Rolling optimisation, stochastic demand modelling and scenario reduction applied to the UK gas market

Mel T. Devine
Department of Mathematics and Statistics
University of Limerick

A thesis submitted for the degree of
Doctor of Philosophy (PhD)
Supervised by: Prof. James P. Gleeson, Dr. John Kinsella and Dr. David M. Ramsey
Submitted to the University of Limerick, August 2012
Declaration

I hereby declare that this thesis is entirely my own work and that it has been not submitted for any other academic awards.

Signature of Author: ________________________________

Mel T. Devine
Abstract

In recent years the daily gas demand in the UK and Ireland has become increasingly uncertain. This due to the changing nature of electricity markets, where intermittent wind energy levels lead to variations in the demand for gas needed to produce electricity. As a result, there is an increasing need for models of natural gas markets that include stochastic demand. In this thesis, a Rolling Optimisation Model (ROM) of the UK natural gas market is introduced. It takes as an input demand scenarios simulated from a stochastic process of UK gas demand which is developed as a part of this work. The model is informed by an analysis of the two main types of natural gas market models: complementarity-based equilibrium models and cost minimisation models. This analysis shows that when market power (i.e. Nash-Cournot competition) is removed from complementarity-based equilibrium models the outputs are equivalent to those from a corresponding cost minimisation model. The outputs of the Rolling Optimisation Model are the flows of gas in the UK, i.e., how the different sources of supply meet demand, as well as how gas flows in to and out of gas storage facilities, and the daily System Average Price of gas in the UK. The model was found to fit reasonably well to historic data (from the UK National Grid) for the years starting on the 1st of April for both 2010 and 2011. This work also investigates the benefit of using scenario reduction techniques on the set of demand scenarios used in ROM. These techniques allow the effects of large sets of stochastically-generated scenarios to be captured in ROM, whilst maintaining a relatively low computational cost for solving the model. In the final chapter of this thesis, ROM is used to predict future flows and prices of gas in the UK and investigate various ‘What-if’ scenarios in the UK natural gas market.
Acknowledgements

- First and foremost, I wish to express my deepest gratitude to my supervisors Prof. James P. Gleeson, Dr. John Kinsella and Dr. David M. Ramsey for their constant support and scholarly advice throughout the entirety of my Ph.D. studies. I appreciate sincerely all the contributions of their time and inspirational expertise, and for instilling in me a greater understanding of applied mathematics.

- Secondly, I would like to thank Bord Gáis Energy, in particular Dr. Gavin Hurley and Dr. Niall Fitzgerald. Their industrial insight proved invaluable to my research.

- Finally, I must thank wholeheartedly everybody in the Department of Mathematics and Statistics at the University of Limerick, particularly those involved in the MACSI (Mathematics Applications Consortium for Science and Industry) group who have provided me with a rich, enjoyable working environment.

- This work was funded by Science Foundation Ireland under programmes MACSI 06/MI/005 and 06/IN.1/I366.
Acronyms

UK  United Kingdom
ROM  Rolling Optimisation Model
SND  Seasonal Normal Demand
GARCH  Generalized Auto-Regressive Conditional Heteroskedasticity
AR  Auto-Regressive
AR(p)  Auto-Regressive model of order p
UKCS  United Kingdom Continental Shelf
LNG  Liquified Natural Gas
BBL  Balgzand Bacton Line
IUK  Interconnector United Kingdom
LRS  Long-Range-Storage
MRS  Medium-Range-Storage
SRS  Short-Range-Storage
SAP  System Average Price
LP  Linear Program
QP  Quadratic Program
KKT  Karush-Kuhn-Tucker
NCP  Non-linear Complementarity Problem
LCP  Linear Complementarity Problem
MCP  Mixed Complementarity Problem
MLCP  Mixed Linear Complementarity Problem
VI  Variational Inequality
WGM  World Gas Model
MCC  Market Clearing Conditions

GAMS  Generic Algebraic Modeling System

mcm  Million Cubic Meters
# Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>List of Figures</td>
<td>xi</td>
</tr>
<tr>
<td></td>
<td>List of Tables</td>
<td>xvii</td>
</tr>
<tr>
<td>1</td>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>UK natural gas market</td>
<td>5</td>
</tr>
<tr>
<td>3</td>
<td>Development of a stochastic process for UK gas demand</td>
<td>11</td>
</tr>
<tr>
<td>3.1</td>
<td>Actual demand and seasonal normal demand</td>
<td>12</td>
</tr>
<tr>
<td>3.2</td>
<td>Log model</td>
<td>17</td>
</tr>
<tr>
<td>3.2.1</td>
<td>Model comparison</td>
<td>23</td>
</tr>
<tr>
<td>3.2.2</td>
<td>GARCH model</td>
<td>24</td>
</tr>
<tr>
<td>3.3</td>
<td>Actual demand and predicted demand</td>
<td>25</td>
</tr>
<tr>
<td>3.3.1</td>
<td>Actual demand and two- to five-day ahead predictions</td>
<td>31</td>
</tr>
<tr>
<td>3.4</td>
<td>Stochastic process for demand</td>
<td>32</td>
</tr>
<tr>
<td>3.4.1</td>
<td>Numerical example</td>
<td>34</td>
</tr>
<tr>
<td>3.5</td>
<td>Summary</td>
<td>36</td>
</tr>
<tr>
<td>4</td>
<td>Gas market models</td>
<td>37</td>
</tr>
<tr>
<td>4.1</td>
<td>Theoretical background</td>
<td>39</td>
</tr>
<tr>
<td>4.1.1</td>
<td>Complementarity-based problems</td>
<td>40</td>
</tr>
<tr>
<td>4.1.2</td>
<td>Non-linear complementarity problem</td>
<td>40</td>
</tr>
<tr>
<td>4.1.3</td>
<td>Mixed complementarity problem</td>
<td>40</td>
</tr>
<tr>
<td>4.1.4</td>
<td>Variational inequality problems</td>
<td>41</td>
</tr>
<tr>
<td>4.1.5</td>
<td>Relationship between a QP and a LCP</td>
<td>41</td>
</tr>
<tr>
<td>4.1.6</td>
<td>Relationship between a QP and a MLCP</td>
<td>42</td>
</tr>
<tr>
<td>4.1.7</td>
<td>Relationship between non-linear complementarity problems and variational inequality problems</td>
<td>43</td>
</tr>
<tr>
<td>4.1.8</td>
<td>Relationship between mixed complementarity models and variational inequality problem</td>
<td>44</td>
</tr>
</tbody>
</table>
## CONTENTS

4.1.9 Summary .................................................. 44

4.2 Complementarity-based models of natural gas markets ....................... 45
  4.2.1 A complementarity-based equilibrium model and its KKT conditions .... 46
  4.2.2 Producer $p$’s problem .................................... 48
  4.2.3 Storage operator $so$’s problem ............................. 49
  4.2.4 Marketer’s problem ........................................... 50
  4.2.5 Market clearing conditions ................................. 52

4.3 Cost minimisation ............................................ 52
  4.3.1 A cost minimisation model and its KKT conditions .................... 53

4.4 Equivalence of models using KKT analysis .................................. 55
  4.4.1 Stationary KKT conditions ................................... 55
  4.4.2 Complementarity KKT conditions ................................ 56
  4.4.3 Constraints and Lagrange Multipliers ............................ 56
  4.4.4 Equivalence of models ......................................... 57

4.5 Equivalence of models using the principle of symmetry ...................... 57

4.6 Non-unique solutions ............................................ 61

4.7 Summary and conclusion .......................................... 63

5 Rolling Optimisation Model ...................................... 65
  5.1 Gas market models with stochastic demand .................................. 67
  5.2 Model .................................................................. 68
    5.2.1 Update rules after each roll ................................ 71
  5.3 ROM fitted to the year beginning in April 2010 ............................ 72
    5.3.1 Parameter estimation using simulated annealing .................... 81
  5.4 ROM tested for the year beginning in April 2011 .......................... 82
  5.5 Varying costs in ROM ............................................ 89
    5.5.1 KKT conditions of ROM .................................... 89
    5.5.2 Scaling the costs in ROM .................................... 92
    5.5.3 Shifting the production costs in ROM .......................... 93
    5.5.4 Proof that gas is emptied from storage .......................... 97
  5.6 Effect of increasing the number of scenarios used in ROM .................. 99
  5.7 Summary .......................................................... 103

6 Scenario reduction .................................................. 105
  6.1 Theoretical background ........................................... 106
  6.2 Algorithms ........................................................ 108
    6.2.1 Performance .................................................. 111
  6.3 Applications of scenario reduction algorithms .............................. 112
  6.4 Accuracy .......................................................... 113
    6.4.1 Alternative distance measure ................................ 115
# CONTENTS

6.5 Computational time ................................................................. 117
   6.5.1 Computational time associated with operations count ............... 121
6.6 Stability ................................................................................. 123
6.7 Comparison of scenario reduction algorithms: effect on ROM .......... 129
6.8 Benefit of using a scenario reduction algorithm with ROM .............. 133
6.9 Summary ................................................................................. 136

7 Applications of Rolling Optimisation Model ....................................... 137
   7.1 Base case .............................................................................. 138
   7.2 Increasing demand scenarios .................................................... 144
   7.3 Decreasing demand scenarios ................................................... 148
   7.4 Shocks to UK gas market ....................................................... 152
      7.4.1 Sudden drop in LNG supplies in July .................................. 152
      7.4.2 Sudden drop in LNG supplies in January ............................. 155
      7.4.3 Cold week in January ....................................................... 158
   7.5 Probability of supply being unable to meet demand ..................... 162
      7.5.1 Varying demand, LNG capacity and Norwegian capacity ........ 163
      7.5.2 Varying the start date ...................................................... 167
      7.5.3 Varying the initial amount of gas in storage ......................... 168
      7.5.4 Start date in January and effect of a particularly cold week ....... 170
   7.6 Summary ................................................................................. 173

8 Summary ....................................................................................... 175
   8.1 Future work .......................................................................... 178

Bibliography .................................................................................... 179

A Statistical analysis of actual demand and predicted demand ............... 185
   A.1 Two-day ahead predictions ...................................................... 185
   A.2 Three-day ahead predictions ................................................... 190
   A.3 Four-day ahead predictions .................................................... 195
   A.4 Five-day ahead predictions .................................................... 200

B Gas market models .................................................................... 205
   B.1 Gabriel et al’s mixed complementarity based model .................... 208
      B.1.1 Pipeline operator’s problem ............................................. 205
      B.1.2 Producer’s problem ....................................................... 208
      B.1.3 Storage operator’s problem ............................................. 206
      B.1.4 Marketer’s problem ....................................................... 207
      B.1.5 Peak gas operator’s problem .......................................... 209
      B.1.6 Market clearing conditions ............................................. 210
B.1.7 Market clearing conditions for pipeline a ........................................ 210
B.1.8 Market clearing conditions at production node n ................................ 210
B.1.9 Storage operator’s market clearing condition for node $n = n_1(a)$ ........... 211
B.1.10 Market clearing conditions for the peak gas operator ......................... 211
B.2 Matrix structure of complementarity-based model ................................. 211

C Rolling Optimisation Model .................................................................. 217
C.1 GAMS code associated with ROM for the year beginning in April 2010 .... 217
C.2 Parameter Estimation ........................................................................ 222
C.3 ROM tested for the year beginning in October 2010 .............................. 222
C.4 Effect of increasing number of scenarios on Rolling Optimisation Model .... 227

D Scenario reduction ............................................................................. 231
D.1 Stochastic process for demand based on 2008 -'09 data ......................... 231
D.2 Computational time associated with operations count ......................... 231
D.3 Comparison of the effect of using scenario reduction algorithms on ROM ... 233
D.4 Benefit of using a scenario reduction algorithm with ROM .................. 235
# List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Actual demand, SND and cold SND for the year starting on the 1st of April 2009.</td>
<td>6</td>
</tr>
<tr>
<td>2.2</td>
<td>Demand profile from the UK gas market starting on the 1st of April 2009.</td>
<td>7</td>
</tr>
<tr>
<td>2.3</td>
<td>Amount of gas in LRS starting on the 1st of April 2009.</td>
<td>8</td>
</tr>
<tr>
<td>2.4</td>
<td>Amount of gas in MRS starting on the 1st of April 2009.</td>
<td>8</td>
</tr>
<tr>
<td>2.5</td>
<td>Amount of gas in SRS starting on the 1st of April 2009.</td>
<td>9</td>
</tr>
<tr>
<td>2.6</td>
<td>SAPs for the year starting on the 1st of April 2009.</td>
<td>10</td>
</tr>
<tr>
<td>3.1</td>
<td>Time series of actual demand and SND.</td>
<td>12</td>
</tr>
<tr>
<td>3.2</td>
<td>Time series of the difference between actual demand and SND.</td>
<td>12</td>
</tr>
<tr>
<td>3.3</td>
<td>Auto-correlations for the difference between actual demand and SND.</td>
<td>13</td>
</tr>
<tr>
<td>3.4</td>
<td>Partial auto-correlations for the difference between actual demand and SND.</td>
<td>14</td>
</tr>
<tr>
<td>3.5</td>
<td>Residuals from the AR(1) process defined in equation (3.2) for the difference between actual demand and SND.</td>
<td>15</td>
</tr>
<tr>
<td>3.6</td>
<td>Auto-correlations for the residuals from the AR(1) process of equation (3.2).</td>
<td>16</td>
</tr>
<tr>
<td>3.7</td>
<td>Histogram for the residuals from the AR(1) process of equation (3.2).</td>
<td>16</td>
</tr>
<tr>
<td>3.8</td>
<td>Residual versus fitted value plot for the residuals from the AR(1) process of equation (3.2).</td>
<td>17</td>
</tr>
<tr>
<td>3.9</td>
<td>Time series of the difference between the natural log of actual demand and the natural log of SND.</td>
<td>18</td>
</tr>
<tr>
<td>3.10</td>
<td>Auto-correlations for the difference between the natural log of actual demand and the natural log of SND.</td>
<td>19</td>
</tr>
<tr>
<td>3.11</td>
<td>Partial auto-correlations for the difference between the natural log of actual demand and the natural log of SND.</td>
<td>19</td>
</tr>
<tr>
<td>3.12</td>
<td>Residuals from the AR(1) process for the difference between the natural log of actual demand and the natural log of SND as defined in equation (3.4).</td>
<td>21</td>
</tr>
<tr>
<td>3.13</td>
<td>Auto-correlations for the residuals from the AR(1) process of equation (3.4).</td>
<td>21</td>
</tr>
<tr>
<td>3.14</td>
<td>Histogram for the residuals from the AR(1) process of equation (3.4).</td>
<td>22</td>
</tr>
<tr>
<td>3.15</td>
<td>Residual versus fitted value plot for the residuals from the AR(1) process of equation (3.4).</td>
<td>22</td>
</tr>
<tr>
<td>3.16</td>
<td>Relative monthly variances for the residuals from the log model and the simple difference model.</td>
<td>24</td>
</tr>
<tr>
<td>3.17</td>
<td>Time series of actual demand and one-day ahead predictions.</td>
<td>26</td>
</tr>
<tr>
<td>3.18</td>
<td>The difference between the natural log of actual demand and the natural log of one-day ahead predictions.</td>
<td>27</td>
</tr>
<tr>
<td>3.19</td>
<td>Auto-correlations for the difference between the natural log of actual demand and the natural log of one-day ahead predictions.</td>
<td>28</td>
</tr>
<tr>
<td>FIGURE</td>
<td>DESCRIPTION</td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>-------------</td>
<td></td>
</tr>
<tr>
<td>3.20</td>
<td>Partial auto-correlations for the difference between the natural log of actual demand and the natural log of one-day ahead predictions.</td>
<td></td>
</tr>
<tr>
<td>3.21</td>
<td>Residuals from the AR(1) process defined in equation (3.15) for the difference between the natural log of actual demand and the natural log of one-day ahead predictions.</td>
<td></td>
</tr>
<tr>
<td>3.22</td>
<td>Auto-correlations for the residuals from the AR(1) process of equation (3.15).</td>
<td></td>
</tr>
<tr>
<td>3.23</td>
<td>Histogram for the residuals from the AR(1) process of equation (3.15).</td>
<td></td>
</tr>
<tr>
<td>3.24</td>
<td>The residual versus fitted value plot for the residuals from the AR(1) process of equation (3.15).</td>
<td></td>
</tr>
<tr>
<td>3.25</td>
<td>Actual demand (red line), seasonal normal demand (blue line), and a simulated demand paths (black line), for the gas year '08 - '09.</td>
<td></td>
</tr>
<tr>
<td>3.26</td>
<td>Actual demand (red line), seasonal normal demand (blue line), and 100 simulated demand paths (black lines), for the first 10 days of the gas year '08 - '09.</td>
<td></td>
</tr>
<tr>
<td>5.1</td>
<td>Rolling horizon of the sets of demand scenarios used in ROM.</td>
<td></td>
</tr>
<tr>
<td>5.2</td>
<td>Inputs and outputs associated with ROM.</td>
<td></td>
</tr>
<tr>
<td>5.3</td>
<td>Actual demand profile from the UK gas market starting on the 1st of April 2010.</td>
<td></td>
</tr>
<tr>
<td>5.4</td>
<td>Demand profile obtained from ROM for the UK gas market starting on the 1st of April 2010.</td>
<td></td>
</tr>
<tr>
<td>5.5</td>
<td>Actual and calibrated ROM SAPs starting on the 1st of April 2010.</td>
<td></td>
</tr>
<tr>
<td>5.6</td>
<td>Actual and simulated amount of LRS starting on the 1st of April 2010.</td>
<td></td>
</tr>
<tr>
<td>5.7</td>
<td>Actual and simulated amount of MRS starting on the 1st of April 2010.</td>
<td></td>
</tr>
<tr>
<td>5.8</td>
<td>Actual and simulated amount of SRS starting on the 1st of April 2010.</td>
<td></td>
</tr>
<tr>
<td>5.9</td>
<td>Actual demand profile from the UK gas market starting on the 1st of April 2011.</td>
<td></td>
</tr>
<tr>
<td>5.10</td>
<td>Demand profile obtained from ROM for the UK gas market starting on the 1st of April 2011.</td>
<td></td>
</tr>
<tr>
<td>5.11</td>
<td>Actual and calibrated predicted SAPs starting on the 1st of April 2011. These prices were calibrated using equation (5.13) with $\alpha = 6.48$ and $\beta = 323.324$.</td>
<td></td>
</tr>
<tr>
<td>5.12</td>
<td>Actual and simulated amount of LRS starting on the 1st of April 2011.</td>
<td></td>
</tr>
<tr>
<td>5.13</td>
<td>Actual and simulated amount of MRS starting on the 1st of April 2011.</td>
<td></td>
</tr>
<tr>
<td>5.14</td>
<td>Actual and simulated amount of SRS starting on the 1st of April 2011.</td>
<td></td>
</tr>
<tr>
<td>5.15</td>
<td>Generated demand paths with $n = 1000$ (red), $n = 100$ (blue) and $n = 10$ (yellow).</td>
<td></td>
</tr>
<tr>
<td>5.16</td>
<td>Mean plus error bars for the total yearly IUK supply from ROM for $n = 1, \ldots, 6, 10, 20, 50$.</td>
<td></td>
</tr>
<tr>
<td>5.17</td>
<td>Mean plus error bars for the total yearly LNG supply from ROM for $n = 1, \ldots, 6, 10, 20, 50$.</td>
<td></td>
</tr>
<tr>
<td>5.18</td>
<td>Mean plus error bars for the total yearly withdrawals from storage for ROM for $n = 1, \ldots, 6, 10, 20, 50$.</td>
<td></td>
</tr>
<tr>
<td>6.1</td>
<td>Average accuracy of scenario reduction algorithms for $N = 100$.</td>
<td></td>
</tr>
<tr>
<td>6.2</td>
<td>Distance between the centroid of the original and reduced probability distributions versus the size of the reduced set for $n = 1, \ldots, 99$ and $N = 100$.</td>
<td></td>
</tr>
<tr>
<td>6.3</td>
<td>Distance between the centroid of the original and reduced probability distributions versus the size of the reduced set for $n = 2, \ldots, 99$ and $N = 100$.</td>
<td></td>
</tr>
<tr>
<td>6.4</td>
<td>Average computational times associated with the backward reduction, simultaneous backward reduction and fast forward selection algorithms for $N = 100$.</td>
<td></td>
</tr>
<tr>
<td>6.5</td>
<td>Average computational times associated with the backward reduction, simultaneous backward reduction and fast forward selection algorithms for $N = 100$.</td>
<td></td>
</tr>
<tr>
<td>6.6</td>
<td>Average computational times associated with the backward reduction, simultaneous backward reduction and fast forward selection algorithms for $N = 1000$.</td>
<td></td>
</tr>
</tbody>
</table>
6.7 Average computational times associated with the backward reduction algorithm for $N = 1000$. .......................................................... 120
6.8 Log-log plot of computational time versus the number of original scenarios ($N$) for $n = 1$. .......................................................... 122
6.9 Log-log plot of computational time versus the number of original scenarios ($N$) for $n = 50$. .......................................................... 123
6.10 Difference between original and perturbed objective functions versus the size of the reduced scenarios sets ($n$) for $\epsilon = 10^{-5}$ and $N = 100$. Scenarios are initially assumed to have equal probabilities. ............................................. 125
6.11 Difference between original and perturbed objective functions versus the size of the reduced scenarios sets ($n$) for $\epsilon = 10^{-5}$ and $N = 100$. Scenarios are initially assumed to have equal probabilities. ............................................. 125
6.12 Difference between original and perturbed objective functions versus the size of the reduced scenarios sets ($n$) for $\epsilon = 10^{-10}$ and $N = 100$. Scenarios are initially assumed to have equal probabilities. ............................................. 126
6.13 Difference between original and perturbed objective functions versus the size of the reduced scenarios sets ($n$) for $\epsilon = 10^{-10}$ and $N = 100$. Scenarios are initially assumed to have equal probabilities. ............................................. 126
6.14 Difference between original and perturbed objective functions versus the size of the reduced scenarios sets ($n$) for $\epsilon = 10^{-5}$ and $N = 100$. Scenarios are initially assumed to have random probabilities. ............................................. 128
6.15 Difference between original and perturbed objective functions versus the size of the reduced scenarios sets ($n$) for $\epsilon = 10^{-5}$ and $N = 100$. Scenarios are initially assumed to have random probabilities. ............................................. 129
6.16 Mean plus error bars for the total yearly supply from Norway for $n = 1,\ldots,6,10$ and $N = 1000$. .......................................................... 130
6.17 Mean plus error bars for the total yearly supply through IUK for $n = 1,\ldots,6,10$ and $N = 1000$. .......................................................... 131
6.18 Mean plus error bars for the total yearly supply through BBL for $n = 1,\ldots,6,10$ and $N = 1000$. .......................................................... 131
6.19 Mean plus error bars for the total yearly supply from LNG for $n = 1,\ldots,6,10$ and $N = 1000$. .......................................................... 132
6.20 Mean plus error bars for the total yearly supply from the UKCS using the fast forward selection algorithm for $N = 100,1000$ and with no reduction. ...................... 133
6.21 Mean plus error bars for the total yearly injections to storage using the fast forward selection algorithm for $N = 100,1000$ and with no reduction. ...................... 134
6.22 Mean plus error bars for the total yearly withdrawals from storage using the fast forward selection algorithm for $N = 100,1000$ and with no reduction. ...................... 134
6.23 Mean plus error bars for the total yearly supply from through IUK using the fast forward selection algorithm for $N = 100,1000$ and with no reduction. ...................... 135

7.1 Demand profile obtained from ROM for the UK gas market starting on the 1st of April 2012. .......................................................... 141
7.2 SAPs produced by ROM for the year starting on the 1st of April 2012. .......................................................... 142
7.3 Amount of gas in LRS produced by ROM for the year starting on the 1st of April 2012. .......................................................... 143
7.4 Amount of gas in MRS produced by ROM for the year starting on the 1st of April 2012. ................................................................. 143
7.5 Seasonal normal demand increased by 0%, 5%, 10%, 15% and 20%. ................................. 144
7.6 Total yearly supplies obtained from ROM for increasing levels of SND. .......................... 145
7.7 SAPs obtained from ROM for increasing levels of SND. .................................................. 146
7.8 The amount of gas in LRS produced by ROM for increasing levels of SND. .................. 147
7.9 The amount of gas in MRS produced by ROM for increasing levels of SND. .................. 147
7.10 Seasonal normal demand decreased by 5%, 10%, 15% and 20%. ................................. 148
7.11 Total yearly supplies predicted by ROM for decreasing levels of SND. .......................... 149
7.12 SAPs predicted by ROM for decreasing levels of SND. ................................................ 150
7.13 The amount of gas in LRS as produced by ROM for decreasing levels of SND. .............. 151
7.14 The amount of gas in MRS as produced by ROM for decreasing levels of SND. .............. 151
7.15 Demand profile obtained from ROM for the UK gas market when there is a sudden drop in LNG supplies in July. ......................................................... 153
7.16 Amount of gas in LRS as produced by ROM for both the base case and when there is a sudden drop in LNG supplies in July. .............................................. 154
7.17 Amount of gas in MRS as produced by ROM for both the base case and when there is a sudden drop in LNG supplies in July. .............................................. 154
7.18 SAPs obtained from ROM for both the base case and when there is a sudden drop in LNG supplies in July. .................................................. 155
7.19 Demand profile obtained from ROM for the UK gas market when there is a sudden drop in LNG supplies in January. ......................................................... 156
7.20 Amount of gas in LRS as produced by ROM for both the base case and when there is a sudden drop in LNG supplies in January. .............................................. 157
7.21 Amount of gas in MRS as produced by ROM for both the base case and when there is a sudden drop in LNG supplies in January. .............................................. 157
7.22 SAPs obtained from ROM for both the base case and when there is a sudden drop in LNG supplies in January. .................................................. 158
7.23 Time series of SND and SND with a cold week at the start of January. ........................ 159
7.24 Demand profile obtained from ROM for the UK gas market when there is a cold week in January. ................................................................. 160
7.25 Amount of gas in LRS as obtained by ROM for both the base case and when there is a cold week in January. ......................................................... 161
7.26 Amount of gas in MRS as obtained by ROM for both the base case and when there is a cold week in January. ......................................................... 161
7.27 SAPs obtained from ROM for both the base case and when there is a cold week in January. ................................................................. 162
7.28 Fraction of infeasible solutions for increased SND and/or changes in LNG capacity. .... 164
7.29 Fraction of infeasible solutions for increased SND and/or changes in Norwegian capacity. ......................................................................................... 165
7.30 Fraction of infeasible solutions with various changes to LNG and Norwegian capacity. ................................................................. 166
7.31 Number of infeasible solutions with for different start days, 30% LNG capacity and 10% Norwegian capacity. ................................................................. 168
7.32 Fraction of infeasible solutions starting on 2 different days with varying initial amounts of gas in storage; 10% LNG capacity and 20% Norwegian capacity. ........................................................................... 169
LIST OF FIGURES

7.33 Time series with cold SND for the first week and SND thereafter and time series of SND. ......................................................... 170
7.34 Fraction of infeasible solutions for changing levels of LNG capacity and Norwegian capacity for a day starting in January based on SND. ......................................................... 172
7.35 Fraction of infeasible solutions for changing levels of LNG capacity and Norwegian capacity for a day starting in January based on cold SND. ................................. 172
A.1 Time series of actual demand and two-day ahead predictions. ......................................................... 185
A.2 The difference between the natural log of actual demand and the natural log of two-day ahead predictions. ......................................................... 186
A.3 Auto-correlations for the difference between the natural log of actual demand and the natural log of two-day ahead predictions. ......................................................... 186
A.4 Partial auto-correlations for the difference between the natural log of actual demand and the natural log of two-day ahead predictions. ......................................................... 187
A.5 Residuals from the AR(1) process defined in equation (3.16) for the difference between the natural log of actual demand and the natural log of two-day ahead predictions. ......................................................... 187
A.6 Auto-correlations for the residuals from the AR(1) process of equation (3.16). ......................................................... 188
A.7 Histogram for the residuals from the AR(1) process of equation (3.16). ......................................................... 188
A.8 The residual versus fitted value plot for the residuals from the AR(1) process of equation (3.16). ......................................................... 189
A.9 Time series of actual demand and three-day ahead predictions. ......................................................... 190
A.10 The difference between the natural log of actual demand and the natural log of three-day ahead predictions. ......................................................... 191
A.11 Auto-correlations for the difference between the natural log of actual demand and the natural log of three-day ahead predictions. ......................................................... 191
A.12 Partial auto-correlations for the difference between the natural log of actual demand and the natural log of three-day ahead predictions. ......................................................... 192
A.13 Residuals from the AR(1) process defined in equation (3.17) for the difference between the natural log of actual demand and the natural log of three-day ahead predictions. ......................................................... 192
A.14 Auto-correlations for the residuals from the AR(1) process of equation (3.17). ......................................................... 193
A.15 Histogram for the residuals from the AR(1) process of equation (3.17). ......................................................... 193
A.16 The residual versus fitted value plot for the residuals from the AR(1) process of equation (3.17). ......................................................... 194
A.17 Time series of actual demand and four-day ahead predictions. ......................................................... 195
A.18 The difference between the natural log of actual demand and the natural log of four-day ahead predictions. ......................................................... 196
A.19 Auto-correlations for the difference between the natural log of actual demand and the natural log of four-day ahead predictions. ......................................................... 196
A.20 Partial auto-correlations for the difference between the natural log of actual demand and the natural log of four-day ahead predictions. ......................................................... 197
A.21 Residuals from the AR(1) process defined in equation (3.18) for the difference between the natural log of actual demand and the natural log of four-day ahead predictions. ......................................................... 197
A.22 Auto-correlations for the residuals from the AR(1) process of equation (3.18). ......................................................... 198
A.23 Histogram for the residuals from the AR(1) process of equation (3.18). ......................................................... 198
A.24 The residual versus fitted value plot for the residuals from the AR(1) process of equation (3.18).......................... 199
A.25 Time series of actual demand and five-day ahead predictions................................................................. 200
A.26 The difference between the natural log of actual demand and the natural log of five-day ahead predictions. .................. 201
A.27 Auto-correlations for the difference between the natural log of actual demand and the natural log of five-day ahead predictions. ................................................................. 201
A.28 Partial auto-correlations for the difference between the natural log of actual demand and the natural log of five-day ahead predictions. ................................................................. 202
A.29 Residuals from the AR(1) process defined in equation (3.19) for the difference between the natural log of actual demand and the natural log of five-day ahead predictions. ................................................................. 202
A.30 Auto-correlations for the residuals from the AR(1) process of equation (3.19) .............................................. 203
A.31 Histogram for the residuals from the AR(1) process of equation (3.19) ...................................................... 203
A.32 The residual versus fitted value plot for the residuals from the AR(1) process of equation (3.19) ................................................................. 204
C.1 Actual demand profile from the UK gas market starting on the 1st of October 2010. ........................................ 224
C.2 Predicted demand profile from the UK gas market starting on the 1st of October 2010 ........................................ 224
C.3 Actual and calibrated predicted SAPs starting on the 1st of October 2010. These prices were calibrated using equation (5.13) with $\alpha = 1.461$ and $\beta = 31.481$. .......................... 225
C.4 Actual and simulated amount of LRS starting on the 1st of October 2010 ....................................................... 225
C.5 Actual and simulated amount of MRS starting on the 1st of October 2010 ....................................................... 226
C.6 Actual and simulated amount of SRS starting on the 1st of October 2010 ....................................................... 226
C.7 Mean plus error bars for the total yearly UKCS supply from ROM for $n = 1, \ldots, 6, 10, 20, 50$. ......................... 227
C.8 Mean plus error bars for the total yearly Norwegian supply from ROM for $n = 1, \ldots, 6, 10, 20, 50$. ................. 228
C.9 Mean plus error bars for the total yearly injections to storage for ROM for $n = 1, \ldots, 6, 10, 20, 50$. ................. 228
C.10 Mean plus error bars for the total yearly BBL supply from ROM for $n = 1, \ldots, 6, 10, 20, 50$. ......................... 229
D.1 Log-log plot of computational time versus the number of original scenarios ($N$) for $n = N/2$. .................. 232
D.2 Log-log plot of computational time versus the number of original scenarios ($N$) for $n = N - 1$. ................. 233
D.3 Mean plus error bars for the total yearly supply from the UKCS for $n = 1, \ldots, 6, 10$ and $N = 1000$. ............... 234
D.4 Mean plus error bars for the total yearly injections into storage for $n = 1, \ldots, 6, 10$ and $N = 1000$. ............... 234
D.5 Mean plus error bars for the total yearly withdrawals from storage for $n = 1, \ldots, 6, 10$ and $N = 1000$. ............... 235
D.6 Mean plus error bars for the total yearly supply from Norway using the fast forward selection algorithm for $N = 100, 1000$ and with no reduction. ........................................ 236
D.7 Mean plus error bars for the total yearly supply through BBL using the fast forward selection algorithm for $N = 100, 1000$ and with no reduction. ........................................ 236
D.8 Mean plus error bars for the total yearly supply from LNG using the fast forward selection algorithm for $N = 100, 1000$ and with no reduction. ........................................ 237
List of Tables

4.1 Parameters and variables associated with the complementarity-based equilibrium model ........................................ 48
4.2 Inputs associated with the producers for both models ......................................................................................... 61
4.3 Inputs associated with demand for both models ................................................................................................. 61
4.4 Inputs associated with the storage operator for both models ............................................................................. 62
4.5 Outputs to both models ........................................................................................................................................ 63

5.1 Outputs associated with ROM. ......................................................................................................................... 70
5.2 Parameters associated with ROM. ................................................................................................................... 70
5.3 Inputs associated with ROM. .......................................................................................................................... 70
5.4 The costs and maximum capacities (in mcm) of the different tranches for each sources of supply. .................. 74
5.5 The total maximum capacity (in mcm) for each source of supply. ................................................................. 75
5.6 Parameters associated with long-, medium- and short-range storage (in mcm) for ROM .................................. 75
5.7 Short- and long-run cost associated with long-, medium- and short-range storage for ROM.......................... 76
5.8 Average daily flows obtained from ROM (in mcm) as well as actual daily average flows for the UK gas market for year beginning on the 1st of April 2010. 78
5.9 The costs and maximum capacities (in mcm) of the different tranches for each sources of supply. .................. 83
5.10 The total maximum capacity (in mcm) for each source of supply. ................................................................. 83
5.11 Parameters associated with long-, medium- and short-range storage (in mcm) for the Rolling Optimisation Model for the year starting on the 1st of April 2012. 84
5.12 Average daily flows obtained from ROM (in mcm) as well as actual daily average flows for the UK gas market for year beginning on the 1st of April 2011. 86
5.13 Mean GAMS solve times (in hours) associated with ROM, for $n = 1, \ldots, 6, 10, 20, 50$. 103

7.1 The costs and maximum capacities (in mcm) of the different tranches for each source of supply for the year starting on the 1st of April 2012. 139
7.2 The total maximum capacity (in mcm) for each source of supply. 139
7.3 Parameters associated with long-, medium- and short-range storage (in mcm) for the Rolling Optimisation Model for the year starting on the 1st of April 2012. 140
7.4 Fraction of infeasible solutions for increased SND and/or changes in LNG capacity. 164
7.5 Fraction of infeasible solutions for increased SND and/or changes in Norwegian capacity. 165
7.6 Fraction of infeasible solutions with various changes to LNG and Norwegian capacity. ................................................. 166
7.7 Fraction of infeasible solutions for changing LNG and Norwegian capacity for a day starting in January based on SND. ................................................. 171
7.8 Fraction of infeasible solutions for changing LNG and Norwegian capacity for a day starting in January based on cold SND. ................................................. 171
B.1 Variables associated with the marketers problem of Gabriel et al’s model ................................................. 209
C.1 The costs and maximum capacities (in mcm) of the different tranches for each source of supply. ................................................. 223
C.2 The total maximum capacity (in mcm) for each source of supply, [1, 2]. ................................................. 223
C.3 Parameters associated with long-, medium- and short-range storage (in mcm) for ROM, [1, 3]. ................................................. 223
Chapter 1

Introduction

Energy markets across the world are constantly changing. The Irish government has committed to increasing the proportion of electricity generated by clean and renewable sources to 40% by 2020 [4, 5]. Similarly, in Britain the government has committed to having 15% of their energy generated by renewable sources of energy in 2020 [6, 7]. For both countries, wind energy is seen as a major contributor to this renewable energy. This has led to huge increases in the amount of electricity generated by wind energy across Ireland and the UK in recent years [5, 7].

