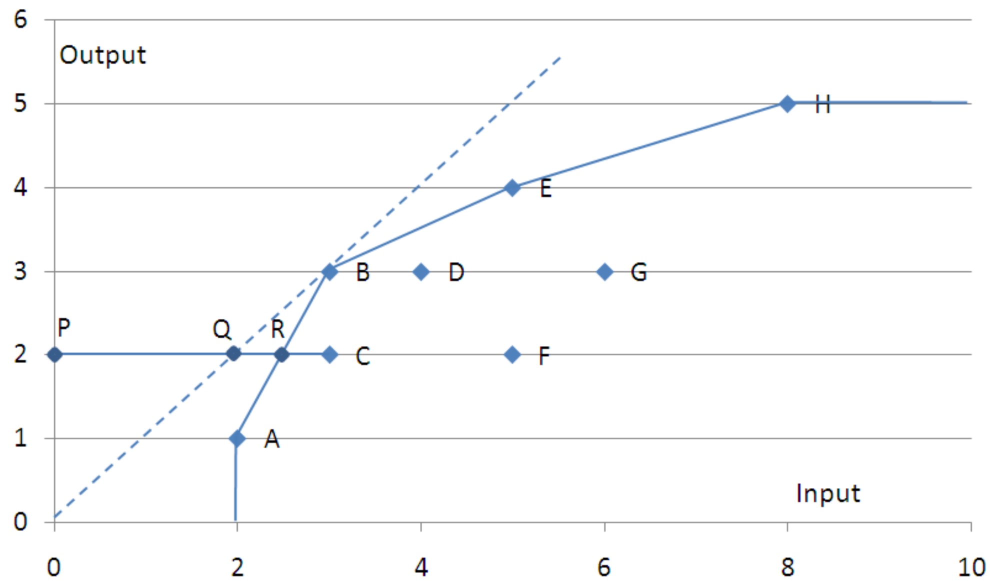


Figure 8.16: CCR and BCC efficiencies

In Figure 8.16, the BCC's efficiency for DMU C is given by the Formula 8.10 (the model is input oriented):

$$DMU_C\ BCC\ efficiency = \frac{PR}{PC} = \frac{5/2}{3} = \frac{5}{6}$$

Formula 8.10: BCC efficiency

While the CCR efficiency is given by Formula 8.11.

$$DMU_C\ CCR\ efficiency = \frac{PQ}{PC} = \frac{2}{3}$$

Formula 8.11: CCR efficiency

Similarly, the output oriented efficiency of DMUc is given by the following ratio BS / CS as illustrated in Figure 8.17.

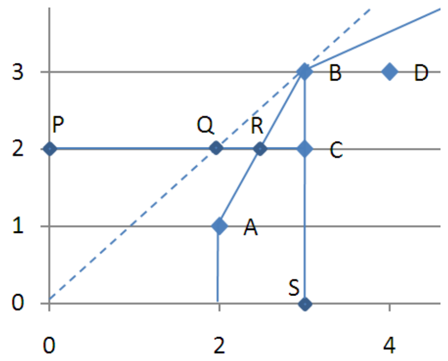


Figure 8.17: BCC output efficiency

Banker Cooper and Charnes define the BCC's production possibility set as in Figure 8.18.

$$P_B = \{(x, y) \mid x \geq X\lambda, y \leq Y\lambda, e\lambda = 1, \lambda \geq 0\}$$
$$where\ X = (x_j) \in R^{m \times n}, Y = (y_j) \in R^{s \times n}\ are\ a\ given\ dataset$$
$$\lambda \in R^n$$
$$and\ e\ is\ a\ row\ vector\ 1$$

Figure 8.18: The BCC production possibility set

The main difference between the BCC and the CCR model resides in the convexity constraint given by $e\lambda = 1$.

The input oriented BCC model is defined as in Figure 8.19.

$$(BCC_O) \;\; min_{\theta,\lambda}\,\theta$$
$$subject\ to$$
$$\theta x_O - X\lambda \geq 0$$
$$Y\lambda \geq y_O$$
$$e\lambda = 1$$
$$\lambda \geq 0$$

Figure 8.19: The BCC model

Its dual is expressed in Figure 8.20.

$$BCC_O \;\; max_{v,u}\, uy_O - \mathrm{u}_0$$
$$subject\ to$$
$$vx_O = 1$$
$$-vX + uY - u_0 \leq 0$$
$$v \geq 0, u \geq 0, u_0 free\ in\ sign$$

Figure 8.20: The dual of the BCC model

The reference set $E_O$ is given from an optimal solution $\lambda^*$ as in Formula 8.12.

$$E_O = \{j|\, \lambda_j^* > 0\}\, (j \in \{1, \ldots, n\})$$

Formula 8.12: The BCC reference set

Formula 8.13 gives the following projections.

$$\widehat{x_O} \Leftarrow \theta^* x_O - s^{-*}$$
$$\widehat{y_O} \Leftarrow y_O + s^{+*}$$

Formula 8.13: The BCC projection formula

The BCC model is solved in two phases: the first phase attempts to maximise $\vartheta$ while the second aims at maximising slacks (while keeping $\vartheta^*$ constant). A DMU is said efficient when $\vartheta = 1$ and there is no non-zero slack.

The BCC model can also be re-written in its output oriented form as in Figure 8.21 (this form is not used in the study):

$$
\begin{aligned}
&(BCC_O - O_O) \quad max_{\eta,\lambda}\, \eta \\
&subject\ to \\
&X\lambda \leq x_O \\
&\eta y_O - Y\lambda \leq 0 \\
&e\lambda = 1 \\
&\lambda \geq 0
\end{aligned}
$$

Figure 8.21: The BCC model in its output oriented form

### 8.3.2. SBM Model

Both the CCR and the BCC model separate the technical inefficiencies (measured by 1 – $\vartheta$ for input oriented models and $\vartheta$ – 1 for output oriented models) from the mix inefficiencies which can only be reduced by changing the respective proportions of inputs and outputs. This approach suits well models which inputs (or outputs) are expected to behave in a radial manner (i.e. where it is possible to reduce (or increase) all inputs (or outputs) simultaneously). However it is less appropriate for models where such an assumption cannot be made. The SBM model, first introduced by Tone (2001) solves this issue by simultaneously considering input reduction and output increase whilst being unit invariant.

The SBM is formulated as in Figure 8.22.

$$
\begin{aligned}
&(SBM) min_{\lambda,s^-,s^+}\, \rho = \frac{1 - \frac{1}{m}\sum_{i=1}^{m} s_i^- / x_{iO}}{1 + \frac{1}{s}\sum_{r=1}^{s} s_r^+ / y_{rO}} \\
&subject\ to: \\
&x_O = X\lambda + s^- \\
&y_O = Y\lambda - s^+ \\
&\lambda \geq 0, s^- \geq 0, s^+ \geq 0
\end{aligned}
$$

Figure 8.22: The SBM model

It is assumed that X ≥ 0, that if $\mathrm{x_{io}} = 0$ the term $\mathrm{s_i^-}/\mathrm{x_{io}}$ is deleted and that if $\mathrm{y_{io}} \leq 0$ $\mathrm{y_{io}}$ is replaced by a very small number so that it plays the role of penalty.