The advantage of harnessing wind energy is that it is extremely clean (i.e., no harmful emissions) and is a sustainable source of energy. The disadvantage is that it is intermittent, i.e., it is not always available. When the wind is not blowing there is no energy to be harnessed; moreover, when the wind is blowing at very high speeds it is also not possible to harness energy from it as damage may be caused to the wind turbines. As a result of this intermittency the different players in the electricity market have been trying to make the market more flexible so it can handle the uncertainty of the wind power i.e., when the energy from wind is not available, there must be a sufficient amount of energy sources available in order to meet demand. Obtaining flexibility in the market is a complex task and there many ways of doing so, such as investing in electricity storage facilities or investing in interconnectors between the national grids of different countries [8]. It is also important to examine the flexibility provided by existing energy sources such as gas. This is particularly true for the industrial partner of this project, Bord Gáis Energy [1].

In both Ireland and the UK, natural gas is the single largest source of energy used to generate electricity, see [9] and [7] respectively. As wind energy is an intermittent source of energy, the increased levels of wind power in electricity markets in both the UK and Ireland has led to, and will continue to cause, variability in the amount of gas required to generate electricity. This is because when wind energy is available to be harnessed for electricity power generation, gas is not needed as much. However, when wind energy is not available demand for gas to generate electricity is increased. With the amount of wind farms increasing, along with the volatile nature of weather, this has led to an increase in uncertainty in the gas markets of the UK and Ireland.

The Irish gas market relies heavily on the UK gas market, as Ireland imports most of its gas from the UK; in fact, about 93% of Ireland’s gas supplies came from imports from the UK [9] in 2011. The Irish gas market is also very small in comparison to the UK, as these imports only make up approximately 6% of the total UK gas supply [1,7].

This thesis details the development of a model of the UK natural gas market. As the Irish market is so small in comparison, it is not considered separately. The aim of the project is to model the flows of gas in the UK whilst incorporating the uncertain nature of demand. The model developed here is called the Rolling Optimisation Model (ROM) and is introduced in Chapter 5. A mathematical model of the UK natural gas market has not previously appeared in the literature. In order to capture the uncertainty of demand, ROM takes as an input demand scenarios simulated from a stochastic process of UK gas demand. This process is developed in Chapter 3. It reflects the daily information that is known in the UK gas market regarding demand. On any given day, demand for that day is exactly known and is scenario-independent. For the next five days the values of demand show increasing uncertainty, i.e., it varies from scenario to scenario. After five days ahead all that is assumed known is seasonal averages. The development of this stochastic process is informed by the statistical analysis of data supplied from the UK National Grid, the owner and operator of the national transmission system throughout Great Britain [10].

Before a detailed model of the UK gas market is introduced, the different methodologies used to model gas markets must be studied. This is the subject of Chapter 4. There are two main methods used to model gas markets, namely complementarity-based equilibrium models and cost minimisation models. Complementarity-based models allow for multiple objective functions. In these models each player in the market maximises their own profits through their own individual objective functions. This allows for market power (i.e., the power for one or more players in the model to influence the level of demand) in the model. In cost minimisation models, instead of each player maximising their own profits separately, there is an imaginary central planner that chooses how the different players in the market meet demand at least cost to the overall system. This is done using a single linear program and hence has only one objective function. As a result, market power cannot be incorporated into these types of models.

In Section 4.2.1 a complementarity-based equilibrium model is introduced. It is a simplified version of Gabriel et al’s model [11] as, among other things, it does not contain market power. Market power is excluded as it is not appropriate, in most reasonable cases, in the context of the UK market, [12]. In Section 4.3.1 a corresponding cost minimisation model is presented while in Sections 4.4 and 4.5 these two models are shown to be equivalent. A detailed analysis of this equivalence has not been previously seen.

As each player acts separately in order to maximise their own profits, complementarity-based equilibrium models may be seen as an economically realistic approach to modelling natural gas markets. In contrast, cost minimisation, with its imaginary central planner, may be seen as unrealistic. Cost minimisation models are, however, algebraically and structurally less complicated than complementarity-based models. The equivalence (of the two models) result in Chapter 4 shows
that while cost minimisation may be unintuitive for natural gas markets, it can be used instead of the more intuitive, but complex, complementarity-based equilibrium approach when modelling markets without market power.

Using the results from Chapters 3 and 4, a detailed model of the UK gas market (i.e., ROM) is introduced in Chapter 5. As mentioned already, the aim of this model is to simulate the flows of gas in the UK whilst incorporating the uncertain nature of demand. The model captures information on how the different sources of supply meet demand, as well as the amount of gas being injected to, and withdrawn from, storage. In addition to this, ROM also models the daily System Average Price (SAP) in the UK. SAP is average price of all gas traded on a given day in the UK [3]. An introduction into these different aspects of the UK gas market is provided in Chapter 2.

ROM decides the amount of gas to be injected to, and withdrawn from, the different storage facilities, as well as how the different sources of supply in the UK meet demand. ROM is based on the cost minimisation model described in Section 4.3.1 which means that these decisions are made in such a way that minimises overall system cost whilst also ensuring possible future demands are met at minimum expected cost. Once these decisions are made the model moves (‘rolls’) to a new day where the same decisions are made using updated demand scenarios. Once these decisions are made the model moves forward again in a similar manner and so on. In Section 5.3 the parameters of ROM are chosen so as to best fit supply data from the UK gas market for the year beginning on the 1st of April 2010. Using similar parameters, ROM is then tested in Section 5.4 for data from the UK gas market for the year starting on the 1st of April 2011. In Section 5.6 the effect of increasing the number of demand scenarios used in the model is examined. This analysis finds that as the number of scenarios used in ROM increases, the amount of gas supplied by the different sources of supply varies.

Ideally ROM should be run with as many demand scenarios as possible. However, as discussed in Section 5.6, the computational cost associated with ROM increases as the number of scenarios increases. This provides the motivation for the introduction of scenario reduction techniques in Chapter 6. The aim of this analysis is to use demand scenarios that accurately represent larger scenario sets whilst at the same time maintaining a relatively low computational time for ROM. Three different scenario reduction techniques, based on the works described in [13, 14], are introduced in Chapter 6. These heuristic algorithms are known as backward reduction, simultaneous reduction and fast forward selection. They are applied to demand scenarios developed from the stochastic process for demand and are examined in terms of their accuracy, stability and computational time. This is the first time that these algorithms are applied to gas demand scenarios; in [13, 14] they were applied to electrical load scenarios. Chapter 6 also examines the effect of using scenarios obtained after scenario reduction on the outputs of ROM.

In the final analysis chapter of the thesis, ROM is used to investigate various stresses on the UK gas market in Chapter 7. The first stress test examines scenarios with increased and decreased levels of demand. Following this, ROM is used to investigate the effect of various potential shocks to the UK gas market. These shocks include an extremely cold winter, as well as sudden drops in
1. INTRODUCTION

the availability in some of the larger sources of supply. Chapter 7 concludes with a ROM-based analysis of the probability of supply being unable to meet demand in the UK.

The remainder of this thesis is organised as follows. Firstly, in Chapter 2 a brief introduction to the various aspects of the UK gas market is presented. In Chapter 3 the stochastic process for UK gas demand is developed. In Chapter 4 the main methodologies used to model natural gas markets are compared. After this, the Rolling Optimisation Model is introduced in Chapter 5 and scenario reduction techniques are applied to gas demand scenarios in Chapter 6. Some of the applications of ROM are examined in Chapter 7. The thesis concludes in Chapter 8 with a summary and a discussion of some possible future work.
Chapter 2

UK natural gas market

As stated in the previous chapter, the goal of this project is to model the UK gas market whilst incorporating the stochastic nature of demand. Before the main work of the thesis is presented, a short introduction to the UK natural gas market is provided in this chapter. In particular, the different sources of supply and demand are described, as well as the different types of storage facilities. A description of wholesale gas prices is also provided.

Demand for gas in the UK comes from four main sectors [2, 15], which are

1. Domestic
2. Industrial & commercial
3. Power generation
4. Exports.

In 2010 these sources made up approximately 32%, 24%, 29% and 15% of the total UK gas demand [2], respectively. The gas demands in the domestic and commercial sectors are heavily influenced by temperature levels [16]. As these are large sources of demand, this means that gas demand in the UK is highly seasonal.
2. UK NATURAL GAS MARKET

Figure 2.1: Actual demand, SND and cold SND for the year starting on the 1st of April 2009.

Figure 2.1 displays actual gas demand (black line) in the UK for each day in the year starting on the 1st of April 2009. They are given in units of million cubic meters (mcm). The data for this time series was obtained from the UK National Grid’s website [3]. Figure 2.1 also shows Seasonal Normal Demand (SND) for the same period. SND is the daily gas demand that one would expect in an average weather year. It is calculated by the UK National Grid and is published approximately one year in advance [17]. SND is used throughout this thesis, particularly in Chapter 3 in the development of a stochastic process of UK gas demand. Figure 2.1 also displays the Cold SND time series over the same period. Cold SND is the gas demand one would expect in cold weather conditions and indicates the variation in demand that may occur throughout any given year. Like SND, it is calculated by the UK National Grid and is available from their website [17]. It is used in Chapter 7 where the Rolling Optimisation Model is used for various stress tests.

There are five main sources of supply in the UK gas market that meet demand [2, 15]. These are

1. UK Continental Shelf (UKCS)
2. Norwegian imports
3. Liquefied Natural Gas (LNG)
4. Balgzand Baction Line (BBL)
5. Interconnector UK (IUK).

Figure 2.2 shows how these different sources of supply met demand for the year starting in April 2009.
Figure 2.2: Demand profile from the UK gas market starting on the 1st of April 2009.

The largest source of supply is the UKCS. This is the UK’s indigenous gas supply that is taken from the seas surrounding Britain. Norwegian imports represent gas imported to the UK from Norway through pipelines. LNG is natural gas in liquid form. It is obtained by cooling natural gas to \(-161\) degrees Centigrade \[18\]. This liquifies the gas and hence makes it easier to ship around the world. The UK imports LNG from places such as Algeria, Trinidad and the Middle East \[18\]. The BBL pipeline links England and the Netherlands and hence represents imports into the UK from that destination \[19\]. Similarly, the IUK pipeline is a pipeline between England and Belgium \[20\]. The IUK pipeline can cater for exports from the UK to Belgium.

As well as the various sources of supply, Figure 2.2 also indicates how the different storage facilities in the UK meet demand. There are eight gas storage facilities in the UK. These facilities can be classified into three main groups namely, long- (LRS), medium- (MRS) and short-range (SRS) storage facilities. Figures 2.3 - 2.5 displays the amount of gas in these three types of storage facilities respectively for the year starting in April 2009.
2. UK NATURAL GAS MARKET

![Graph showing the amount of gas in LRS starting on the 1st of April 2009.](image1.png)

**Figure 2.3:** Amount of gas in LRS starting on the 1st of April 2009.

![Graph showing the amount of gas in MRS starting on the 1st of April 2009.](image2.png)

**Figure 2.4:** Amount of gas in MRS starting on the 1st of April 2009.
Figure 2.5: Amount of gas in SRS starting on the 1st of April 2009.

Figure 2.3 shows the amount of gas in LRS for this period. It indicates that LRS captures the seasonal variation in demand in the UK gas market whereby gas is injected in the relatively low-demand summer and withdrawn in the relatively high-demand winter. There is only one LRS in the UK and that is the Rough storage facility [1,7]. Figure 2.4 displays a similar plot for the amount of gas in MRS. In contrast to Figure 2.3, it indicates that MRS captures the weekly/monthly variations in gas demand in the UK, as well as the seasonal variation in demand – note the difference in the vertical scales between Figures 2.3 and 2.4. There are six MRS facilities in the UK: Hornsea, Holehouse Farm, Hatfield Moor, Humbly Grove, Aldbrough and Holford [1,21]. Figure 2.5 displays a similar graph for SRS. SRS is also known as LNG storage [1,7] as it involves storing gas by freezing it into its liquid form [22]. Prior to May 2011, there were three SRS facilities in the UK: Avonmouth, Glenmavis and Partington. However, in May 2011, the Glenmavis and Partington facilities stopped offering commercial services [2]. As a result, SRS has become a minor part of the UK gas market and is of relatively little significance to this project.

\[\text{The Holford MRS facility only became operational in 2012 and is, hence, not included in Figure 2.4.}\]
Figure 2.6 displays the daily System Average Price (SAP) for the year starting in April 2009. The unit in which these are given is pence per therm\(^1\). SAP is the average price of all gas traded on a given day in the UK gas market \[^3\]. The plot indicates that there is a somewhat seasonal trend to SAPs whereby prices are relatively high in winter and relatively low in summer. This suggests that seasonal varying demand strongly affects gas prices in the UK gas market. This is true, however there are many other factors to be considered. For example, many gas pricing contracts in the UK, and particularly in Europe, are index-linked to oil prices \[^{16}\]. This means that when oil prices increase or decrease significantly so do gas prices. Other factors that have affected UK gas prices in the past include major international events such as the Russia-Ukraine dispute in 2009, as well as the Japanese nuclear power plant outages in 2011 \[^2\].

Having outlined some of the details of the UK natural gas market, the remaining chapters focus on developing on a model that captures the qualitative features of this market.

\(^1\)One mcm is approximately 360000 therms \[^{23}\].
Chapter 3

Development of a stochastic process for UK gas demand

In this chapter a stochastic process that captures the uncertainty of gas demand in the UK is developed. This is informed by analysis of historic UK natural gas demand data. In particular, the following relationships are examined:

1. The difference between actual demand & seasonal normal demand,
2. The difference between actual demand & predicted demand.

Actual demand is the time series of the daily figures given by the UK National Grid for overall system gas demand in the UK. Seasonal normal demand (SND) is the daily time series for an average gas demand year. In other words, it is the daily gas demand one would expect in an average year. It is calculated by the UK National Grid and is published approximately one year in advance. Predicted demand is the time series of the predicted daily demand given by the UK National Grid. There are five types of predicted demand supplied by the UK National Grid, one- to five-day ahead predictions. One-day ahead predictions are the daily demands predicted one day beforehand. Similarly two- to five-day ahead predictions are the daily demands predicted two to five days beforehand. The time series for each of these predictions are all considered in the analysis of this chapter. The data for each of these time series was obtained from the UK National Grid’s website.

This chapter is organised as follows: firstly, in Section 3.1 a statistical analysis of the difference between actual demand and SND is presented. In Section 3.2 a similar analysis is provided for the difference between the natural log of these two series. Following this, in Section 3.3 the relationship between actual demand and predicted demand is analysed. Finally, in Section 3.4 the results from the rest of the chapter are used to develop a stochastic process of UK gas demand. The chapter concludes in Section 3.5 with a summary.
3. DEVELOPMENT OF A STOCHASTIC PROCESS FOR UK GAS DEMAND

3.1 Actual demand and seasonal normal demand

Figure 3.1: Time series of actual demand and SND.

Figure 3.2: Time series of the difference between actual demand and SND.
In this section the difference between actual demand and SND is considered. Both time series start on the 1st of October 2007 and end on the 31st of March 2012 and have been obtained from the UK National Grid’s website, [3]. Figure 3.1 shows the time series of actual demand and SND in million cubic metres (mcm) of gas. It indicates that the gas demand in the UK is highly seasonal. In the summer demand is relatively low, while in winter demand is relatively high. This is intuitive given that demand for natural gas is heavily dependent on weather. Figure 3.2 displays the difference between the two time series. It shows that the difference tends to be larger in the winter than in the summer.

Figure 3.3 gives the autocorrelations for the difference between the two time series while Figure 3.4 shows the partial autocorrelations. The auto-correlations at each lag in Figure 3.3 are highly statistically significant. This figure also shows that the auto-correlations decrease as the lag increases. This, along with the fact that only at lag 1 is there a highly statistically significant partial auto-correlation, suggests that this time series follows an auto-regressive process of order one (AR(1)), [24].

Figure 3.3: Auto-correlations for the difference between actual demand and SND.
3. DEVELOPMENT OF A STOCHASTIC PROCESS FOR UK GAS DEMAND

Figure 3.4: Partial auto-correlations for the difference between actual demand and SND.

An auto-regressive model of order $p$ (AR($p$)), for the time series $r_t$, is defined in [24] as follows:

$$r_t = \phi_0 + \sum_{i=1}^{p} \phi_i r_{t-i} + \epsilon_t,$$

(3.1)

where $\phi_i$ is the parameter associated with the $i$th lag and $\epsilon_t$ is a white noise series with mean 0 and variance $\sigma^2$. The constant in the model is $\phi_0$. An AR($p$) model is deemed to be an adequate model if its residuals behave as white noise, [24]. A white noise process is a sequence of independent and identically distributed random variables with finite mean and variance. This means that the auto-correlations of the residuals must be zero. However, in practice, if the auto-correlations are close to zero, then the series can assumed to be white noise, [24]. Throughout this chapter the simplest time series model found to be adequate is used to model the time series in question. This, invariably, is an AR(1) model. However, a GARCH model is considered in Section 3.2.2.

Using the statistical software package R, an AR(1) process can be fitted to the difference between actual demand and SND as follows:

$$\text{ActDem}_t - \text{SND}_t = 0.93(\text{ActDem}_{t-1} - \text{SND}_{t-1}) + \epsilon_t,$$

(3.2)

where $\text{ActDem}_t$ represents actual demand at time $t$ and $\text{SND}_t$ represents seasonal normal demand at time $t$. The noise term $\epsilon_t$ represents Gaussian white noise with a mean of zero and variance of one while $\sigma^2 = 163$ represents the variance of the residuals. The p-value associated with the t-test test statistic is $< 0.001$ which suggests that the AR(1) coefficient of 0.93 is highly statistically
Figure 3.5 displays the residuals from this model. It shows that the variance of the residuals are larger in the winter (October to April) than in the summer (April to October). Figure 3.6 indicates that there are no highly statistically significant auto-correlations in the residuals from this process, which suggests that the residuals are uncorrelated in time.

Figure 3.5: Residuals from the AR(1) process defined in equation (3.2) for the difference between actual demand and SND.
Figure 3.6: Auto-correlations for the residuals from the AR(1) process of equation (3.2).

Figure 3.7: Histogram for the residuals from the AR(1) process of equation (3.2).
Figure 3.8: Residual versus fitted value plot for the residuals from the AR(1) process of equation (3.2).

Figure 3.7 shows that the histogram of the residuals is relatively bell-shaped and symmetric. This suggests that the residuals are approximately normally distributed. The Kolmogorov-Smirnov test gives a p-value of 0.153, which is above the significance threshold of 0.05, thus confirming that the residuals are normally distributed. Figure 3.8 displays the residuals versus fitted values plot. It shows that the residuals are scattered randomly around zero which suggests that the residuals are randomly distributed.

Figures 3.6 - 3.8 suggest that the residuals of this model behave as a Gaussian white noise process. This implies that the model is an adequate model, [24]. The model assumes that, for a given day, actual demand is equal to SND for that day plus some random noise plus some dependence on the size of the error from the previous day. Thus, if actual demand is low (or high) on a given day, then it is likely be low (or high) again on the next day.

3.2 Log model

Figure 3.2 implies that the difference between actual demand and SND tends to vary more in the winter time (i.e., when demand is high) than in the summer time. This suggests that the model described above may not be a suitable choice as it assumes that the residuals have constant variance. As a result, the difference between the natural log of actual demand and the natural log

---

1Fitted values are the values that the model predicts for the difference between actual demand and SND.
of SND, i.e.,

$$\ln(\text{ActDem}_t) - \ln(\text{SND}_t),$$

(3.3)

is now considered. Taking the difference of the natural logs allows for the variance of residuals to vary with demand. Figure 3.9 shows that this difference does not vary as much as in Figure 3.2.

Figure 3.9: Time series of the difference between the natural log of actual demand and the natural log of SND.
3.2 Log model

Figure 3.10: Auto-correlations for the difference between the natural log of actual demand and the natural log of SND.

Figure 3.11: Partial auto-correlations for the difference between the natural log of actual demand and the natural log of SND.
3. DEVELOPMENT OF A STOCHASTIC PROCESS FOR UK GAS DEMAND

Figure 3.10 displays the autocorrelations for this difference, while Figure 3.11 shows the partial autocorrelations. As before, the auto-correlations in Figure 3.10 are highly statistically significant with the size of the auto-correlations decreasing as the lag increases. Figure 3.11 shows that the only highly statistically significant partial auto-correlation occurs at lag 1. These plots suggest that this time series, as above, follows an auto-regressive process of order one (AR(1)). Again, using the statistical software package R, the following AR(1) process can be fitted to the data:

\[
\ln(ActDem_t) - \ln(SND_t) = -0.003 + 0.92(\ln(Act_{t-1}) - \ln(SND_{t-1})) + \sigma \epsilon_t, \quad (3.4)
\]

which can be rewritten as

\[
ActDem_t = e^{-0.003} SND_t \left( \frac{Act_{t-1}}{SND_{t-1}} \right)^{-0.92} e^{\sigma \epsilon_t}. \quad (3.5)
\]

As before, \(ActDem_t\) represents actual demand for day \(t\), while \(SND_t\) represents seasonal normal demand for day \(t\). The noise term \(\epsilon_t\) is again a Gaussian white noise process with a mean of zero and variance of one while \(\sigma^2 = 0.002\) represents the variance of the residuals. The value of the AR(1) coefficient is 0.92 while the intercept of the model is -0.003. The p-values for the test of significance for these parameters are < 0.001 and 0.02, respectively. This suggests that the parameters are highly statistically significant.

Figure 3.12 displays the residuals from this model. It shows that these residuals do not vary as much over time as those in Figure 3.5. Figure 3.13 indicates that there are no highly statistically significant auto-correlations in the residuals from this process, suggesting no time dependence among the residuals.
3.2 Log model

Figure 3.12: Residuals from the AR(1) process for the difference between the natural log of actual demand and the natural log of SND as defined in equation (3.4).

Figure 3.13: Auto-correlations for the residuals from the AR(1) process of equation (3.4).
Figure 3.14: Histogram for the residuals from the AR(1) process of equation (3.4).

Figure 3.15: Residual versus fitted value plot for the residuals from the AR(1) process of equation (3.4).

Figure 3.14 shows the histogram of the residuals. The p-value associated with the Kolmogorov-
Smirnov test statistic is 0.018, which is below the significance threshold of 0.05. This suggests that the residuals are not normally distributed. However, as the distribution is relatively bell-shaped and symmetric, the rejection of the Kolmogorov-Smirnov test may be due to a small number of positively skewed outliers. Figure 3.15 displays the residuals versus fitted values plot. It shows that the residuals are scattered randomly around zero which suggests that the residuals are randomly distributed. As before, Figures 3.13 - 3.15 suggest that the residuals of this model behave as a Gaussian white noise process. This implies that taking the difference between the natural log of actual demand and the natural log of SND also provides an adequate model, [24].

The model assumes that, for a given day, actual demand is equal to SND for that day times some random noise times some dependence on the size of the error from the previous day. Thus, if actual demand is low (or high) then it is likely be low (or high) again on the next day.

### 3.2.1 Model comparison

The above analysis presented two models for the difference between actual demand and SND; namely a simple difference model and a log model. Both models were found to fit adequately to the data. However, there was a suggestion that the variance of the residuals was not constant, particularly for the simple difference model (see Figure 3.5). This subsection compares these models by analysing the variance of their residuals over time. Figure 3.16 displays the relative monthly variance of the residuals from both the simple difference model and the log model. It also shows the relative monthly variance for a single realisation of a Gaussian white noise process with mean 0, variance 1 and length equal to that of both the residuals’ time series. This measure can be defined as follows

$$\frac{\sigma^2_{\text{month}_i}}{\sigma^2},$$

where $\sigma^2$ is the overall variance for the residuals and $\sigma^2_{\text{month}_i}$ is the variance calculated from the residuals found in month $i$. 


Figure 3.16: Relative monthly variances for the residuals from the log model and the simple difference model.

Figure 3.16 shows that the residuals from the simple difference model vary a lot more over time than those of the log model. The relative monthly variance for the simple difference model ranges from 0.4 to 1.6. In contrast, the relative monthly variance for the log model ranges only from 0.8 to 1.4, which is much closer to the range of a Gaussian white noise process. This confirms that the log model is a better model, as both models should ideally have residuals with constant variance, [24].

3.2.2 GARCH model

The above analysis suggested that the variance of the residuals from the log model change over time. As a result, a Generalized Auto-Regressive Conditional Heteroskedasticity (GARCH) model is now examined for the residuals ($\epsilon_t$) of the log model. A GARCH model of order $(m, s)$ is defined in [24] as follows

\begin{align*}
\epsilon_t &= \sigma_t z_t, \\
\sigma_t^2 &= \alpha_0 + \sum_{i=1}^{m} \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^{s} \beta_j \sigma_{t-j}^2,
\end{align*}

where $z_t$ is a Gaussian white noise process with mean 0 and variance 1. The constant of the model is $\alpha_0$. The parameters associated with the error terms are $\alpha_i$, while the parameters associated with the time-dependent standard deviations are $\beta_j$. Using the software package R, the difference be-
3.3 Actual demand and predicted demand

Between the natural log of actual demand and the natural log of SND were fitted to a joint estimation AR(1)-GARCH(1,1) model as follows:

\[
\ln(\text{ActDem}_t) - \ln(SND_t) = -0.003 + 0.92(\ln(\text{ActDem}_{t-1}) - \ln(SND_{t-1})) + \epsilon_t, \quad (3.9)
\]

\[
\epsilon_t = \sigma_t z_t, \quad (3.10)
\]

\[
\sigma_t^2 = 0.0007 + 0.01\epsilon_{t-1}^2 + 0.65\sigma_{t-1}^2, \quad (3.11)
\]

where \(\text{ActDem}_t\) and \(SND_t\) are as described before. The p-values associated with the test statistic for the test of significance for the constant and AR(1) coefficient in equation (3.9) are 0.02 and < 0.001, respectively. This suggests that these parameters are statistically significant. These parameters also agree with those seen for the AR(1) model developed above (equation (3.4)). The constant term in equation (3.11) is 0.0007, while the error term coefficient is 0.01. However, both of these parameters were found to be statistically insignificant, as the t-test gives p-values of 0.17 and 0.50, respectively. In contrast, the coefficient of the time-dependent standard deviations term (0.65) was found to be highly statistically significant (p-value = 0.01). When the insignificant \(\epsilon_{t-1}^2\) term is removed, the model becomes

\[
\epsilon_t = \sigma_t z_t \quad (3.12)
\]

\[
\sigma_t^2 = 0.0007 + 0.65\sigma_{t-1}^2 \quad (3.13)
\]

which means that

\[
\sigma^2 = 0.002 \quad (3.14)
\]

as \(t \to \infty\). This variance is approximately equal to the variance of the residuals from the log model (0.002). This suggests that the behaviour of the log model with this GARCH process included is the same as when it is not included and thus there are no clear benefits to modelling the residuals of the log model with a GARCH process.

3.3 Actual demand and predicted demand

Having established the relationship between actual demand and seasonal normal demand, the relationship between actual demand and predicted demand is now considered. As mentioned above, the UK National Grid supplies five types of predicted demand on a daily basis, one- to five-day ahead predictions. One-day ahead predictions are the time series of gas demand predictions for the following day. Similarly two- to five-day ahead predictions are the time series of gas demand predictions given for the next two- to five-days ahead. These predictions are available on the UK
National Grid’s website [3]. This section describes the relationship between the natural log of actual demand and the natural log of one-day ahead predictions. The natural log of data was chosen following the analysis of Section 3.2.

Figure 3.17: Time series of actual demand and one-day ahead predictions.
Figure 3.18: The difference between the natural log of actual demand and the natural log of one-day ahead predictions.

Figure 3.17 displays the time series of actual demand and one-day ahead predictions, while Figure 3.18 shows the difference between the natural logs of the two time series. Both time series are from the 1st of October 2008 to the 31st of March 2012. The one-day ahead predictions were issued on each day from the 30th of September 2008 to the 30th March 2012. Figure 3.17 shows that there is little difference between the two series, indicating that one-day ahead predictions are quite accurate.
3. DEVELOPMENT OF A STOCHASTIC PROCESS FOR UK GAS DEMAND

Figure 3.19: Auto-correlations for the difference between the natural log of actual demand and the natural log of one-day ahead predictions.

Figure 3.20: Partial auto-correlations for the difference between the natural log of actual demand and the natural log of one-day ahead predictions.
3.3 Actual demand and predicted demand

Figure 3.19 displays the autocorrelations for the difference between the two time series while Figure 3.20 shows the partial autocorrelations. The correlations in both figures are statistically significant at lag 1. This suggests that this time series follows an auto-regressive process of order one (AR(1)). Using R, it can be shown that an AR(1) process fits the difference between the natural log of actual demand and the natural log of one-day ahead predictions as follows:

\[
\ln(\text{ActDem}_t) - \ln(\text{PreDem}_{t-1,t}) = -0.0067 + 0.22(\ln(\text{ActDem}_{t-1}) - \ln(\text{PreDem}_{t-2,t-1})) + \sigma_1\epsilon_t, \tag{3.15}
\]

where \(\text{ActDem}_t\) represents actual demand on day \(t\) and \(\text{PreDem}_{t-1,t}\) represents the predicted demand made on day \(t - 1\) for day \(t\), i.e., the one-day ahead prediction for day \(t\). As before, \(\epsilon_t\) represents Gaussian white noise with a mean of zero and variance of one. The variance of the residuals is \(\sigma^2_1 = 0.0015\). The constant and AR(1) coefficient of the model are both highly statistically significant (p-values are < 0.001 in both cases). Figure 3.21 shows the time series of the residuals from this process, while Figure 3.22 shows the auto-correlations of the residuals. It displays no statistically significant correlation at any lag, which suggests that the residuals are uncorrelated.

Figure 3.21: Residuals from the AR(1) process defined in equation (3.15) for the difference between the natural log of actual demand and the natural log of one-day ahead predictions.
Figure 3.22: Auto-correlations for the residuals from the AR(1) process of equation (3.15).

Figure 3.23: Histogram for the residuals from the AR(1) process of equation (3.15).
3.3 Actual demand and predicted demand

Figure 3.23 shows that the histogram of the residuals is relatively bell-shaped and symmetric. The Kolmogorov-Smirnov test gives a p-value of 0.724, which is well above the significance threshold of 0.05, thus confirming that the residuals are normally distributed. Figure 3.24 shows the residuals versus fitted values plot. It suggests that the residuals are randomly distributed around zero.

Figures 3.22 - 3.24 demonstrate that the residuals of this model behave as a Gaussian white noise process. This confirms that the model is an adequate model, [24]. The model assumes that for a given day the error in predicting demand for one day ahead is correlated with the previous day’s error.

3.3.1 Actual demand and two- to five-day ahead predictions

Having shown the relationship between one-day ahead predictions and actual demand, the difference between two- to five-day ahead predictions and actual demand are now considered. Each of the time series used in this section run from the 1st of October 2008 to the 31st of March 2012. As before, they were all obtained from the UK National Grid’s website [3]. In a similar manner to the previous section, it was found that the difference between the natural log of actual demand and the natural log of two-, three-, four- and five-day ahead predictions each follows an auto-regressive
3. DEVELOPMENT OF A STOCHASTIC PROCESS FOR UK GAS DEMAND

process of order one (AR(1)). These processes can be defined as follows:

\[ \ln(ActDem_t) - \ln(PreDem_{t-2,t}) = \\ - 0.009 + 0.48(\ln(ActDem_{t-1}) - \ln(PreDem_{t-3,t-1})) + \sigma_2 \varepsilon_t, \]

(3.16)

\[ \ln(ActDem_t) - \ln(PreDem_{t-3,t}) = \\ - 0.010 + 0.57(\ln(ActDem_{t-1}) - \ln(PreDem_{t-4,t-1})) + \sigma_3 \varepsilon_t, \]

(3.17)

\[ \ln(ActDem_t) - \ln(PreDem_{t-4,t}) = \\ - 0.012 + 0.65(\ln(ActDem_{t-1}) - \ln(PreDem_{t-5,t-1})) + \sigma_4 \varepsilon_t, \]

(3.18)

\[ \ln(ActDem_t) - \ln(PreDem_{t-5,t}) = \\ - 0.012 + 0.68(\ln(ActDem_{t-1}) - \ln(PreDem_{t-6,t-1})) + \sigma_5 \varepsilon_t, \]

(3.19)

where \( ActDem_t \) represents actual demand on day \( t \), \( PreDem_{t-i,t} \) represents the demand that was predicted on day \( t - i \) for day \( t \), i.e., \( i \) day ahead predictions. Again, \( \varepsilon_t \) represents a Gaussian white noise term with a mean of zero and a variance of one. The variances of the residuals (\( \sigma_i^2 \)) associated with each of the equations (3.16) - (3.19) are 0.0028, 0.0032, 0.0031 and 0.0031 respectively. The graphs associated with the analysis of each of these relationships can be found in appendices A.1, A.2, A.3 and A.4 respectively. As above, these models assume that for a given day the error in predicting demand for two to five days ahead is correlated with the previous day’s error.

3.4 Stochastic process for demand

The above analysis shows the statistical relationships between actual demand, SND and predicted demand for the UK gas market. In this section these relationships are used to develop a stochastic process for demand. Consider the following process with \( S \) scenarios

\[ dem_{t,t} = ActDem_t, \]

(3.20)

\[ \ln(dem_{t,t+1}^s) = \ln(ActDem_{t+1}) \\ - 0.007 + 0.22(\ln(dem_{t-1,t}^s) - \ln(ActDem_t)) + \sigma_1 \varepsilon_t, \]

(3.21)
\[
\ln(dem_{t,t+2}^s) = \ln(ActDem_{t+2}) \\
- 0.009 + 0.28(\ln(dem_{t-1,t+1}^s) - \ln(ActDem_{t+2})) + \sigma_2 \epsilon_t,
\] (3.22)

\[
\ln(dem_{t,t+3}^s) = \ln(ActDem_{t+3}) \\
- 0.010 + 0.57(\ln(dem_{t-1,t+2}^s) - \ln(ActDem_{t+2})) + \sigma_3 \epsilon_t,
\] (3.23)

\[
\ln(dem_{t,t+4}^s) = \ln(ActDem_{t+4}) \\
- 0.012 + 0.65(\ln(dem_{t-1,t+3}^s) - \ln(ActDem_{t+3})) + \sigma_4 \epsilon_t,
\] (3.24)

\[
\ln(dem_{t,t+5}^s) = \ln(ActDem_{t+5}) \\
- 0.012 + 0.68(\ln(dem_{t-1,t+4}^s) - \ln(ActDem_{t+4})) + \sigma_5 \epsilon_t,
\] (3.25)

\[
\ln(dem_{t,t+d}^s) = \ln(SND_{t+d}) - 0.0027 + \\
0.92(\ln(dem_{t,t+d-1}^s) - \ln(SND_{t+d-1})) + \sigma \epsilon_t, \quad \forall \; d > 5,
\] (3.26)

where \(dem_{t,t+d}^s\) is the simulated demand for gas on day \(t\) for day \(t+d\), associated with scenario \(s\). These simulated demands are used as inputs to the Rolling Optimisation Model described in Chapter 5 while the rest of the variables in equations (3.20) - (3.26) are as defined above.

On the first day of the stochastic process, equation (3.20) states that the simulated demand \((dem_{t,t})\) equals actual demand for that day. It also shows that the simulated demand is scenario-independent on this first day (note: there is no superscript \(s\) for this day). For the next five days, the simulated demand is actual demand for that day plus some randomly simulated noise with a dependence on the error for predicting demand from the previous day \(1\). Equations (3.21) - (3.25) follow from the statistical analysis described in Section 3.3 for the difference between the natural log of actual demand and the natural log of predicted demand (equations (3.15) - (3.19)). For \(d > 5\), equation (3.26) indicates that the simulated demand follows a mean-reverting process where Seasonal Normal Demand (SND) is the mean. This process and its parameters are based on the log model described in Section 3.2.

This overall stochastic process replicates the information that those in the UK gas market have on a given day whereby on the first day (today), the market knows demand exactly as in equation (3.20). For the next five days, market players have an idea what the demand is going to be, but with some uncertainty, as in equations (3.21)-(3.25). Beyond this, all they know is seasonal normal

---

\(^1\)When simulating stochastic demand for the first day of this process, the error from the previous day is assumed to be zero.
demand and that if demand is low (or high) on one day, it is likely to be around that level again the next day, as described by equation (3.26).

3.4.1 Numerical example

Figure 3.25: Actual demand (red line), seasonal normal demand (blue line), and a simulated demand paths (black line), for the gas year '08 - '09.
Having shown how the stochastic process for gas demand was developed, a numerical example is now considered. The historical actual demand time series used for $\text{ActDem}_t$ is from the gas year ’08 - ’09. This time series, and its corresponding seasonal normal demand time series, can be found on the UK National Grid’s website, [3]. Both time series start on the 1st of October 2008 and end on the 30th of September 2009.

Figure 3.25 displays actual demand, seasonal normal demand and a simulated demand path for the gas year ’08 -’09. The simulated demand path was developed using the methodology described in equations (3.21) - (3.26) above. Figure 3.25 indicates that for the first few days of the process, actual demand and the simulated demand are similar. After about a month, it can be seen that the simulated demand does not have any dependence on actual demand. It can also be seen in Figure 3.25 that if the simulated demand is low (or high) it tends to stay like that for a couple days. This reflects actual gas demand trends, as cold and warm weather periods affect gas demand patterns in a similar and intuitive manner.

Figure 3.26 shows the time series over the first 10 days of the process and with a total of 100 simulated demand paths. It highlights that for the first few days the simulated paths only vary around actual demand and not around seasonal normal demand. As time goes on however, the simulated paths (black lines) start to move towards, and vary around, seasonal normal demand.
3. DEVELOPMENT OF A STOCHASTIC PROCESS FOR UK GAS DEMAND

3.5 Summary

In this chapter a stochastic process capturing the daily uncertainty of UK gas demand was developed. This was done firstly in Section 3.1 by examining the statistical relationship between actual demand and SND. This analysis found that the difference between these two time series follows an auto-regressive process of order one (AR(1)). This result is intuitive because if demand for gas in the UK is high (low) on one day, then it is likely to be relatively high (low) again on the following day. However, the residuals from the fitted model established in Section 3.1 were found not to have constant variance. As a result, in Section 3.2 a similar analysis was presented for the difference between the natural log of the two series. An AR(1) model was again shown to provide a reasonable fit to the data. The residuals from this model were found not to vary as much and hence provided a better fit than the model described in Section 3.1.

In Section 3.3 the statistical relationships between actual demand and predicted demand were examined. This analysis found that the difference between actual demand and one-day ahead predictions follows an AR(1) process. This result means that the error in predicting UK gas demand for one day ahead is correlated with yesterday’s error. Section 3.3 also established that the differences between actual demand and two- to five-day ahead predictions also follow AR(1) processes.

Using the results from Sections 3.2 and 3.3 a stochastic process for UK gas demand was developed in Section 3.4. On the first day of this process, demand is exactly known. The next five days of the process is based on the relationships established in Section 3.3 for the difference between actual demand and one- to five-day ahead predictions. After the sixth day, the process is based on the relationship between actual demand and SND established in Section 3.2. This stochastic process for demand reflects the daily information that is known in the UK gas market regarding demand and is used as an input for a more detailed model of the UK gas market to be described in Chapter 5.
Chapter 4

Gas market models

In this chapter the two main methodologies used to model natural gas markets are reviewed and examined. This is done by comparing examples of simplified models from the two methodologies. The results of this analysis are used to develop a more complicated model of the UK natural gas market in Chapter 5.

The first class of model used to model natural gas markets is that of complementarity-based equilibrium models, e.g., \[11, 25\]. This methodology involves each player in the market separately optimising their own position in an attempt to maximise their profits. Each player has their own objective function. These objective functions are connected through market clearing conditions and are simultaneously optimised according to a complementarity-based equilibrium model. Market clearing conditions ensure conservation of gas, i.e., every unit of gas produced is either consumed or stored. Examples of the different players used in these models include producers, storage operators, Liquified Natural Gas (LNG) operators and marketers. Models that use this methodology include Nash-Cournot competition among some of the players. This means that these players have market power and can hence influence the level of gas demand by adjusting their price of gas in order to maximise their profits. This results in demand being determined endogenously.

The second type of natural gas market model found in the literature are cost minimisation models, e.g., \[26, 27\]. Instead of each player in the market separately maximising their own profits, cost minimisation involves an imaginary central planner choosing the different sources of supply to meet demand at minimum overall cost. Using this methodology, a gas market model can be formulated using a single objective function. As a result, cost minimisation models tend to be either linear or quadratic programs. Unlike complementarity-based equilibrium models, Nash-Cournot competition cannot be included in the formulation of a cost minimisation model, as there is only one objective function. This means that these models assume that none of the players have market power and that demand is determined exogenously.

Because it assumes an imaginary central planner, it might be argued that the cost minimisation approach is unintuitive for modelling natural gas markets. Generally, central planners do not exist in natural gas markets, particularly the UK market. The complementarity-based equilibrium
approach, with each player separately maximising their own profits, is more plausible. However, cost minimisation models are not as complex as complementarity-based equilibrium models. They are formulated with a single objective function, whereas complementarity-based equilibrium models are formulated with multiple objective functions. The cost minimisation approach is therefore algebraically simpler and structurally less complicated.

In this chapter it is shown that when market power (i.e., Nash-Cournot competition) is removed from all players in a complementarity-based equilibrium model, the model is equivalent to a corresponding cost minimisation model. This is done in two different ways. Firstly, in Section 4.4 it is shown that the (Karush-Kuhn-Tucker) KKT conditions of optimality for the two models reduce to the same set of equations while in Section 4.5 the principle of symmetry [28] is used to show that the two models are equivalent. This result means that when a more detailed model of the UK natural gas market is being developed in Chapter 5, the simpler cost minimisation model can be used instead of the complementarity-based equilibrium model.

The complementarity-based equilibrium model used in this analysis is based on Gabriel et al’s model [11]. It contains $P$ producers, who maximise their profits by selling gas to either $SO$ storage operators or a marketer. The storage operators maximise their profits by selling the gas on to the marketer. The marketer maximises their profits by buying gas from either the producers or the storage operators, so as to meet consumer demand. As explained above, none of the players in this model exhibit market power. According to a report made by Ofgem for the UK Parliament, this is a reasonable assumption to make about the UK natural gas market [12]. They base this assumption on the fact that there are many players in the UK market and hence “there are limited signs of one source being able to influence market price to a significant extent”. The corresponding cost minimisation model used in this analysis is formulated with $P$ producers and $SO$ storage operators. Instead of each of the players optimising their own positions through their own objective functions, an imaginary central planner chooses how these different sources of supply meet demand at minimum overall system cost.

This chapter is organised as follows: firstly, in Section 4.1 relevant background theory for cost minimisation and complementarity-based equilibrium models is presented. Secondly, in Section 4.2 the complementarity-based model of this chapter is introduced, while in Section 4.3 its corresponding cost minimisation model is introduced. Following this, the models are shown to be equivalent. This is done using KKT analysis and the principle of symmetry in Sections 4.4 and 4.5 respectively. After this, a discussion on non-unique solutions is provided in Section 4.6. The chapter concludes with a summary in Section 4.7.

\footnote{Ofgem (Office of the Gas and Electricity Markets) regulates the gas and electricity markets in the UK. See: http://www.ofgem.gov.uk/About%20us/Pages/AboutUsPage.aspx}
4.1 Theoretical background

In this section various optimisation problems relevant to cost minimisation and complementarity-based models are defined. The relationships between the various problems are also established. These relationships are necessary for the analysis in Section 4.5. As mentioned in the previous section, cost minimisation models are generally presented as either quadratic or linear programs. A Quadratic Program (QP) can be stated as:

\[
\begin{align*}
\min & \quad c^T x + \frac{1}{2} x^T H x, \\
\text{subject to:} & \\
Ax & \geq b, \\
Dx & = e.
\end{align*}
\]

When the matrix \( H \) is zero the problem becomes a Linear Program (LP). The cost minimisation model presented in Section 4.3.1 is an example of a LP. As mentioned above, the KKT optimality conditions are used for the presentation and comparison of each of the models introduced in this chapter. These conditions are necessary and in some cases sufficient for optimality. Consider the following generic optimisation problem:

\[
\min_{x \in \mathbb{R}^n} f(x)
\]

subject to:

\[
\begin{align*}
c_i(x) &= 0 \quad i \in \mathcal{E}, \\
c_i(x) &\geq 0 \quad i \in \mathcal{J},
\end{align*}
\]

where \( i \in \mathcal{E} \) are the set of equality constraints and \( i \in \mathcal{J} \) are the set of inequality constraints. The KKT conditions [29] for this problem are as follows:

\[
\begin{align*}
\nabla_x \mathcal{L}(x^*, \lambda^*) &= 0, \\
c_i(x^*) &= 0, \quad \forall \ i \in \mathcal{E}, \\
c_i(x^*) &\geq 0, \quad \forall \ i \in \mathcal{J}, \\
\lambda_i^* &\geq 0, \quad \forall \ i \in \mathcal{J}, \\
\lambda_i^* c_i(x^*) &= 0, \quad \forall \ i \in \mathcal{E} \cup \mathcal{J},
\end{align*}
\]

where \( x^* \) is the vector of optimal solutions, \( \lambda^* \) is the vector of optimal Lagrange multipliers and

\[
\mathcal{L}(x, \lambda) = f(x) - \sum_{i \in \mathcal{E} \cup \mathcal{J}} \lambda_i c_i(x),
\]
which is known as the Lagrangian equation. Equation (4.6) is known as the set of stationary KKT conditions. Equations (4.7) and (4.8) ensure that the constraints of the problem are satisfied, while equation (4.10) is known as the set of complementarity conditions. Each constraint $i$ in the problem has a Lagrange multiplier $(\lambda_i)$ associated with it. Equation (4.9) requires that the Lagrange multipliers associated with inequality constraints must be non-negative. For quadratic programs, and trivially linear programs, the KKT conditions are both necessary and sufficient for optimality as long as the matrix $H$ in equation (4.1) is positive semi-definite, [29].

4.1.1 Complementarity-based problems

Having presented the definitions for quadratic and linear programs, the different types of complementarity-based problems found in the literature are now considered. In Sections 4.1.2 - 4.1.8 these different problems are defined and it is explained how they are related to each other.

4.1.2 Non-linear complementarity problem

A Non-linear Complementarity Problem (NCP) is defined in [11],[28] and [30] as follows: given a mapping $F(z) = (F_i(z)) : \mathbb{R}^n \rightarrow \mathbb{R}^n$, find a $z \in \mathbb{R}^n$ satisfying:

\[
\begin{align*}
z & \geq 0, \\
F(z) & \geq 0, \\
z^T F(z) & = 0.
\end{align*}
\]

This problem is denoted as $NCP(F)$. When $F(z) = Mz + q$ (i.e., when $F(z)$ is linear), then the $NCP(F)$ becomes linear and is an instance of a Linear Complementarity Problem (LCP), denoted $LCP(q, M)$, [30].

4.1.3 Mixed complementarity problem

A Mixed Complementarity Problem (MCP)\(^1\), denoted $MCP(F, L, U)$, is defined in [11] as follows: find a vector $z \in \mathbb{R}^n$ such that

\[
\begin{align*}
z_i = U_i & \Rightarrow F_i(z) \leq 0, \\
L_i < z_i < U_i & \Rightarrow F_i(z) = 0, \\
z_i = L_i & \Rightarrow F_i(z) \geq 0,
\end{align*}
\]

where $L$ and $U$ are the lower and upper bound $n$-dimensional vectors for $z$, respectively, and

\[
L_i, U_i \in \mathbb{R} \cup \{-\infty, +\infty\}, L_i < U_i \quad \forall \ i = 1, 2...n.
\]

\(^1\)A MCP may also be referred to as a Mixed Non-linear Complementarity Problems (MNCP)
As with the NCPs (see Section 4.1.2), when \( F(z) = Mz + q \) (i.e., when \( F(z) \) is linear), then the MCP becomes linear and becomes known as a Mixed Linear Complementarity Problem (MLCP), denoted \( MLCP(F, L, U) \) [30]. When the lower and upper bounds take the form \( L_i = 0 \) and \( U_i = +\infty \), it is clear that, \( MCP(F, L, U) \) is equivalent to \( NCP(F) \), [11]. Similarly, if the same conditions on the lower and upper bounds hold and \( F(z) = Mz + q \), then \( MLCP(F, L, U) \) reduces to \( LCP(q, M) \).

### 4.1.4 Variational inequality problems

Variational Inequality problems, denoted \( VI(X, F) \), are defined in [31] as follows: let \( X \) be a nonempty subset of \( \mathbb{R}^n \) and let \( F \) be a mapping \( z \to F(z) \in \mathbb{R}^n \). Then the problem \( VI(X, F) \) is to find a vector \( z \in X \) such that

\[
(y - z)^T F(z) \geq 0 \quad \forall \ y \in X.
\]

When \( X = \mathbb{R}^n_+ \), then \( VI(\mathbb{R}^n_+, F) \) is equivalent to \( NCP(F) \). Similarly, when \( X = [L, U] \), where \(-\infty \leq l_i < u_i \leq +\infty \), then the variational inequality \( VI([L, U], F) \) is equivalent to \( MCP(F) \). Sections 4.1.7 and 4.1.8 establish these equivalences.

### 4.1.5 Relationship between a QP and a LCP

The relationship between a quadratic program and a linear complementarity problem is demonstrated in [30] and [32] as follows. Consider the following QP:

\[
\min \ c^T x + \frac{1}{2} x^T H x,
\]

subject to:

\[
Ax \geq b, \quad x \geq 0.
\]

The KKT conditions for this QP are:

\[
\begin{align*}
v &= c + Hx - A^T u \geq 0, \\ y &= -b + Ax \geq 0, \\ x^T v &= 0, \\ u^T y &= 0, \\ x, u &\geq 0,
\end{align*}
\]
where the vectors \( u \) and \( v \) represent the Lagrange multipliers associated with constraints (4.21) and (4.22) respectively. The vector \( y \) denotes the vector of slack variables. By letting

\[
M = \begin{bmatrix} H & -A^T \\ A & 0 \end{bmatrix},
\]

(4.28)

\[
q = \begin{bmatrix} c \\ -b \end{bmatrix},
\]

(4.29)

\[
w = \begin{bmatrix} v \\ y \end{bmatrix},
\]

(4.30)

and

\[
z = \begin{bmatrix} x \\ u \end{bmatrix},
\]

(4.31)

the KKT conditions of this QP can be re-written as a LCP where \( F(z) = w = Mz + q \). Thus solving this LCP is equivalent to solving the QP defined above. Note: this result holds for linear programs with the same constraints as linear programs are trivially quadratic programs.

4.1.6 Relationship between a QP and a MLCP

Following [30], it is now shown how a QP can also be described as (an instance of) a MLCP. Consider the QP defined in equations (4.1) - (4.3) above. Its KKT conditions are:

\[
0 = c + Hx - A^Tu - D^Tv,
\]

(4.32)

\[
0 = -e + Dx,
\]

(4.33)

\[
y = -b + Ax \geq 0,
\]

(4.34)

\[
u^Ty = 0,
\]

(4.35)

\[
u \geq 0,
\]

(4.36)

where the elements of the vectors \( x \) and \( v \) are free. Analogously to equations (4.28) - (4.31), letting

\[
M = \begin{bmatrix} H & -D^T & -A^T \\ D & 0 & 0 \\ A & 0 & 0 \end{bmatrix},
\]

(4.37)

\[
q = \begin{bmatrix} c \\ -e \\ -b \end{bmatrix},
\]

(4.38)
and

\[
\mathbf{z} = \begin{bmatrix} \mathbf{x} \\ \mathbf{v} \\ \mathbf{u} \end{bmatrix},
\]

means that the KKT conditions for this QP can be formulated as a \( \text{MLCP}(F, L, U) \), where \( F(\mathbf{z}) = M\mathbf{z} + \mathbf{q} \). The lower and upper bounds for each element in both the vectors \( \mathbf{x} \) and \( \mathbf{v} \) are \( L_i = -\infty \) and \( U_i = +\infty \), while the lower and upper bounds for each element in the vector \( \mathbf{u} \) are \( L_i = 0 \) and \( U_i = +\infty \). Note: as with LCPs, this result holds for linear programs with the same constraints as linear programs are trivially as quadratic programs.

### 4.1.7 Relationship between non-linear complementarity problems and variational inequality problems

In this subsection, it is demonstrated that non-linear complementarity problems are special cases of variational inequalities. Let \( \mathbf{z}^* \) be a solution to \( \text{NCP}(F) \). This means

\[
\begin{align*}
\mathbf{z}^* &\geq 0, \\
F(\mathbf{z}^*) &\geq 0, \\
z^{*T}F(\mathbf{z}^*) &= 0.
\end{align*}
\]

Equation (4.43) is satisfied by equation (4.40). If \( y \geq \mathbf{z}^* \), equation (4.44) is satisfied as \( F(\mathbf{z}^*) \geq 0 \) by equation (4.41). If \( 0 \leq y < \mathbf{z}^* \), then equation (4.44) is also satisfied as \( F(\mathbf{z}^*) = 0 \) by equation (4.42). As both equations (4.43) and (4.44) are satisfied, this means that if \( \mathbf{z}^* \) is a solution to \( \text{NCP}(F) \), then it is also a solution to \( \text{VI}(\mathbb{R}_+^n, F) \).
4.1.8 Relationship between mixed complementarity models and variational inequality problem

In this subsection, it is demonstrated that mixed complementarity problems can be rewritten as a certain variational inequality problems. Let \( z^* \) be a solution to a MCP. This means that

\[
\begin{align*}
z_i^* &= U_i \quad \Rightarrow \quad F_i(z^*) \leq 0, \\
L_i < z_i^* < U_i \quad \Rightarrow \quad F_i(z^*) = 0, \\
z_i^* &= L_i \quad \Rightarrow \quad F_i(z^*) \geq 0,
\end{align*}
\]  

(4.45)

(4.46)

(4.47)

where \( L \) and \( U \) are lower and upper bound \( n \)-dimensional vectors of \( z^* \), respectively, and

\[
L_i, U_i \in R \cup \{-\infty, +\infty\}, L_i < U_i \quad \forall \quad i = 1, 2...n.
\]  

(4.48)

It is now shown that \( z^* \) is also a solution to the variational inequality \( VI([L, U], F) \), which is defined as:

\[
\begin{align*}
z^* &\in [L, U], \\
(y - z)^T F(z^*) &\geq 0 \quad \forall \quad y \in [L, U].
\end{align*}
\]  

(4.49)

(4.50)

Trivially, \( z^* \) satisfies equation (4.49). Now consider the following three cases:

1. If \( z_i^* = L_i \), then \( (y_i - z_i^*) \geq 0 \quad \forall \quad y \in [L, U] \) and \( F_i(z^*) \geq 0 \) using equation (4.47). Therefore \( (y_i - z_i^*)F_i(z^*) \geq 0 \).

2. If \( z_i^* = U_i \), then \( (y_i - z_i^*) \leq 0 \quad \forall \quad y \in [L, U] \) and \( F_i(z^*) \leq 0 \) using equation (4.45). Therefore \( (y_i - z_i^*)F_i(z^*) \geq 0 \).

3. If \( L_i < z_i^* < U_i \) then \( F_i(z^*) = 0 \) by equation (4.46) which means that \( (y_i - z_i^*)F_i(z^*) = 0 \).

Using these three cases it can be seen that \( (y_i - z_i^*)F_i(z^*) \geq 0 \quad \forall \quad i, y \in [L, U] \). Therefore \( z^* \) satisfies equation (4.50). Hence, if \( z^* \) is a solution to a MCP then it is also a solution to the variational inequality \( VI([L, U], F) \). This means that mixed complementarity problems can be re-written as variational inequalities.

4.1.9 Summary

In this section the various optimisation problems relevant to this chapter are introduced; these are quadratic and linear programs, non-linear complementarity problems, mixed complementarity problems and variational inequality problems. In Section 4.2.1 a complementarity-based model of a natural gas market is introduced, while in Section 4.3.1 a cost minimisation model of a natural gas market is introduced. These models are instances of a MCP and a LP, respectively. This section also details how the various problems defined are related to other. In particular, Sections 4.1.5...
and [4.1.6] explain how quadratic programs, with and without equality constraints, can be reformed as mixed and linear complementarity problems respectively. The analysis also described, in Section [4.1.3] how non-linear complementarity problems can be considered as mixed complementarity problems while Sections [4.1.7] and [4.1.8] show how both of these types of problems are special cases of variational inequality problems. These relationships are important to this project as they are used, in Section [4.5], to prove that the aforementioned models of natural gas markets are equivalent to each other.

4.2 Complementarity-based models of natural gas markets

Having defined the different types of complementarity-based problems in the previous section, the different types of equilibrium models of natural gas markets are now considered. In particular, Gabriel et al’s mixed complementarity-based equilibrium model [11] is firstly discussed. This model is of particular interest to this project, as the model described in Section [4.2.1] is based on it. The model consists of a number of players, who each maximise their own profit through their own separate objective functions. The different types of players are:

1. Pipeline operators,
2. Producers,
3. Storage operators,
4. Marketers,
5. Peak gas operators.

Each of the players in Gabriel et al’s model interact with the others so as the demand from different sectors is met. They do this over three different types of demand seasons for each year, low, high and very high.

Gabriel et al assumes perfect competition between the pipeline operator, who is responsible for transporting the gas between the rest of the players. Similarly, Gabriel et al also assume perfect competition among the producers, storage operators, and peak gas operators, who are all, hence, price takers. In contrast however, Nash-Cournot competition is assumed among the marketers who adjust their price of gas in order to influence the level of demand and hence maximise their profits.

In each season the producers supply gas to marketers, who in turn supply the gas to the different demand sectors. In season 1, the low demand season, the producers also supply gas to the storage operators. In seasons 2 and 3, the high and very high demand seasons respectively, the storage operator then supplies this gas to the marketers. The peak gas operator only supplies the marketer with gas and this is done in season 3 only.

Each of the different player’s problems are connected through market clearing conditions. Gabriel et al formulate the model by combining these market clearing conditions with the KKT optimality conditions for each of the different player’s problems to form a mixed complementarity-based equilibrium model. In Appendix [B.1] detailed mathematical formulations for each of the different problems are presented.
4. GAS MARKET MODELS

While the model described in Section 4.2.1 may be based on Gabriel et al’s model, there are other complementarity-based models to be found in the literature. For example, in [33], Gabriel et al use the model described in [11] and apply it to the North American natural gas market.

Gabriel et al mention in [11] how this model is closely related to Boots et al’s GASTALE model, [34]. Boots et al’s model is applied to the European gas market and only has two players namely, producers and traders. These traders are equivalent to the marketers in Gabriel et al’s model, as they buy gas off the producers in order to meet demand. Both sets of players in Boots et al’s model are modeled as Nash-Cournot players. In contrast to Gabriel et al’s model in [11] however, there are no storage operators, pipeline operators or peak gas operators. As well as this, Boots et al’s model only has a single overall demand sector instead of separate demand sectors as in [11].

Holz et al’s GASMOD model, [25], is very similar to Boots et al’s model, [34], in that only two sets of players (traders and producers) are modeled and the model is also applied to the European gas market. The difference in these two models is in their representations of producers.

In [35], Egging and Gabriel developed a mixed complementarity equilibrium model for the European gas market. It is an extension to Boots et al’s GASTALE model, [34], in that it additionally considers storage operators, seasonally varying demand and LNG terminal capacities. In [36] a similar model is used. However, producers and traders are modelled separately. The formulations of both [35] and [36] are similar to Gabriel et al’s model, [11]. However, they are both applied to the European gas market.

In [37], Lise and Hobbs extend Boots et al’s GASTALE model, [34], in a similar manner to Egging and Gabriel. In contrast to Egging and Gabriel however, Lise and Hobbs’s model allows for long-term investment decisions on things such as production and storage capacities. It is used to model the European gas market over a long period of time (e.g., 25 years) in [38, 39].

In [40], Egging et al describe a world gas model (WGM), which is based on [11] but also allows for endogenous capacity expansions. The WGM is used to test a range of different scenarios in [41, 42]. Similarly, Gabriel et al use an improved version of the WGM in [43] to study the effect of cartelisation in the global gas market.

4.2.1 A complementarity-based equilibrium model and its KKT conditions

Having reviewed the different types of complementarity-based models of natural gas markets seen in the literature, the complementarity-based equilibrium model of this project is now considered. It is based on the model introduced in [11] and described in Section 4.2.

As in [11], there are $P$ producers, who produce and then sell gas to a marketer and/or $SO$ storage operators. In contrast to [11], there are no pipeline operators. In their model pipeline operators charge other players a fixed price for using the pipelines. In the model presented here, this charge can be accounted for in the costs of the other players and is hence a simplification.

Similarly, there are no peak gas operators modelled in the current analysis. In [11], peak gas operators are producers who are restricted to supplying gas only in peak demand seasons.
Also, in contrast to [11], there is no Nash-Cournot competition modelled in this complementarity-based equilibrium model. Nash-Cournot competition allows different players to have market power and hence influence the level of supply in order to maximise their profits. As discussed in the introduction of this chapter, Market power (i.e., Nash-Cournot competition) is removed from this model because, in most reasonable cases, it is not suitable in the context of the UK natural gas market [12].

Without Nash-Cournot competition, marketers are not allowed to affect the supply and hence the price of gas. This means that marketers have no effect on the price they receive for selling gas to consumers. This also means that marketers are not able to distinguish themselves from each other in this model. As a result, there is only one marketer modelled in this analysis. This marketer buys gas from either the storage operators or the producers, in order to ensure demand is met.

As well as not having Nash-Cournot competition, the model presented in this section considers demand as a whole, in contrast to the different demands from separate sectors as in [11]. As there is only one marketer in the current work, they are required to meet all demand and thus there is no need to consider separate demand sectors.

While the model of [11] is defined over seasons and years, this model allows for daily demands, as it is defined over \( D \) days. The model is broken up into \( P + SO + 1 \) problems, one for each of the \( P \) producers, one for each of the \( SO \) storage operators, and one for the marketer. Each player has their own objective function, and the various objective functions are linked through the market clearing conditions.

Each of the players’ problems, as well as the market clearing conditions, are defined below. In each case the index \( p \) runs over the range 1, ..., \( P \). Similarly the indices \( so \) and \( d \) over the ranges 1, ..., \( SO \) and 1, ..., \( D \), respectively. As each of the different problems are linear programs, the KKT conditions are both necessary and sufficient for optimality [29]. By combining the market clearing conditions with the KKT conditions of the different sub-problems, as seen with Gabriel et al’s model [11], this model can be defined as a mixed complementarity-based equilibrium model (see Section 4.5 for details). Table 4.1 displays the parameters and variables of the model.
Table 4.1: Parameters and variables associated with the complementarity-based equilibrium model

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>$c_p$</td>
<td>Daily marginal production cost for producer $p$</td>
</tr>
<tr>
<td>$Q_p^{max}$</td>
<td>Maximum daily production rate for producer $p$</td>
</tr>
<tr>
<td>$a_{so}$</td>
<td>Daily marginal cost of injection for storage operator $so$</td>
</tr>
<tr>
<td>$b_{so}$</td>
<td>Daily marginal cost of withdrawal for storage operator $so$</td>
</tr>
<tr>
<td>$I_{so}^{max}$</td>
<td>Maximum daily injection rate for storage operator $so$</td>
</tr>
<tr>
<td>$W_{so}^{max}$</td>
<td>Maximum daily withdrawal rate for storage operator $so$</td>
</tr>
<tr>
<td>$MinCap_{so}$</td>
<td>Minimum amount of gas allowed in storage for storage operator $so$</td>
</tr>
<tr>
<td>$IntCap_{so}$</td>
<td>Initial amount of gas in storage for storage operator $so$</td>
</tr>
<tr>
<td>$MaxCap_{so}$</td>
<td>Maximum amount of gas allowed in storage for storage operator $so$</td>
</tr>
<tr>
<td>$\tau$</td>
<td>The marginal price the marketer receives for selling gas</td>
</tr>
<tr>
<td>$Demand_{d}$</td>
<td>Demand for gas on day $d$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variable</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\pi_{p,d}$</td>
<td>The marginal selling price of gas received by producer $p$ on day $d$</td>
</tr>
<tr>
<td>$Q_{p,d}$</td>
<td>Producer $p$’s production rate for day $d$</td>
</tr>
<tr>
<td>$\gamma_{so,d}$</td>
<td>The marginal selling price of gas received by storage operator $so$ on day $d$</td>
</tr>
<tr>
<td>$I_{so,p,d}$</td>
<td>Amount of gas injected by storage operator $so$ from producer $p$ on day $d$</td>
</tr>
<tr>
<td>$W_{so,d}$</td>
<td>Amount of gas withdrawn by storage operator $so$ on day $d$</td>
</tr>
<tr>
<td>$H_{p,d}$</td>
<td>Amount of gas bought by the marketer from producer $p$ on day $d$</td>
</tr>
<tr>
<td>$U_{so,d}$</td>
<td>Amount of gas bought by the marketer from storage operator $so$ on day $d$</td>
</tr>
</tbody>
</table>

### 4.2.2 Producer $p$’s problem

Producer $p$’s problem is defined as follows:

$$\text{Max} \sum_{d=1}^{D} (\pi_{p,d} Q_{p,d} - c_p Q_{p,d}),$$

subject to:

$$0 \leq Q_{p,d} \leq Q_{p,d}^{max}, \quad (\lambda Q_{p,d}), \quad (4.51)$$

where $\pi_{p,d}$ is the marginal price of gas received by producer $p$ on day $d$, $Q_{p,d}$ is the amount producer $p$ produces on day $d$, and $c_p$ is the unit cost of production for producer $p$. The daily maximum production rate for producer $p$ is $Q_{p,d}^{max}$. Both $c_p$ and $Q_{p,d}^{max}$ are fixed pre-defined inputs to the model, while $\pi_{p,d}$ and $Q_{p,d}$ are outputs of the model. $\pi_{p,d}$ is exogenous to the producer’s problem and is determined according to the optimality problem defined for the system as a whole. Inequality (4.51) provides upper and lower bounds on the production rate of producer $p$. The variable in brackets alongside this constraint is its corresponding Lagrange multiplier. The stationary
KKT condition for producer $p$’s problem is as follows:

$$-\pi_{p,d} + c_p + \lambda_{Q_{p,d}} - \bar{\lambda}_{Q_{p,d}} = 0,$$

(4.52)

where $\lambda_{Q_{p,d}}$ and $\bar{\lambda}_{Q_{p,d}}$ are the Lagrange multipliers associated with the upper and lower bounds of constraint (4.51), respectively. As these Lagrange multipliers are associated with inequality constraints, they are required by the KKT conditions to be non-negative. The complementary KKT conditions for this problem are as follows:

$$\lambda_{Q_{p,d}} (Q_{p,d} - Q_{p,d}^{\text{max}}) = 0,$$

(4.53)

$$\bar{\lambda}_{Q_{p,d}} Q_{p,d} = 0.$$

(4.54)

The KKT conditions also require that constraint (4.51) is satisfied.

### 4.2.3 Storage operator $so$’s problem

Storage operator $so$’s problem is:

$$\text{Max} \sum_{d=1}^{D} (\gamma_{so,d} W_{so,d} - b_{so} W_{so,d} - \sum_{p=1}^{P} (\pi_{p,d} I_{so,p,d} + a_{so} I_{so,p,d})), $$

subject to:

$$0 \leq I_{so,p,d} \leq I_{so,p,d}^{\text{max}}, \quad (\lambda_{I_{so,p,d}}),$$

(4.55)

$$0 \leq W_{so,d} \leq W_{so,d}^{\text{max}}, \quad (\lambda_{W_{so,d}}),$$

(4.56)

$$\text{MinCap}_{so} \leq \text{IntCap}_{so} + \sum_{e=1}^{d} (\sum_{p=1}^{P} I_{so,p,e} - W_{so,e}) \leq \text{MaxCap}_{so}, \quad (\lambda_{\text{Cap}_{so,d}}),$$

(4.57)

where $\gamma_{so,d}$ and $\pi_{p,d}$ are the marginal prices of gas received by storage operator $so$ and producer $p$ on day $d$ respectively, $W_{so,d}$ is the withdrawal rate of storage operator $so$ on day $d$, and $I_{so,p,d}$ is the injection rate of storage operator $so$ from producer $p$ on day $d$. Both $\gamma_{so,d}$ and $\pi_{p,d}$ are exogenous to the storage operator’s problem and are determined according to the optimality problem defined for the system as a whole, while $W_{so,d}$ and $I_{so,p,d}$ are endogenous to the storage operator $so$’s problem.

All the other values used above are pre-determined inputs to the model. The costs per unit of withdrawal and injection for storage operator $so$ are $a_{so}$ and $b_{so}$ respectively. $W_{so,d}^{\text{max}}$ and $I_{so,d}^{\text{max}}$ are the maximum withdrawal and injection rates from/to the storage facility $so$. $\text{MaxCap}_{so}$ and $\text{MinCap}_{so}$ are the maximum and minimum amounts of gas the facility can hold on a given day.
The initial amount of gas storage operator \( so \) holds is \( IntCap_{so} \).

Inequalities (4.55) and (4.56) ensure upper and lower bounds on the daily injection and withdrawal rates respectively. Inequality (4.57) ensures the physical constraints of the storage facility are met for each day. The variables in brackets alongside these constraints are their corresponding Lagrange multipliers. The stationary KKT conditions for this problem are as follows:

\[
\pi_{p,d} + a_{so} + \lambda_{I_{so,p,d}} - \sum_{e=d}^{D} (\lambda_{Cap_{so,e}} - \lambda_{Cap_{so,e}}) = 0, \tag{4.58}
\]

\[
-\gamma_{so,d} + b_{so} + \lambda_{W_{so,d}} - \sum_{e=d}^{D} (\lambda_{Cap_{so,e}} - \lambda_{Cap_{so,e}}) = 0, \tag{4.59}
\]

where \( \lambda_{I} \) and \( \lambda_{W} \) are the Lagrange multipliers associated with the upper and lower bounds of the constraints above. As these Lagrange multipliers are associated with inequality constraints they are required by the KKT conditions to be non-negative. The complementary KKT conditions for this problem are:

\[
\lambda_{I_{so,p,d}} (I_{so,p,d} - I_{so}^{max}) = 0, \tag{4.60}
\]

\[
\frac{\lambda_{I_{so,p,d}}}{I_{so,p,d}} I_{so,p,d} = 0, \tag{4.61}
\]

\[
\lambda_{W_{so,d}} (W_{so,d} - W_{so}^{max}) = 0, \tag{4.62}
\]

\[
\frac{\lambda_{W_{so,d}}}{W_{so,d}} W_{so,d} = 0, \tag{4.63}
\]

\[
\lambda_{Cap_{so,d}} (IntCap_{so} + \sum_{e=1}^{d} \sum_{p=1}^{P} I_{so,p,e} - W_{so,e}) - MaxCap_{so} = 0, \tag{4.64}
\]

\[
\frac{\lambda_{Cap_{so,d}}}{MaxCap_{so}} (MinCap_{so} - IntCap_{so} - \sum_{e=1}^{d} \sum_{p=1}^{P} I_{so,p,e} - W_{so,e}) = 0. \tag{4.65}
\]

The KKT conditions also require that constraints (4.55)–(4.57) are satisfied.

### 4.2.4 Marketer’s problem

The final subproblem defined by the model is the marketer’s. The marketer buys gas from both the storage operators and the producers. He/she then sells this gas on at a fixed price \( \tau \), in order to
4.2 Complementarity-based models of natural gas markets

meet demand. The problem is:

\[
\text{Max } \sum_{d=1}^{D} \left( \sum_{p=1}^{P} \left( \tau H_{p,d} + \sum_{so=1}^{SO} U_{so,d} \right) - \sum_{p=1}^{P} \pi_{p,d} H_{p,d} - \sum_{so=1}^{SO} \gamma_{so,d} U_{so,d} \right),
\]

subject to:

\[
\sum_{p=1}^{P} H_{p,d} + \sum_{so=1}^{SO} U_{so,d} = \text{Demand}_d, \quad (\lambda_{\text{Demand}_d}), \quad (4.66)
\]

\[
\gamma_{so,d} \geq 0, \quad (\lambda_{\gamma_{so,d}}), \quad (4.67)
\]

where \( \gamma_{so,d} \) is storage operator so’s charge per unit of gas on day \( d \) and \( \pi_{p,d} \) is producer \( p \)’s charge on day \( d \). Both \( \gamma_{so,d} \) and \( \pi_{p,d} \) are exogenous to the marketer’s problem and are determined according to the optimality problem defined for the system as a whole. The amount the marketer buys from producer \( p \) on day \( d \) is \( H_{p,d} \). The amount he buys from storage operator so on day \( d \) is \( U_{so,d} \). Both \( H_{p,d} \) and \( U_{so,d} \) are endogenous to the marketer’s problem.