The Fractional SBM is transformed into a linear problem by the introduction of a small positive scalar *t* as in Figure 8.23.

$$SBM_t \; min_{t,\lambda,s^-,s^+} \; \tau = t - \frac{1}{m}\sum_{i=1}^{m} t\, s_i^-/x_{io}$$

$$subject\ to:$$

$$1 = t + \frac{1}{s}\sum_{r=1}^{s} t\, s_r^+/y_{ro}$$

$$x_o = X\lambda + s^-$$

$$y_o = Y\lambda - s^+$$

$$\lambda \geq 0, s^- \geq 0, s^+ \geq 0, t > 0$$

**Figure 8.23: The SBM model with scalar t**

The previous model can be transformed to the linear model as in Figure 8.24.

$$SBM_{LP} \; min \; \tau = t - \frac{1}{m}\sum_{i=1}^{m} S_i^-/x_{io}$$

$$subject\ to:$$

$$1 = t + \frac{1}{s}\sum_{r=1}^{s} S_r^+/y_{ro}$$

$$tx_o = X\Lambda + S^-$$

$$ty_o = Y\Lambda - S^+$$

$$\Lambda \geq 0, S^- \geq 0, S^+ \geq 0, t > 0$$

**Figure 8.24: The SBM model in its linear form**

Where $S^-$ = $ts^-$, $S^+$ = $ts^+$ and Λ = tλ.

Equivalence between the two models can be found in Cooper *et al* (2007, p. 102). The SBM model is solved in one phase only as technical and mix inefficiency are not separated in this model.

A DMU is SBM efficient if and only if $\rho = 1$. This only happens when there is no input and output slack.

The projections are given by Formula 8.14.

$$\widehat{x_o} \Leftarrow x_o + s^{-*}$$
$$\widehat{y_o} \Leftarrow y_o - s^{+*}$$

Formula 8.14: SBM projections

The reference set can be expressed as in Formula 8.15.

$$R_o = \{j \mid \lambda_j^* > 0\} \ (j \in \{1, \dots, n\})$$

Formula 8.15: SBM Reference set

The SBM model can be transformed to its input (or output) oriented form by ignoring the nominator (or numerator). The input oriented form is illustrated as in Figure 8.25.

$$SBM - I \ \ \rho_I^* = min_{\lambda,s^-} \quad 1 - \frac{1}{m}\sum_{i=1}^{m} s_i^- / x_{io}$$
$$subject\ to:$$
$$x_o = X\lambda + s^-$$
$$y_o \leq Y\lambda$$
$$\lambda \geq 0, s^- \geq 0$$

Figure 8.25: The SBM model in its input oriented form

And the output oriented form is illustrated as in Figure 8.26.

$$SBM - O \;\; \rho_O^* = min_{\lambda,s^-} \quad \frac{1}{1 + \frac{1}{s}\sum_{r=1}^{s} s_r^+ / y_{ro}}$$

$$subject\ to:$$

$$x_o \geq X\lambda$$

$$y_o = Y\lambda - \;\; s^+$$

$$\lambda \geq 0, s^+ \geq 0$$

**Figure 8.26: The SBM model in its output oriented form**

Comparing the SBM score to the CCR and BCC score give indications on each DMU's scale and mix efficiencies. These ratios are important determining whether a unit's efficiency levels are due to its position in the production possibility set (under increasing, constant or decreasing RTS), or to poor performance only.

### 8.3.3. Summary

Several important model characteristics were discussed while introducing the CCR, BCC and SBM models. It was for example explained that the CCR model was developed under a semi-positive assumption which assumes that there is at least one element of the dataset which is (strictly) positive. This section will briefly summarise each model's characteristics.

Translation invariance is another fundamental property of (some) DEA models. It allows an axis to be shifted which is particularly useful some variables in a dataset are negative (a scalar is then added to all the variables; thus making the negative variable positive). As illustrated below, not all DEA models are translation invariant.

Unit invariance is another essential property of DEA models. A unit invariant model will provide the same results regardless of which unit is used. In this particular example this means that the same fuel efficiency level will be measured should kilometres be used instead of miles or should litres per hundred kilometres be used instead of mpg.

Finally, the CCR and BCC model principally measure technical efficiency (although the slacks can be calculated in the second phase) while the SBM model considers mix efficiency.

Although these characteristics will not be often used in this study, it was essential to mention them. Table 8.4 (Cooper et al., 2007, p. 115) summarises each model's characteristics in regards to these properties.

| Model | | CCR-I | CCR-O | BCC-I | BCC-O | SBM |
|---|---|---|---|---|---|---|
| Data | X | Semi-p | Semi-p | Semi-p | Free | Semi-p |
| | Y | Free | Free | Free | Semi-p | Free |
| Trans. Invariance | X | No | No | No | Yes | No |
| | Y | No | No | Yes | No | No |
| Units invariance | | Yes | Yes | Yes | Yes | Yes |
| Tech. Or mix | | Tech. | Tech. | Tech. | Tech. | Mix |
| RTS | | CRS | CRS | VRS | VRS | CRS/VRS |

Table 8.4: Summary table of models' properties

## 8.4. Appendix 4: The Charnes Cooper transformation

The Charnes Cooper transformation (Charnes and Cooper, 1962) enables to transform a fractional problem to a linear problem.

Starting with the fractional problem illustrated in Figure 8.27.

$$CCR_{FP_O} \quad max_{v,u}\, \theta = \frac{\sum_{r=1}^{s} u_r y_{r_O}}{\sum_{i=1}^{m} v_i x_{i_O}}$$

$$subject\ to\ (s.t.):$$

$$\frac{\sum_{r=1}^{s} u_r y_{r_j}}{\sum_{i=1}^{m} v_i x_{i_j}} \leq 1\ (j = 1, \dots, n)$$

$$u_r, v_r, \geq 0\ \forall\, r, i$$

Figure 8.27: The CCR fractional problem

It is possible to take a variable t as illustrated in Figure 8.28.

$$t \sum_{i=1}^{m} v_i x_{i_O} = 1$$

Figure 8.28: Defining the variable *t*

Because $\nu_i \geq 0$ and $x_{iO} \geq 0\ \forall$ i (semi-positive assumption) t has to be positive so that it is possible to multiply both numerator and denominator of a ratio without changing its value. This is illustrated in Figure 8.29.

$$\frac{\sum_{r=1}^{s} u_r y_{r_O}}{\sum_{i=1}^{m} v_i x_{i_O}} = \frac{\mathrm{t}\ \sum_{r=1}^{s} u_r y_{r_O}}{\mathrm{t}\ \sum_{i=1}^{m} v_i x_{i_O}} = \frac{\sum_{r=1}^{s} t u_r y_{r_O}}{\sum_{i=1}^{m} t v_i x_{i_O}}$$

$$and\ that$$

$$\mu_r = t u_r, r = 1, \dots, s$$

$$\nu_i = t v_i, i = 1, \dots, m$$

Figure 8.29: Multiplying the ratio by *t*

The original problem as consequently been replaced by the equivalent problem in Figure 8.30.

$$\max \theta = \sum_{r=1}^{s} \mu_r y_{rO}$$

$$subject\ to:$$

$$\sum_{i=1}^{m} v_i x_{iO} = 1$$

$$\sum_{r=1}^{s} \mu_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} \leq 0, j = 1, \ldots, n$$

$$\mu_r v_i \geq 0 \; \forall \; r, i$$

Figure 8.30: Transforming the CCR problem

Which is the $CCR_{LP}$ model introduced in section 8.2 and illustrated in Figure 8.31.

$$CCR_{LP_O} \quad max_{v,\mu}\, \theta = \mu_1 y_{1_O} + \mu_2 y_{2_O} + \cdots + \mu_s y_{s_O}$$

$$s.t.$$

$$v_1 x_{1_O} + v_2 x_{2_O} + \cdots + v_m x_{m_O} = 1$$

$$\mu_1 y_{1_O} + \cdots + \mu_s y_{s_O} \leq v_1 x_{1_O} + \cdots + v_m x_{m_O} \; (j = 1, \ldots, n)$$

$$v_1, v_2, \ldots, v_m \geq 0$$

$$\mu_1, \mu_2, \ldots, \mu_s \geq 0$$

Figure 8.31: The CCR problem in its linear form

## 8.5. Appendix 5: Smoothing Algorithm Calculations Example

This appendix illustrates the behaviour of the smoothing algorithm with real data.