Equation (4.66) ensures that demand is met on a daily basis. The demand for gas on day \( d \) is \( \text{Demand}_d \). It is a pre-determined fixed input into the model. Inequality (4.67) ensures \( H_{p,d} \) is non-negative.

Using constraint (4.66), the objective function becomes:

\[
\text{Max } \sum_{d=1}^{D} \left( \text{Demand}_d \tau - \sum_{p=1}^{P} \pi_{p,d} H_{p,d} - \sum_{so=1}^{SO} \gamma_{so,d} U_{so,d} \right).
\]

As daily demands in this model are fixed inputs, the constant \( \tau \sum_{d=1}^{D} \text{Demand}_d \) can be removed from the objective function. This means that the marketer’s problem can be reformulated as:

\[
\text{Max } -\sum_{d=1}^{D} \left( \sum_{p=1}^{P} \pi_{p,d} H_{p,d} + \sum_{so=1}^{SO} \gamma_{so,d} U_{so,d} \right),
\]

subject to conditions (4.66) and (4.67). The resulting stationary KKT conditions for this problem are:

\[
\pi_{p,d} - \lambda_{\text{Demand}_d} - \lambda_{H_{p,d}} = 0, \quad (4.68)
\]

\[
\gamma_{so,d} - \lambda_{\text{Demand}_d} = 0. \quad (4.69)
\]

The complementary KKT condition is

\[
H_{p,d} \lambda_{H_{p,d}} = 0. \quad (4.70)
\]

The KKT conditions also require that constraints (4.66) and (4.67) are satisfied. As \( \lambda_{H_{p,d}} \) is asso-
4. GAS MARKET MODELS

associated with an inequality constraint, it is required by the KKT conditions to be non-negative.

4.2.5 Market clearing conditions

In order to solve the different subproblems defined by the model simultaneously, the market clearing conditions must be satisfied. The market clearing conditions are:

\[ H_{p,d} + \sum_{so=1}^{SO} I_{so,p,d} = Q_{p,d}, \]  
\[ (4.71) \]

\[ W_{so,d} = U_{so,d}, \]  
\[ (4.72) \]

where each of the variables are defined as above. Equation (4.71) ensures that producer \( p \) sells all the gas he has produced on day \( d \) to either the storage operators or the marketer. Equation (4.72) ensures that the amount of gas withdrawn by the storage operator \( so \) on day \( d \) is equal to the amount of gas bought by the marketer from storage operator \( so \) on day \( d \).

4.3 Cost minimisation

Having presented the complementarity-based equilibrium model of this project in the previous section, the second methodology (cost minimisation) used to model natural gas markets is now considered. Firstly, some models that use this methodology are discussed.

Cost minimisation involves an imaginary central planner choosing the different sources of supply to meet demand at least overall cost. Using this methodology, a gas market model can be formulated using a single linear program. Unlike complementarity-based equilibrium models, Nash-Cournot competition cannot be included in the formulation of a cost minimisation model as there is only one objective function. As a result, demand must be determined exogenously. There are a number of previous models than have used this approach.

In [26], Bopp et al. developed a cost minimisation model for optimising the flows of gas for a particular company. The model includes stochastic demand and ensures that the demand for every scenario is met at least cost. In [27], Butler developed a similar model. The two models differ in that Bopp’s model is optimised over yearly quarters, while Butler’s is optimised over days, weeks and months. They also differ in their application. Bopp’s model is applied to a natural gas local distribution company, while Butler’s model uses data from an electric utility company. In both models, demand is met at least cost by producers and a storage operator.

In [44, 45], Perner uses the EUGAS model he developed in [46] to model the European gas market. Like Bopp’s and Butler’s models it optimises the flows of gas by ensuring that demand is met at least cost. Unlike these models however, the EUGAS model does not include stochastic demand. It does, however, allow for long-term investment decisions on things such as production capacities. Also, in contrast to [26, 27], it is optimised over five year periods.
4.3 Cost minimisation

In [47], Lochner developed a model which is based on the MAGELEN model, [48]. The MAGELEN model is a global extension to Perneer’s EUGAS model, [45]. It too optimises natural gas supply, whilst also including long-term decisions on investments such as production capacities.

4.3.1 A cost minimisation model and its KKT conditions

Having discussed the models seen in the literature that used the cost minimisation methodology to model natural gas markets, the cost minimisation model of this project is now introduced. As with the model presented in Section 4.2.1, it consists of \( P \) producers and \( SO \) storage operators over \( D \) days. However, instead of each player in the market separately maximising their profits, this model involves an imaginary central planner choosing how these different sources of supply meet demand at minimum cost. In the complementarity-based equilibrium model, the marketer decides between the producers and storage operators when it comes to buying gas, in order to meet consumer demand. As the imaginary central planner does this in the cost minimisation model, there is no need to include a marketer in this formulation.

As previously, the indices \( p, so \) and \( d \) run over the ranges \( 1, \ldots, P, 1, \ldots, SO \) and \( 1, \ldots, D \), respectively. The model is defined as follows:

\[
\min \sum_{d=1}^{D} \sum_{p=1}^{P} \sum_{so=1}^{SO} (b_{so} W_{so,d} + a_{so} I_{so,p,d} + c_{p} Q_{p,d}),
\]

subject to:

\[
0 \leq Q_{p,d} \leq Q_{p,d}^{\max}, \quad (\lambda_{Q_{p,d}}), \quad (4.73)
\]

\[
0 \leq I_{so,p,d} \leq I_{so,p,d}^{\max}, \quad (\lambda_{I_{so,p,d}}), \quad (4.74)
\]

\[
0 \leq W_{so,d} \leq W_{so,d}^{\max}, \quad (\lambda_{W_{so,d}}), \quad (4.75)
\]

\[
\text{MinCap}_{so} \leq \text{IntCap}_{so} + \sum_{e=1}^{d} \left( \sum_{p=1}^{P} I_{so,p,e} - W_{so,e} \right) \leq \text{MaxCap}_{so}, \quad (\lambda_{\text{Cap}_{so,d}}), \quad (4.76)
\]

\[
\text{Demand}_{d} = \sum_{so=1}^{SO} W_{so,d} + \sum_{p=1}^{P} H_{p,d}, \quad (\lambda_{\text{Demand}_{d}}), \quad (4.77)
\]

\[
H_{p,d} + \sum_{so=1}^{SO} I_{so,p,d} = Q_{p,d}, \quad (\lambda_{MCC_{p,d}}), \quad (4.78)
\]

\[
H_{p,d} \geq 0, \quad (\lambda_{H_{p,d}}), \quad (4.79)
\]

where \( c_{p} \) is the cost per unit of gas associated with producer \( p \), \( Q_{p,d} \) is the amount of gas produced
4. GAS MARKET MODELS

by producer $p$ on day $d$, $W_{so,d}$ is the amount withdrawn from storage facility $so$ on day $d$ and $I_{so,p,d}$ is the amount injected on day $d$ by producer $p$ to storage operator $so$. The costs per unit associated with injecting and withdrawing gas to and from storage for storage operator $so$ are $a_{so}$ and $b_{so}$ respectively. $Q_{p,d}^{\text{max}}$ is the daily maximum withdrawal rate for producer $p$. $\text{MaxCap}_{so}$ and $\text{MinCap}_{so}$ are the maximum and minimum capacities for storage operators. The initial amount of gas in storage for storage operator $so$ is $\text{IntCap}_{so}$. $W_{so}^{\text{max}}$ and $I_{so}^{\text{max}}$ are the maximum daily withdrawal and injection rates for storage operator $so$.

$H_{p,d}$ is the amount that producer $p$ produces on day $d$ that does not go into storage. The demand on day $d$ is $\text{Demand}_d$. It is a fixed pre-defined input to the model along with $c_p$, $a_{so}$, $b_{so}$, $Q_{p,d}^{\text{max}}$, $W_{so}^{\text{max}}$ and $I_{so}^{\text{max}}$. The outputs to be determined by the model are $Q_{p,d}$, $H_{p,d}$, $W_{so,d}$ and $I_{so,p,d}$. These inputs and outputs are identical to those used in the complementarity-based equilibrium model and hence those in Table 4.1.

Inequalities (4.125)–(4.127) ensure upper and lower bounds on the daily production, withdrawal and injection rates. Inequality (4.128) ensures the physical constraints of the storage facility are met for each day. Constraint (4.130) is the demand constraint, which ensures that the amount produced on day $d$, plus the amount withdrawn from storage, less the amount injected, is equal to demand for that day. Equation (4.78) is a market clearing condition, which ensures that the amount of gas produced by producer $p$ on day $d$ either goes to satisfy demand or to storage, while inequality (4.129) ensures $H_{p,d}$ is non-negative.

The variable in brackets alongside each of the constraints above is the Lagrange multiplier associated with that constraint. The stationary KKT conditions for this model are as follows:

$$\lambda_{\text{MCC}_{p,d}} + c_p + \lambda_{Q_{p,d}} - \overline{\lambda}_{Q_{p,d}} = 0,$$

(4.80)

$$-\lambda_{\text{MCC}_{p,d}} + a_{so} + \lambda_{I_{so,p,d}} - \overline{\lambda}_{I_{so,p,d}} + \sum_{e=d}^{D} \left( \lambda_{\text{Cap}_{so,e}} - \overline{\lambda}_{\text{Cap}_{so,e}} \right) = 0,$$

(4.81)

$$-\lambda_{\text{Demand}_d} + b_{so} + \lambda_{W_{so,d}} - \overline{\lambda}_{W_{so,d}} - \sum_{e=d}^{D} \left( \lambda_{\text{Cap}_{so,e}} - \overline{\lambda}_{\text{Cap}_{so,e}} \right) = 0,$$

(4.82)

$$-\lambda_{\text{Demand}_d} - \lambda_{\text{MCC}_{p,d}} - \lambda_{H_{p,d}} = 0,$$

(4.83)

where $\lambda_{\cdot}$ represents the multiplier associated with the upper bound, while $\overline{\lambda}_{\cdot}$ represents the multiplier associated with the lower bound. The complementarity KKT conditions are:

$$\lambda_{Q_{p,d}} (Q_{p,d} - Q_{p,d}^{\text{max}}) = 0,$$

(4.84)

$$\overline{\lambda}_{Q_{p,d}} Q_{p,d} = 0,$$

(4.85)

\(^1\)The values $\tau_p$, $\pi_{p,d}$, $\gamma_{so,d}$ and $U_{so,d}$ only appear in the complementarity-based equilibrium model.
4.4 Equivalence of models using KKT analysis

\[ \lambda_{I_{so,p,d}} (I_{so,p,d} - I_{so}^{max}) = 0, \]  
\[ \overline{\lambda}_{I_{so,p,d}} I_{so,p,d} = 0, \]  
\[ \lambda_{W_{so,d}} (W_{so,d} - W_{so}^{max}) = 0, \]  
\[ \overline{\lambda}_{W_{so,d}} W_{so,d} = 0, \]

\[ \lambda_{Cap_d} (IntCap_{so} + \sum_{p=1}^{P} \sum_{e=1}^{d} I_{so,p,e} - W_{so,e}) - MaxCap_{so} = 0, \]  
\[ \overline{\lambda}_{Cap_d} (MinCap_{so} - IntCap_{so} - \sum_{p=1}^{P} \sum_{e=1}^{d} I_{so,p,e} - W_{so,e}) = 0, \]  
\[ H_{p,d} \lambda_{H_{p,d}} = 0. \]

Each of the Lagrange multipliers associated with an inequality constraint is required to be non-negative by the KKT conditions. The only Lagrange multipliers not required to be non-negative are \( \lambda_{Demand_d} \) and \( \lambda_{MCC_{p,d}} \) as they are associated with equality constraints, equations (4.130) and (4.78) respectively. The KKT conditions also require that constraints (4.125)–(4.129) are satisfied. As this model is a linear program, the KKT conditions are both necessary and sufficient for optimality [29].

4.4 Equivalence of models using KKT analysis

Having presented the complementarity-based equilibrium model in Section 4.2.1 and the cost minimisation model in Section 4.3.1, it remains to show that the two models are equivalent. This is done by showing the KKT conditions for both models reduce to the same set of equations. In Section 4.4.1 the stationary KKT conditions for both models are considered, followed by a comparison of the complementarity KKT conditions in Section 4.4.2. Finally, in Section 4.4.3 the constraints and Lagrange multipliers are considered.

4.4.1 Stationary KKT conditions

The stationary KKT conditions for the complementarity-based equilibrium model are equations (4.52), (4.58), (4.59) and (4.68)–(4.69). From equations (4.68) and (4.69), it can be seen that \( \pi_{p,d} = \lambda_{Demand_d} + \lambda_{H_{p,d}} \) and \( \gamma_{so,d} = \lambda_{Demand_d} \). Using these, the remaining KKT conditions for this
model can be reduced to:

\[-\lambda_{\text{Demand}_d} - \lambda_{H_{p,d}} + c_p + \lambda_{Q_{p,d}} - \overline{\lambda_{Q_{p,d}}} = 0, \tag{4.93}\]

\[\lambda_{\text{Demand}_d} + \lambda_{H_{p,d}} + a_{so} + \lambda_{I_{so,p,d}} - \overline{\lambda_{I_{so,p,d}}} + \sum_{e=d}^D (\lambda_{\text{Cap}_{so,e}} - \overline{\lambda_{\text{Cap}_{so,e}}}) = 0, \tag{4.94}\]

\[-\lambda_{\text{Demand}_d} + b_{so} + \lambda_{W_{so,d}} - \overline{\lambda_{W_{so,d}}} - \sum_{e=d}^D (\lambda_{\text{Cap}_{so,e}} - \overline{\lambda_{\text{Cap}_{so,e}}}) = 0. \tag{4.95}\]

The stationary KKT conditions for the cost minimisation model are equations \((4.80)\)–\((4.83)\). From equation \((4.83)\), it can be seen that \(\lambda_{MCC_{p,d}} = -\lambda_{\text{Demand}_d} - \lambda_{H_{p,d}}\). Using this result, equations \((4.80)\)–\((4.82)\) reduce to:

\[-\lambda_{\text{Demand}_d} - \lambda_{H_{p,d}} + c_p + \lambda_{Q_{p,d}} - \overline{\lambda_{Q_{p,d}}} = 0, \tag{4.96}\]

\[\lambda_{\text{Demand}_d} + \lambda_{H_{p,d}} + a_{so} + \lambda_{I_{so,p,d}} - \overline{\lambda_{I_{so,p,d}}} + \sum_{e=d}^D (\lambda_{\text{Cap}_{so,e}} - \overline{\lambda_{\text{Cap}_{so,e}}}) = 0, \tag{4.97}\]

\[-\lambda_{\text{Demand}_d} + b_{so} + \lambda_{W_{so,d}} - \overline{\lambda_{W_{so,d}}} - \sum_{e=d}^D (\lambda_{\text{Cap}_{so,e}} - \overline{\lambda_{\text{Cap}_{so,e}}}) = 0. \tag{4.98}\]

When the stationary KKT conditions for the complementarity-based equilibrium model, equations \((4.93)\)–\((4.95)\), are compared with those with from the cost minimisation problem, equations \((4.96)\)–\((4.98)\), it can be seen that the two sets of equations are identical. This shows that the stationary KKT conditions from the two models reduce to the same set of equations.

### 4.4.2 Complementarity KKT conditions

The complementarity KKT conditions for the complementarity-based equilibrium model are given by equations \((4.53)\)–\((4.54)\), \((4.60)\)–\((4.65)\) and \((4.70)\). These are the same as those from the cost minimisation model, equations \((4.84)\)–\((4.92)\). This shows that the complementarity KKT conditions for the complementarity-based equilibrium model and the cost minimisation model are identical.

### 4.4.3 Constraints and Lagrange Multipliers

The constraints of the complementarity-based equilibrium model include the constraints from each of the different subproblems in that model, plus the market clearing conditions. They are given by equations \((4.51)\), \((4.55)\)–\((4.57)\), \((4.66)\)–\((4.67)\) and \((4.71)\)–\((4.72)\). Combining equations \((4.72)\) and
4.4.4 Equivalence of models

As the stationary KKT conditions, complementarity KKT conditions, constraints and Lagrange multipliers in both models reduce to an identical set of equations, the KKT conditions for optimality are the same for both models. In this sense, both models are equivalent. At this point it must also be noted that the two models may produce different solutions, given identical inputs, if there are non-unique optimal solutions. An example of non-unique solutions is provided in Section 4.6.

4.5 Equivalence of models using the principle of symmetry

In the previous section, it was shown that the complementarity-based equilibrium and the cost minimisation models are equivalent. This was done by comparing the KKT conditions of the two problems. In this section, the same result is proven using the principle of symmetry described in [28]. This is done by reconsidering the KKT conditions of the complementarity-based equilibrium model. Its stationary KKT conditions are

\[ -\pi_{p,d} + \lambda_{Q,p,d} - \overline{\lambda_{Q,p,d}} + c_p = 0, \]

\[ \pi_{p,d} + \lambda_{I_{so,p,d}} - \overline{\lambda_{I_{so,p,d}}} + \sum_{e=d}^{D} (\lambda_{Cap_{so,e}} - \overline{\lambda_{Cap_{so,e}}}) + a_{so} = 0, \]

\[ -\gamma_{so,d} + \lambda_{W_{so,d}} - \overline{\lambda_{W_{so,d}}} - \sum_{e=d}^{D} (\lambda_{Cap_{so,e}} - \overline{\lambda_{Cap_{so,e}}}) + b_{so} = 0, \]

\[ \pi_{p,d} - \lambda_{Demand_d} - \lambda_{H_{p,d}} = 0, \]

\[ \gamma_{so,d} - \lambda_{Demand_d} = 0, \]
while the constraints are:

\[ H_{p,d} + \sum_{so=1}^{SO} I_{so,p,d} = Q_{p,d}, \quad (4.105) \]

\[ W_{so,d} = U_{so,d}, \quad (4.106) \]

\[ \sum_{p=1}^{P} H_{p,d} + \sum_{so=1}^{SO} U_{so,d} = Demand_d, \quad (4.107) \]

\[ 0 \leq Q_{p,d} \leq Q_{p,d}^{\text{max}}, \quad (4.108) \]

\[ 0 \leq I_{so,p,d} \leq I_{so,d}^{\text{max}}, \quad (4.109) \]

\[ 0 \leq W_{so,d} \leq W_{so}^{\text{max}}, \quad (4.110) \]

\[ MinCap_{so} \leq IntCap_{so} + \sum_{e=1}^{d} (\sum_{p=1}^{P} I_{so,p,e} - W_{so,e}) \leq MaxCap_{so}, \quad (4.111) \]

\[ H_{p,d} \geq 0. \quad (4.112) \]

Using these equations, the problem presented above can be represented as a mixed linear complementarity problem (MLCP(\(F; L; U\))) and hence a mixed complementarity problem where \(F = Mz + q\) and \(L\) and \(U\) are the lower and upper bounds for the variables in the vector \(z\) respectively. \(F = Mz + q\) takes the form:

\[ F = \begin{bmatrix} 0 & -D^T & -A^T \\ -D & 0 & 0 \\ -A & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ v \\ u \end{bmatrix} + \begin{bmatrix} c \\ e \\ b \end{bmatrix}, \quad (4.113) \]
where

\[ M = \begin{bmatrix}
0 & -D & -A^T \\
-D & 0 & 0 \\
-A & 0 & 0
\end{bmatrix}, \quad (4.114) \]

\[ z = \begin{bmatrix}
x \\
v \\
u
\end{bmatrix}, \quad (4.115) \]

\[ q = \begin{bmatrix}
c \\
e \\
b
\end{bmatrix}, \quad (4.116) \]

The matrix \( D \) represents the constant coefficients associated with the equality constraints (equations (4.105)-(4.107)), while the matrix \( A \) represents those associated with the inequality constraints (equations (4.108)-(4.112)). The vector \( x \) represents the decision variables while the vectors \( v \) and \( u \) represent the Lagrange multipliers associated with equality and inequality constraints, respectively. Note: the variables \( \pi_{p,d} \) and \( \gamma_{so,d} \) are included in the vector \( v \) as they are multipliers associated with the MCC. The components of the vector \( q \), the vectors \( c \), \( e \), and \( b \), represent the fixed terms associated with the stationary KKT conditions, equality constraints and inequality constraints, respectively. Details of the structure of the matrices \( D \) and \( A \), as well as the vectors \( x, v, u, c, e \) and \( b \) can be found in Appendix B.2.

The upper bounds for each of the variables in \( z \) are equal to \( \infty \) i.e., \( U_i = \infty \) \( \forall \ i \). The lower bounds for each of the variables in \( z \) are 0 for variables that are Lagrange multipliers associated with inequality constraints (i.e., the vector \( u \)) and \(-\infty \) otherwise.

As this problem is a \( MLCP(F, L, U) \), it can be re-written as the variational inequality problem \( VI([L, U], F) \) (see [11] and Section 4.1.1). A variational inequality problem, \( VI(X, F) \), can be represented by the following mathematical program:

\[ \min \ f(x) \ \text{subject to } x \in X, \quad (4.117) \]

where \( F = \nabla f \). This result is known as the principle of symmetry [11, 28] and holds if and only if the Jacobian matrix, \( \nabla F \), is symmetric. As mentioned above, the \( F \) function for the \( MLCP(F, L, U) \) is \( F = Mz + q \). As a result, \( \nabla F = M \), is clearly symmetric from equation (4.114). It also means (using the fact that \( F = \nabla f \)) that the mathematical program, (4.117), can be written in the form:

\[ \min \ \frac{1}{2} z^T M z + q^T z \ \text{subject to } z \in X, \quad (4.118) \]

where \( X = [L, U] \) and where \( L \) and \( U \) are as stated above. The objective function in equation

59
4. GAS MARKET MODELS

(4.118) can be re-written as follows:

\[
\min \frac{1}{2} \begin{bmatrix} x \\ v \\ u \end{bmatrix} \begin{bmatrix} 0 & -D^T & -A^T \\ -D & 0 & 0 \\ -A & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ v \\ u \end{bmatrix} + \begin{bmatrix} c \\ e \\ b \end{bmatrix} \begin{bmatrix} x \\ v \\ u \end{bmatrix},
\]

(4.119)

and hence,

\[
\min \ c^T x + e^T v + b^T u - v^T D x - u^T Ax,
\]

(4.120)

subject to \( z^T = [x, v, u] \in [L, U] \). This leads to:

\[
\min \ c^T x - u^T (Ax - b) - v^T (Dx - e),
\]

(4.121)

subject to \( z^T = [x, v, u] \in [L, U] \). The mathematical program, (4.121), minimises the Lagrangian of, and hence is equivalent to, the following linear program:

\[
\min \ c^T x,
\]

(4.122)

subject to:

\[
Ax \geq b, \quad \text{ (4.123)}
\]

\[
Dx = e. \quad \text{ (4.124)}
\]

Re-using the definitions for each of these matrices/vectors given above, this linear program can be re-written as follows:

\[
\min \sum_{d=1}^{D} \sum_{p=1}^{P} \sum_{so=1}^{SO} (b_{so}W_{so,d} + a_{so}I_{so,p,d} + c_P Q_{p,d}),
\]

subject to:

\[
0 \leq Q_{p,d} \leq Q_{p,d}^{\text{max}}, \quad \text{ (4.125)}
\]

\[
0 \leq I_{so,p,d} \leq I_{so,p,d}^{\text{max}}, \quad \text{ (4.126)}
\]

\[
0 \leq W_{so,d} \leq W_{so,d}^{\text{max}}, \quad \text{ (4.127)}
\]

\[
\text{MinCap}_{so} \leq \text{IntCap}_{so} + \sum_{e=1}^{d} \left( \sum_{p=1}^{P} I_{so,p,e} - W_{so,e} \right) \leq \text{MaxCap}_{so}, \quad \text{ (4.128)}
\]

\[
H_{p,d} \geq 0, \quad \text{ (4.129)}
\]
4.6 Non-unique solutions

\[ Demand_d = \sum_{so=1}^{SO} U_{so,d} + \sum_{p=1}^{P} H_{p,d}, \]  
(4.130)

\[ H_{p,d} + \sum_{so=1}^{SO} I_{so,p,d} = Q_{p,d}, \]  
(4.131)

\[ W_{so,d} = U_{so,d}. \]  
(4.132)

By combining and rewriting equations (4.130) - (4.132) with the following single equation:

\[ Demand_d = \sum_{so=1}^{SO} W_{so,d} + \sum_{p=1}^{P} H_{p,d}, \]  
(4.133)

this linear program is the same as the cost minimisation problem presented in Section 4.3.1. Hence, as already shown using KKT analysis in Section 4.4, the complementarity-based equilibrium model and the cost minimisation model are equivalent models.

4.6 Non-unique solutions

It is stated in Section 4.4.4 that despite having been proved equivalent, the complementarity-based equilibrium model and the cost minimisation model may provide different results, given identical inputs. This is due to non-unique solutions. A simple numerical example is used in this section to demonstrate this. Consider each model with two producers (i.e., \( P = 2 \)), one storage operator (i.e., \( SO = 1 \)) and three days (i.e., \( D = 3 \)) with the following inputs:

<table>
<thead>
<tr>
<th></th>
<th>Producer 1</th>
<th>Producer 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>( Q_{p}^{max} )</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>( c_{p} )</td>
<td>20</td>
<td>40</td>
</tr>
</tbody>
</table>

Table 4.2: Inputs associated with the producers for both models

<table>
<thead>
<tr>
<th></th>
<th>Day 1</th>
<th>Day 2</th>
<th>Day 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>( Demand_d )</td>
<td>120</td>
<td>125</td>
<td>215</td>
</tr>
</tbody>
</table>

Table 4.3: Inputs associated with demand for both models
4. GAS MARKET MODELS

<table>
<thead>
<tr>
<th>$W_{\text{max}}$</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>$I_{\text{max}}$</td>
<td>20</td>
</tr>
<tr>
<td>MaxCap</td>
<td>500</td>
</tr>
<tr>
<td>IntCap</td>
<td>0</td>
</tr>
<tr>
<td>MinCap</td>
<td>0</td>
</tr>
<tr>
<td>$a_1$</td>
<td>1</td>
</tr>
<tr>
<td>$b_1$</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 4.4: Inputs associated with the storage operator for both models

On the first two days producers 1 and 2 are both needed to meet demand. The first producer (i.e., the cheapest producer) will be at their maximum production rate (i.e., $Q_{1,1} = Q_{1,2} = 100$) and the second producer will meet the rest of the demand. On the third day the two producers will be at their maximum production rate and storage will be needed to meet demand. As a result, gas will have to be injected into storage on either the first or the second day in order to ensure that there is gas in storage for the third day. In the equilibrium model, the prices associated with the two producers on the first two days are the exact same (i.e., $\pi_{1,1} = \pi_{1,2} = \pi_{2,1} = \pi_{2,2} = 40$) as both producers, but not storage, are needed. This means it does not matter on which day the gas is injected into storage, as the profits of the two players will be not affected. Similarly, it does not matter which producer injects the gas, as the producers will be either meeting demand or injecting and thus their profits will not be affected. This means that there is a range of optimal solutions for this example. The same can be said for the cost minimisation model, as the cost of the system is equally minimised regardless of whether the gas is injected on the first or second day or by producer 1 or 2. The following outputs are optimal for each of the models for this example.
Table 4.5: Outputs to both models

<table>
<thead>
<tr>
<th></th>
<th>Solution 1</th>
<th>Solution 2</th>
<th>Solution 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Q_{1,1}$</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>$Q_{1,2}$</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>$Q_{1,3}$</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>$Q_{2,1}$</td>
<td>20</td>
<td>35</td>
<td>27</td>
</tr>
<tr>
<td>$Q_{2,2}$</td>
<td>40</td>
<td>25</td>
<td>33</td>
</tr>
<tr>
<td>$Q_{2,3}$</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>$I_{1,1,1}$</td>
<td>0</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>$I_{1,1,2}$</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$I_{1,1,3}$</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$I_{1,2,1}$</td>
<td>0</td>
<td>15</td>
<td>2</td>
</tr>
<tr>
<td>$I_{1,2,2}$</td>
<td>15</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>$I_{1,2,3}$</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$W_{1,1}$</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$W_{1,2}$</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$W_{1,3}$</td>
<td>15</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>$Q_{1,1}$</td>
<td>100</td>
<td>100</td>
<td>95</td>
</tr>
<tr>
<td>$H_{1,2}$</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>$H_{1,3}$</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>$H_{2,1}$</td>
<td>20</td>
<td>20</td>
<td>25</td>
</tr>
<tr>
<td>$H_{2,2}$</td>
<td>25</td>
<td>25</td>
<td>25</td>
</tr>
<tr>
<td>$H_{2,3}$</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>$U_{1,1}$</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$U_{1,2}$</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$U_{1,3}$</td>
<td>15</td>
<td>15</td>
<td>15</td>
</tr>
</tbody>
</table>

The prices ($p, d, \gamma, \lambda_{\text{demand}}$) associated with each of the days 1-3 are 40, 40 and 42, respectively. Table 4.5 shows that, according to solution 1, producer 2 injects 15 units of gas into storage on the second day while, according to solution 2 they do this on day 1. According to solution 3, producer 1 injects 5 units on day 1 while producer 2 injects 2 units on day 1 and 8 on day 2. Each of these solutions are optimal for both the complementarity-based equilibrium model and the cost minimisation model. These results show that while the cost minimisation and complementarity-based models may be equivalent, they can produce different outputs given identical inputs.

4.7 Summary and conclusion

In this chapter, models of a natural gas market without market power are examined. This is done by comparing examples of two different types of natural gas market models seen in the literature. In
4. GAS MARKET MODELS

Section 4.2.1, a complementarity-based equilibrium model is introduced. It is based on the model described in [11], but does not allow market power amongst its players (i.e., no Nash-Cournot competition). It contains \( P \) producers, \( SO \) storage operators and a marketer. Each of these players act separately in this model by attempting to maximise their own individual profits.

In Section 4.3.1, the corresponding cost minimisation model is presented. This model involves an imaginary central planner choosing how the different sources of supply (i.e., the producers and storage operators) meet demand at minimum overall system cost. In Sections 4.4 and 4.5 these two models are shown to be equivalent. In Section 4.4 this is done by showing that the KKT conditions for optimality under both models reduce to the same set of equations while, in Section 4.5 the same result is shown using the principle of symmetry.

Cost minimisation, with its imaginary central planner, may be seen as an unintuitive approach to modelling natural gas markets. Complementarity-based equilibrium models are more intuitive as each player acts separately in order to maximise their own profits. They are, however, algebraically and structurally more complicated. This result shows that while the cost minimisation model may be unintuitive for natural gas markets, it can be used instead of the more intuitive, but complex, complementarity-based equilibrium approach when modelling markets without market power. This is the approach followed in the development of the Rolling Optimisation Model in Chapter 5.
Chapter 5

Rolling Optimisation Model

In the previous chapter it was shown that the outputs from the cost minimisation model are equivalent to those from the complementarity-based equilibrium model when market power is not included. As the cost minimisation model is structurally less complicated (i.e., it has fewer variables and constraints), it is used in this chapter as a basis for developing a more complicated model. This new model is referred to as the Rolling Optimisation Model (ROM). It is an extension of the cost minimisation model described in Section 4.3.1 as it incorporates stochastic demand. The outputs of ROM are the collection of outputs obtained from a linear program solved numerous times, each time with updated inputs. Some of the inputs are updated using the results of the previous optimisation. This is in contrast to the simple cost minimisation model of Chapter 4, where the outputs are determined by solving a single linear program.

This new model attempts to replicate the daily decisions made in the UK gas market under uncertain information about future demand. It takes as an input the exogenously determined stochastic process for demand developed in Section 3.4. This process generates multiple demand time series, each of length $D$ days, from actual demand and Seasonal Normal Demand (SND) time series supplied by the National Grid. Each generated time series represents a demand scenario in this model. As explained in Section 3.4, the demand on the first day is the same for each of these time series. This means that the stochastic demand on the first day of the model is scenario-independent. The model decides the amount of gas to be injected to (or withdrawn from) the different storage facilities of the model, as well as how the various sources of supply meet this scenario-independent demand on the first day ($d_1$). This is done at minimum cost, whilst also ensuring all possible future demands (over the set of scenarios) are also met at minimum expected cost.

Once the initial optimisation problem of the model is solved, a new set of demand scenarios are developed using the process described in Section 3.4 but now with actual demand and SND time series that have shifted forward by one day. For example, if the first set of demand scenarios was generated using actual demand and SND for the year starting on the 1st of April 2010, then the second set would be generated with the same time series but starting on the 2nd of April 2010. The length ($D$) of the time series remains the same. Using these updated demand scenarios and
the updated amount of gas in storage (obtained from the results of the previous optimisation), the optimisation problem is solved again so that the new scenario-independent demand on the first day (2nd of April) is now met at minimum cost, whilst again ensuring all possible future demands are met at minimum expected cost. Once this demand is met, the demand scenarios are updated again in a similar manner using actual demand and SND time series that have shifted forward one day once more. Following this the optimisation problem is solved again and so on until $R$ optimisations have been performed. Throughout this thesis, solving one optimisation of ROM is defined as one ‘roll’ of the model. Figure 5.1 describes the rolling horizon of the sets of demand scenarios used in this model.

![Figure 5.1: Rolling Horizon of the Sets of Demand Scenarios Used in ROM.](image)

Each roll of the model represents a single day’s decisions of how demand in the UK gas market is met. When one day’s demand is met a new day arrives (i.e., the model moves forward to the next optimisation) where new decisions have to be made on how to meet demand again. It is envisaged that the model will be used by those in the UK gas market on a day-to-day basis, whereby one roll of the model is solved each day. It is anticipated that the stochastic demand scenarios needed for this day would be generated using the daily demand information that is made available by the UK National Grid. This information includes actual demand for the given day, predicted demand for the subsequent five days ahead and SND thereafter.

In this thesis, ROM is specifically applied to the UK gas market. However, it should be noted that, in principle, the model may be applied to any gas market where market power is not an important factor. The outputs of the model are the flows of gas in the UK gas market as well as the
5.1 Gas market models with stochastic demand

In comparison to deterministic models, there are relatively few gas market models that can handle stochastic demand. In this section, however, a sample of such models are discussed. For example, Zhuang’s model [50] is a stochastic complementarity-based equilibrium model. It follows from previous work in [51, 52] and [53] and is an extension of Gabriel et al’s model [11] described in Appendix 4.2, which is a deterministic model. Demand for each season in Zhuang’s model has two possible levels, low or high, with certain probabilities. This stochastic demand is treated through the probabilistic inverse demand functions of the Nash-Cournot players. Players have long-term decisions on how demand is met (which are scenario-independent) while they also have short-term decisions on how demand is met for every scenario. In contrast to Zhuang’s model, the model introduced in Section 5.2 treats demand exogenously and hence has no Nash-Cournot players.

While Zhuang’s model is an example of a complementarity-based equilibrium model that uses stochastic demand, Bopp’s and Butler’s models, [26] and [27] respectively, are examples of single linear program cost minimisation models that use stochastic demand. Both Bopp’s and Butler’s models optimise the flows of gas by minimising the cost of meeting demand. They do this whilst ensuring the demand for each of the scenarios is met.

The Rolling Optimisation Model presented in Section 5.2 also ensures that demand for each day for each of the scenarios is met. This constraint means that as the number of scenarios included in the model increases so does the size of extreme demands required to be met by supply. Further detail on this and its effects are provided in Section 5.6. In contrast to both Bopp’s and Butler’s
models, ROM is formulated with multiple storage operators and the demand on the first day is scenario-independent. As well as this, the current work is applied to the UK gas market as a whole and not from the perspective of a single company.