The Smoothing Algorithm behaviour can be illustrated by reproducing each step of the algorithm calculations with real vehicle's data. The measurement period starts on the 2009-04-01 00:00 and stops on the 2009-06-30 23:59:59. During this period, the test vehicle (a Citroen Berlingo of 1650kg) refills as illustrated in Table 8.5 (transactions are in litres):

| VehCode | TransactionDate | RefuelQuantity |
|---|---|---|
| 1 | 2009-04-15 00:00:00 | 34.57 |
| 1 | 2009-04-27 00:00:00 | 46.94 |
| 1 | 2009-05-11 18:03:00 | 17.91 |
| 1 | 2009-05-18 07:40:00 | 38.91 |
| 1 | 2009-05-27 15:34:00 | 36.89 |
| 1 | 2009-06-08 06:53:00 | 42.06 |
| 1 | 2009-06-22 07:42:00 | 32.44 |
| 1 | 2009-06-29 09:24:00 | 35.54 |

Table 8.5: Fuel Transactions

**Calculating the 'Bad MPG'**

The refill quantity sums up to 285.26 litres for a distance travelled of 3367.75 miles. Formula 8.16 gives the calculations for the 'Bad MPG'.

$$Bad\ MPG = \frac{3367.75}{285.26/4.54609188} = \frac{3367.75}{62.74} = 53.67$$

Formula 8.16: Bad mpg calculation

**Calculating the accurate mpg**

The amount of fuel used between the first and the last transaction is 250.69 (285.26 – 34.57) which is equal to 54.14 gallons.

The distance travelled between the first and last transaction is 2823.22 miles.

The 'actual mpg' is consequently 52.14.

**Calculating the smoothed volume**

The vehicle has travelled 602.56 miles between the start of the period and the first refill. Similarly, the vehicle has travelled 123.57 miles between the last refill and the end of the period.

The smoothed volume calculations are as in Formula 8.17.

$$Smoothed\ Volume = \sum_{i=2}^{n} Volume\ of\ refill_i + \frac{D_1}{mpg_a} + \frac{D_{n+1}}{mpg_a}$$
$$Smoothed\ Volume = 54.14 + \frac{602.56}{52.14} + \frac{123.57}{52.14}$$
$$Smoothed\ Volume = 54.14 + 11.55 + 2.37$$
$$Smoothed\ Volume = 68.06\ gallons\ (309.4\ litres)$$

Formula 8.17: Smoothed Volume formula

The smoothed distance can finally be calculated as in Formula 8.18.

$$Smoothed\ mpg = \frac{Total\ Distance\ Travelled}{Smoothed\ Volume}$$
$$Smoothed\ mpg = \frac{3367.75}{68.06} = 49.482$$

Formula 8.18: Smoothed mpg formula

The C# algorithm found a similar result as illustrated in Table 8.6.

| Difference | VehCode | Bad MPG | Smoothed MPG | Fuel Drawn | Fuel Used |
|---|---|---|---|---|---|
| 4.18 | 1 | 53.671 | 49.49 | 285.26 | 309.356 |
| 1.81 | 2 | 53.025 | 51.211 | 248.76 | 257.569 |
| -0.51 | 3 | 51.452 | 51.963 | 433.45 | 429.187 |
| 0.25 | 4 | 43.252 | 42.999 | 121.39 | 122.104 |
| 0.43 | 5 | 50.554 | 50.124 | 248.74 | 250.876 |
| -1.37 | 6 | 42.361 | 43.726 | 170.45 | 165.129 |

Table 8.6: Smoothed mpg results

This vehicle probably did refill just before the start of the period (this information was unfortunately not recorded). It was thus able to travel many miles (600) before needing refilling which artificially increased its (incorrect) mpg value.

## 8.6. Appendix 6: Details on non-controllable and non-discretionary models

This section lists a few non-controllable and non discretionary models in various forms. Transformations from fractional to linear forms use a similar approach as the Charnes Cooper transformation of fractional programming (see 8.4 Appendix 4: The Charnes Cooper transformation and (Charnes and Cooper, 1962)).

The Fractional SBM-NC model can be illustrated as in Figure 8.32.

$$(FSBM - NC)\min\rho = \frac{1 - \frac{1}{m}\sum_{i=1}^{m} s_i^- / x_{iO}}{1 + \frac{1}{s}\sum_{r=1}^{s} s_r^+ / y_{rO}} \quad i, r\ \in C$$

$$m\ and\ s\ are\ respectively\ the\ inputs\ and\ outputs\ in\ C$$

$$subject\ to:$$

$$x_O^C \geq X^C\lambda$$

$$x_O^N = X^N\lambda$$

$$y_O^C \leq Y^C\lambda$$

$$y_O^N = Y^N\lambda$$

$$\lambda\ \geq 0$$

Figure 8.32: The SBM-NC model in its fractional form

This formulation is similar to the Slack Based Model except that non-controllable variables were discarded from the objective and that the constraint matrix was partitioned to prevent any slacks for non-controllable variables.

The SBM-ND model can be found in Figure 8.33.

$$(FSBM - ND)\ \min \rho = \frac{1 - \frac{1}{m}\sum_{i=1}^{m} {s_i^-}/{x_{iO}}}{1 + \frac{1}{s}\sum_{r=1}^{s} {s_r^+}/{y_{rO}}} \quad i, r\ \in D$$

$m\ and\ s\ are\ respectively\ the\ inputs\ and\ outputs\ in\ D$

$subject\ to:$

$x_O \geq X\lambda$

$y_O \leq Y\lambda$

$\lambda\ \geq 0$

**Figure 8.33: The SBM-ND model in its fractional form**

## 8.7. Appendix 7: Subsidiary information about the companies

The following table summarises the different contacts at the different companies.

| Company Name | Contact Name | Position |
|---|---|---|
| **Avonline** | Alan Thatcher | Fleet manager |
| **Avonline** | Gary Woodhouse | Fleet support |
| **Carillion** | Patrick Nolan | Fleet manager |
| **FSH Maintenance** | Martin Smith | Stores |
| **FSH Maintenance** | Stuart Welburn | Senior finance manager |

All companies were first emailed on the 17th of June 2009. The companies which replied positively were all further contacted in regards to the data collection. The results were communicated on the 31st of August 2009 and were discussed during September 2009 with the different companies. Discussions were made over the phone as this was the most convenient alternative for the different fleet managers.

# 9. Glossary of Terms and Abbreviations

| Term or notion | Definition or explanation |
|---|---|
| Aerodynamic drag | The total resistance to an object through air (Slater, 2010). |
| Algorithm | Formula for problem solving (Slater, 2010). |
| Anti-Isotonic factor | Undesirable output or inhibiting input (Dyson et al., 2001). See isotonic factor q.v. |
| BCC | Banker Charnes Cooper model. See (Banker et al., 1984) and section 8.3.1 BCC Model. |
| BTAC | Abbreviation. British Transport Advisory Consortium |
| CAN | Abbreviation. Controlled Area Network. An electronic network as per defined by Bosch's specification (Bosch, 1991). CAN is made of a pair of twisted wire to which compliant Electronic Control Units (ECU, see ECU below q.v.) can connect to in order to send information across the same shared network. |
| CANbus | Controlled Area Network Bus. A CAN electronic bus or network. CANbus technology is now extensively used on vehicles as it enables different vehicle's electronic units to share data across a single or a limited number of networks |
| Car | Powered road vehicle designed to carry a driver and a small number of passengers (7 or under) (Slater, 2010). |
| CCR | Charnes Cooper and Rhode model. See (Charnes et al., 1978) and section 8.2. |
| Cobb Douglas functional form | The Cobb–Douglas functional form of production functions represents the relationship between some inputs and an output. It was proposed by Knut Wicksell (1851–1926), and tested against statistical evidence by Charles Cobb and Paul Douglas in 1900–1928 (Cobb and Douglas, 1928). |
| Data Envelopment Analysis | A non-parametric method for the estimation of the efficient production frontiers and measurement of efficiency by the mean of ratio and comparison to this efficient frontier. |