While each of the models mentioned in this section include stochastic demand, none of them use the rolling optimisation technique whereby the results of the model are obtained by solving a series of linked optimisation problems with a rolling horizon of demand scenarios. This technique is, however, used in the WILMAR model for electricity markets [54, 55]. In WILMAR, decisions are made on how to meet electricity demand in the scenario-independent day ahead market, whilst respecting all possible demands in the intra-day market. Once these decisions are made, the model rolls forward to the next planning period where the same decisions are made for the new period.

5.2 Model

Having discussed models of natural gas markets that included stochastic demand in the previous section, ROM is now introduced. As described in the introduction of this chapter, ROM is applied by optimising a linear program numerous times, each time with updated inputs.

Each roll of the model consists of $P$ sources of supply, $SO$ storage facilities, $n$ different demand scenarios and a time horizon of $D$ days. The $P$ sources of supply provide gas that is used either to meet demand or injected to storage for each day within the time horizon. The model constrains the daily amount of gas each source $P$ can supply. These sources of supply are equivalent to producers in Chapter 4 and hence, similar notation is used. On each day, the $SO$ storage facilities either inject gas coming from the sources of supply to its facility or withdraw gas from its facility to be used in meeting demand, or do nothing. ROM constrains the maximum amount of gas injected to, or withdrawn from, storage on any given day. It also constrains the maximum and minimum amount of gas allowed to be held in storage on any given day. These storage facilities are equivalent to storage operators modelled in Chapter 4.

As mentioned in the introduction of this chapter, each of the $n$ demand scenarios is a time series generated using the stochastic process for demand developed in Section 3.4. Each scenario $s$ has a probability, $Prob^s$, associated with it. As explained in Section 3.4, all these demand scenario time series have an identical value on the first day. As a result, the first day of ROM is scenario-independent.

As with the cost minimisation model (see Section 4.3.1), ROM involves an imaginary central planner choosing how the different sources of supply and storage facilities meet demand on this scenario-independent first day, whilst ensuring all possible future demands are also met at minimum expected cost. Throughout this chapter, the index $p$ (for producers) runs from 1 to $P$; the index $so$ (for storage operators) runs from 1 to $SO$; the index $s$ (for scenarios) runs from 1 to $n$ while the index $d$ runs from 2 to $D$, unless otherwise stated. The index $d_1$ represents the scenario-independent first day of the model. Each optimisation (or roll) of ROM may be defined as follows:
\[
\begin{align*}
\min \quad & \sum_{p=1}^{P} (c_p Q_{p,d_1}) + \sum_{s_0=1}^{SO} (a_{s_0,d_1} I_{s_0,d_1} + b_{s_0,d_1} W_{s_0,d_1}) + \\
& \quad + \sum_{s=1}^{n} \text{Prob}^{s} \left[ \sum_{d=2}^{D} \left( \sum_{p=1}^{P} (c_p Q_{p,d}^s) + \sum_{s_0=1}^{SO} (a_{s_0,d} I_{s_0,d}^s + b_{s_0,d} W_{s_0,d}^s) \right) \right] \\
\text{subject to:} \\
\text{Demand}_{d_1} &= \sum_{p=1}^{P} Q_{p,d_1} + \sum_{s_0=1}^{SO} (W_{s_0,d_1} - I_{s_0,d_1}), \quad (\lambda_{\text{Demand}_{d_1}}), \quad (5.2) \\
\text{Demand}_{d_1}^s &= \sum_{p=1}^{P} Q_{p,d}^s + \sum_{s_0=1}^{SO} (W_{s_0,d}^s - I_{s_0,d}^s), \quad (\lambda_{\text{Demand}_{d_1}^s}), \quad (5.3) \\
0 &\leq Q_{p,d_1} \leq Q_{p,d_1}^{\text{max}}, \quad (\lambda_{Q_{p,d_1}}), \quad (5.4) \\
0 &\leq I_{s_0,d_1} \leq I_{s_0,d_1}^{\text{max}}, \quad (\lambda_{I_{s_0,d_1}}), \quad (5.5) \\
0 &\leq W_{s_0,d_1} \leq W_{s_0,d_1}^{\text{max}}, \quad (\lambda_{W_{s_0,d_1}}), \quad (5.6) \\
0 &\leq Q_{p,d}^s \leq Q_{p,d}^{\text{max}}, \quad (\lambda_{Q_{p,d}^s}), \quad (5.7) \\
0 &\leq I_{s_0,d}^s \leq I_{s_0,d}^{\text{max}}, \quad (\lambda_{I_{s_0,d}^s}), \quad (5.8) \\
0 &\leq W_{s_0,d}^s \leq W_{s_0,d}^{\text{max}}, \quad (\lambda_{W_{s_0,d}^s}), \quad (5.9) \\
\text{MinStor}_{s_0} \leq \text{IntStor}_{s_0} + I_{s_0,d_1} - W_{s_0,d_1} \leq \text{MaxStor}_{s_0}, \quad (\lambda_{\text{Stor}_{s_0,d_1}}), \quad (5.10) \\
\text{MinStor}_{s_0} \leq \text{IntStor}_{s_0} + I_{s_0,d_1} - W_{s_0,d_1} + \sum_{e=d}^{\varepsilon} (I_{s_0,e}^s - W_{s_0,e}^s) \leq \text{MaxStor}_{s_0}, \quad (\lambda_{\text{Stor}^s_{s_0,d}}). \quad (5.11)
\end{align*}
\]

Tables 5.1 and 5.2 name the output variables and parameters associated with the model, respectively, while Table 5.3 describes the input variables that are updated before each roll of the model. As in Chapter 4, the variable in the parentheses, alongside constraints (5.2) – (5.11), represent the Lagrange multiplier of that constraint. These Lagrange multipliers are needed for the Karush-
Kuhn-Tucker (KKT) analysis used in Section 5.5. Some of these Lagrange multipliers are also used as outputs of the model. In particular $\lambda_{\text{Demand}_d}$ and $\lambda_{\text{Demand}_d^s}$ are the Lagrange Multipliers associated with the demand constraints (5.2) and (5.3), and hence represent the marginal cost of meeting demand. As a result, they are used throughout this chapter to represent the price of gas.

The objective function (equation (5.1)) minimises the total cost of production, the cost of injection to storage and the cost of withdrawal from storage for the scenario-independent day $d_1$, as well as the expected cost for all other days in the model. This expected cost is calculated as a weighted average over all possible demand scenarios with the weight associated with each scenario being determined by the probability $\text{Prob}^s$. Equation (5.2) ensures that demand on the first day is met, while equation (5.3) ensures that demand is met on every other day $d$ for each scenario $s$. Equations

<table>
<thead>
<tr>
<th>Variable</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Q_{p,d_1}$</td>
<td>Amount supplied by source $p$ for the first day</td>
</tr>
<tr>
<td>$I_{so,d_1}$</td>
<td>Amount injected by storage facility $so$ on the first day</td>
</tr>
<tr>
<td>$W_{so,d_1}$</td>
<td>Amount withdrawn by storage facility $so$ on the first day</td>
</tr>
<tr>
<td>$Q_{p,d}$</td>
<td>Amount supplied by source $p$ for day $d$ associated with scenario $s$</td>
</tr>
<tr>
<td>$I_{so,d}$</td>
<td>Amount injected by storage facility $so$ on day $d$ associated with scenario $s$</td>
</tr>
<tr>
<td>$W_{so,d}$</td>
<td>Amount withdrawn by storage facility $so$ on day $d$ associated with scenario $s$</td>
</tr>
<tr>
<td>$\lambda_{\text{Demand}_d}$</td>
<td>Marginal price of gas associated with the first day</td>
</tr>
<tr>
<td>$\lambda_{\text{Demand}_d^s}$</td>
<td>Marginal price of gas associated with day $d$ and scenario $s$</td>
</tr>
</tbody>
</table>

Table 5.1: Outputs associated with ROM.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>$c_p$</td>
<td>Cost associated with supply source $p$</td>
</tr>
<tr>
<td>$a_{so,d}$</td>
<td>Unit cost of injection for storage facility $so$ on day $d$</td>
</tr>
<tr>
<td>$b_{so,d}$</td>
<td>Unit cost of withdrawal for storage facility $so$ on day $d$</td>
</tr>
<tr>
<td>$Q_p^{\text{max}}$</td>
<td>Daily maximum production rate for producer $p$</td>
</tr>
<tr>
<td>$I_{so}^{\text{max}}$</td>
<td>Daily maximum injection rate for storage facility $so$</td>
</tr>
<tr>
<td>$W_{so}^{\text{max}}$</td>
<td>Daily maximum withdrawal rate for storage facility $so$</td>
</tr>
<tr>
<td>$\text{MaxStor}_{so}$</td>
<td>Maximum storage capacity for storage facility $so$</td>
</tr>
<tr>
<td>$\text{MinStor}_{so}$</td>
<td>Minimum storage capacity for storage facility $so$</td>
</tr>
</tbody>
</table>

Table 5.2: Parameters associated with ROM.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\text{Demand}_{d_1}$</td>
<td>Gas demand on the first day</td>
</tr>
<tr>
<td>$\text{Demand}_{d}^s$</td>
<td>Gas demand on day $d$ associated with scenario $s$</td>
</tr>
<tr>
<td>$\text{Prob}^s$</td>
<td>Probability associated with scenario $s$</td>
</tr>
<tr>
<td>$\text{IntStor}_{so}$</td>
<td>Initial amount of gas held by storage facility $so$</td>
</tr>
</tbody>
</table>

Table 5.3: Inputs associated with ROM.

---

1KKT conditions are the conditions that are necessary and sufficient for optimality for linear programs. They, along with Lagrange multipliers, are described in detail in Section 4.1.
and (5.7) provide upper and lower bounds on the production rate for producer \( p \). Equations (5.5), (5.8), (5.6) and (5.9) provide upper and lower bounds for the injection and withdrawal rates for storage facility \( so \). Equation (5.10) ensures that storage facility \( so \) does not withdraw more gas on the first day than is already in the facility. It also ensures that storage facility \( so \) does not inject more gas than the facility can hold. Equation (5.11) does the same for each other day \( d > d_1 \) and for each scenario \( s \).

The decisions made in ROM are those made in the real world UK gas market. Each day decisions are made in the real market on how today’s (exactly-known) demand is met, as in equation (5.2). These decisions must take into account all possible futures demands, as in equation (5.3), and are based on the information that is known today about demand, i.e., the stochastic process for demand. Once these decisions are made in the real market, a new day arrives, which brings improved information about demand. This is replicated in the model, as after each optimisation the inputs (i.e., demands and amount of gas in storage) of the model are updated before a new optimisation problem is solved in the next roll.

The outputs listed in Table 5.1 represent the outputs associated with one roll of ROM. Hence, the total output of ROM is the collection of these variables from each roll. In Sections 5.3 and 5.4 only the output variables associated with the scenario-independent first day (\( d_1 \)) are used for comparison with actual data. The reason for this is that the decision variables associated with \( d_1 \) (i.e., \( Q_{p,d_1}, I_{so,d_1}, W_{so,d_1} \) and \( \lambda_{Demand_{d_1}} \)) replicate actual daily decisions of how daily known demand is met. The rest of the variables presented in Table 5.1 are associated hypothetical decisions of how possible future demands might be met.

In Section 5.5 it is proven that the production costs of ROM can be shifted by adding a constant \( \beta \) to the value of each cost parameter \( c_p \) without affecting the optimal flows according to the model \( (Q_{p,d_1}, Q^{s}_{p,d_1}, I_{so,d_1}, I^{s}_{so,d_1}, W_{so,d_1} \) and \( W^{s}_{so,d_1} \)). Similarly, the production and storage costs can be both scaled by multiplying \( c_p, a_{so,d} \) and \( b_{so,d} \) by a positive constant \( \alpha \), again without affecting the optimal flows. When this is done, the prices of ROM \( (\lambda_{Demand_{d_1}} \) and \( \lambda_{Demand_{d_1}}^{s} \)) become similarly scaled and shifted by the same values of \( \alpha \) and \( \beta \) respectively. This allows the prices produced by ROM to be calibrated to actual SAPs, as is shown in Sections 5.3 and 5.4.

### 5.2.1 Update rules after each roll

After each optimisation (or roll) the inputs of ROM are updated as follows:

1. The actual demand and SND times series move forward one day. For example, if for the first optimisation the actual demand and SND time series were from the 1st October 2008 to the 30th of September 2009, then they would now be from the 2nd of October 2008 to the 1st of October 2009.
2. A new set of demand scenarios are developed using the updated actual demand time series, updated SND time series and the methodology described in Section 3.4.
3. Using the injections to, and withdrawals from, storage for the first day in the previous op-
5. ROLLING OPTIMISATION MODEL

Optimisation, the initial amount of gas in storage, for each storage facility $so$, is updated as follows:

$$IntStor_{so} = IntStor_{so} + I_{so,d1} - W_{so,d1}. \quad (5.12)$$

After each optimisation, the exactly known demand on the first day, $demand_{d1}$, as well as the demand for the following days changes. This reflects what happens in the real UK gas market: for each new day, those in the market have improved information on demand to make their decisions with. Figure 5.2 describes the inputs and outputs of each roll of the model for an example starting on the 1st of April 2012 with 365 rolls.

![Diagram of ROM](image)

Figure 5.2: Inputs and outputs associated with ROM.

5.3 ROM fitted to the year beginning in April 2010

Having introduced ROM in the previous section, the model is now calibrated by fitting it to the UK gas market for the year starting on the 1st of April 2010. The aim of this analysis is to choose the parameters of the model that produce results that best fit actual data. These results include the flows of gas used to meet demand, the amount of gas in storage and the daily SAPs. Each of these are described in further detail below. In order to obtain these results, ROM was formulated with 18 sources of supply ($P = 18$), 3 storage facilities ($SO = 3$) and 3 scenarios ($n = 3$). The model was run over a horizon of $D = 365$ days and for 365 rolls.

$^1$For the first roll of the model, the initial amount of gas in storage is a parameter typically determined using actual storage data.
The sets of three demand scenarios were generated using the stochastic process for demand described in Section 3.4. Actual demand and SND for the year starting on the 1st of April 2010 were used for the first roll (or optimisation) of the model. For the second roll of the model, the stochastic process for demand moved forward one day, i.e., actual demand and SND for the year starting on the 2nd of April 2010 was used to develop the scenarios. For subsequent rolls, SND moved forward in a similar pattern. The choice of 3 for the size of the scenario sets was taken following the analysis in Chapter 6. In Section 5.6 however, the effect of changing the number of scenarios used in the model is examined in detail. As each of the 3 scenarios were randomly generated for each roll of the model, they were assigned equal probabilities, i.e., \( Pr\{s\} = \frac{1}{3} \forall s \).

The 3 storage facilities were chosen to represent the 3 different types of storage facilities in the UK, namely long-, medium- and short-range-storage (see Chapter 2). As described in Chapter 2, the five sources of supply in the UK gas market are

1. UK Continental Shelf (UKCS),
2. Norwegian imports,
3. LNG imports,
4. Balgzand Bacton Line (BBL) pipeline,
5. Interconnector UK (IUK) pipeline.

In ROM these five sources of supply were split up into multiple tranches giving 18 sources of supply in total. Each tranche had a separate cost associated with it. Splitting the five sources of supply into 18 tranches provides greater variability to price of gas modelled in ROM. If \( P = 5 \) instead of \( P = 18 \), then there would only five price levels associated with gas in ROM, which is not comparable with actual prices as detailed in Figure 5.5. The costs \( (c_p) \) and maximum capacities \( (Q_{p}^{max}) \) associated with the different tranches are given in Table 5.4. The maximum capacities are in units of million cubic meters (mcm). As it is proven in Section 5.5 that the costs of the model can be varied, in certain ways, without affecting the optimal flows of the model, these costs are assigned no particular unit.
### Rolling Optimisation Model

<table>
<thead>
<tr>
<th>Tranche</th>
<th>( c_p )</th>
<th>( Q_{p}^{max} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>UKCS_1</td>
<td>58.464</td>
<td>48</td>
</tr>
<tr>
<td>UKCS_2</td>
<td>57.611</td>
<td>61</td>
</tr>
<tr>
<td>UKCS_3</td>
<td>56.552</td>
<td>64</td>
</tr>
<tr>
<td>UKCS_4</td>
<td>69.985</td>
<td>10</td>
</tr>
<tr>
<td>Norway_1</td>
<td>60.985</td>
<td>21</td>
</tr>
<tr>
<td>Norway_2</td>
<td>64.322</td>
<td>54</td>
</tr>
<tr>
<td>Norway_3</td>
<td>48.375</td>
<td>30</td>
</tr>
<tr>
<td>Norway_4</td>
<td>58.637</td>
<td>23</td>
</tr>
<tr>
<td>LNG_1</td>
<td>55.242</td>
<td>38</td>
</tr>
<tr>
<td>LNG_2</td>
<td>58.865</td>
<td>36</td>
</tr>
<tr>
<td>LNG_3</td>
<td>71.458</td>
<td>12</td>
</tr>
<tr>
<td>BBL_1</td>
<td>64.189</td>
<td>11</td>
</tr>
<tr>
<td>BBL_2</td>
<td>55.084</td>
<td>16</td>
</tr>
<tr>
<td>BBL_3</td>
<td>59.352</td>
<td>14</td>
</tr>
<tr>
<td>IUK_1</td>
<td>61.626</td>
<td>11</td>
</tr>
<tr>
<td>IUK_2</td>
<td>69.543</td>
<td>16</td>
</tr>
<tr>
<td>IUK_3</td>
<td>70.977</td>
<td>27</td>
</tr>
<tr>
<td>IUK_4</td>
<td>60.336</td>
<td>18</td>
</tr>
</tbody>
</table>

Table 5.4: The costs and maximum capacities (in mcm) of the different tranches for each source of supply.

The number and size of these tranches, as well as the costs associated with them were determined following consultations with Bord Gáis Energy and by using the simulated annealing algorithm. This algorithm ensured that the parameters chosen produced results that are close to actual data. More details on the algorithm and how it was used to determine the parameters of ROM are described in Section 5.3.1.

Table 5.5 displays the total maximum capacities for the five different sources of supply (i.e., the sum of the tranches’ capacities). The total maximum capacities for Norway, the BBL pipeline and the IUK pipeline were obtained from the UK’s Department of Energy and Climate Change [1]. The total maximum capacities for the UKCS and LNG were estimated using the UK National Grid’s 10 year statement [2]. This statement provides peak supply availability for the gas years beginning on the 1st of October 2009 and 1st of October 2010, neither of which are applicable to the analysis of this section, as this analysis is fitted to data for the year beginning on the 1st of April 2010. As a result, the maximum UKCS and LNG capacities displayed in Table 5.5 are the average of the October 2009 and October 2010 values.

The model’s total maximum capacity of LNG was also assumed to be only at 70% of its true maximum capacity. This is in accordance with the levels assumed by the UK National Grid [56, 57]. 100% LNG capacity would correspond to a situation where a LNG ship is constantly of-
flooding gas into the UK gas network (i.e., when one ship is finished another always automatically begins offloading; this assumption is highly unrealistic and is not supported by the available data). While the size and number of tranches of the different sources of supply were varied in the model’s development, the total maximum capacity of each source of supply was fixed at levels determined from the references above.

<table>
<thead>
<tr>
<th>Source of supply</th>
<th>Total maximum capacity</th>
</tr>
</thead>
<tbody>
<tr>
<td>UKCS</td>
<td>183</td>
</tr>
<tr>
<td>Norway</td>
<td>128</td>
</tr>
<tr>
<td>LNG</td>
<td>86</td>
</tr>
<tr>
<td>BBL</td>
<td>41</td>
</tr>
<tr>
<td>IUK</td>
<td>72</td>
</tr>
</tbody>
</table>

Table 5.5: The total maximum capacity (in mcm) for each source of supply.

The 3 storage operators represent the aggregate of the 3 different types of storage facilities in the UK gas market, namely long- (LRS), medium- (MRS) and short-range-storage (SRS). Table 5.6 provides the initial (IntStor$_{so}$), maximum (MaxStor$_{so}$) and minimum storage (MinStor$_{so}$) levels for each of the 3 facilities. The initial amount of gas in storage is only a parameter for the first roll of the model. For subsequent rolls, IntStor$_{so}$ is determined using the amount of gas injected ($I_{so,d_i}$) to and withdrawn ($W_{so,d_i}$) from storage from the previous roll of the model, as described in Section 5.2.1.

<table>
<thead>
<tr>
<th></th>
<th>LRS</th>
<th>MRS</th>
<th>SRS</th>
</tr>
</thead>
<tbody>
<tr>
<td>IntStor$_{so}$</td>
<td>440</td>
<td>295</td>
<td>56</td>
</tr>
<tr>
<td>MaxStor$_{so}$</td>
<td>3300</td>
<td>810</td>
<td>180</td>
</tr>
<tr>
<td>MinStor$_{so}$</td>
<td>440</td>
<td>169</td>
<td>39</td>
</tr>
<tr>
<td>$I_{so}^{max}$</td>
<td>43</td>
<td>43</td>
<td>35</td>
</tr>
<tr>
<td>$W_{so}^{max}$</td>
<td>43</td>
<td>43</td>
<td>35</td>
</tr>
</tbody>
</table>

Table 5.6: Parameters associated with long-, medium- and short-range storage (in mcm) for ROM.

The maximum storage levels, daily injection and daily withdrawals rates were obtained from the UK’s Department of Energy and Climate Change [1], while the initial levels of gas in storage were obtained from the actual levels on the 1st of April 2010, as recorded by the UK National Grid, [3]. While the theoretical minimum level of gas in storage is zero, actual stocks levels suggest that this is never the case. As a result the minimum levels of gas in storage were estimated from the minimum actual storage levels observed in the UK gas market from the 1st of April 2010 to 31st of March 2011. This data was again obtained from the UK National Grid, [3]. Table 5.6 also displays the daily maximum injection ($I_{so}^{max}$) and withdrawal rates ($W_{so}^{max}$), while Table 5.7 presents the long- and short-run unit cost of storage injection and withdrawal, $a_{so,d}$ and $b_{so,d}$ respectively. Short-run costs are the storage costs associated with the first 26 days of ROM, while
the long-run costs are those associated with all other days. As before, the units of these costs are arbitrary.

<table>
<thead>
<tr>
<th></th>
<th>LRS</th>
<th>MRS</th>
<th>SRS</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a_{so,d}$ (Short-run)</td>
<td>0.069</td>
<td>0.069</td>
<td>2</td>
</tr>
<tr>
<td>$a_{so,d}$ (Long-run)</td>
<td>0.664</td>
<td>0.664</td>
<td>2</td>
</tr>
<tr>
<td>$b_{so,d}$ (Short-run)</td>
<td>0.064</td>
<td>0.019</td>
<td>2</td>
</tr>
<tr>
<td>$b_{so,d}$ (Long-run)</td>
<td>0.019</td>
<td>0.664</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 5.7: Short- and long-run cost associated with long-, medium- and short-range storage for ROM.

When $d \leq 26$ the injection and withdrawal costs take the short-run values, while when $d > 26$ (called the lag) they take the long-run costs. The long- and short-run costs for MRS encourage gas to be injected quickly and withdrawn quickly, as the short-run costs are cheaper than the long-run costs. As mentioned in Chapter 2, MRS is used to capture short term (i.e., weekly/monthly) variations in demand and prices in the UK gas market. The long- and short-run costs for MRS encourage this behaviour. Similar to MRS, the long-run and short-run injection costs for LRS also encourages gas to be injected quickly. In contrast, the long-run and short-run withdrawal costs for LRS discourage this behaviour, as the long-run costs are cheaper than the short-run costs. These costs allow LRS to capture the seasonal variation of natural gas demand as discussed in Chapter 2. For SRS, the long- and short-run costs are the same. Section 5.3.1 details how the parameters in Table 5.7 as well as the short-run lag of 26 were obtained by measuring the goodness of fit of the results they produce from ROM against actual data. It is also proven in Section 5.5 that these costs can be scaled without affecting the optimal flows of the model.

The rest of this section qualitatively compares the outputs of ROM with actual flows and prices. Appendix C.1 contains the GAMS code associated with these results. As described in the previous section, the outputs of the model were obtained from the production rates ($Q_{p,d_i}$), injections rates ($I_{so,d_i}$) and withdrawal rates ($W_{so,d_i}$) of the scenario-independent first days, which were obtained from each roll of the model. Figure 5.3 displays the actual demand profile for the UK gas market for the year beginning on the 1st of April 2010. It indicates how the different sources of supply meet demand for each day of the year. The data for this graph was obtained from the UK National Grid’s website, [3]. Figure 5.4 shows a similar plot, but for the flows produced by ROM.
5.3 ROM fitted to the year beginning in April 2010

Figure 5.3: Actual demand profile from the UK gas market starting on the 1st of April 2010.

Figure 5.4: Demand profile obtained from ROM for the UK gas market starting on the 1st of April 2010.
### 5. ROLLING OPTIMISATION MODEL

<table>
<thead>
<tr>
<th>Source of supply</th>
<th>Actual</th>
<th>ROM</th>
</tr>
</thead>
<tbody>
<tr>
<td>UKCS</td>
<td>152.6</td>
<td>152.3</td>
</tr>
<tr>
<td>Norway</td>
<td>47.0</td>
<td>47.2</td>
</tr>
<tr>
<td>BBL</td>
<td>19.6</td>
<td>19.2</td>
</tr>
<tr>
<td>IUK</td>
<td>3.0</td>
<td>1.9</td>
</tr>
<tr>
<td>LNG</td>
<td>56.0</td>
<td>58.4</td>
</tr>
<tr>
<td>SRS</td>
<td>0.1</td>
<td>0.0</td>
</tr>
<tr>
<td>MRS</td>
<td>6.4</td>
<td>5.2</td>
</tr>
<tr>
<td>LRS</td>
<td>8.7</td>
<td>6.9</td>
</tr>
</tbody>
</table>

Table 5.8: Average daily flows obtained from ROM (in mcm) as well as actual daily average flows for the UK gas market for year beginning on the 1st of April 2010.

Figures 5.3 and 5.4 plus Table 5.8 indicate that the order and relative size of the actual flows obtained from the ROM are qualitatively similar to observed data. The largest source of supply is clearly the UKCS source for both the actual data and the results produced from ROM. Norwegian and LNG imports are the next two largest sources of supply and are at roughly the same level while, imports through the BBL and IUK pipelines are the smallest sources of supply. In terms of storage, the actual data and ROM also both indicate that LRS is the largest storage supplier, while SRS is the smallest.

![Figure 5.5](image.png)

Figure 5.5: Actual and calibrated ROM SAPs starting on the 1st of April 2010.

---

1The values for UKCS exclude gas that is injected to storage.
5.3 ROM fitted to the year beginning in April 2010

Figure 5.5 displays the daily actual SAP, as well as SAPs produced by ROM. In Section 5.5 it is that proven that the costs of ROM (and hence \( \lambda_{\text{demand}d_1} \)) can be both scaled and shifted without affecting the optimal flows of the model. As a result, the modelled prices shown in Figure 5.5 were obtained using the following formula:

\[
CalSAP = \alpha \lambda_{\text{demand}d_1} - \beta, \quad \forall r,
\]

where \( \alpha \) is the positive scaling parameter and \( \beta \) is the shifting parameter. The values \( \alpha = 6.878 \) and \( \beta = 359.052 \) were obtained following the minimisation of the \( L_2 \)-norm of the difference between actual SAPs and SAPs produced by ROM. Figure 5.5 indicates that the calibrated prices provide a reasonable fit to the actual data. It also shows that the prices produced by ROM are highly seasonal. The reason for this is that these modelled SAPs are obtained from Lagrange multipliers associated with demand constraints. Thus, when demand is high (low) in the winter (summer), so is the predicted SAP. In reality, other factors affect natural gas prices in the UK. For example, many gas pricing contracts in the UK and particularly in Europe are indexed linked to oil prices, 16. Thus, when oil prices increase or decrease significantly so do gas prices.

Figures 5.6 - 5.8 display the actual and simulated amount of gas in storage for long-, medium- and short-range-storage, respectively, for the year beginning on the 1st of April 2010. As above, the actual data was obtained from the UK National Grid 3, while the simulated results were obtained from the withdrawal \( W_{so,d_1} \) and injection \( I_{so,d_1} \) rates from the scenario-independent first days from each roll of ROM.

![Graph of Actual and Simulated LRS]

Figure 5.6: Actual and simulated amount of LRS starting on the 1st of April 2010.
Figure 5.7: Actual and simulated amount of MRS starting on the 1st of April 2010.

Figure 5.8: Actual and simulated amount of SRS starting on the 1st of April 2010.

Figure 5.6 indicates that the amount of gas in LRS, as produced by ROM, is qualitatively similar to the actual amount. The upward and downward slopes of both time series in the plot...
are particularly similar, which suggests that the withdrawal and injection rates of the model are correct. The time when LRS starts injecting and withdrawing is also almost identical for both sets of time series.

Figure 5.7 shows the actual and simulated amount of gas in MRS. It too indicates that the results obtained from ROM provide a reasonably good fit to the actual data, as it captures the weekly and monthly fluctuations in the amount of gas in MRS. Figure 5.8 displays the actual and simulated amount of gas in SRS. Both time series indicate relatively little flows into and out of SRS for the year starting in April 2010. This is because, in May 2011, the Glenmavis and Partington short-range storage facilities stopped offering commercial services, [7]. This left only one remaining SRS facility (Avonmouth). As a result, SRS has become a minor part of the UK gas market and is of minimal interest to this project. At this point, it must also be noted that the vertical scale of Figure 5.8 is relatively small in comparison to that of Figure 5.6. Hence, the importance of SRS is relatively minor compared with that of LRS.

To summarise this section, ROM was applied and fitted to the UK gas market for data starting on the 1st of April 2010. Figures 5.3 - 5.8 show that the flows and prices produced by the model fit reasonably well to actual data. The parameters and actual data used in this section are of particular interest to this thesis, as more detailed analysis (Sections 5.6 and Chapter 6) are based on them. In the Section 5.4 below, the model is tested using actual data for the year starting on the 1st of April 2011.

5.3.1 Parameter estimation using simulated annealing

As mentioned above, the parameters presented for ROM were obtained by measuring the goodness-of-fit of model predictions to actual data. This was done using simulated annealing, which is an optimisation algorithm that locates a good approximation to the global optimum given a large search space [58, 59]. The optimum of interest here is the minimum error between actual data and the results obtained from ROM. This error is detailed in the following equation:

$$Error = \sum_{l=1}^{L} \sqrt{\sum_{r=1}^{R} (Actual_{lr} - Simulated_{lr})^2 / R},$$

(5.14)

where $Actual_{lr}$ and $Simulated_{lr}$ are the actual and simulated flows from source $l$ on roll $r$ respectively. The $L = 11$ different sources of supply appearing in the sum of equation (5.14) were UKCS, Norway, LNG, BBL, IUK, as well as the injections and withdrawals into and out of LRS, MRS and SRS. The simulated annealing algorithm allows for uphill movements in the search for a global minimum and hence prevents the search from being trapped in a local minimum. It is described by Kirkpatrick et al. [58] following earlier work by Metropolis et al. in [59]. Simulated annealing is analogous to the physical annealing of glass or metals, whereby the atoms of the material are displaced randomly by an initially high temperature initially. The temperature is then slowly reduced, allowing the material to find an equilibrium. Details of the algorithm are provided...
5. ROLLING OPTIMISATION MODEL

in Appendix C.2.

At each stage of simulated annealing, the parameters of interest were perturbed. The parameters of ROM perturbed were the size and costs associated with each tranche of the different sources of supply (Table 5.4), the costs associated with storage (Table 5.7) and the lag of 26 associated with the short-run costs of storage. The total maximum capacities of the different sources of supply (Table 5.5), as well as the capacities associated with storage (Table 5.6) were not perturbed as they are fixed values obtained from either the UK National Grid or the UK’s Department of Energy and Climate Change.

An initial guess at parameter values was based on information from industrial partner Bord Gáis Energy. Using simulated annealing, and some manual intervention, the parameters described in Section 5.3 were located. Further attempts to reduce the error metric, using simulated annealing, were unsuccessful. This suggests that the parameters detailed in Section 5.3 provide a good approximation to the minimum distance (as measured by equation (5.14)) between the actual flows observed in the UK gas market and those simulated by ROM.

5.4 ROM tested for the year beginning in April 2011

In Section 5.3, ROM was applied to data for the UK gas market for the year starting on the 1st of April 2010. This analysis established the parameters of the model. In this section, ROM is tested with similar parameters using data for the year starting on the 1st of April 2011. As with section 5.3, each roll of the model contained 3 storage facilities ($SO = 3$), 3 demand scenarios ($n = 3$) and was run over a horizon of $D = 365$ days. The number of rolls (or optimisations) was 365. In contrast to Section 5.3 however, the model was formulated with 17 sources of supply ($P = 17$), a decrease of one. This loss of a source of supply takes into account the decreased level of UKCS supply for this time period as detailed below.

For each roll of the model, the sets of three demand scenarios were generated using the stochastic process for demand described in Section 5.4. Actual demand and SND for the year starting on the 1st of April 2011 were used for the first roll. For the second roll of the model, the stochastic process for demand moved forward one day, i.e., actual demand and SND for the year starting on the 2nd of April 2011 was used to develop the scenarios. For subsequent rolls, SND moved forward in a similar pattern. As in Section 5.3, the choice of 3 for the size of the scenario sets was based on the analysis in Chapter 6. In Section 5.6, however, the effect of changing the number of scenarios used in the model is examined in detail. As each of the 3 scenarios were again randomly generated for each roll of the model, they were assigned equal probabilities, i.e., $Prob^s = \frac{1}{3}$ for $s = 1, 2, 3$.

As in Section 5.3, the 17 sources represent the different sources of supply in the UK gas market, broken up into multiple tranches representing varying costs. The 3 storage operators represent the 3 different types of storage facilities in the UK, namely long-, medium- and short-range-storage. Table 5.9 shows the costs and maximum capacities associated with each tranche, while the total
maximum capacities of the five different sources of supply (i.e., the sum of the tranches’ capacities) are given in Table 5.10.

<table>
<thead>
<tr>
<th>Tranche</th>
<th>$c_p$</th>
<th>$Q^\text{max}_p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>UKCS₁</td>
<td>58.464</td>
<td>48</td>
</tr>
<tr>
<td>UKCS₂</td>
<td>57.611</td>
<td>61</td>
</tr>
<tr>
<td>UKCS₃</td>
<td>56.552</td>
<td>48</td>
</tr>
<tr>
<td>Norway₁</td>
<td>60.985</td>
<td>21</td>
</tr>
<tr>
<td>Norway₂</td>
<td>64.322</td>
<td>54</td>
</tr>
<tr>
<td>Norway₃</td>
<td>48.375</td>
<td>40</td>
</tr>
<tr>
<td>Norway₄</td>
<td>58.637</td>
<td>30</td>
</tr>
<tr>
<td>LNG₁</td>
<td>55.242</td>
<td>38</td>
</tr>
<tr>
<td>LNG₂</td>
<td>58.865</td>
<td>36</td>
</tr>
<tr>
<td>LNG₃</td>
<td>71.458</td>
<td>28</td>
</tr>
<tr>
<td>BBL₁</td>
<td>64.189</td>
<td>11</td>
</tr>
<tr>
<td>BBL₂</td>
<td>55.084</td>
<td>16</td>
</tr>
<tr>
<td>BBL₃</td>
<td>59.352</td>
<td>14</td>
</tr>
<tr>
<td>IUK₁</td>
<td>61.626</td>
<td>11</td>
</tr>
<tr>
<td>IUK₂</td>
<td>69.543</td>
<td>16</td>
</tr>
<tr>
<td>IUK₃</td>
<td>70.977</td>
<td>27</td>
</tr>
<tr>
<td>IUK₄</td>
<td>60.336</td>
<td>18</td>
</tr>
</tbody>
</table>

Table 5.9: The costs and maximum capacities (in mcm) of the different tranches for each source of supply.

<table>
<thead>
<tr>
<th>Source of supply</th>
<th>Total maximum capacity</th>
</tr>
</thead>
<tbody>
<tr>
<td>UKCS</td>
<td>157</td>
</tr>
<tr>
<td>Norway</td>
<td>145</td>
</tr>
<tr>
<td>LNG</td>
<td>102</td>
</tr>
<tr>
<td>BBL</td>
<td>41</td>
</tr>
<tr>
<td>IUK</td>
<td>72</td>
</tr>
</tbody>
</table>

Table 5.10: The total maximum capacity (in mcm) for each source of supply.