| | |
|---|---|
| DEA | Abbreviation. See Data Envelopment Analysis q.v. |
| Deterministic | To cause to occur in a particular manner (Slater, 2010). |
| DfT | Abbreviation. Department for Transport. The United Kingdom government department for transport. |
| ECU | Abbreviation. Electronic Control Unit |
| Effectiveness | The production of a required result (Slater, 2010). |
| Efficiency | The differential between outputs and inputs of a purposely conducted action or process.<br>See the **Key Concepts and Definitions** section for further information q.v. |
| Efficient Frontier | The piecewise linear set spanned by the collection of efficient DMUs. |
| Electronic Control Unit (ECU) | A CAN compliant device which can be connected to a Controlled Area Network. On modern vehicles, ECU are generally connected to electronic sensors and can retrieve, process and share key vehicle information such as rpm, fuel used or vehicle distance. |
| Fuel Card | A special credit card given to drivers or employees which allows them to buy fuel or goods at petrol stations. |
| Fuel efficiency | The differential between the vehicle's outputs and inputs in relation to fuel performance.<br>See also Traditional fuel Efficiency and efficiency q.v.<br>A new fuel efficiency measure is defined in the section The fuel efficiency model. |
| Heavy Goods Vehicle (HGV) | The old legal term for goods vehicles exceeding 7.5 tonnes permissible maximum weight and used in driver licensing and operator licensing rules. The term was replaced under the EU unified driver licensing scheme by "Large Goods Vehicle" (LGV) (Slater, 2010). |
| Heuristics | Proceeding to a solution by means of trial and error of alternative scenarios (Slater, 2010). |
| HGV | Abbreviation. Heavy Goods Vehicle |
| Improvement | When something gets better or when you make it better (2008). |
| Interval scale | A scale divided in intervals. |

| | |
|---|---|
| Isotonic factor | Desirable output or input<br>See anti-isotonic factor q.v. |
| Key Performance Indicator | A Key Performance Indicator is a measure of performance defined by an organisation to evaluate how successful it is. |
| KPI | Abbreviation. See Key performance indicator q.v. |
| LGV | Abbreviation. Large Goods Vehicle (over 3,500 kilograms). The term replaced HGV (see HGV above q.v.) under the EU unified driver licensing scheme (Slater, 2010). |
| Litres per 100 km (lp100k) | A measure of a vehicle's fuel efficiency in regards to the litres used to cover a 100 kilometres. This measure has the advantage to be consistent in regards to the amount of fuel used; i.e. improving fuel efficiency from 13 l/100km to 15 l/100km saves as much fuel as improving fuel efficiency from 34 to 36 l/100km. mpg does not have this characteristic. |
| Measurement | [C or U] the act or process of measuring (2008). |
| Miles per gallon (mpg) | A measure of a vehicle's fuel efficiency in regards to the distance travelled (measured in miles) with a single gallon of fuel. |
| Mix efficiency | Efficient allocation of inputs and outputs.<br>See mix inefficiencies q.v. |
| Mix inefficiency | Mix inefficiencies are caused by sub-optimal allocation of inputs or outputs in production. This term is used in Frontier Analysis methods (e.g. DEA or SFA).<br>See mix efficiency q.v. |
| Modern vehicle | Vehicles manufactured after 2000. Most of these will have CAN technology. |
| mpg | Abbreviation. Miles per gallons. The number of miles that can be done with a single gallon. This fuel efficiency measure is used in the UK and other countries such as the US (although the US gallon differs from the UK gallon). This measure is sometimes used along with the pence per mile measure (ppm).<br>See ppm q.v. |
| Non-parametric | This antonym of parametric.<br>See parametric q.v. |

| | |
|---|---|
| Non-sampling error | Is a generic term to refer to deviations around a true value. |
| Objective function | In a linear optimisation problem, this is the function to optimise. Optimisation can be done in two direction, maximisation or minimisation. |
| Over-acceleration | Over-acceleration relates to excessive acceleration. These are generally defined by an excessive pressure on the speed pedal and an rpm higher than a pre-defined value. |
| Over-revving | Over-revving relates to reaching unnecessarily high rpm while pressing the speed pedal and with a low torque (i.e. not going uphill). |
| Parametric | Which assumes the data come from a type of probability distribution and (which potentially) makes inferences about the parameters of this distribution. |
| Performance Measurement | The qualification and/or quantification of a purposefully executed action.<br>See the **Key Concepts and Definitions** section for further information q.v. |
| Piecewise linear | Define in linear pieces (corresponding to segments in two dimensional spaces). |
| PM | Abbreviation. Performance Measurement |
| ppl | Abbreviation. Pence per litre. The amount in pence of a litre of fuel. |
| ppm | Abbreviation. Pence per mile. The cost of fuel per mile, expressed in pence. This performance measure is often used along with mpg.<br>See mpg q.v. |
| Producer | The term producer refers to the entity in Stochastic Frontier Analysis. Producers are called Decision Making Unit (DMU) in DEA, another Frontier Analysis method.<br>See SFA, DEA, DMU q.v. |

| | |
|---|---|
| Rigid | An independent vehicle on which the driver's cab and the load carrying compartment are mounted on the same rigid chassis - defined under the Construction and Use Regulations as a vehicle not constructed or adapted to form part of an articulated vehicle. Distinguished from a van by means of carrying capacity and weight and that an HGV licence is the requirement to drive the vehicle (Slater, 2010). |
| Routing | The practice of planning routes for vehicles. (Slater, 2010). |
| SBM | Slack Based Model. See (Tone, 2001) and section 8.3.2 SBM Model. |
| Scheduling | The planning of vehicles and drivers to match cargo delivery or collection requirements or the passenger transport timetables (Slater, 2010). |
| SFA | Abbreviation. See Stochastic Frontier Analysis q.v. |
| Stochastic | Which behaviour is non-deterministic (i.e. random).<br>See deterministic q.v. |
| Stochastic Frontier Analysis | The estimation of productive efficiency through the use of stochastic methods. |
| Technical efficiency | Efficient production in regards to the production methods.<br>See Technical inefficiency q.v. |
| Technical inefficiency | Inefficient production in regards to the production methods.<br>See Technical efficiency q.v. |
| Tobit model | The Tobit Model is a model originally developed by James Tobin (1958) which describes the relationship between a non-negative dependent variable yi and an independent variable (or vector) xi. |
| Total factor productivity | 'Total productivity includes intermediate goods in the measure of output as well as their inclusion in adding up inputs. Intermediate goods include purchased material and energy' (Christopher and Thor, 1993, p. 6–1.5). |
| Total productivity | 'This measure looks at the ratio of outputs to labor and capital inputs' (Christopher and Thor, 1993, p. 6–1.5). |
| Traditional fuel efficiency | Traditional fuel efficiency here refers to the miles per gallon (mpg) measure.<br>See also efficiency q.v. |

| | |
|---|---|
| Translog | Abbreviation. Transcendental logarithmic function. A generalised adaptation of the Cobb Douglas production function.<br>See Coob Douglas Transform q.v. |
| Van | An independent small vehicle on which the driver's cab and the load carrying compartment are mounted on the same (rigid) chassis – defined under the Construction and Use Regulations as a vehicle not constructed or adapted to form part of an articulated vehicle. Distinguished from a rigid vehicle by means of carrying capacity and weight (and that a car licence is the requirement to drive the vehicle) There are a number of other descriptions, including: (a) Car Derived Van: based upon a car chassis up to 1.5 tonne GVW. (b) Small Van: rigid box bodied vehicle up to 3.5 tonne GVW. (c) Medium Van: rigid box bodied vehicle up to 7.5 tonne GVW. (d) Large Van: rigid box bodied vehicle up to 18 tonne GVW. (e) Articulated Van: box bodied trailer up to a maximum length. (f) Car Derived Van: based upon a car chassis up to 1.5 tonne GVW. (g) Drawbar Van: rigid box bodied vehicle towing a rigid box bodied trailer (Slater, 2010). |

## 10. References

(2008) Cambridge Advanced Learner's dictionary. *Cambridge Advanced Learner's dictionary.* Third Edition ed. Cambridge, Cambridge University Press.