The capacities of the BBL and IUK pipelines, for the year beginning in April 2011, are the same as in Table 5.5 and are in line with figures supplied by the UK’s Department for Energy and Climate Change, [1, 7]. The Norwegian total maximum capacity has increased to 145 mcm, again in accordance with values supplied by the UK’s Department for Energy and Climate Change, [7]. In contrast, the total maximum capacity for the UKCS has decreased to 157 mcm. This is line with figures supplied by the UK National Grid and the general downward trend seen in UKCS supplies since 2000, [2, 16]. The LNG capacity has increased to 102 mcm. This increase takes into account
the improved LNG infrastructure in the UK and is again in accordance with figures supplied by the UK National Grid. As explained in the previous section, the total maximum capacity of LNG was assumed to be at only 70% of its true maximum capacity.

<table>
<thead>
<tr>
<th></th>
<th>LRS</th>
<th>MRS</th>
<th>SRS</th>
</tr>
</thead>
<tbody>
<tr>
<td>$IntStor_{so}$</td>
<td>1040</td>
<td>440</td>
<td>39</td>
</tr>
<tr>
<td>$MaxStor_{so}$</td>
<td>3500</td>
<td>850</td>
<td>80</td>
</tr>
<tr>
<td>$MinStor_{so}$</td>
<td>440</td>
<td>169</td>
<td>15</td>
</tr>
<tr>
<td>$I_{so}^{max}$</td>
<td>43</td>
<td>43</td>
<td>13</td>
</tr>
<tr>
<td>$W_{so}^{max}$</td>
<td>43</td>
<td>43</td>
<td>13</td>
</tr>
</tbody>
</table>

Table 5.11: Parameters associated with long-, medium- and short-range storage (in mcm) for ROM.

Table 5.11 provides the initial ($IntStor_{so}$), maximum ($MaxStor_{so}$) and minimum storage ($MinStor_{so}$) levels for each of the 3 storage facilities. It also displays the daily maximum injection ($I_{so}^{max}$) and withdrawal rates ($W_{so}^{max}$). As previously, the initial and minimum amount of gas in storage were obtained from data made available from the UK National Grid. The maximum storage levels, the maximum daily injection rate and the maximum daily withdrawal rate are updated in accordance with the UK’s Department of Energy Climate Change. Most noticeably, they were updated to take into account the de-commissioning of the SRS facilities, Glenmavis and Partington. The costs associated with injection and withdrawal to and from storage remained the same as detailed in Table 5.7.

As described in the previous section, the outputs of the model were obtained from the production rates ($Q_{p,d}$), injection rates ($I_{so,d}$) and withdrawal rates ($W_{so,d}$) of the scenario-independent first days, which were obtained from each roll of the model. Figure 5.9 displays the actual demand profile for the UK gas market for the year beginning on the 1st of April 2011. The data for this graph was obtained from the UK National Grid’s website. Figure 5.10 shows a similar plot, but for the flows produced by ROM. Table 5.12 shows the actual and simulated daily average flows of gas in the UK gas market over the same period.
5.4 ROM tested for the year beginning in April 2011

Figure 5.9: Actual demand profile from the UK gas market starting on the 1st of April 2011.

Figure 5.10: Demand profile obtained from ROM for the UK gas market starting on the 1st of April 2011.
Table 5.12: Average daily flows obtained from ROM (in mcm) as well as actual daily average flows for the UK gas market for year beginning on the 1st of April 2011.

Figures 5.9 and 5.10 plus Table 5.8 indicate that the order and relative size of the actual flows obtained from ROM are again qualitatively similar to observed data. The largest source of supply is again the UKCS for both the actual data and the results produced from ROM. The second largest source of supply is now clearly LNG, while Norwegian supplies are the third largest. Imports through the BBL and IUK pipelines are the smallest sources of supply. In terms of storage, the results obtained from ROM indicate that MRS is the largest storage supplier, while SRS is the smallest. This is in contrast to Section 5.3 where LRS was the largest storage supplier.

Figure 5.11: Actual and calibrated predicted SAPs starting on the 1st of April 2011. These prices were calibrated using equation (5.13) with $\alpha = 6.48$ and $\beta = 323.324$.
Figure 5.11 displays the daily actual SAPs, as well as SAPs produced by ROM. In a similar manner to the previous section, the SAPs produced from ROM were calibrated using equation 5.13. The values of $\alpha = 6.48$ and $\beta = 323.324$ were again obtained following the minimisation of the $L_2$-norm of the difference between the actual SAPs and SAPs produced by ROM. Figure 5.11 indicates that the calibrated prices provide a reasonable fit to the actual data. However, the prices produced by ROM fail to capture the magnitude of some of the price spikes, in particular the large spike seen in February for the actual data.

Figure 5.12: Actual and simulated amount of LRS starting on the 1st of April 2011.
Figures 5.12 - 5.14 display the actual amount of gas in each of the three different storage facilities, as well as the amount in storage as predicted by ROM. Figure 5.12 shows that the simulated
results are similar to the actual data for LRS. As the upward and downward slopes are again relatively similar it indicates that the injection and withdrawal rates of ROM are correct. The time of withdrawals on both the actual data and simulated results are also almost identical. In particular, ROM successfully models the time when withdrawals stop in winter. However, ROM underestimates the amount of withdrawals from LRS. Figure 5.13 shows the actual and simulated amount of gas in MRS. It too indicates that the results obtained from ROM are qualitatively similar to the actual data, as it captures the weekly and monthly fluctuations in the amount of gas in MRS.

Figure 5.14 displays the actual and simulated amount of gas in SRS. As with the actual data for the year starting on the 1st of April 2010, the actual data in this graph indicates relatively few injections to SRS with some withdrawals. ROM successfully models this lack of injections into SRS, but does not capture the withdrawals from SRS. In fact, ROM simulates no flows into or out of SRS for the year beginning in April 2011. As mentioned in the previous section, the Glenmavis and Partington short-range storage facilities stopped offering commercial services in May 2011. As a result, SRS has become a minor part of the UK gas market and is of minimal interest to this project.

To summarise this section, Figures 5.9 - 5.14 show that the results produced by ROM provide a reasonably good fit to actual data for the year starting on the 1st of April 2011. In Section 5.3 the parameters of the model were fitted for the year beginning on the 1st of April 2010. In this section similar parameters were used again, thus showing the robustness of ROM to changes in the data. In appendix C.3 similar outputs and analysis are shown again for the year starting on the 1st of October 2010. The results provided in this appendix demonstrate that ROM is not dependent on the model starting on the 1st of April.

5.5 Varying costs in ROM

In Section 5.3, it was stated that the costs associated with ROM can be varied, in certain ways, without affecting the optimal flows of the model (i.e., $Q_{p,d}$, $I_{so,d}$, $W_{so,d}$, $Q_{p,d}$, $I_{so,d}$, and $W_{so,d}$). The variation can occur in two different ways, both of which are considered in this section. Firstly, the costs associated with each producer ($c_p$), as well as the costs associated with injections ($a_{so,d}$) and withdrawals ($b_{so,d}$) from storage, can be scaled by multiplying the objective function in ROM (equation (5.1)) by a positive constant $\alpha$. The second way this can be done is by adding another positive constant $\beta$ to the each of the production costs, $c_p$. Both proofs involve the comparison of KKT conditions. See Section 4.1 for a detailed description of KKT conditions. Examples of these production and storage costs can be seen in Tables 5.4 and 5.7 respectively.

5.5.1 KKT conditions of ROM

In order to prove that the costs in ROM can be varied without affecting the optimal flows of the model, its KKT conditions need to be stated. The stationary KKT conditions of the model
5. ROLLING OPTIMISATION MODEL

described in equations (5.1) - (5.11) are

\[
\frac{\partial L}{\partial Q_{p,d_1}}: c_p - \lambda_{Demand_{d_1}} + \lambda Q_{p,d_1} - \lambda Q_{p,d_1} = 0, \quad (5.15)
\]

\[
\frac{\partial L}{\partial Q_{p,d}}: \text{Prob}^s c_p - \lambda_{Demand_{d}}^s + \lambda Q_{p,d}^s - \lambda Q_{p,d}^s = 0, \quad (5.16)
\]

\[
\frac{\partial L}{\partial I_{so,d_1}}: a_{so,d_1} + \lambda_{Demand_{d_1}} + \lambda I_{so,d_1} - \lambda I_{so,d_1} + \lambda Stor_{so,d_1} - \lambda Stor_{so,d_1} +
\sum_{s=1}^{S} \sum_{e=1}^{D} (\lambda_{Stor_{so,e}}^s - \lambda_{Stor_{so,e}}^s) = 0,
\quad (5.17)
\]

\[
\frac{\partial L}{\partial I_{so,d}}: \text{Prob}^s a_{so,d} + \lambda_{Demand_{d}}^s + \lambda I_{so,d}^s - \lambda I_{so,d}^s + \sum_{e=d}^{D} (\lambda_{Stor_{so,e}}^s - \lambda_{Stor_{so,e}}^s) = 0, \quad (5.18)
\]

\[
\frac{\partial L}{\partial W_{so,d_1}}: b_{so,d_1} - \lambda_{Demand_{d_1}} + \lambda W_{so,d_1} - \lambda W_{so,d_1} - \lambda Stor_{so,d_1} + \lambda Stor_{so,d_1} -
\sum_{s=1}^{S} \sum_{e=1}^{D} (\lambda_{Stor_{so,e}}^s - \lambda_{Stor_{so,e}}^s) = 0,
\quad (5.19)
\]

\[
\frac{\partial L}{\partial W_{so,d}}: \text{Prob}^s b_{so,d} - \lambda_{Demand_{d}}^s + \lambda W_{so,d}^s - \lambda W_{so,d}^s - \sum_{e=d}^{D} (\lambda_{Stor_{so,e}}^s - \lambda_{Stor_{so,e}}^s) = 0, \quad (5.20)
\]

where \( \lambda \) is the Lagrange multiplier associated with the upper bound constraint, while \( \overline{\lambda} \) is associated with the lower bound constraint. The complementarity KKT conditions are:

\[
\lambda Q_{p,d_1} (Q_{p,d_1} - Q_{p}^{max}) = 0, \quad (5.21)
\]

\[
\overline{\lambda Q_{p,d_1}} (Q_{p,d_1}) = 0, \quad (5.22)
\]

\[
\lambda Q_{p,d}^s (Q_{p,d}^s - Q_{p}^{max}) = 0, \quad (5.23)
\]

\[
\overline{\lambda Q_{p,d}^s} (Q_{p,d}^s) = 0, \quad (5.24)
\]

\[
\lambda I_{so,d_1} (I_{so,d_1} - I_{so}^{max}) = 0, \quad (5.25)
\]

90
5.5 Varying costs in ROM

\[ \lambda_{I_{so,d_1}}(I_{so,d_1}) = 0, \]  
(5.26)

\[ \lambda_{I_{so,d}}(I_{so,d}^s - I_{so}^{max}) = 0, \]  
(5.27)

\[ \lambda_{I_{so,d}}^{\ast}(I_{so,d}^s) = 0, \]  
(5.28)

\[ \lambda_{W_{so,d_1}}(W_{so,d_1} - W_{so}^{max}) = 0, \]  
(5.29)

\[ \lambda_{W_{so,d_1}}^{\ast}(W_{so,d_1}) = 0, \]  
(5.30)

\[ \lambda_{W_{so,d}}^{\ast}(W_{so,d}^s - W_{so}^{max}) = 0, \]  
(5.31)

\[ \lambda_{W_{so,d}}^{\ast}(W_{so,d}^s) = 0, \]  
(5.32)

\[ \lambda_{Stor,d_1}(IntStor_{so} + I_{so,d_1} - W_{so,d_1} - MaxStor_{so}) = 0, \]  
(5.33)

\[ \lambda_{Stor,d}(MinStor_{so} - IntStor_{so} - I_{so,d_1} + W_{so,d_1}) = 0, \]  
(5.34)

\[ \lambda_{Stor,d}^{\ast}(IntStor_{so} + I_{so,d_1} - W_{so,d_1} + \sum_{e=1}^{e=d} (I_{so,e}^s - W_{so,e}^s) - MaxStor_{so}) = 0, \]  
(5.35)

\[ \lambda_{Stor,d}^{\ast}(MinStor_{so} - IntStor_{so} - I_{so,d_1} + W_{so,d_1} - \sum_{e=1}^{e=d} (I_{so,e}^s - W_{so,e}^s)) = 0. \]  
(5.36)

Each of the Lagrange multipliers associated with inequality constraints are required to be non-negative by the KKT conditions. The only Lagrange multipliers not required to be non-negative are \( \lambda_{Demand,d_1} \) and \( \lambda_{Demand,d} \) as they are associated with equality constraints, equations (5.2) and (5.3) respectively. The KKT conditions also require that constraints (5.2)-(5.11) are all satisfied. As this model is a linear program, the KKT conditions are both necessary and sufficient for optimality, [29].
5. ROLLING OPTIMISATION MODEL

5.5.2 Scaling the costs in ROM

Consider the following linear program with $I$ inequality constraints and $J$ equality constraints

$$\min f(x) \quad (5.37)$$

subject to:

$$g_i(x) \leq 0, \quad (5.38)$$
$$h_j(x) = 0. \quad (5.39)$$

The KKT conditions for this problem are

$$\nabla f(x) + \sum_{i=1}^{I} \lambda_i \nabla g_i(x) + \sum_{j=1}^{J} \lambda_j \nabla g_j(x) = 0, \quad (5.40)$$
$$g_i(x) \leq 0, \quad \forall i, \quad (5.41)$$
$$h_j(x) = 0, \quad \forall j, \quad (5.42)$$
$$\lambda_i \geq 0, \quad \forall i, \quad (5.43)$$
$$\lambda_i g_i(x) = 0 \quad \forall i. \quad (5.44)$$

As this is a linear program, the KKT conditions are both necessary and sufficient for optimality, [29]. Let $x^*$ and $\lambda^*$ be the vectors of optimal primal and dual variables, respectively, for this problem. Now consider the same constraints, but with the objective function multiplied by a positive constant $\alpha$. The problem becomes

$$\min \alpha f(x) \quad (5.45)$$

subject again to conditions (5.38) and (5.39). The KKT conditions of this problem are again equations (5.41) - (5.44) with equation (5.40) being replaced by

$$\alpha \nabla f(x) + \sum_{i=1}^{I} \lambda_i \nabla g_i(x) + \sum_{j=1}^{J} \lambda_j \nabla g_j(x) = 0. \quad (5.46)$$
By letting \( x = x^* \) and \( \lambda = \alpha \lambda^* \), where \( x^* \) and \( \lambda^* \) are as defined above, the KKT conditions of the problem with \( \alpha \) become

\[
g_i(x^*) \leq 0, \quad \forall i, \quad (5.47)
\]
\[
h_j(x^*) = 0, \quad \forall j, \quad (5.48)
\]
\[
\alpha \nabla f(x^*) + \alpha \sum_{i=1}^{I} \lambda_i^* \nabla g_i(x^*) + \alpha \sum_{j=1}^{J} \lambda_j^* \nabla g_j(x^*) = 0, \quad (5.49)
\]
\[
\alpha \lambda_i^* \geq 0, \quad \forall i, \quad (5.50)
\]
\[
\alpha \lambda_i^* g_i(x^*) = 0 \quad \forall i. \quad (5.51)
\]

As each of the \( \alpha \) terms, can be cancelled out of this system, these KKT conditions are the same as equations (5.40) - (5.44) above and are thus satisfied. This means that the vector \( x^* \) is the optimal primal solution to both problems. As the rolling optimisation problem is a linear program, this proof can be applied to it. This means that the optimal flows of ROM (\( Q_{p,d_1}, Q_{s,d}, I_{so,d_1}, I_{so,d}, W_{so,d_1}, \text{ and } W_{so,d} \)) are also optimal when each of the costs of the model (\( c_p, a_{so,d}, b_{so,d} \)) are multiplied by a positive constant. Multiplying the objective function of ROM by a positive constant \( \alpha \) is equivalent to multiplying each of the costs of the model by the same constant. When this is done, the Lagrange multipliers of ROM and hence the prices of the model (\( \lambda_{Demand_{d_1}} \) and \( \lambda_{Demand_{d_2}} \)) become scaled by the same constant \( \alpha \).

### 5.5.3 Shifting the production costs in ROM

As mentioned above, the second way the costs of this model can be varied is by adding a positive constant \( \beta \) to the costs of each of the different sources of supply (\( c_p \)). This changes the objective function of ROM, equation (5.1), as follows:

\[
\min \sum_{p=1}^{P} ((\beta+c_p)Q_{p,d_1}) + \sum_{so=1}^{SO} (a_{so}I_{so,d_1} + b_{so}W_{so,d_1}) +
\]
\[
+ ES[ \sum_{d=d_1+1}^{D} (\sum_{p=1}^{P} ((\beta+c_p)Q^s_{p,d}) + \sum_{so=1}^{SO} (a_{so}I^s_{so,d} + b_{so}W^s_{so,d}))]
\]

The constraints of the model (equations (5.2) - (5.11)) remain the same, while the stationary KKT conditions become

\[
\frac{\partial L}{\partial Q_{p,d_1}} : c_p + \beta - \lambda_{Demand_{d_1}} + \lambda Q_{p,d_1} - \lambda Q_{p,d_1} = 0, \quad (5.52)
\]
\[
\frac{\partial L}{\partial Q^s_{p,d}} : Prob^s c_p + Prob^s \beta - \lambda_{Demand^s_{d}} + \lambda Q^s_{p,d} - \lambda Q^s_{p,d} = 0, \quad (5.53)
\]
Note that the Lagrange multipliers associated with the demand constraints and the storage constraints for the last day (day $D$) of the model have modified labels, denoted with the symbol $\sim$. As the constraints remain unchanged, all the complementarity KKT conditions (equations (5.22) - (5.36)) also remain unchanged, except for the complementarity condition that includes $\lambda_{Stor_{so,D}}$. This becomes

$$\lambda_{Stor_{so,D}}(MinStor_{so} - IntStor_{so} - I_{so,d1} + W_{so,d1} - \sum_{e=1}^{D} (I_{so,e} - W_{so,e})) = 0,$$  \hspace{1cm} (5.58)
5.5 Varying costs in ROM

with the following adjustments:

\[
\lambda_{\text{Demand}_{d}}^\star + \beta = \tilde{\lambda}_{\text{Demand}_{d}}, \tag{5.59}
\]
\[
\lambda_{\text{Demand}_{d}}^* + \text{Prob}^\beta \beta = \tilde{\lambda}_{\text{Demand}_{d}}, \tag{5.60}
\]
\[
\lambda_{\text{Stor}_{s,d}}^* + \text{Prob}^\beta \beta = \tilde{\lambda}_{\text{Stor}_{s,d}}, \tag{5.61}
\]

where \(\tilde{\lambda}_{\text{Demand}_{d}}, \tilde{\lambda}_{\text{Demand}_{d}}^*\) and \(\tilde{\lambda}_{\text{Stor}_{s,d}}^*\) are Lagrange Multipliers associated with the modified model as explained previously. These adjusted optimal solutions can be shown to satisfy the KKT conditions of the model with the modified costs. Firstly consider the constraints of both models. As the primal variables and the constraints have remained unchanged, the optimal solutions defined above satisfy the constraints of the \(\beta\) model. Now consider the stationary KKT conditions. By using the adjusted optimal solutions defined in equations (5.59) - (5.61), each of the \(\beta\) terms in equations (5.52) - (5.57) cancel out. This means that these stationary KKT conditions become identical to those of the original Rolling Optimisation Model (equations (5.15) - (5.20)) and are thus satisfied. This means that the adjusted optimal solutions defined above satisfy the KKT conditions of the modified model. Now consider the complementarity KKT conditions using the adjusted optimal solutions, which are

\[
\lambda_{Q_{p,d1}}^*(Q_{p,d1}^* - Q_{p}^{\text{max}}) = 0, \tag{5.62}
\]
\[
\overline{\lambda}_{Q_{p,d1}}^*(Q_{p,d1}^*) = 0, \tag{5.63}
\]
\[
\lambda_{Q_{p,d1}}^*(Q_{p,d1}^* - Q_{p}^{\text{max}}) = 0, \tag{5.64}
\]
\[
\overline{\lambda}_{Q_{p,d1}}^*(Q_{p,d1}^*) = 0, \tag{5.65}
\]
\[
\lambda_{I_{so,d1}}^*(I_{so,d1}^* - I_{so}^{\text{max}}) = 0, \tag{5.66}
\]
\[
\overline{\lambda}_{I_{so,d1}}^*(I_{so,d1}^*) = 0, \tag{5.67}
\]
\[
\lambda_{I_{so,d1}}^*(I_{so,d1}^* - I_{so}^{\text{max}}) = 0, \tag{5.68}
\]
\[
\overline{\lambda}_{I_{so,d1}}^*(I_{so,d1}^*) = 0, \tag{5.69}
\]
5. ROLLING OPTIMISATION MODEL

\[ \lambda_{W_{so,d}}(W_{so,d}^* - W_{so}^{max}) = 0, \]  
(5.70)

\[ \lambda_{\bar{W}_{so,d}}(W_{so,d}^*) = 0, \]  
(5.71)

\[ \lambda_{W_{so,d}}(W_{so,d}^* - W_{so}^{max}) = 0, \]  
(5.72)

\[ \lambda_{\bar{W}_{so,d}}(W_{so,d}^*) = 0, \]  
(5.73)

\[ \lambda_{Stor^*_d}(IntStor_{so} + I_{so,d_1}^* - W_{so,d_1}^* - MaxStor_{so}) = 0, \]  
(5.74)

\[ \lambda_{\bar{Stor}^*_d}(MinStor_{so} - IntStor_{so} - I_{so,d_1}^* + W_{so,d_1}^*) = 0, \]  
(5.75)

\[ \lambda_{Stor^*_d}(IntStor_{so} + I_{so,d_1}^* - W_{so,d_1}^* + \sum_{e=1}^{e=d}(I_{so,e}^* - W_{so,e}^{*,s}) - MaxStor_{so}) = 0, \]  
(5.76)

\[ \lambda_{\bar{Stor}^*_d}(MinStor_{so} - IntStor_{so} - I_{so,d_1}^* + W_{so,d_1}^*) 
- \sum_{e=1}^{e=d}(I_{so,e}^* - W_{so,e}^{*,s})) = 0, \quad \forall d = 1, ..., D - 1, \]  
(5.77)

\[ (\lambda_{Stor^*_D} + Prob^* \beta)(MinStor_{so} - IntStor_{so} - I_{so,d_1}^* + W_{so,d_1}^*) - \sum_{e=1}^{e=d}(I_{so,e}^* - W_{so,e}^{*,s})) = 0. \]  
(5.78)

Equations (5.65) - (5.77) are all included in the original rolling optimisation problem and are thus satisfied. The only equation not included is equation (5.78). However, this condition is satisfied, because the constraint associated with it is always binding. This is because no storage facility ever contains more gas in storage on the final day \( D \) than the minimum amount, \( MinCap_{so} \). The proof of this, under some reasonable assumptions (with justification), can be seen through KKT analysis in Section 5.5.4. As a result, the complementarity KKT conditions of the \( \beta \) model are all satisfied. The final KKT conditions to consider are those that require the Lagrange multipliers associated with inequality constraints to be non-negative. Nearly all of the Lagrange multipliers satisfy this condition, as they are required by the original Rolling Optimisation Model to be non-negative. The only Lagrange multiplier associated with an inequality constraint not found in the original Rolling Optimisation Model is \( \lambda_{Stor^*_so,D} = \lambda_{Stor^*_so,D} + Prob^* \beta \). However, this Lagrange multiplier satisfies the non-negativity condition as \( \lambda_{Stor^*_so,D} \) is required to be non-negative by the original Rolling Optimisation Model plus \( \beta \) and \( Prob^* \) are positive.
This means that the adjusted optimal solutions defined above satisfy all the KKT conditions of
the model with $\beta$. As a result, these solutions are optimal for the modified model. As $Q_{p,d}, Q_{p,d}^s, I_{so,d}, I_{so,d}^s, W_{so,d},$ and $W_{so,d}^s$ were the same for both models, this means that the optimal flows for ROM are also optimal when a positive constant is added to the costs of each of the different sources of supply ($c_p$). When this is done, the Lagrange multipliers associated with the demand for scenario-independent first day of the model ($\lambda_{Demand_{d1}}$) also become shifted by the positive constant. These are the multipliers that are used for calibration with actual SAPs in Sections 5.3 and 5.4.

5.5.4 Proof that gas is emptied from storage

In this section, it is shown that for each roll of ROM, gas in storage is at its minimum level on the final day $D$, for each scenario $s$. This is done for two sets of assumptions. Both sets are reasonable for any natural gas market. If, however, one of the sets of assumptions does not hold, then the others will.

**Proof A**

The first set of assumptions are

1. For each scenario $s$, there is at least one supplier active on day $d,$
2. For each scenario $s$, there is at least one day in the model when there are some injections to storage facility $so$.

The first of these assumptions is reasonable, as otherwise demand would never be met. Similarly, the second assumption is also reasonable, because if there was never any injections to storage, then it would become pointless to include storage in the model.

Consider the first assumption that there is at least one active source of supply, i.e., $Q_{p,d}^s > 0$, for some $p \in P$. This means that the stationary KKT condition associated with this variable (equation (5.16)) becomes:

$$\frac{\partial L}{\partial Q_{p,d}^s} : Prob^s c_{p,d} - \lambda_{Demand_{d1}}^s + \lambda_{Q_{p,d}^s}^s = 0, \ \forall s.$$  \hspace{1cm} (5.79)

Note: $\lambda_{Q_{p,d}^s}^s = 0$ in order to satisfy the complementarity KKT condition, equation (5.24). Equation (5.79) requires that $\lambda_{Demand_{d1}}^s > 0$, as $Prob^s > 0$, $c_{p,d} > 0$ and $\lambda_{Q_{p,d}^s}^s \geq 0$.

Following this, now consider the second assumption that on some day $d' \in D$, under scenario $s$, there are some injections to storage facility $so$, i.e., $I_{so,d'}^s > 0$. Also assume that day $d'$ is the last day when there are injections to storage. This means that the stationary KKT condition for this
variable, using equation (5.18), becomes

\[ \frac{\partial L}{\partial I_{so,d}} : \text{Prob}_s a_{so,d} + \lambda_{Demand_{so,d}} + \lambda_{I_{so,d}} + \sum_{e=d'}^{D} (\lambda_{Stor_{so,e}} - \lambda_{Stor_{so,e}}) = 0, \quad \forall s, so. \] (5.80)

Again, \( \lambda_{I_{so,d}} = 0 \) in order to satisfy the KKT condition, equation (5.28). Now assume that not all of this gas is withdrawn from storage, which means that storage facility \( so \) is not at its minimum on day \( D \). This means equation (5.80) becomes

\[ \text{Prob}_s a_{so,d} + \lambda_{Demand_{so,d}} + \lambda_{I_{so,d}} + \sum_{e=d'}^{D} \lambda_{Stor_{so,e}} = 0, \quad \forall s, so. \] (5.81)

Here, \( \lambda_{Stor_{so,d}} = 0 \) in order to satisfy the complementarity KKT condition, equation (5.36), as there is always more than the minimum amount of gas in storage. As \( \text{Prob}_s > 0, a_{so,d'} > 0, \lambda_{Demand_{d'}} > 0 \) and \( \lambda_{Stor_{so,d'}} \geq 0 \), equation (5.81) can never be satisfied. This means it is infeasible to inject gas to storage facility \( so \) and not withdraw it all again. As a result, all gas in storage must be withdrawn until the minimum amount of gas in storage is reached. This holds under each scenario \( s \), for each roll of ROM\(^{1}\).

Proof B

The second set of assumptions are

1. For each scenario \( s \), there is at least one supplier active on day \( d \),
2. For each scenario \( s \), there is at least one day after the final day of injection, where each storage facility \( so \) is not withdrawing at its maximum rate.
3. The unit cost of withdrawing is always less than the production costs, i.e., \( b_{so} < c_{p} \forall so, p \).

The first of these assumptions is reasonable, as otherwise demand would never be met. The figures associated with storage in Chapter 2, Sections 5.3 and 5.4 and Appendix C.3 show that the second assumption is reasonable. The final assumption is reasonable as, otherwise storage would never be profitable.

Consider the first assumption that there is at least one active source of supply, i.e., \( Q^{s}_{p',d'} > 0 \), for some \( p' \in P \). This means that the stationary KKT condition associated with this variable (equation (5.16)) becomes:

\[ \frac{\partial L}{\partial Q^{s}_{p',d}} : \text{Prob}_s c_{p'} - \lambda_{Demand_{d'}} + \lambda_{Q^{s}_{p',d}} = 0, \quad \forall s. \] (5.82)

\(^{1}\)Note: this does not mean that storage must empty to the minimum amount for the entire Rolling Optimisation Model as the amount of gas in storage for the entire model is obtained from the flows of the scenario-independent day \( d_1 \) of each roll of the model.
Note: $\lambda_{Q^{s}_{p',d'}} = 0$ in order to satisfy the complementarity KKT condition, equation (5.24). Following this, now consider the second assumption that on some day $d' \in D$, after the final day of injections under scenario $s$, storage facility $so$ is not withdrawing at its maximum capacity, i.e., $W^{s}_{so,d'} < W^{Max}_{so}$. This means that the stationary KKT condition for this variable, using equation (5.20), becomes

$$\frac{\partial L}{\partial W^{s}_{so,d'}} : Prob^{s}b^{s}_{so,d'} - \lambda_{Demand^{s}_{d'}} - \lambda_{W^{s}_{so,d'}} - \sum_{e=d'}^{D} (\lambda^{s}_{Stor^{so,e}_{so}} - \lambda^{s}_{Stor^{so,e}_{so}}) = 0,$$ (5.83)

Here, $\lambda_{W^{s}_{so,d'}} = 0$ in order to satisfy the KKT condition, equation (5.31). Now assume that storage facility $so$ is not at its minimum on day $D$. This means equation (5.83) becomes

$$\frac{\partial L}{\partial W^{s}_{so,d'}} : Prob^{s}b^{s}_{so,d'} - \lambda_{Demand^{s}_{d'}} - \lambda_{W^{s}_{so,d'}} - \sum_{e=d'}^{D} \lambda^{s}_{Stor^{so,e}_{so}} = 0,$$ (5.84)

Here, $\lambda^{s}_{Stor^{so,e}_{so}} = 0$ for $e = d'..D$ in order to satisfy the complementarity KKT condition, equation (5.36). This is because if there is always more than the minimum amount of gas in storage on the final day $D$ and no more injections to storage on, or after, day $d'$ then there will always be more than the minimum in storage from day $d'$ to $D$. Combining equations (5.82) and (5.84) leads to

$$Prob^{s}(c^{s}_{p'} - b^{s}_{so,d'}) + \lambda_{Q^{s}_{p',d'}} + \lambda_{W^{s}_{so,d'}} + \sum_{e=d'}^{D} \lambda^{s}_{Stor^{so,e}_{so}} = 0$$ (5.85)

As $Prob^{s} > 0$, $c^{s}_{p'} > b^{s}_{so,d'}$, $\lambda_{W^{s}_{so,d'}} \geq 0$ and $\lambda^{s}_{Stor^{so,d'}_{so}} \geq 0$, equation (5.85) can never be satisfied. This means that all gas in storage must be withdrawn until the minimum amount of gas in storage is reached. This holds under each scenario $s$, for each roll of ROM.

5.6 Effect of increasing the number of scenarios used in ROM

In this section, the effect of increasing the number of scenarios ($n$) used in ROM is examined. Figure 5.15 displays generated demand paths for scenario sets of various sizes.

\[\text{If there are no injections then, any } d \in D \text{ is suitable.}\]
This plot shows that when the number of scenarios increases, the size of extreme demand also increases. It is clear that when $n = 1000$ the size of the extreme demands are greater than those when $n = 100$, with these in turn exceeding the extremes when $n = 10$. Equations (5.2) and (5.3) of ROM state that demand for each day and for each scenario must be met. This means that the size of the scenario set used in the model has an effect on the results, as larger extreme demands are more likely to appear in the scenario set as $n$ increases. Figures 5.16 - 5.18 examine this effect for the supply of gas through the IUK pipeline, Norwegian imports and withdrawals from storage. This was done by running ROM over a range of different scenario sets of various sizes. For each value of $n$, the experiment was repeated ten times. As all scenarios were randomly generated for each roll of the model, they were assigned equal probabilities for all values of $n$, i.e., $\text{Prob}^s = \frac{1}{n}$ $\forall$ $s$. All other parameters were as described in Section 5.3 for the April 2010 data.
5.6 Effect of increasing the number of scenarios used in ROM

Figure 5.16: Mean plus error bars for the total yearly IUK supply from ROM for \( n = 1, \ldots, 6, 10, 20, 50 \).

Figure 5.17: Mean plus error bars for the total yearly LNG supply from ROM for \( n = 1, \ldots, 6, 10, 20, 50 \).
5. ROLLING OPTIMISATION MODEL

Figure 5.16 shows the mean yearly IUK supply (with error bars) from ROM for \( n = 1, \ldots, 6, 10, 20, 50 \). Throughout this chapter, error bars are defined as \( \bar{x} \pm 1.96 \frac{\sigma}{\sqrt{n}} \), where \( \bar{x} \) and \( \sigma \) are the sample mean and sample standard deviation, respectively.\( ^1 \) The graph indicates that as the size of the scenario set increases, the mean yearly IUK supply decreases. Figure 5.17 shows a similar graph for the mean yearly LNG flows. In contrast to Figure 5.16 it indicates that LNG supply increases as the size of the scenario set increases.

![Figure 5.16: Mean yearly IUK supply from ROM for various scenario set sizes.](image)

Figure 5.18: Mean plus error bars for the total yearly withdrawals from storage for ROM for \( n = 1, \ldots, 6, 10, 20, 50 \).

Figure 5.18 displays a similar graph for the total yearly withdrawals from each of the storage facilities combined. For \( n = 1, \ldots, 5 \) there is no obvious trend in the data shown in this plot. However, for \( n = 6, 10, 20, 50 \) and in comparison to Figure 5.17 there is an obvious decreasing trend in the mean level of withdrawals to storage. Appendix C.4 contains similar plots to Figures 5.16-5.18 but for the total yearly supply of gas from the UKCS, Norway, BBL and injections to storage. The results for UKCS supply and injections to storage indicate similar trends to Figures 5.16-5.18 while there are no obvious trends for BBL and Norwegian supply.

Ideally ROM should be run with as many scenarios as possible. However, Table 5.13 shows the increasing computational cost of running the model as \( n \) increases.\( ^1 \)If the data is normally distributed then these error bars correspond to a 95% confidence interval.
5.7 Summary

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Coefficient of variation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.04</td>
<td>11%</td>
</tr>
<tr>
<td>2</td>
<td>0.10</td>
<td>12%</td>
</tr>
<tr>
<td>3</td>
<td>0.18</td>
<td>7%</td>
</tr>
<tr>
<td>4</td>
<td>0.25</td>
<td>7%</td>
</tr>
<tr>
<td>5</td>
<td>0.32</td>
<td>9%</td>
</tr>
<tr>
<td>6</td>
<td>0.50</td>
<td>5%</td>
</tr>
<tr>
<td>10</td>
<td>0.96</td>
<td>6%</td>
</tr>
<tr>
<td>20</td>
<td>2.05</td>
<td>13%</td>
</tr>
<tr>
<td>50</td>
<td>9.72</td>
<td>7%</td>
</tr>
</tbody>
</table>

Table 5.13: Mean GAMS solve times (in hours) associated with ROM, for \( n = 1, \ldots, 6, 10, 20, 50 \).

It displays the mean times for solving ROM in GAMS for the experiments presented in Figures 5.16 - 5.18. For each value of \( n \), these times were averaged over ten experiments. All test runs were performed on a PC with a 3.4 GHz frequency processor and with 16GB of RAM.