ABBOTT, C. (1994) *Performance measurement in library and information services,* London, Aslib.

AIGNER, D. J., LOVELL, C. A. K. & SCHMIDT, P. (1977) Formulation and Estimation of Stochastic Frontier Production Function Models. *Journal of Econometrics,* 6**,** pp. 21-37.

AMARATUNGA, D. & BALDRY, D. (2002) Moving from performance measurement towards performance management. *Facilities,* 20**,** 217-223.

ANAND, G. & KODALI, R. (2008) Benchmarking the benchmarking models. *Benchmarking: An International Journal,* 15**,** 34.

ANON (1985) Survey of Distributions Costs. *Institute of Physical Distribution Management.*

ANON (1989) Oxford English Dictionary. *English Dictionary.* Second Edition ed. Oxford, Oxford University Press.

ANON (2005a) DfT sponsors plan to educate White Van Man. *Commercial Motor,* 202**,** 19-19.

ANON (2005b) Michelin man calls for tyre efficiency rating. *Commercial Motor,* 202**,** 12-12.

ANON (2007a) Air kit cuts fuel by 6%. *Commercial Motor,* 205**,** 7-7.

ANON (2007b) Serial bus systems in motor vehicles. IN VECTOR (Ed.).

ANON (2007c) Untitled. *Commercial Motor,* 206**,** 25-25.

ANON (2008a) Benchmarking pilot seeks haulier support. *Motor Transport***,** 7-7.

ANON (2008b) Cambridge Advanced Learner's dictionary. *Cambridge Advanced Learner's dictionary.* Third Edition ed. Cambridge, Cambridge University Press.

ANON (2008c) *EOBD - a brief history*, [Online], Available: http://www.omitec.com/en/support/technology-briefs/brief-history-of-eobd/

ANON (2009a) *Lagging Indicator*.

ANON (2009b) NHTSA proposes tyre labels. *European Rubber Journal,* 191**,** 11-11.

ANON (2009c) Revised fuel card from Esso launches. *Commercial Motor,* 210**,** 15-15.

AVKIRAN, N. K. & THORANEENITIYAN, N. (2009) Purging data before productivity analysis. *Journal of Business Research,* In Press, Corrected Proof.

AXINN, W. G. & PEARCE, L. D. (2006) *Mixed Method Data Collection Strategies,* Cambridge, Cambridge University Press.

BAGCHI, P. K. & NAG, B. N. (1991) Dynamic Vehicle Scheduling: An Expert Systems Approach. *International Journal of Physical Distribution and Logistics Management,* 21**,** pp. 10-18.

BALCI, O. (1998) Chapter 10 - Verification, Validation, and Testing. IN BANKS, J. (Ed.) *Handbook of Simulation: Principles, Methodology, Advances, Applications and Practice.* John Wiley & Sons, Inc.

BAND, W. (1990) Performance metrics keep customer satisfaction programmes on track. *Marketing news*.

BANKER, R. D., CHARNES, A. & COOPER, W. W. (1984) SOME MODELS FOR ESTIMATING TECHNICAL AND SCALE INEFFICIENCIES IN DATA ENVELOPMENT ANALYSIS. *Management Science,* 30**,** 1078-1092.

BARBA-ROMERO, S. & POMEROL, J. C. (2000) *Multicriterion Decision Management: Principles and Practice,* Dordrecht, Kluwer.

BAUMGARTNER, M., LÉONARDI, J. & KRUSCH, O. (2008) Improving computerized routing and scheduling and vehicle telematics: A qualitative survey. *Transportation Research Part D: Transport and Environment,* 13**,** 377-382.

BEARNE, S. (2010) BP puts natural and paid search accounts out to tender as it looks to increase spend. *New Media Age***,** 06-06.

BIGGS, N. L., K., L. E. & J., W. R. (1976) *Graph Theory 1736 - 1936*, Oxford University Press.

BITITCI, U. & MACBRYDE, J. (Eds.) (2002) *Performance Measurement*, Emerald Group Publishing.

BOSCH, R. (1991) CAN Specification Version 2.0. Bosch.

BOURNE, M., MILLS, J., NEELY, A., PLATTS, K. & RICHARDS, H. (2002) *Getting the measure of your business,* Cambridge, Cambridge University Press.

BOUYSSOU, D. & ROY, B. (1993) *Aide Multicritere a la decision: Methodes et Cas,* Paris, Economica.

BRACKSTONE, M., WATERSON, B. & MCDONALD, M. (2009) Determinants of following headway in congested traffic. *Transportation Research Part F: Traffic Psychology and Behaviour,* 12**,** 131-142.

BRADY, F. N. (1985) A Janus-headed model of ethical theory: looking two ways at business/society issues. *The academy of Management review,* 10**,** pp. 568-576.

BROWN, R. (2008) Tyre management saves cash. *Commercial Motor,* 208**,** 7-7.

BRYMAN, A. (2001) Social research methods - triangulation. Loughborough, Loughborough University.

BUCHANAN, J. & VANDERPOOTEN, D. (2007) Ranking projects for an electricity utility using ELECTRE III. *International Transactions in Operational Research,* 14**,** 309-323.

BUCHANAN, W. (2000) CAN bus. *Computer Busses.* Oxford, Butterworth-Heinemann.

CAMP, R. C. (1989) *Benchmarking: The search for industry Best practices that lead to superior performance,* Milwaukee, Wisconsin, ASQC Quality Press.

CHARNES, A. & COOPER, W. W. (1961) *Management Models and Industrial Applications of Linear Programming,* New York, Wiley.

CHARNES, A. & COOPER, W. W. (1962) Programming with Linear Fractional Functionals. *Naval Research Logistics Quarterly,* 9**,** pp. 181-185.

CHARNES, A. & COOPER, W. W. (1968) Structural Sensitivity Analysis in Linear Programming and an Exact Product Form Left Inverse. *Naval Research Logistics Quarterly,* 15**,** pp. 517-522.

CHARNES, A., COOPER, W. W. & RHODES, E. (1978) Measuring the efficiency of decision making units. *European Journal of Operational Research,* 2**,** 15.

CHARNES, A., ROUSSEAU, J. J. & SEMPLE, J. H. (1996) Sensitivity and Stability of Efficiency Classification in Data Envelopment Analysis. *Journal of Productivity Analysis,* 7**,** pp. 5-18.

CHATCHAI, U., KEN, P. & KIM HUA, T. (2007) Predictive performance measurement system: A fuzzy expert system approach. *Benchmarking: An International Journal,* 14**,** 77-91.

CHRISTOPHER, W. F. & THOR, C. G. (1993) *Handbook for Productivity Measurement and Improvement*, Productivity Press.

CLARKE, S. (2008) Same old, same old. *Motor Transport.*

COBB, C. W. & DOUGLAS, P. H. (1928) A Theory of Production. *American Economic Review,* 18 (supplement)**,** pp. 139-165.

COELLI, T. J., RAO, D. S. P., O'DONNELL, C. J. & BATTESE, G. E. (2005) *An introduction to efficiency and productivity analysis*, Springer.

COELLO COELLO, C., A., LAMONT, G. B. & VAN VELDHUIZEN, D. A. (2007) *Evolutionary Algorithms for Solving Multi-Objective Problems*.

COLE, L. (2009) Smart cards? (cover story). *Commercial Motor,* 209**,** 46-47.

COOPER, J. (1987) Physical distribution in the UK. *international Journal of Physical Distribution & Logistics,* 17.

COOPER, W. W., SEIFORD, L. M. & TONE, K. (2007) *Data Envelopment Analysis - A Comprehensive Text with Models, Applications References and DEA-Solver Software*, Springer.

COOPER, W. W. & TONE, K. (1997) Measures of inefficiency in data envelopment analysis and stochastic frontier estimation. *European Journal of Operational Research,* 99**,** 72-88.

COWIE, J. & ASENOVA, D. (1999) Organisation form, scale effects and efficiency in the British bus industry. *Transportation,* 26**,** 231-248.

CULLEN, R. (1999) Does performance measurement improve organisational effectiveness? A postmodern analysis. *Performance measurement and metrics,* 1**,** 21.

CULLINANE, K., WANG, T.-F., SONG, D.-W. & JI, P. (2006) The technical efficiency of container ports: Comparing data envelopment analysis and stochastic frontier analysis. *Transportation Research Part A: Policy and Practice,* 40**,** 354-374.

CURTIS, D. A. (1985) The modern accounting system. *Financial Executive***,** pp. 91-93.

CURWIN, J. & SLATER, R. (2002) *Quantitative methods for business decisions*, Thompson Learning.

DANTZIG, G. B., FULKERSON, D. R. & JOHNSON, S. M. (1959) ON A LINEAR-PROGRAMMING, COMBINATORIAL APPROACH TO THE TRAVELING-SALESMAN PROBLEM. *Operations Research,* 7**,** 58.

DE KOCK, J. (2009) Look at the whole picture. *Motor Transport.*

DEBREU, G. (1951) The Coefficient of Resource Utilization. *Econometrica,* 19**,** 273-292.

DENCE, R. (1995) Best practice benchmarking (Chapter 7). IN HOLLOWAY, J., LEWIS, J. & MALLORY, G. (Eds.) *Performance measurement and evaluation.* London, Sage Publication.

DFT (2009) Transport Statistics Great Britain 2009. IN DFT (Ed.) *Transport Statistics Great Britain.* 35 ed. London, DfT.

DOUMPOS, M. & ZOPOUNIDIS, C. (2002) *Multicriteria Decision Aid Classification Methods*, Kluwer Academic Publisher.

DYSON, R. G., ALLEN, R., CAMANHO, A. S., PODINOVSKI, V. V., SARRICO, C. S. & SHALE, E. A. (2001) Pitfalls and protocols in DEA. *European Journal of Operational Research,* 132**,** 245-259.

ECCLES, R., G. (1995) The performance measurement manifesto. IN HOLLOWAY, J., LEWIS, J. & MALLORY, G. (Eds.) *Performance measurement and evaluation.* London, Sage Publication.

EDWARDS, W. (1977) How to use multiattributes utility measurement for social decision making. *IEEE Transactions on Systems, Man and Cybernetics,* 7**,** pp. 326-340.

ELLINGER, A. D., WATKINS, K. E. & MARSICK, V. J. (2005) Case Study Research Methods. San Francisco, Berrett-Koehler Publishers Inc.

ERIKSON, S. G. (2009) Performance indicators. *Safety Science,* 47**,** 468-468.

FÄRE, R. & GROSSKOPF, S. (2004) Modeling undesirable factors in efficiency evaluation: Comment. *European Journal of Operational Research,* 157**,** 242-245.

FARRELL, M. J. (1957) The Measurement of Productive Efficiency. *Journal of the Royal Statistical Society. Series A (General),* 120**,** 253-290.

FBP (2005) In Fleet Trials of Fuel Saving Interventions. IN TRANSPORT, D. F. (Ed.). Freight Best Practice - Department of Transport.

FBP (2007a) Aerodynamics For efficient Road Freight operations. IN TRANSPORT, D. F. (Ed.). Freight Best Practice - Department of Transport.

FBP (2007b) Fuel Saving Devices. IN TRANSPORT, D. F. (Ed.). Freight Best Practice - Department of Transport.

FBP (2008a) Fuel efficiency trials research. IN TRANSPORT, D. F. (Ed.). Freight Best Practice - Department of Transport.

FBP (2008b) Performance Management for Efficient Road Freight Operations. IN FBP (Ed.). Freight Best Practice.

FBP (2009) Fuel Efficiency Interventions Trials - How to Test and Save. IN TRANSPORT, D. F. (Ed.). Freight Best Practice - Department of Transport.

FENNER, D. D. (2002) Performance is your reality, forget everything else. *Urban Water,* 4**,** 119-121.

FERRIER, G. D. & HIRSCHBERG, J. G. (1999) Can We Bootstrap DEA Scores. *Journal of Productivity Analysis,* 11**,** pp. 81-92.

FMS-STANDARD (2002) FMS-Standard Interface description Vers. 01.00. Working group FMS-Standard.

FOLAN, P. & BROWNE, J. (2005) A review of performance measurement: Towards performance management. *Computers in Industry,* 56**,** 663-680.

FRANKE, R. H. & KAUL, J. D. (1978) The Hawthorne Experiments: First Statistical Interpretation. *American Sociological Review,* 43**,** 623-643.

GASS, G. L. & PRINCE, C. (1993) Family of Measures: A Tool for Continuous Improvement. IN WILLIAM, F. C. & THOR, C. G. (Eds.) *Handbook for Productivity Measurement and Improvement.* Portland, Productivity Press Inc.

GOH, S. & RICHARDS, G. (1997) Benchmarking the learning capability of organizations. *European Management Journal,* 15**,** 575-583.

GOVERNMENT OF CANADA, C. H., CANADIAN CONSERVATION INSTITUTE (2002) *Analytical Hierarchy Process (AHP) Program*, [Online], Available: http://www.cci-icc.gc.ca/tools/ahp/index_e.asp [2010-01-07]

GREENROAD (2010) *Our solutions*, [Online], Available: http://www.greenroad.com/our_solution.html [2010-07-09]

GROSSKOPF, S. (1993) Efficiency and productivity. IN FRIED, H. O., C. A., K. L. & SHELTON, S. (Eds.) *Measurement of Productive Efficiency: Techniques and Applications.* Oxford, Oxford University Press Incorporated.

HADI VENCHEH, A., KAZEMI MATIN, R. & TAVASSOLI KAJANI, M. (2005) Undesirable factors in efficiency measurement. *Applied Mathematics and Computation,* 163**,** 547-552.

HAHN, F. (2007) Environmental determinants of banking efficiency in Austria. *Empirica,* 34**,** 231-245.

HALACHMI, A. (2002a) Performance measurement and government productivity. *Work Study,* 51**,** 63-73.

HALACHMI, A. (2002b) PERFORMANCE MEASUREMENT, ACCOUNTABILITY, AND IMPROVED PERFORMANCE. *Public Performance & Management Review,* 25**,** 370-374.

HALACHMI, A. (2005) Performance measurement is only one way of managing performance. *Performance measurement and performance management.* Emerald.

HARBOUR, J. L. (2009) *The Basics of Performance measurement,* New York, Productivity Press - taylor & Francis.

HARZING, A.-W., BALDUEZA, J., BARNER-RASMUSSEN, W., BARZANTNY, C., CANABAL, A., DAVILA, A., ESPEJO, A., FERREIRA, R., GIROUD, A., KOESTER, K., LIANG, Y.-K., MOCKAITIS, A., MORLEY, M. J., MYLONI, B., ODUSANYA, J. O. T., O'SULLIVAN, S. L., PALANIAPPAN, A. K., PROCHNO, P., CHOUDHURY, S. R., SAKA-HELMHOUT, A., SIENGTHAI, S., VISWAT, L., SOYDAS, A. U. & ZANDER, L. (2009) Rating versus ranking: What is the best way to reduce response and language bias in cross-national research? *International Business Review,* 18**,** 417-432.