At this stage, it is important to note that the scales in the plots presented in Figures 5.16 - 5.18 are all different. For example, the difference between the maximum and minimum means in Figures 5.16 - 5.18 are approximately 27%, 1% and 4% respectively of their maximum means. However, the differences are all approximately of the order of 100 mcm. The significance of these results lies, not in the size of these differences but, in the trends observed. This suggests that the model does not converge as \( n \to \infty \). As a result, even if there was infinite computational power available, there is no obvious choice of \( n \). In the next chapter of this thesis, different scenario reduction techniques are introduced and studied. These techniques allow large scenario sets to be reduced to smaller ones. This analysis allows ROM to be solved with demand scenarios that accurately incorporate information from larger scenarios sets, whilst maintaining a relatively low computational time.

5.7 Summary

In this chapter, the Rolling Optimisation Model (ROM) was introduced. This was motivated, firstly, by a review of different gas market models that incorporate stochastic demand in Section 5.1. Following this, a detailed description of the model was provided in Section 5.2. Thirdly, in Section 5.3 the parameters of ROM were chosen so as to best fit data from the UK gas market for the year beginning on the 1st of April 2010. The results of this analysis found that flows of gas simulated by ROM were qualitatively similar to actual flows. These flows describe how the different sources of supply meet demand, as well as the amount of gas in the different types of storage facilities. Using similar parameters, ROM was then tested in Section 5.4 using data from the UK gas market.
for the year starting on the 1st of April 2011. This analysis again showed that results obtained from 
ROM fitted reasonably well to actual data. Following this, in Section 5.5, it was shown that the 
costs in the model can be varied without affecting the resulting optimal flows. This can done in two 
ways. Firstly, the production costs can be shifted by a constant $\beta$. Secondly, both the production 
costs and the storage costs can be scaled by a positive constant $\alpha$. When this is done, the prices in 
ROM become similarly scaled and shifted by the same values of $\alpha$ and $\beta$ respectively. The chapter 
concluded in Section 5.6 by examining the effect of increasing the number of scenarios used in the 
model. This analysis showed that as the number of demand scenarios used in the model increased, 
the total yearly flows of some of the different sources of supply increased, while for others they 
decreased. Consequences of these observations are considered in Chapter 6.
Chapter 6

Scenario reduction

In the previous Chapter, the Rolling Optimisation Model (ROM) was introduced. Using stochastically generated demand scenarios, this model was shown to be able to capture some of the qualitative aspects of the UK natural gas market. For example, ROM is able to model how gas demand in the UK is met by the different sources of supply. The analysis in Section 5.6 however, showed that as the number of demand scenarios used increased (i.e., as \( n \) increased) there was a clear upward trend in total yearly supplies for some of the different sources of supply (see, e.g., Figure 5.17), while for others there was a clear downward trend (see, e.g., Figure 5.16). Ideally ROM should be applied with as many scenarios as possible. However, Section 5.6 also notes the increasing computational cost associated with ROM as \( n \) increases. This chapter attempts to rectify this issue by introducing various scenario reduction techniques. The aim of this analysis is to allow large scenario sets to be replaced by smaller reduced sets, whilst maintaining the effects of the diversity contained in the larger sets.

The scenario reduction techniques studied here are the heuristic algorithms developed in [13] and [14]. These algorithms determine a subset of the initial scenario set and assign new probabilities to the preserved scenarios. They have no requirements on the structure or the dimension of the stochastic data process (i.e., the scenarios). This chapter describes the background theory for the algorithms, the algorithms themselves, and some of the papers seen in the literature that use them. Other scenario reduction techniques, such as cluster analysis, were also studied during this project but were found to be less efficient than those examined here and are hence not described in this thesis.

A selection of the algorithms developed in [13] and [14] are also analysed with respect to scenarios obtained from the stochastic process for demand developed in this project. This is done by examining the accuracy, computational time and stability of the algorithms. Finally, the effects of using the algorithms on ROM are investigated.
6.1 Theoretical background

Let $P$ be the initial probability distribution of the set of $k$-dimensional scenarios $w_i$ with weights $p_i > 0$, where $i = 1, \ldots, N$, $\sum_{i=1}^{N} p_i = 1$ and $P = \sum_{i=1}^{N} p_i \delta_{w_i}$. Here, $\delta_{w_i}$ denotes the Dirac measure. Now let $Q$ be the probability distribution of another $k$-dimensional set of scenarios $w_j$ with weights $q_j > 0$, where $j = 1, \ldots, \tilde{N}$, $\sum_{j=1}^{\tilde{N}} q_j = 1$ and $Q = \sum_{j=1}^{\tilde{N}} q_j \delta_{w_j}$. In both [13] and [14], the Monge-Kantorovich distance $D$ of (multivariate) probability distributions, [60], is used to measure the distance between the distributions $P$ and $Q$. For continuous probability distributions, the Monge-Kantorovich distance is

$$D(P, Q) := \inf \int_{\Omega \times \Omega} c(w, \bar{w}) \eta(d(w, \bar{w}))$$

subject to:

$$\pi_1 \eta = P,$$

$$\pi_2 \eta = Q,$$

$$\eta \in \mathcal{P}(\Omega \times \Omega),$$

where $\mathcal{P}(\Omega)$ is the set of all Borel probability measures on $\Omega$, a closed subset of $\mathbb{R}^k$. The projections onto the first and second components are denoted by $\pi_1$ and $\pi_2$, respectively. For discrete probability distributions, however, the Monge-Kantorovich distance is the optimal value of the following linear transportation problem:

$$D(P, Q) = \min \sum_{i=1}^{N} \sum_{i=1}^{\tilde{N}} n_{ij} c(w_i, w_j)$$

subject to:

$$\sum_{i=1}^{N} n_{ij} = q_j, \quad \forall \ j,$$

$$\sum_{j=1}^{\tilde{N}} n_{ij} = p_i, \quad \forall \ i,$$

$$n_{ij} \geq 0, \quad \forall \ i, j,$$

The cost function $c(w, \bar{w})$ measures the distance between the scenarios $w$ and $\bar{w}$ and takes the form

$$c(w, \bar{w}) = \max \{(1, h(\|w - w_0\|), h(\|\bar{w} - w_0\|)) \|w - \bar{w}\|,$$

where $w_0$ is some fixed element in $\mathbb{R}^k$, $h : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ is continuous and non-decreasing and $\| \cdot \|$ is some norm on $\mathbb{R}^k$. When this cost function is applied to gas demand scenarios in this chapter, $h$
is a constant function with a value of 1 (i.e., $h = 1$) and the norm used is the Euclidean norm.

Let us now consider $Q$ to be the distribution of the reduced set of scenarios $w_j$, with $j = \{1, \ldots, N\}/J$, where $J$ denotes the index set of deleted scenarios. For a fixed $J$, the minimal distance $D$ may be computed explicitly using the result from Theorem 3.1 in [13]:

$$D(P, Q) := \sum_{i \in J} p_i \min_{j \notin J} c(w_i, w_j).$$

(6.10)

Moreover, the probabilities $q_j$ of the preserved scenarios $w_j, j \notin J$ of $Q$ are given by the following rule:

$$q_j = p_j + \sum_{i \in J_j} p_i,$$

(6.11)

where

$$J_j := \{i \in J : j = j(i)\}, j(i) \in \arg \min_{j \notin J} c(w_i, w_j), \forall i \in J).$$

(6.12)

The index set $J_j$ is the subset of the deleted scenario index ($J$) which represents deleted scenarios that were closest to the kept scenario $w_j$. Formula (6.11) is known as the optimal distribution rule. It defines the probability of a preserved scenario to be equal to the sum of its former probability and of all the probabilities of deleted scenarios that are closest to it with respect to the distance $c$. Deleted scenarios are assigned a probability of zero, i.e., $(q_j = 0 \forall j \in J)$. For scenario reduction with fixed cardinality, $\#J$, the optimal choice of an index set $J$ is given by the following mathematical program:

$$\min \sum_{i \in J} p_i \min_{j \notin J} c(w_i, w_j),$$

(6.13)

subject to:

$$J \subset \{1, \ldots, N\},$$

(6.14)

$$\#J = N - n,$$

(6.15)

where the preserved number of scenarios is $n = N - \#J > 0$. This program is a well known set-covering problem [14, 62] and may be formulated as a 0-1 integer programming problem which is difficult to solve. As a result, heuristic algorithms exploiting the structure of $D$ were developed in [13, 14] and are described in Section 6.2. It is noted in [14] that these algorithms only provide approximate solutions to (6.13).

At this stage, it must also be noted that there are some variants to the background theory described in this section. For example, in [63] the Kantorovich-Rubinstein distance is used instead of the Monge-Kantorovich distance to measure the distance between $P$ and $Q$. This leads to the
6. SCENARIO REDUCTION

cost function (equation 6.9) being replaced by the more complicated

\[ \hat{c}_r(w_i, w_j) = \sum_{\tau=1}^{m-1} c_{r}(w_{l_{\tau}}, w_{l_{\tau}}), \quad (6.16) \]

where \( m \in \mathbb{N}, l_{r} \in \{1, \ldots, N\}, l_1 = i \) and \( l_m = j \not\in J \), while the cost function \( c_r(w, \bar{w}) \) measures the distance between the scenarios \( w \) and \( \bar{w} \) as follows

\[ c_r(w, \bar{w}) = \max \{(1, ||w - w_0||^{r-1}, ||\bar{w} - w_0||^{r-1})\} ||w - \bar{w}||, \quad (6.17) \]

for some \( r \geq 1 \) and \( w_0 \in \Omega \).

6.2 Algorithms

The first scenario reduction algorithm that takes advantage of the theory described in the previous section was developed in [13] and is known as backward reduction. It involves the deletion of one scenario from the scenario set at each step. This is done by finding the two scenarios that are closest to each other. Out of these two scenarios, the scenario with the lower probability is then removed from the scenario set. If the two scenarios have equal probability then it is optimal to remove either.

By letting \( \#J = 1 \), the program (6.13) above takes the form:

\[ \min_{l \in 1, \ldots, N} p_l \min_{j \neq l} c(w_l, w_j), \quad (6.18) \]

The scenario \( w_{l_*} \) is deleted if the minimum is attained at \( l_* \in \{1, \ldots, N\} \). The redistribution rule (6.11) is then used to obtain the reduced probability measure \( Q \). This procedure can then be repeated recursively until the desired number \( N - n \) scenarios are deleted.

The second scenario algorithm developed in [13] is known as backward reduction of scenario sets. It is a close relation of the backward reduction algorithm. However, in contrast, it allows multiple deletions at each time-step. Let the indices \( l_i \) be selected such that

\[ l_i \in \arg \min_{l \in \{1, \ldots, N\} \setminus \{l_1, \ldots, l_{i-1}\}} p_l \min_{j \neq l} c(w_l, w_j), \quad i = 1, \ldots, N - n. \quad (6.19) \]

Then,

\[ lb := \sum_{i=1}^{N-n} p_{l_i} \min_{j \neq l_i} c(w_{l_i}, w_j), \quad (6.20) \]

can be shown to be a lower bound for the optimal value of (6.13), [13, 14]. In addition, if the set \( \arg \min_{j \neq l_i} c(w_{l_i}, w_j) \setminus \{l_1, \ldots, l_{i-1}, l_{i+1}, \ldots, l_{N-n}\} \) is nonempty for all \( i = 1, \ldots, N - n \), then the index set \( \{l_1, \ldots, l_{N-n}\} \) is a solution of equation (6.13), [13, 14]. As a result, the backward reduction of scenario sets algorithm takes the following form:
Backward reduction of scenario sets algorithm

*In the first step, an index \( n_1 \) with \( n \leq n_1 < N \) is determined using formula (6.19) such that \( J_1 = \{l_1, \ldots, l_{N-n_1}\} \) is a solution of (6.13) for \( n = n_1 \). Next, the redistribution rule (6.11) is used. This leads to the reduced probability measure \( P_1 \) containing all scenarios indexed by \( \{1, \ldots, N\}\setminus J_1 \). If \( n < n_1 \), the measure \( P_1 \) is further reduced by deleting all scenarios belonging to some index set \( J_2 \) with \#\( J_2 = n_1 - n_2 \) and \( n \leq n_2 < n_1 \), which is obtained in the same way by using formula (6.19). This procedure is continued until, in step \( r \), \( n_r = n \) and \( J = \bigcup_{i=1}^{r} J_i \). Finally, the redistribution rule (6.11) is used again for the index set \( J \).

The final scenario algorithm developed in [13] is known as *forward selection*. While the two previous algorithms described in this section involve the recursive deletion of scenarios, this new algorithm involves the recursive selection of scenarios to be kept in the reduced scenario set. This is done, at each stage of the algorithm by selecting the scenario that is nearest to all other scenarios. By letting \#\( J = N - 1 \), the mathematical program (6.13) takes the form

\[
\min_{u \in 1, \ldots, N} \sum_{i=1}^{N} p_i c(w_i, w_u).
\]

(6.21)

The scenario \( w_u \) is retained if the minimum is attained at \( u \in \{1, \ldots, N\} \). The redistribution rule (6.11) provides the reduced probability measure as follows: \( q_U = p_{u*} + \sum_{i \neq u*} p_i = 1 \). Like the backward reduction algorithm, it can be used recursively until the desired number of \( n \) scenarios are selected.

Following the development of these algorithms, two further algorithms were developed in [14]. The first is a modification of the backward reduction algorithm and is known as the *simultaneous backward reduction* algorithm. The major difference with this new algorithm is that all deleted scenarios are included into each backward step simultaneously as described in the following:
### Simultaneous backward reduction algorithm

**Step 0:** Compute the distances between scenario pairs:
\[ c_{kj} := c(w_k, w_j) \quad k, j = 1, ..., N. \]
Sort the records \( \{c_{kj} : j = 1, ..., N\}, k = 1, ..., N. \)

**Step 1:** Compute
\[ c^{[1]}_{ll} := \min_{j \neq l} c_{lj}, \quad l = 1, ..., N \text{ and} \]
\[ z^{[1]}_l := p_l c^{[1]}_{ll}, \quad l = 1, ..., N. \]
Choose \( l_1 \in \arg \min_{l \in \{1, ..., N\}} z^{[1]}_l. \)
Set \( J^{[1]} := \{l_1\}. \)

**Step i:** Compute
\[ c^{[i]}_{kl} := \min_{j \notin J^{[i-1]} \cup \{l\}} c_{kj} \]
for \( l \notin J^{[i-1]}, \ k \in J^{[i-1]} \cup \{l\} \text{ and} \]
\[ z^{[i]}_l := \sum_{k \in J^{[i-1]} \cup \{l\}} p_k c^{[i]}_{kl}, \ l \notin J^{[i-1]} \text{.} \]
Choose \( l_i \in \arg \min_{l \notin J^{[i-1]}} z^{[i]}_l. \)
Set \( J^{[i]} := J^{[i-1]} \cup \{l_i\}. \)

**Step N - n + 1:** \( J := J^{[N-n]} \) becomes the index set of deleted scenarios.
Compute the probabilities of the preserved scenarios using the redistribution rule (6.11).

The second algorithm developed in [14] is a modification of the forward selection algorithm and is known as the fast forward selection algorithm. The major difference with this new algorithm is that the cost matrix \( c \) is updated after each time-step as the following describes:
6.2 Algorithms

### Fast forward selection algorithm

**Step 0:** Compute the distances between scenario pairs:
\[ c_{ku} := c(w_k, w_u) \quad k, u = 1, \ldots, N. \]

**Step 1:** Compute
\[ z_u^{[1]} := \sum_{k=1, k \neq u}^{N} p_k c_{ku}^{[1]} \quad u = 1, \ldots, N. \]
Choose \( u_1 \in \arg \min_{u \in 1, \ldots, N} z_u^{[1]} \).
Set \( J^{[1]} := \{1, \ldots, N\} \setminus \{u_1\} \).

**Step i:** Compute
\[ c_{ku}^{[i]} := \min\{c_{ku}^{[i-1]}, c_{ku}^{[i-1]}\} \quad k, u \in J^{[i-1]} \]
\[ z_u^{[i]} := \sum_{k \in J^{[i-1]} \setminus \{u\}} p_k c_{ku}^{[i]} \quad u \in J^{[i-1]} \]
Choose \( u_i \in \arg \min_{u \in J^{[i-1]}} z_u^{[i]} \).
Set \( J^{[i]} := J^{[i-1]} \setminus \{u_i\} \).

**Step \( N - n + 1 \):** \( J := J^{[N-n]} \) becomes the index set of deleted scenarios.
Compute the probabilities of the preserved scenarios using the redistribution rule (6.11).

#### 6.2.1 Performance

In [13], the backward reduction, backward reduction of scenario sets and forward selection algorithms are all compared in terms of both accuracy and computational speed. This was done for the reduction of a scenario tree that represents an approximation of the electrical load process in a power management model under uncertainty. Accuracy was measured using the following formula:
\[
D_{rel}(P, Q) = \frac{D(P, Q)}{D(P, w_i)}, \tag{6.22}
\]
where \( D(P, Q) \) represents the minimal distance between the original scenario set, with probability distribution \( P \), and the reduced scenario with probability distribution \( Q \), as in equation (6.10). The distance \( D(P, \delta_{w_i}) \) is defined by:
\[
D(P, \delta_{w_i}) = \min\{D_J : \#J = N - 1\} = \min_{i \in \{1, \ldots, N\}} D(P, \delta_{w_i}), \tag{6.23}
\]
which is a solution to the forward selection algorithm (equation (6.21)) for \( n = 1 \). This solution is known to be optimal. Essentially, \( D_{rel}(P, Q) \) is a relative distance measure and it is assumed in [13] that smaller values of \( D_{rel}(P, Q) \) correspond to more accurate reductions.

It was found in [13] that all three algorithms performed well in terms of accuracy with for-
ward selection being slightly better than the two backward reduction algorithms. With respect to computing times, the two backward reduction algorithms were found to be much faster than the forward selection algorithm when the reduced tree had a large number of scenarios. Conversely, when the reduced tree had a small number, the opposite result was observed. When comparing the two backward reduction algorithms, backward reduction of scenario sets was found to be slightly slower.

In a similar manner, the backward reduction of scenario sets, simultaneous backward reduction algorithm and fast forward selection algorithms were all compared in terms of both accuracy and computational speed in [14]. This was done for three different examples of scenario trees where the last one was the same tree used in [13]. Again, accuracy was compared by examining minimal distances using equation (6.22) above. It was found that the simultaneous backward reduction and fast forward selection algorithms produce more accurate results than the backward reduction of scenario sets algorithm, but this came at the expense of longer running times. The results also indicated that the fast forward selection algorithm was slightly more accurate than the simultaneous backward reduction algorithm.

In comparison to [13], it was also found that the simultaneous backward reduction algorithm had much faster running times than the fast forward selection algorithm when the reduced number of scenarios was large, while the opposite was seen when the reduced number was small. More specifically, the theoretical results in [14] showed that the simultaneous backward reduction and fast forward selection algorithms were preferable when \( n > \frac{N}{T} \) and \( n < \frac{N}{T} \) respectively. As mentioned already, the computational times associated with the backward reduction of scenario sets algorithm are smaller than those for the simultaneous backward reduction algorithm. When the numerical results from [13] and [14] were compared, it was also found that the computational times associated with the fast forward selection algorithm were much smaller than the simpler forward selection algorithm.

In Sections 6.4 and 6.5 the backward reduction, simultaneous backward reduction and fast forward selection algorithms will all be compared with each other in terms of accuracy and computational speed, respectively. In contrast to [13] and [14], this will be done for new scenarios, which are the gas demand scenarios developed in this project. Sections 6.7 and 6.8 will also examine how the reduced scenario sets from these three algorithms affects ROM.

### 6.3 Applications of scenario reduction algorithms

As described above, some of the scenario reduction algorithms have been applied in a number of different papers. These applications are mainly in the context of electricity markets. As mentioned already for example, a scenario tree that represents an electrical load process in a power management model is reduced in [13]. In [14] and [63] the same tree is also reduced. Likewise, in [62], the simultaneous backward reduction and fast forward selection algorithms are applied to a scenario tree representing an uncertain electrical load process for a hydro-thermal generation system of a
German utility. These algorithms are also applied in this paper to a set of scenarios representing uncertain load and market spot prices.

In [64], the fast forward selection algorithm has also been applied to an electricity market. However, in this case, it is not used for an uncertain electrical load process, but rather in the context of an electricity futures market. In [65] the fast forward selection algorithm is again used. It is applied, firstly, to reduce a scenario set of stochastic electricity prices and secondly to scenarios that represent the availability of production units. For both cases in [65], the scenario reduction algorithms are applied in the context of a risk-averse electricity trader. In both [55] and [66], two more models of electricity markets are described. In contrast however, the scenario reduction algorithms used in both of these papers reduce scenarios of both uncertain wind power and electrical load.

While most of these papers apply the different algorithms in the context of electricity markets, there are other examples. For instance, in [67], scenario reduction algorithms are applied to a multi-asset financial optimisation problem. In [68], such algorithms are applied to a model of the Danish mortgage market. As well as these financial examples, the simultaneous backward reduction algorithm is used in [69] for a stochastic optimisation model of a pharmaceutical multi-period project selection problem.

The next section of this chapter examines some of the different scenario reduction algorithms applied to scenarios derived from the stochastic process of natural gas demand developed in this project. This is done in a number of different ways. Firstly, the accuracy of some of the algorithms are considered.

### 6.4 Accuracy

As described above, some of the different scenario reduction algorithms are applied to scenarios representing an electrical load process in both [13] and [14]. As explained in Section 6.2.1, accuracy is measured by calculating the relative distance between the original scenario set and the reduced scenario set. This automatically leads them to conclude that the larger the value of $n$, the more the accurate the solution.

In this project however, a different formula is used to measure the accuracy of the backward reduction, simultaneous backward reduction and the fast forward selection algorithms. As shown above, these algorithms provide approximate solutions to the mathematical program, (6.13). In both [13] and [14], it is stated that this mathematical program is difficult to solve, as it is an mixed-integer programming problem. For this project, however, it was found that when $N = 100$ it was quite tractable to find a solution of this mixed-integer program and hence the optimal solution to reducing a set of scenarios of size $N = 100$ to $n$ for $n < 100$. As a result, accuracy in this project was measured as the ratio between the optimal solution obtained from the mathematical program, (6.13) and the solution obtained from the scenario reduction algorithm as follows:

$$A(P, Q) = \frac{D(P, Q_{optimal})}{D(P, Q)}.$$  

(6.24)
Here, $D(P, Q_{\text{optimal}})$ represents the minimal distance between the original probability distribution $P$, and the optimally reduced distribution $Q$, obtained from the mixed-integer program (6.13). In order to solve the mixed-integer program, it was reformulated as follows:

$$\min \sum_{i=1}^{N} \sum_{j=1}^{N} n_{i,j} c(w_i, w_j)$$

(6.25)

subject to:

$$\sum_{j=1}^{N} k_j = n,$$

(6.26)

$$\sum_{j=1}^{N} n_{i,j} = p_i, \ \forall \ i,$$

(6.27)

$$\sum_{i=1}^{N} n_{i,j} \leq k_j, \ \forall \ j,$$

(6.28)

$$n_{i,j} \geq 0, \ \forall \ i, j,$$

(6.29)

where all variables are defined as previously in this chapter, except for $k_j$ which is a 0-1 binary variable indicating whether the scenario $j$ is in the reduced set or not.

Figure 6.1: Average accuracy of scenario reduction algorithms for $N = 100$.

Figure 6.1 shows the average accuracy of the backward reduction (red), simultaneous backward
6.4 Accuracy

reduction (black) and the fast forward selection (red) algorithms. These algorithms were used in 30 experiments to reduce scenario sets of size $N = 100$ to size $n$ in each experiment. The 100 scenarios were obtained from the stochastic process for demand described in Appendix D.1. The optimal solutions of equation 6.24 were found by solving the above mixed-integer program using the CPLEX solver [70] in GAMS while, the solutions for the heuristic algorithms were obtained by coding the algorithms in MATLAB. The computational time associated with these are analysed in the following section.

Figure 6.1 shows that the fast forward selection algorithm is the most accurate. It maintains an average accuracy of over 99.5% for all sizes of $n$ considered while its accuracy decreases as $n$ increases. Conversely, the two backward reduction algorithms increase in accuracy as $n$ gets bigger. Both these algorithms have the same accuracy for $n$ larger than approximately 60. For $n < 60$, Figure 6.1 shows that the simultaneous backward reduction algorithm is more accurate than the backward reduction. In conclusion however, it can be said that each of these heuristic algorithms are highly accurate at obtaining a near optimal reduction.

When this analysis was attempted with $N = 1000$, the computational times associated with finding the optimally reduced solutions from the mixed-integer program were very large. This reinforced the need for heuristic algorithms and is discussed in more detail in Section 6.5.

6.4.1 Alternative distance measure

In order to calculate the above accuracies, the Monge-Kantorovich distance was used to measure the distance between the original and reduced probability distributions (see Section 6.1). In this project an alternative distance measure was tested. It involved finding the relative distance between the centroid of the original probability distribution and that of the reduced distribution. The centroid for the original probability distribution $P$ was calculated as follows:

\[ Centroid^P = \sum_{i=1}^{N} p_i w_i, \]  

(6.30)

while for the reduced probability distribution it was

\[ Centroid^Q = \sum_{j=1}^{n} q_j w_j, \]  

(6.31)

where all the variables are defined as above. The relative distance was then calculated by using the following formula:

\[ distance = \frac{\|Centroid^P - Centroid^Q\|_2}{\|Centroid^P\|_2}, \]  

(6.32)

where $\|.\|_2$ represents the Euclidean norm.
6. SCENARIO REDUCTION

Figure 6.2: Distance between the centroid of the original and reduced probability distributions versus the size of the reduced set for $n = 1, \ldots, 99$ and $N = 100$.

Figure 6.3: Distance between the centroid of the original and reduced probability distributions versus the size of the reduced set for $n = 2, \ldots, 99$ and $N = 100$. 

116
Figure 6.2 shows the difference between the centroids obtained for the original and reduced probability distributions for the backward reduction, simultaneous backward reduction and fast forward selection algorithms. The size of the reduced sets vary from \( n = 1, \ldots, 99 \) while the size of the original scenario set is fixed at \( N = 100 \). The scenarios were obtained from the stochastic process for demand described in Appendix D.1. Figure 6.3 shows the same graph with \( n = 1 \) excluded.

Figure 6.3 shows that there is a very small \( \mathcal{O}(10^{-2}) \) relative difference between the centroids for \( n \geq 2 \), for each of the three algorithms. This was because the centroids of both the original and reduced sets was always found to be approximately equal to Seasonal Normal Demand (SND). The only case when the centroid was found not to be SND was when \( n = 1 \), which explains the large relative difference seen at this point in Figure 6.2. In this case, the centroid of the reduced set was the one remaining scenario itself. Figure 6.3 also shows that the relative difference between the centroids decreases as \( n \) increases. This is because the centroid of the reduced scenario set becomes closer to SND when the size of that scenario set increases.

Figures 6.2 and 6.3 show that this (centroid) distance measure is not a prudent distance measure and hence why it is redundant in terms of analysing the accuracy of the scenario reduction algorithms.

### 6.5 Computational time

In the previous section, the accuracy of the backward reduction, simultaneous backward reduction and the fast forward selection algorithms were all examined. In this section the computational time associated with the three algorithms is analysed. This is done for gas demand scenarios developed using the process presented in Appendix D.1. As mentioned previously, these three algorithms were all coded in MATLAB. This means that the following results were obtained using MATLAB’s `cputime` function. All test runs were performed on a PC with a 3.4 GHz frequency processor and with 16GB of RAM.
Figure 6.4: Average computational times associated with the backward reduction, simultaneous backward reduction and fast forward selection algorithms for $N = 100$.

Figure 6.5: Average computational times associated with the backward reduction, simultaneous backward reduction and fast forward selection algorithms for $N = 100$. 
Figure 6.4 shows the average computational time required for the backward reduction (red), simultaneous backward reduction (blue) and the fast forward selection (black) algorithms. It also shows the average computational time required for finding the optimal solutions of the mixed-integer program (green). These results were averaged over the same 30 experiments described in Section 6.4 to reduce scenario sets of size $N = 100$ to size $n$. Figure 6.4 shows that the average computational time associated with solving the mixed-integer problem is u-shaped in its dependence on $n$. For $n = 1$ and $n = 99$ a solution is found quite quickly. In contrast, large increases in computational time occur at $n = 2$, $n = 68$ and $n = 96$. The small computational times at $n = 1$ and $n = 99$ can be explained by the relatively simple form the mathematical program (6.13) takes for these values. As explained in Section 6.2 when $n = 1$ and $n = N - 1$ (i.e., in this case $n = 99$), the mixed-integer problem reduces to using the backward reduction (equation (6.18)) and the forward selection (equation (6.21)) algorithms, respectively. For the rest of the features seen in the green line there is no obvious explanation. However, as the average computational time associated with solving the mixed-integer problem is always greater than those associated with each of the three different algorithms, these features are not of great interest to this project. This fact also illustrates the benefit of using the heuristic algorithms over the mixed-integer program.

Figure 6.5 shows the same times as Figure 6.4. However, the times associated with the mixed-integer program are excluded. It shows that the time associated with the backward reduction algorithm is relatively constant before decreasing towards zero as $n$ approaches $N$. It is the quickest algorithm for $n > 10$. For $n \leq 10$, the fast forward selection algorithm is the quickest. However, there is little difference between it and backward reduction. Figure 6.5 also shows that the time associated with the fast forward selection increases as $n$ increases. Conversely, the time associated with the simultaneous backward reduction algorithm decreases as $n$ increases.

These results are consistent with the experiments performed in [14] and detailed in Section 6.2.1 stating that simultaneous backward reduction is faster when $n$ is large, while the fast forward selection algorithm is faster when $n$ is small.

---

The computational times for solving the mixed-integer problem were obtained using the Cplex solver in GAMS, while all other times were obtained using MATLAB.
6. SCENARIO REDUCTION

Figure 6.6: Average computational times associated with the backward reduction, simultaneous backward reduction and fast forward selection algorithms for $N = 1000$.

Figure 6.7: Average computational times associated with the backward reduction algorithm for $N = 1000$. 
6.5 Computational time

Figure 6.6 shows the average computational time required for the backward reduction (blue), simultaneous backward reduction (red) and the fast forward selection (green) algorithms when $N = 1000$. As before, these results were averaged over 30 experiments. The graph shows that the backward reduction algorithm is by far the fastest algorithm for most values of $n$. Not visible in this figure, however, is that the fast forward selection algorithm is faster for $n \leq 3$. As concluded above for $N = 100$ and in [14], fast forward selection is preferable (in terms of computational cost) over simultaneous backward when $n$ is small, while the opposite is true when $n$ is large.

As the times associated with the backward reduction algorithm are not clearly visible in Figure 6.6, they are shown on their own in Figure 6.7. This graph shows that the times associated with this algorithm are relatively constant when $n$ is small. However, as $n$ increases towards $N$, the computational time decreases to almost zero.

As mentioned in Section 6.4 it was attempted to find the optimal reduced solutions by solving the mixed-integer program (equation 6.13). However, it became highly intractable to find these solutions when $N = 1000$. For example, when $N = 1000$ and $n = 2$ it took approximately 5.5 days to obtain a solution to the mixed-integer problem. Again, this shows the benefit of using the much faster heuristic algorithms.

6.5.1 Computational time associated with operations count

In [14] the operations count associated with the backward reduction, simultaneous backward reduction and fast forward selection algorithms are presented. These counts were obtained by counting the steps required using the different algorithms. In this subsection, computational times from these three algorithms, applied to scenarios generated from the stochastic process for demand in Appendix D.1 are compared with the theoretical operations count. This is done by comparing the log of the computational time with the log of the original size of the scenario sets ($N$) for four different fixed values of $n$. The operations count for the simultaneous backward reduction algorithm is given in [14] as follows:

$$b_N(n) = n^3 - n^2\left(\frac{3}{2}N + \frac{1}{2}\right) - n\left(\frac{3}{2}N + 1\right) + \frac{N^3}{2} + 2N^2 + \frac{3}{2}N + \Theta(N^2 \log N),$$

(6.33)

while for the fast forward selection algorithm it is:

$$f_N(n) = \frac{3}{2}n^3 - n^2(2N + 1) + n(2N^2 + 2N + \frac{1}{3}).$$

(6.34)

Throughout this chapter each $\Theta(g(N))$ describes the limiting behaviour of the function $g(N)$ as $N \to \infty$. The operations count for the backward reduction algorithm is given in [14] to be of $\Theta(N^2)$, regardless of the size of $n$. 

121
Figure 6.8 shows a log-log plot of the computational time for the backward reduction (red), simultaneous backward reduction (blue) and the fast forward selection (black) algorithms versus the size of the original scenario sets \((N)\) for \(n = 1\). For each algorithm, a least squared regression line was fitted. This provides an estimate to the slopes of the lines, which should correspond to the order of the operations count for the different algorithms. When \(n = 1\), equations (6.33) and (6.34) are of \(O(N^3)\) and \(O(N^2)\) respectively. These correspond to Figure 6.8, where the slopes associated with the backward reduction, simultaneous backward reduction and fast forward selection algorithms are 1.8, 2.7 and 1.8, which are all slightly below the expected orders of 2, 3 and 2, respectively. The difference between these observed and expected values can be explained by the fact that the largest value of \(N\) used in these experiments is \(N = 10000\), while the expected orders are determined as \(N \rightarrow \infty\).
Figure 6.9: Log-log plot of computational time versus the number of original scenarios \((N)\) for \(n = 50\).

Figure 6.9 shows a similar graph to Figure 6.8 except with \(n = 50\) instead of \(n = 1\). Again, for each algorithm, a least squared regression line was fitted in order to estimate the slopes of the lines. When \(n = 50\), equations (6.33) and (6.34) are of \(O(N^3)\) and \(O(N^2)\), respectively. In Figure 6.9 the slopes associated with the backward reduction, simultaneous backward reduction and fast forward selection algorithms are 1.8, 2.8 and 2.0, respectively. The backward reduction and simultaneous backward reduction algorithms are slightly below their expected order of 2 and 3, respectively. The slope of the fast forward selection algorithm however, matches its expected order. In appendix D.2 similar results are shown for \(n = N/2\) and \(n = N - 1\).

In summary, Figures 6.4 - 6.9 show that when the three different heuristic algorithms are applied to scenarios developed from the stochastic process, the computational times correspond reasonably closely with the results expected from \([14]\). This section also shows the computational attractiveness of the three algorithms, particularly over finding the optimal solutions of the mixed-integer program equation (6.13).

### 6.6 Stability

In the previous sections, the accuracy and computational times of the backward reduction, simultaneous backward reduction and fast forward selection algorithms were examined. In this section, the stability of the same three algorithms, when applied to gas demand scenarios, are analysed. As before the process used to develop these scenarios is described in Appendix D.1. The stability
of the algorithms is analysed by firstly making small perturbations to both the demands and to the probabilities associated with the different scenarios. If the scenario reduction algorithms are unstable, one would expect these perturbations to significantly affect the results, i.e., the reduced scenario sets. The three algorithms are used on both the original (unperturbed) data and the perturbed data. The two sets of solutions are then compared by considering the following relative difference:

$$\frac{D(P, Q) - D(P^*, Q^*)}{D(P, Q)},$$

(6.35)

where $D(P, Q)$ represents the minimal distance between the original scenario set and the reduced scenario, as in equation (6.10). It also equals the objective function in the mixed-integer problem, equation (6.13). Here, $P$ and $Q$ represent the original and reduced probability distributions of the unperturbed data, respectively. The probability distributions of the perturbed data are represented by $P^*$ and $Q^*$. The demands were perturbed as follows:

$$demand^d_i = demand^d_i + \epsilon X^d_i demand^d_i, \forall d, i,$$

(6.36)

where $\epsilon$ represents the size of the perturbations and where each $X^d_i$ is a random variable obtained from a normal distribution with mean 0 and standard deviation 1. Demand for natural gas in the UK for day $d$ under scenario $i$ is denoted by $demand^d_i$. The perturbed probabilities ($\bar{p}_i$) were obtained in a similar manner in the following order:

$$\bar{p}_i = p_i + \epsilon X_i p_i, \forall i,$$

(6.37)

$$\bar{\bar{p}}_i = \max(0, \bar{p}_i), \forall i,$$

(6.38)

$$\hat{p}_i = \frac{\bar{\bar{p}}_i}{\sum_i \bar{\bar{p}}_i}, \forall i,$$

(6.39)

where $p_i$, $\epsilon$ and $X_i$ are all as defined previously. As the scenarios were randomly generated, the initial probabilities ($p_i$) associated with each scenario were set to be equal. This in comparison with the analysis of Section 5.6. Equation (6.38) ensured that none of the probabilities became negative after the perturbation, while equation (6.39) ensured that the probabilities summed to one.
Figure 6.10: Difference between original and perturbed objective functions versus the size of the reduced scenarios sets \((n)\) for \(\epsilon = 10^{-5}\) and \(N = 100\). Scenarios are initially assumed to have equal probabilities.