HAYES, R. H., WHEELWRIGHT, S. C. & CLARK, K. B. (1988) Measuring manufacturing performance. *Dynamic manufacturing: creating the learning organisation*, pp. 130-160.

HEALTH AND SAFETY EXECUTIVE (2010) *HSE: Corporate Manslaughter*, [Online], Available: [2010-07-16]

HJALMARSSON, L. & ODECK, J. (1996) Efficiency of trucks in road construction and maintenance: An evaluation with data envelopment analysis. *Computers & Operations Research,* 23**,** 393-404.

HUSAIN, N., ABDULLAH, M. & KU, S. (2000) Evaluating public sector efficiency with data envelopment analysis (DEA): a case study in Road Transport Department. *Total Quality Management,* 11**,** pp. 830-836.

ISORAITE, M. (2005) Benchmarking methodology in a transport sector. *Transport,* 19**,** 269-275.

JAHANSHAHLOO, G. R., LOTFI, F. H., SHOJA, N., TOHIDI, G. & RAZAVYAN, S. (2005) Undesirable inputs and outputs in DEA models. *Applied Mathematics and Computation,* 169**,** 917-925.

JESSOP, A. (2009) A multicriteria blockmodel for performance assessment. *Omega,* 37**,** 204-214.

JONDROW, J., KNOX LOVELL, C. A., MATEROV, I. S. & SCHMIDT, P. (1982) On the estimation of technical inefficiency in the stochastic frontier production function model. *Journal of Econometrics,* 19**,** 233-238.

JOURNEY DYNAMICS (2010) *Journey Dynamics - the products*, [Online], Available: http://www.journeydynamics.com/products/products.htm [2010-07-09]

KAPLAN, R. S. & NORTON, D. P. (1996) *Translating strategy into action - The Balanced Scorecard,* Boston, Massachusetts, Harvard Business School Press.

KERSTENS, K. (1996) Technical efficiency measurement and explanation of French urban transit companies. *Transportation Research Part A: Policy and Practice,* 30**,** 431-452.

KING, D. C., ZECKHAUSER, R. J. & KIM, M. T. (2002) The management performance of the US States. *Working paper*.

KONTIO, J. (2005) Thesis Topic Definition Template. *Software Business Laboratory.* Helsinki, Helsinki University of Technology.

KOOPMANS, T. C. (1951) Analysis of production as an efficient combination of activities. IN KOOPMANS, T. C. (Ed.) *Activity Analysis of Production and Allocation.* John Wiley & Sons Inc

Chapman & Hall.

KOSTER, M. B. M. D., BALK, B. M. & NUS, W. T. I. V. (2009) On using DEA for benchmarking container terminals. *International Journal of Operations & Production Management,* 29**,** 1140-1155.

KUMBHAKAR, S. C. & KNOX LOVELL, C. A. (2000) *Stochastic Frontier Analysis,* Cambridge, Cambridge University Press.

LABECKAS, G. & SLAVINSKAS, S. (2005) Influence of fuel additives on performance of direct-injection Diesel engine and exhaust emissions when operating on shale oil. *Energy Conversion and Management,* 46**,** 1731-1744.

LAISE, D. (2004) Benchmarking and learning organizations: ranking methods to identify "best in class". *Benchmarking: An International Journal,* 11**,** 9.

LANDY, F. & ZEDECK, S. (1983) Introduction. IN LANDY, F., ZEDECK, S. & CLEVELAND, J. (Eds.) *Performance measurement and theory.* Hillsdale, New Jersey ; London, Lawrence Erlbaum Associates Publishers.

LAWSON, P. (1995) Performance management: an overview. IN WALTERS, M. (Ed.) *The Performance Management.* Institute of Personnel and Development.

LINDO (2010) *LINDO SYSTEMS - Optimization Software: Integer Programming, Linear Programming, Nonlinear Programming, Global Optimization*, [Online], Available: http://www.lindo.com/index.php?option=com_content&view=article&id=28&Itemid=4 [2010-02-20]

LONGENECKER, C. & LUDWIG, D. (1990) Ethical dilemnas in performance appraisal revisited. *Journal of Business Ethics,* 9.

LONGENECKER, C. & LUDWIG, D. (1995) Ethical dilemnas in performance appraisal revisited. IN HOLLOWAY, J., LEWIS, J. & MALLORY, G. (Eds.) *Performance measurement and evaluation.* Sage Publications. Published in association with the Open University.

LOW CARBON VEHICLE PARTNERSHIP (2009) *LowCVP | Technology Challenge | Winners 2009*, [Online], Available: http://www.lowcvp.org.uk/technologychallenge/announcements.asp [2010-05-25]

MADEN, W., EGLESE, R. & BLACK, D. (2010) Vehicle routing and schedulign with time-varying data: A case study. *Journal of the Operational Research Society,* 61**,** pp. 515-522.

MALIK, W., RATHINAM, S. & DARBHA, S. (2007) An approximation algorithm for a symmetric Generalized Multiple Depot, Multiple Travelling Salesman Problem. *Operations Research Letters,* 35**,** 747-753.

MASKELL, B. (1992) Performance measurement for world class manufacturing. *Corporate Controller (COP).*

MASTERNAUT THREE X (2010a) *Drive for Life*, [Online], Available: http://www.masternaut.co.uk/driveforlife/ [2010-05-25]

MASTERNAUT THREE X (2010b) *Masternaut Alert Panel - Masternaut Three X*, [Online], Available: http://www.masternaut.co.uk/vehicle-tracking/masternaut-alert-panel.aspx [2010-07-09]

MCCLELLAND, D. & MCKINNON, A. (2004) Use of telematics systems for the colletion of Key Performance Indicators Data in Road Freight Transport, Summary report of a survey of suppliers and users of telematics systems undertaken as part of a future integrated transport research project funded by the Department of Transports. *FIT 022.* Heriot Watt University.

MCKINNON, A. (1993) Improving the fuel efficiency of road freight operations. *International Journal of Physical Distribution & Logistics Management,* 23**,** pp.3-11.

MCNALLY, M. G. (2008) The Four-step Model. IN HENSHER, D. A. & BUTTON, K. J. (Eds.) *Handbook of Transport Modelling.* Oxford, Elsevier.

MEEUSEN, W. & VAN DEN BROECK, J. (1977) EFFICIENCY ESTIMATION FROM COBB-DOUGLAS PRODUCTION FUNCTIONS WITH COMPOSED ERROR. *International Economic Review,* 18**,** 435.

MENGER, K. (1932) Das botenproblem. *Ergebnisse Eines Mathematischen Kolloqiums,* 2**,** pp. 11-12.

MILNES, J. (2009) Operators reap benefits from fuel-saving device. *Motor Transport***,** 7-7.

MIX TELEMATICS (2010) *MiX Telematics - RIBAS | Overview*, [Online], Available: http://www.mixtelematics.com/product/ribas/ [2010-07-09]

MPG MARATHON (2009) MPG marathon results.

MURPHY, K. M. & CLEVELAND, J. (1995) *Understanding performance appraisal - Social, Organizational, and Goal Based Perspectives*, Sage Publications.

ODECK, J. & ALKADI, A. (2001) Evaluating efficiency in the Norwegian bus industry using data envelopment analysis. *Transportation,* 28**,** pp. 211-232.

ODECK, J. & HJALMARSSON, L. (1996) The performance of trucks - an evaluation using data envelopment analysis. *Transportation Planning and Technology,* 20**,** 49 - 66.

PALMER, A. (2008) The Development of an Integrated Routing and Carbon Dioxide Emissions Model for Goods Vehicles.

PESTANA BARROS, C. & DIEKE, P. U. C. (2007) Performance evaluation of Italian airports: A data envelopment analysis. *Journal of Air Transport Management,* 13**,** 184-191.

PETERS, M. & ZELEWSKI, S. (2008) Pitfalls in the application of analytic hierarchy process to performance measurement. *Management Decision,* 46**,** pp. 1039-1051.

PUNNEN, A. P. (2002) The travelling salesman problem: applications, formulation and variations. IN GUTIN, G. & PUNNEN, A. P. (Eds.) *The Travelling Salesman Problem and Its Variation.* Kluwer Academic Publishers.

RAMANATHAN, R. (2000) A holistic approach to compare energy efficiencies of different transport modes. *Energy Policy,* 28**,** 743-747.

ROCHAT, Y. & SEMET, F. (1994) A Tabu Search Approach for Delivering Pet Food and Flour in Switzerland. *Journal of the Operational Research Society,* 45**,** pp. 1233-1246.

SAATY, T. L. (1980) *The Analytic Hierarchy Process: Planning, Pritority Setting, Resource Allocation (Decision Making Series)*, Mcgraw-Hill.

SAEN, R. F. (2005) Developing a nondiscretionary model of slacks-based measure in data envelopment analysis. *Applied Mathematics and Computation,* 169**,** 1440-1447.

SANGHYUN, C. (2009) Impact of global terminal operators on port efficiency: a tiered data envelopment analysis approach. *International Journal of Logistics: Research & Applications,* 12**,** 85-101.

SARAVANAN, N. & NAGARAJAN, G. (2008) An experimental investigation of hydrogen-enriched air induction in a diesel engine system. *International Journal of Hydrogen Energy,* 33**,** 1769-1775.

SAUNDERS, M., LEWIS, P. & THORNHILL, A. (2009) *Research methods for business students,* Harlow, Pearson Education Limited.

SEIFORD, L. M. & ZHU, J. (2002) Modeling undesirable factors in efficiency evaluation. *European Journal of Operational Research,* 142**,** 16-20.

SEIFORD, L. M. & ZHU, J. (2005) A response to comments on modeling undesirable factors in efficiency evaluation. *European Journal of Operational Research,* 161**,** 579-581.

SHARIF, A. M. (2002) Benchmarking performance management systems. *Benchmarking: An International Journal,* 9**,** 23.

SHEPHARD, R. W. (1953) *Cost and Production Functions,* Princeton, Princeton University Press.

SIMAR, L. & WILSON, P. W. (1999a) Of Course We Can Bootstrap DEA Scores! But Does It Mean Anything? Logic Trumps Wishful Thinking. *Journal of Productivity Analysis,* 11**,** pp. 93-97.

SIMAR, L. & WILSON, P. W. (1999b) Some Problems with the Ferrier / Hirschberg Bootstrap Idea. *Journal of Productivity Analysis,* 11**,** pp. 67-80.

SLATER, A. (2002) Specification for a dynamic vehicle routing and scheduling system. *International Journal of Transport Management,* 1**,** 29-40.

SLATER, A. (2010) A Dictionary of Terms Encountered in Transport Management. Altrincham, Added Value Logistics Publications.

SMITH, P. & GODDARD, M. (2002) Performance management and operational research: a marriage made in heaven? *Journal of the Operational Research Society,* 53**,** 247-255.

SURESHCHANDAR, G. S. & LEISTEN, R. (2006) A framework for evaluating the criticality of software metrics: an analytic hierarchy process (AHP) approach. *Measuring Business Excellence,* 10**,** pp. 22-23.

TANIGUCHI, E. & SHIMAMOTO, H. (2004) Intelligent transportation system based dynamic vehicle routing and scheduling with variable travel times. *Transportation Research Part C: Emerging Technologies,* 12**,** 235-250.

THOMPSON, R. G., DHARMAPALA, P. S. & THRALL, R. M. (1994) Sensitivity Analysis of Efficiency Measures with Applications to Kansas Farming and Illinois Coal Mining. IN CHARNES, A., COOPER, W. W., LEWIN, A. Y. & SEIFORD, L. M. (Eds.) *Data Envelopment Analysis: Theory, Methodology and Applications.* Norwell, Kluwer Academic Publishers.

THOR, C. G. (1993) The Family of Measures Method for Improving Organizational Performance. IN WILLIAM, F. C. & THOR, C. G. (Eds.) *Handbook for Productivity Measurement and Improvement.* Portland, Productivity Press Inc.

TINGLEY, D., PASCOE, S. & COGLAN, L. (2005) Factors affecting technical efficiency in fisheries: stochastic production frontier versus data envelopment analysis approaches. *Fisheries Research,* 73**,** 363-376.

TONE, K. (2001) A slacks-based measure of efficiency in data envelopment analysis. *European Journal of Operational Research,* 130**,** 498-509.

TONKIN, I. (2009a) Pre-purchase trials pay off. *Motor Transport***,** 8-8.

TONKIN, I. (2009b) Vauxhall Vivaro leads the way in fuel economy contest. *Commercial Motor,* 210**,** 8-8.

TYAGI, M. S. (1983) A practical method for despatching problems. *Journal of the Operations Research Society of Japan,* 10**,** pp. 351-357.

VINCENT, S. (1998) Input data analysis. IN BANKS, J. (Ed.) *Handbook of Simulation: principles, methodology, advances applications and practice.* Engineering & Management Press.

VIRTOS, H. P. B. (2010) *DEA : Open Source DEA (Data Envelopment Analysis)*, [Online], Available: http://www.opensourcedea.org/dea.html [2010-02-20]

VIRTOS, H. P. B., SAVAGE, C. J., MADEN, W., SLATER, A. & BAMFORD, C. (2009) A NOVEL APPLICATION OF DATA ENVELOPMENT ANALYSIS TO TRANSPORT OPERATION. IN NETWORK, L. R. (Ed.) *Logistics Research Network 2009.* Hull.

VIRTOS, H. P. B., SAVAGE, C. J., MADEN, W., SLATER, A. & BAMFORD, C. (2010) VAN FUEL EFFICIENCY MEASUREMENT - A SUCCESSFUL APPLICATION OF DATA ENVELOPMENT ANALYSIS. IN NETWORK, L. R. (Ed.) *Logistics Research Network 2010.* Leeds.

VOSS, W. (2005) *A comprehensive guide to Controller Area Network,* Greenfield, Copperhill Technology Corporation.

WALTERS, M. (Ed.) (1995) *The performance management handbook,* London, Institute of Personnel and Development.

WEATHERLEY, B. (2009) Born on the wind. *Commercial Motor,* 209**,** 44-47.

WILSON, P. (1987) Fleet costing. *Management Research News,* 10**,** pp.17-18.

WINSTON, W. L. (2004) *Operations Research - Applications and Algorithms*, Thompson.

WISNER, J. & FAWCETT, S. (1991) Linking firm strategy to operating decisions through performance measurement. *Production and inventory Management Journal,* 32**,** pp.5-11.

YIN, R. K. (1994) *Case Study Research - Design and Methods*, Sage Publications, Inc.

YOSHIDA, Y. & FUJIMOTO, H. (2004) Japanese-airport benchmarking with the DEA and endogenous-weight TFP methods: testing the criticism of overinvestment in Japanese regional airports. *Transportation Research Part E: Logistics and Transportation Review,* 40**,** 533-546.

YU, M.-M. (2004) Measuring physical efficiency of domestic airports in Taiwan with undesirable outputs and environmental factors. *Journal of Air Transport Management,* 10**,** 295-303.

YUN, Y. B., NAKAYAMA, H. & TANINO, T. (2004) A generalized model for data envelopment analysis. *European Journal of Operational Research,* 157**,** 87-105.

ZAIRI, M. (1996) *Benchmarking for Best practice,* Oxford, Reed Educational and Professional Publishing Ltd.

ZHU, J. (2010) *DEAFrontier.com, DEA Software*, [Online], Available: http://www.deafrontier.com/software.html [2010-02-20]