Figure 6.11: Difference between original and perturbed objective functions versus the size of the reduced scenarios sets \((n)\) for \(\epsilon = 10^{-5}\) and \(N = 100\). Scenarios are initially assumed to have equal probabilities.
Figure 6.12: Difference between original and perturbed objective functions versus the size of the reduced scenarios sets \( (n) \) for \( \epsilon = 10^{-10} \) and \( N = 100 \). Scenarios are initially assumed to have equal probabilities.

Figure 6.13: Difference between original and perturbed objective functions versus the size of the reduced scenarios sets \( (n) \) for \( \epsilon = 10^{-10} \) and \( N = 100 \). Scenarios are initially assumed to have equal probabilities.

Figure 6.10 shows the relative difference (equation (6.35)) between the unperturbed and per-
sturbed solutions for the three different algorithms for varying sizes of the reduced scenario sets \( n \). The original size of the scenario sets is fixed at \( N = 100 \), while the size of the perturbations is fixed at \( \epsilon = 10^{-5} \). Scenarios with the unperturbed data in this experiment are assumed to be equally probable. Figure 6.10 shows that relative difference for both backward reduction and simultaneous backward reduction is of \( O(10^{-2}) \). Figure 6.11 shows the relative difference for the fast forward selection algorithm is of \( O(10^{-5}) \).

Figure 6.12 shows the relative difference for \( \epsilon = 10^{-10} \) for the same original (unperturbed) data. Despite \( \epsilon \) decreasing, the relative difference for both backward reduction algorithms remains of \( O(10^{-2}) \). The order of the relative difference for the fast forward Section algorithm decreases however to \( O(10^{-10}) \), (Figure 6.13).

Figures 6.10 and 6.12 suggest that the relative difference (as defined in equation (6.35)) for the two backward reduction algorithms is of \( O(10^{-2}) \) regardless of the size of \( \epsilon \), assuming \( \epsilon \) is small. Figures 6.11 and 6.13 however, suggest that the relative difference for the fast forward selection algorithm is approximately of \( O(10^{-k}) \) when \( \epsilon \) is of \( O(10^{-k}) \), again, assuming \( \epsilon \) is small.

A reason for these observations is that equal probabilities are assumed in the unperturbed data. Equal probabilities lead to numerous optimal solutions when using either backward reduction algorithms. This is because if there are two scenarios with minimal distance to each other and equal probability, then it is optimal to delete either. This means that when the probabilities are perturbed, regardless of size, it makes a difference, albeit only of \( O(10^{-2}) \). As a result, the same analysis was performed with the initial probabilities of the unperturbed data being randomly generated. They were generated as follows:

\[
p_i = X_i, \quad \forall i, \quad (6.40)
\]

\[
\hat{p}_i = \frac{p_i}{\sum_i p_i}, \quad \forall i, \quad (6.41)
\]

where each \( X_i \) is a random variable from the uniform distribution on the interval \((0, 1)\). As before, equation (6.41) ensures that these random probabilities sum to one.
Figure 6.14: Difference between original and perturbed objective functions versus the size of the reduced scenarios sets \( n \) for \( \epsilon = 10^{-5} \) and \( N = 100 \). Scenarios are initially assumed to have random probabilities.

Figure 6.14 shows the relative difference between the unperturbed and perturbed solutions for the three different algorithms for varying sizes of the reduced scenario sets \( n \). The original size of the scenario sets is fixed at \( N = 100 \), while the size of the perturbations is fixed at \( \epsilon = 10^{-5} \). Unlike Figure 6.10, the initial probabilities of the unperturbed data are not assumed equal. This graph shows that the relative difference for each of the three algorithms is of \( \mathcal{O}(10^{-5}) \) for \( \epsilon = 10^{-5} \).
6.7 Comparison of scenario reduction algorithms: effect on ROM

Figure 6.15 shows the same relative difference for $\epsilon = 10^{-10}$ for the same original (unperturbed) data. As with Figure 6.14, the initial probabilities of the unperturbed data are randomly generated. This graph shows that the relative difference for each of the three algorithms is of $O(10^{-10})$ for $\epsilon = 10^{-10}$. These two figures suggest that the relative difference for each of the three algorithms is approximately of $O(10^{-k})$ when $\epsilon$ is of $O(10^{-k})$, assuming $\epsilon$ is small and the probabilities are randomly generated.

Overall, Figures 6.10 - 6.15 all suggest that each of the three heuristic algorithms are stable with respect to small perturbations in both demands and probabilities. This is true regardless of the values of the initial probabilities in the unperturbed data. When the initial probabilities are assumed to be equal however, the three algorithms are more sensitive to small changes in data compared with unequal initial probabilities.

6.7 Comparison of scenario reduction algorithms: effect on ROM

Having analysed the stability of the three scenario reduction algorithms, their effect on ROM is now compared. This is done by comparing the results given by the model after applying the three different algorithms for $n = 1, \ldots, 6, 10$ and for $N = 1000$. As in Section 5.6, the demand scenarios were developed using the process described in Section 3.4 using actual demand and Seasonal Normal Demand (SND) for the year beginning on the 1st of April 2010. As these scenarios were randomly generated, the initial probabilities associated with them were all set to be equal. All
other parameters are as described in Section 5.3 while the experiments were repeated ten times. Figures 6.16 - 6.18 display the mean yearly Norway, IUK and BBL supplies according to ROM, with error bars, for each of the heuristic algorithms. As in Section 5.6 the error bars throughout this chapter are defined to be $\bar{x} \pm 1.96 \frac{\sigma}{\sqrt{N}}$, were $\bar{x}$ and $\sigma$ are the sample mean and sample standard deviation, respectively.\(^1\)

---

\(1\)If the data is normally distributed, then these error bars correspond to a 95% confidence interval.
Figure 6.17: Mean plus error bars for the total yearly supply through IUK for $n = 1, \ldots, 6, 10$ and $N = 1000$.

Figure 6.18: Mean plus error bars for the total yearly supply through BBL for $n = 1, \ldots, 6, 10$ and $N = 1000$. 
Figure 6.16 indicates that there is little difference between the three algorithms according to the results from ROM. The largest difference between the means in this plot is only approximately 0.3% of the corresponding mean. Figures 6.17 and 6.18 also indicate little difference between the algorithms. For \( n = 6 \) in Figure 6.18, the results obtained after using the backward reduction algorithm are significantly different from those from the other two algorithms. However, the largest differences between the means in Figures 6.17 and 6.18 are 4% and 0.4% respectively.

![Figure 6.19: Mean plus error bars for the total yearly supply from LNG for \( n = 1, \ldots, 6, 10 \) and \( N = 1000 \).](image)

Figure 6.19 shows the mean yearly LNG supplies as produced by ROM after using the three different scenario reduction algorithms. It indicates a clear difference between the backward reduction algorithm and the other two algorithms. This suggests that the heuristic algorithms do not always produce the same results. However, the maximum difference between the means displayed in Figure 6.19 is only 0.7%. Appendix D.3 contains similar graphs for the mean yearly UKCS flows, as well as the mean yearly injections and withdrawals associated with storage. As with Figures 6.16 - 6.19, they again indicate little difference between the three algorithms on the results of ROM. Overall, Figures 6.16 - 6.19 indicate some differences between the mean results obtained from ROM for the three different scenario reduction algorithms. However, these differences are relatively small.
6.8 Benefit of using a scenario reduction algorithm with ROM

Having shown that there is little difference between the backward reduction, simultaneous backward reduction and fast forward selection algorithms on the results using ROM, this section now looks at the benefit of using the fast forward selection algorithm on the model. This algorithm is chosen as it is the most accurate (see Section 6.4). Section 5.6 demonstrates the effect of increasing the number of scenarios used in ROM. It also details the increasing computational cost associated with ROM as the number of scenarios increases. This section shows that using scenario reduction techniques allows the effects of using a relatively large number of scenarios to be incorporated into ROM. This ensures significant savings in GAMS solve times. Figure 6.20 displays the mean plus error bars for the total yearly UKCS supply from ROM for scenarios sets produced from the fast forward selection algorithm for \( N = 100, 1000 \) and for scenarios sets produced without reduction. Figures 6.21 - 6.22 show similar results for the mean yearly injections into, and withdrawals out of, storage. As before, the parameters based on the April 2010 data described in Section 5.3 were used for each experiment, while the scenarios were developed from the stochastic process derived in Section 3.4. For each experiment, the model was run over a horizon of 365 days and for 365 rolls, while for each combination of \( n \) and \( N \), the experiments were repeated ten times to obtain error bars.

Figure 6.20: Mean plus error bars for the total yearly supply from the UKCS using the fast forward selection algorithm for \( N = 100, 1000 \) and with no reduction.
6. SCENARIO REDUCTION

Figure 6.21: Mean plus error bars for the total yearly injections to storage using the fast forward selection algorithm for $N = 100, 1000$ and with no reduction.

Figure 6.22: Mean plus error bars for the total yearly withdrawals from storage using the fast forward selection algorithm for $N = 100, 1000$ and with no reduction.
Figures 6.20 - 6.22 show that the results from using scenarios produced randomly (i.e., without scenario reduction) are significantly different to those produced after scenario reduction. Because the scenarios obtained after scenario reduction represent a large number of scenarios and hence more information, this shows that the effect of using a large number of scenarios can be usefully incorporated into ROM. This ensures that the computational cost associated with the model remains relatively low and hence shows the benefit of using scenario reduction techniques with ROM.

![Plot](image)

**Figure 6.23:** Mean plus error bars for the total yearly supply from through IUK using the fast forward selection algorithm for $N = 100, 1000$ and with no reduction.

Figure 6.23 displays similar results again, but for the mean yearly supplies through the IUK pipeline. In contrast to Figures 6.20 - 6.22, it indicates that the results obtained from the scenarios without scenario reduction are similar to those produced after scenario reduction. As a result, this graph suggests no obvious benefit to using scenario reduction techniques. Appendix D.4 shows similar plots for the mean yearly supplies from Norway, BBL and LNG. In comparison to Figures 6.20 - 6.22, the results for Norwegian and BBL supplies show the benefit of using scenario reduction. In contrast (but consistent with Figure 6.23), the results for LNG supplies show no obvious benefit from using the algorithms.

Figures 6.20 - 6.22 also show that the results produced from scenarios with $N = 100$ are different to those when $N = 1000$. This suggests that $N = 1000$ is the better choice for the initial number of scenarios, as more information is captured by $N = 1000$ than $N = 100$ scenarios. Moreover, because the results in the $N = 1000$ and $n = 3$ case are relatively similar to the results...
for $N = 1000$ and $n \geq 4$, this suggests that $n = 3$ is the best choice for the reduced number of scenarios. These choices of $N$ and $n$ are used in Chapter 7.

### 6.9 Summary

In this chapter, the various scenario reduction techniques developed in [13] and [14] were studied. Firstly, the theoretical background behind the derivation of the algorithms was considered. Secondly, the algorithms themselves were presented in Section 6.2 while in Section 6.3 applications of the algorithms seen in the literature were described. Following this, the backward reduction, simultaneous backward reduction and fast forward selection algorithms were all analysed in terms of accuracy, computational time and stability. Section 6.4 showed each of the heuristic algorithms to be highly accurate with the fast forward selection algorithm being the most accurate. Section 6.5 then analysed the computational times for the three algorithms with respect to stochastic gas demand scenarios. The results agreed with the theoretical results derived in [14]. The algorithms were also found to be stable with respect to small perturbations in the initial scenario sets in Section 6.6.

In Section 6.7 the effects of using these three algorithms with ROM were compared. The results showed minimal differences between the algorithms; consequently the fast forward selection algorithm was used in later analyses. Finally, in Section 6.8 the benefit of using scenario reduction in the model was analysed. These results indicated that using the fast forward selection algorithm allows the effect of using a large number of scenarios to be incorporated into ROM, whilst maintaining a relatively low GAMS solve time for the model.
Chapter 7

Applications of Rolling Optimisation Model

In this chapter, some of the applications of the Rolling Optimisation Model (ROM) are analysed. In particular, the model is used to examine the impact of different stresses that might occur in the UK gas market. These includes the effects of increasing and decreasing demand and the effects of three different possible shocks to the UK gas market. The first two of these shocks involve a sudden withdrawal of Liquified Natural Gas (LNG) supplies from the UK. The motivation for this comes from the UK National Grid: in their Development of Energy Scenarios document, [56], they state that LNG flows to the UK are “subject to high levels of uncertainty”. This is “due to the global nature of LNG and the options to flow to alternative markets”. The third shock to the UK gas market examined in this chapter, is a particularly cold week in January.

The chapter finishes with a section on the probability of supply being unable to meet demand in the UK gas market. This probability is obtained by running ROM numerous times in GAMS and counting the number of times the model is found to be infeasible. Infeasible solutions mean that it was impossible to satisfy all the constraints of the model. In ROM, there are three types of constraints: supply, demand and storage constraints. In this analysis, the level of demands, as well as LNG and Norwegian capacity, are placed under stress (i.e., varied). As a result, infeasible solutions in this analysis correspond to supply being unable to meet demand. The probability of supply being unable to meet demand is calculated by dividing the number of infeasible solutions by the total number of experiments.

In Chapter 5, the model was applied to parameters for the years beginning on the 1st of April 2010 and 2011. These analyses are not used in this chapter as they are both retrospective analyses and were compared with actual data. In this chapter, ROM is applied to parameters for the year beginning on the 1st of April 2012. This analysis is predictive and is hence suitable for testing stresses that may or may not occur in the UK gas market in the future.

This chapter is organised as follows: firstly, in Section 7.1, ROM is analysed under no stresses. The results of this analysis are used as a comparative base case for other analyses in this chapter.
7. APPLICATIONS OF ROLLING OPTIMISATION MODEL

Secondly, in Sections 7.2 and 7.3 ROM is used to examine increasing and decreasing demand scenarios, while in Section 7.4 the model is used to investigate the three different shocks that are modelled in the UK gas market. Following this, ROM is used in Section 7.5 to determine the probability of supply being insufficient for demand in the UK gas market. The chapter concludes in Section 7.6 with a summary.

7.1 Base case

In order to consider the different stress tests performed in this chapter, a base case must be analysed first. This was done by applying ROM to parameters for the year beginning on the 1st of April 2012. As in Sections 5.3 and 5.4 each roll of the model contained 3 scenarios \((n = 3)\) and was run over a horizon of \(D = 365\) days and for 365 rolls. The sets of three demand scenarios were generated using the stochastic process for demand described in Section 3.4. Seasonal Normal Demand (SND) for the year starting on the 1st of April 2010 was used for both actual demand and SND for the first roll (or optimisation) of the model. For the second roll of the model, the stochastic process for demand moved forward one day, i.e., SND for the year starting on the 2nd of April was used to develop the scenarios. For subsequent rolls, SND moved forward in a similar pattern. Initially \(N = 1000\) demand scenarios were generated. These scenarios were then reduced to \(n = 3\) using the fast forward selection algorithm. As a result, the probabilities associated with the scenarios were determined in accordance with the description of this algorithm in Section 6.2. This choice of the initial and reduced scenario set sizes follows from the analysis in Chapter 6.

The reason for using SND instead of actual demand in this analysis is that actual demand for the year starting on the 1st of April 2012 was not available at the time of writing. As mentioned in the introduction of Chapter 5 it is anticipated that ROM will be used by those in the UK gas market on a day-to-day basis, whereby one roll of the model will be optimised each day. Each day, those using the model will have available actual demand for that day, predicted demand for the next five days and SND thereafter. When actual demand is not available, predicted demand would be the first choice replacement. However, in this example predicted demand is also not available for the analysis. The use of SND instead of actual demand in the development of the demand scenarios in this chapter is in contrast to Section 5.3 and 5.4 where actual demand was available.

As in Section 5.4 the model was formulated with 17 sources of supply \((P = 17)\) and 3 storage facilities \((SO = 3)\). The 17 sources of supply represent the various sources of supply in the UK gas market broken up into multiple tranches representing varying costs, while the 3 storage operators represent the 3 different types of storage facilities in the UK, namely long, medium and short range storage (see Chapter 2). Table 7.1 shows the costs and maximum capacities associated with each tranche.

\(^1\)SND for the year starting on the 1st of April 2012 is not fully available at the time of writing.
Table 7.1: The costs and maximum capacities (in mcm) of the different tranches for each source of supply for the year starting on the 1st of April 2012.

<table>
<thead>
<tr>
<th>Tranche</th>
<th>$c_p$</th>
<th>$Q^{\text{max}}_p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>UKCS_1</td>
<td>58.464</td>
<td>48</td>
</tr>
<tr>
<td>UKCS_2</td>
<td>57.611</td>
<td>61</td>
</tr>
<tr>
<td>UKCS_3</td>
<td>56.552</td>
<td>37</td>
</tr>
<tr>
<td>Norway_1</td>
<td>60.985</td>
<td>21</td>
</tr>
<tr>
<td>Norway_2</td>
<td>64.322</td>
<td>54</td>
</tr>
<tr>
<td>Norway_3</td>
<td>48.375</td>
<td>40</td>
</tr>
<tr>
<td>Norway_4</td>
<td>58.637</td>
<td>30</td>
</tr>
<tr>
<td>LNG_1</td>
<td>55.242</td>
<td>38</td>
</tr>
<tr>
<td>LNG_2</td>
<td>58.865</td>
<td>36</td>
</tr>
<tr>
<td>LNG_3</td>
<td>71.458</td>
<td>28</td>
</tr>
<tr>
<td>BBL_1</td>
<td>64.189</td>
<td>11</td>
</tr>
<tr>
<td>BBL_2</td>
<td>55.084</td>
<td>16</td>
</tr>
<tr>
<td>BBL_3</td>
<td>59.352</td>
<td>14</td>
</tr>
<tr>
<td>IUK_1</td>
<td>61.626</td>
<td>11</td>
</tr>
<tr>
<td>IUK_2</td>
<td>69.543</td>
<td>16</td>
</tr>
<tr>
<td>IUK_3</td>
<td>70.977</td>
<td>27</td>
</tr>
<tr>
<td>IUK_4</td>
<td>60.336</td>
<td>18</td>
</tr>
</tbody>
</table>

Table 7.2: The total maximum capacity (in mcm) for each source of supply.

<table>
<thead>
<tr>
<th>Source of supply</th>
<th>Total maximum capacity</th>
</tr>
</thead>
<tbody>
<tr>
<td>UKCS</td>
<td>146</td>
</tr>
<tr>
<td>Norway</td>
<td>145</td>
</tr>
<tr>
<td>LNG</td>
<td>102</td>
</tr>
<tr>
<td>BBL</td>
<td>41</td>
</tr>
<tr>
<td>IUK</td>
<td>72</td>
</tr>
</tbody>
</table>

The total maximum capacities of the five different sources of supply (i.e., the sum of the tranches’ capacities) are given in Table 7.2. The capacities of Norway, the BBL and IUK pipelines are the same as in Table 5.10 and are in line with the figures supplied by the UK’s Department for Energy and Climate Change [7]. The total maximum capacity of LNG is supplied by the UK National Grid [2] and is also the same as in Table 5.10. In contrast, the total maximum capacity for the UKCS has decreased from 157 mcm, in Table 5.10, to 146 mcm in Table 7.2. This is consistent with the general downward trend seen in UKCS supplies since 2000 [16].
Table 7.3: Parameters associated with long-, medium- and short-range storage (in mcm) for the Rolling Optimisation Model for the year starting on the 1st of April 2012.

Table 7.3 provides the initial ($IntStor_{so}$), maximum ($MaxStor_{so}$) and minimum storage ($MinStor_{so}$) levels for each of the 3 facilities. It also displays the daily maximum injection ($I_{so}^{max}$) and withdrawal rates ($W_{so}^{max}$). The initial amounts of gas in storage are the actual observed levels for the 1st of April 2012, while the minimum levels are the actual minimum levels of storage observed in the UK from April 2010 to April 2012. These values were obtained from data made available from the UK National Grid [3]. The maximum storage levels, the maximum daily injection rate and the maximum daily withdrawal rate are taken from information provided in [7, 21]. They are updated to take into account the new MRS facility, Holford, as well as the de-commissioning of the SRS facilities, Glenmavis and Partington. The de-commissioning of these two facilities, in May 2011, has meant SRS has become insignificant to the UK gas market. As a result the amount of gas in SRS, as modelled by ROM, is not discussed in this chapter. The long- and short-run unit costs of storage injection and withdrawal, $a_{so,d}$ and $b_{so,d}$ respectively, are the same as in Table 5.7.
Figure 7.1: Demand profile obtained from ROM for the UK gas market starting on the 1st of April 2012.

Figure 7.1 displays how the different sources of supply meet demand in the UK gas market for the year beginning on the 1st of April 2012, as modelled by ROM. As with Figures 5.4 and 5.10 in Chapter 5, it predicts that the UKCS will be the largest source of supply in the UK gas market. Imports from Norway and LNG Imports will be the next two biggest sources, at roughly the same level, while imports through the BBL and IUK will be the fourth and fifth largest sources respectively.

Figure 7.2 shows the System Average Prices (SAPs) obtained from ROM for the year starting on the 1st of April 2012. As in Sections 5.3 and 5.4, these prices were calculated using equation 5.13 with $\alpha = 6.48$ and $\beta = 323.324$. These values of $\alpha$ and $\beta$ were determined following the calibration of predicted and actual SAPs (see Section 5.4). As with Figures 5.5 and 5.11, Figure 7.2 predicts that prices will follow a similar pattern to demand and be highly seasonal.
Figures 7.2, 7.3, and 7.4 display the amount of gas in storage, according to ROM, for the year beginning on the 1st of April 2012. Figure 7.2 follows a similar pattern to Figures 5.6 and 5.12 whereby gas is injected in the summer and withdrawn in the winter. However, in contrast to Figures 5.7 and 5.13, Figure 7.4 does not show much weekly or monthly variation in the amount of gas in MRS. This is because seasonal normal demand was used as the actual demand time series when the demand scenarios for this example were developed. Seasonal normal demand is a relatively smooth time series when compared to actual demand (see Figure 2.1 from Chapter 2). As it is unlikely that actual demand for the year starting on the 1st of April 2012 will be as smooth as seasonal normal demand, it is expected that the actual MRS will contain much more weekly/monthly variation.
7.1 Base case

Figure 7.3: Amount of gas in LRS produced by ROM for the year starting on the 1st of April 2012.

Figure 7.4: Amount of gas in MRS produced by ROM for the year starting on the 1st of April 2012.
7. APPLICATIONS OF ROLLING OPTIMISATION MODEL

7.2 Increasing demand scenarios

The first stress test examined in this chapter is for increasing demand scenarios. This was done by applying the model to the parameters for the year beginning on the 1st of April 2012. The first set of demand scenarios used in this section are from Section 7.1 and are known as the ‘base case’. The second set of demand scenarios were developed using a time series whereby each value in SND was increased by 5%. The subsequent demand scenarios were developed using time series whereby each value in SND was increased by 10%, 15% and 20%. Figure 7.5 shows the SND time series used to develop the different sets of scenarios for this analysis. All parameters used in this analysis are as described in Section 7.1.

Figure 7.5: Seasonal normal demand increased by 0%, 5%, 10%, 15% and 20%.

Figure 7.6 displays the total yearly supplies for the five different experiments obtained from ROM. It shows that supplies from the UK Continental Shelf (UKCS) remain relatively constant as SND increases. In contrast, the supplies of Norwegian imports, LNG and imports through the BBL and IUK pipelines are increased in order to meet the increased demand. These four sources of supply increase by 37%, 22%, 38% and 2105% respectively as SND increases by 20% of the base case. The large percentage increase in IUK supplies is because IUK supplies are small compared to the others in the base case.
Figure 7.6: Total yearly supplies obtained from ROM for increasing levels of SND.

Figure 7.7 shows the System Average Prices (SAPs) forecasted by ROM for various levels of increasing SND. It indicates that as SND increases, the SAP remains lower in the summer and higher the winter. When this plot is compared to Figure 7.5 it is noticeable that the level of demand is highly positively correlated with the SAPs. Figure 7.7 also shows that as demand increases, the increase in SAPs is much more sizeable in the winter than in the summer. As previously mentioned in Section 5.3, SAPs from ROM are obtained from Lagrange multipliers associated with demand constraints. Thus, when demand is high (low) in the winter (summer) so is the predicted SAP. As commented on in Chapter 5, the model is limited in predicting SAPs due to the neglect of other factors affecting natural gas prices in the UK [16].
Figures 7.8 and 7.9 display the amount of gas in LRS and MRS as obtained by ROM for increasing levels of SND. Figure 7.8 indicates that as demand increases, the timing and rate of withdrawals (from October onwards) remains very similar to the base case. This is because, in the base case, LRS is already withdrawing at its maximum rate. Thus, no more can be done to reduce the cost of meeting the increased demand. In contrast however, Figure 7.8 also shows that the timing of injections changes from around April to June as demand increases. Figure 7.9 indicates similar results for MRS.
Figure 7.8: The amount of gas in LRS produced by ROM for increasing levels of SND.

Figure 7.9: The amount of gas in MRS produced by ROM for increasing levels of SND.
7.3 Decreasing demand scenarios

In the previous section, ROM was used to stress test the model for increasing levels of SND. In this section, the obvious alternative of decreasing levels of SND is considered. As in the previous section, this is done by applying the model to the parameters for the year beginning on the 1st of April 2012, while the first set of demand scenarios (i.e., the base case) were again generated using SND for the year starting on the 1st of April 2010 for both actual demand and SND. Following this, the experiment was repeated four more times for stochastic processes of demand generated using time series whereby each value in SND was decreased by 5%, 10%, 15% and 20%. Figure 7.10 shows the different time series used to develop the different sets of scenarios for this analysis. All parameters used in this analysis are as described in Section 7.1.

![Figure 7.10: Seasonal normal demand decreased by 5%, 10%, 15% and 20%.

Figure 7.11 displays the total yearly supplies produced according to ROM for various levels of decreasing SND. It shows that supply from each of the sources decreases as SND decreases. The supplies from the UKCS, Norway, LNG and through the BBL pipeline decrease by 15%, 23%, 32% and 21%, respectively, as SND decreases by 20%. The amount of gas supplied through the IUK pipeline decreases to zero once SND decreases by 5%.
Figure 7.11: Total yearly supplies predicted by ROM for decreasing levels of SND.

Figure 7.12 displays SAPs obtained from ROM for decreasing levels of demand. When compared with Figure 7.10, it is again noticeable that these prices are heavily correlated with the level of demand, i.e., when the level of SND decreases so does the SAP. This is consistent with the findings in Figures 7.5 and 7.7. The other dramatic feature of Figure 7.12 is that when SND decreases by 20%, SAPs become relatively constant. As demand reduces, the number of different tranches of supply needed to meet demand also reduces. Once SND decreases by 20%, demand can be met solely by a small number of tranches, which means that the cost of meeting demand (and hence the SAP) does not vary greatly. As discussed previously in Section 5.3, gas prices in the UK are affected by many factors. As a result, the likelihood of SAPs remaining relatively constant, under any assumption about demand, is small. This suggests that if there was a significant decrease in demand (i.e., 20% or more) the model should be re-calibrated appropriately.
Figures 7.13 and 7.14 display the amount of gas in LRS and MRS as produced by ROM for decreasing levels of SND. Figure 7.13 indicates that LRS maintains its pattern of injecting in the relatively low-demand summer time and withdrawing in the relatively high-demand winter time. Consistent with Figure 7.8 the timing of withdrawals remains similar as demand decreases. Figure 7.13 also indicates that the timing of injections to LRS varies as demand decreases. In the base case, injections begin to occur in April while in the case of 20% decreased demand, injections begin to occur in July.

Figure 7.14 shows that as demand decreases the amount of gas injected into MRS decreases also. When demand decreases by 15%, ROM predicts that there will be very little gas injected into MRS. This is consistent with the prices displayed in Figure 7.12. In the cases of 15% and 20% decreased demand, prices remain relatively constant throughout the year. This means that there is little incentive (in terms of cost reduction) in using MRS. As discussed above, these prices, and hence the amount of gas in MRS displayed in Figure 7.14 are highly unrealistic. This suggests that if there was a large decrease in demand the model should be re-calibrated appropriately.
7.3 Decreasing demand scenarios

Figure 7.13: The amount of gas in LRS as produced by ROM for decreasing levels of SND.

Figure 7.14: The amount of gas in MRS as produced by ROM for decreasing levels of SND.
7.4 Shocks to UK gas market

Having stress tested ROM for both increasing and decreasing levels of demand, the model is now examined for three different shocks to the market. These are

1. A sudden drop in LNG supplies in the middle of July,
2. A sudden drop in LNG supplies at the start of January,
3. A cold week in January.

7.4.1 Sudden drop in LNG supplies in July

The first shock to the UK gas market to be analysed using ROM is a sudden drop in LNG supplies. The sudden shock is an unanticipated complete cessation of LNG supplies for one month beginning on the 15th of July. This is to model events such as a dockers strike in the UK or a sharp increase in the demand of LNG from other countries (e.g., in Japan after the 2011 earthquake). Once the shock occurs, the model assumes the cessation of LNG supplies will last for one month. The results for this analysis were obtained with, as in Sections 7.2 and 7.3 parameters for the year beginning on the 1st of April 2012. The demand scenarios were generated as in Section 7.1. For the first 105 rolls of the model (i.e., until the 15th of July) all parameters were as above and as described in Section 7.1. For the next 31 rolls the maximum capacities of the LNG tranches were set to zero for each day from the 15th of July to the 15th of August 2012. Once the 15th of August had passed, the maximum capacities of the LNG tranches returned to the original levels. Figure 7.15 shows the demand profile for this analysis.
When compared with Figure 7.15, Figure 7.15 shows that with a sudden loss of LNG supplies in July, Norway, BBL and MRS supplies make up the difference. Figures 7.16 and 7.17 show the amount of gas in LRS and MRS respectively for both the base case and the shocked LNG case. Figures 7.16 indicates that the shock has no effect on the amount of gas in LRS. In contrast, Figure 7.17 shows that when the drop in LNG occurs, MRS changes from injecting to withdrawing. Once the LNG supply levels return to normal, MRS begins injecting gas again and retains a similar pattern to the base case for the winter months.
Figure 7.16: Amount of gas in LRS as produced by ROM for both the base case and when there is a sudden drop in LNG supplies in July.

Figure 7.17: Amount of gas in MRS as produced by ROM for both the base case and when there is a sudden drop in LNG supplies in July.
Figure 7.18 displays the SAPs obtained from ROM for both the base case and the case where there is a sudden drop in LNG supplies in July. It shows that before the 15th of July, the SAPs are identical for both cases. Once the LNG supplies are removed however, there is an increase in the SAP. The prices return to the base case level once the LNG supplies return to the normal level.

7.4.2 Sudden drop in LNG supplies in January

The second shock to the UK gas market analysed in this section is also a sudden drop in LNG supplies. However, instead of a month in summer, the drop in LNG occurs for a week starting on the 1st of January. This is done to examine the effect of the cessation in the relatively high-demand winter time which is in contrast to Section 7.4.1 where the cessation occurred in the relatively low-demand summer time. The results for this analysis were obtained by applying the model to the parameters for the year beginning on the 1st of April 2012. For the first 276 rolls of the model (i.e., until the 1st of January), all the parameters were as described in Section 7.1. The maximum capacities of the LNG tranches were set to zero for the next seven rolls, i.e., for each day from the 1st of January to the 7th of January. After that, the maximum capacities of the LNG tranches returned to the original levels. The market has no prior knowledge of the loss, which is in contrast to Section 7.4.3 below.
Figure 7.19: Demand profile obtained from ROM for the UK gas market when there is a sudden drop in LNG supplies in January.

Figure 7.19 shows the demand profile predicted by ROM for this case. When compared with Figure 7.1, this indicates that withdrawals from MRS and IUK supplies replace the removed LNG supplies in the affected week in January. Figures 7.20 and 7.21 show the amount of gas in LRS and MRS, respectively, for both the base case and the shocked case. Figure 7.21 shows that there is no dramatic change in the amount of gas in LRS between the two cases. The reason for this is that LRS is withdrawing at (or near) its maximum capacity before the loss of LNG. Thus when the sudden drop in LNG occurs, LRS can do no more but continue to withdraw at its maximum rate. In contrast, Figure 7.21 shows that MRS is not withdrawing at its maximum capacity in the base case. As a result, when the sudden drop in LNG occurs, withdrawals from MRS increase sharply for that week.
Figure 7.20: Amount of gas in LRS as produced by ROM for both the base case and when there is a sudden drop in LNG supplies in January.

Figure 7.21: Amount of gas in MRS as produced by ROM for both the base case and when there is a sudden drop in LNG supplies in January.
Figure 7.22 displays the SAPs obtained from ROM for the base case and the case with a sudden drop in LNG capacity for a week in January. As expected, the SAPs are identical in the two cases up until the 1st of January. When the sudden loss in LNG capacity occurs, the SAPs rise. This increase persists past the 7th of January and lasts approximately for three months. The reason for this prolonged increase is the reduced levels of gas available from MRS resulting from the loss of LNG for the first week in January. The reduced amount of gas in MRS means that more expensive sources of supply are needed to meet demand in the months following the reduction in LNG supplies, thus increasing SAPs.

![Figure 7.22: SAPs obtained from ROM for both the base case and when there is a sudden drop in LNG supplies in January.](image)

**7.4.3 Cold week in January**

The final shock to the UK gas market examined in this chapter is a particularly cold week at the start of January. This cold week is modelled by assuming a sharp rise in the level of demand. As in Sections 7.2 - 7.4.2, the results for these analyses were obtained by applying the model to the parameters for the year beginning on the 1st of April 2012. In contrast to the two shocks considered in Sections 7.4.1 and 7.4.2, the demand scenarios used in the analysis of this section were developed using a time series consisting of Cold SND for a week in January and SND at all other times. By definition, SND is the demand expected in seasonal normal weather conditions, while Cold SND is the demand expected in particularly cold weather conditions. It is generated by the UK National Grid and is available from their website, [17]. Figure 7.23 shows the time series
used to develop the demand scenarios used in this analysis, along with SND. All the parameters used in this analysis are as described in Section 7.1.

![Graph showing SND and SND with cold SND for a week in Jan](image)

Figure 7.23: Time series of SND and SND with a cold week at the start of January.
Figure 7.24: Demand profile obtained from ROM for the UK gas market when there is a cold week in January.

Figure 7.24 displays the demand profile obtained from ROM with a cold week in January. When compared with Figure 7.1, it indicates that withdrawals from MRS are the main source of supply to increase as demand increases for the cold week. Figures 7.25 and 7.26 display the amount of gas in LRS and MRS, respectively, for the base case and the case with a cold week in January. Figure 7.25 shows no dramatic difference between the two cases. As with Figure 7.20, the reason for this is that LRS is withdrawing at (or near) its maximum capacity throughout December and January in the base case. Thus when the cold week arrives, LRS can do no more but continue to withdraw at its maximum rate. In contrast, Figure 7.26 shows a sharp rise in the rate of withdrawals from MRS once the cold snap arrives. This is similar to Figure 7.21, where there was a sudden drop in LNG supplies. However, in contrast to Figure 7.21, Figure 7.26 also shows a sharp rise in injections to storage before the cold snap. This occurs because the stochastic process for demand has a limited foresight of five days ahead (see Section 3.4). As a result, the model starts to take into account the increased demand five days before the cold snap. This means that MRS can prepare for the cold week five days before it happens, hence the sharp rise in injections to MRS before the 1st of January.
7.4 Shocks to UK gas market

Figure 7.25: Amount of gas in LRS as obtained by ROM for both the base case and when there is a cold week in January.

Figure 7.26: Amount of gas in MRS as obtained by ROM for both the base case and when there is a cold week in January.
Figure 7.27: SAPs obtained from ROM for both the base case and when there is a cold week in January.

Figure 7.27 displays SAPs obtained from ROM for the base case and the case with a particularly cold week in January. It indicates that the SAPs in the two cases are very similar up until the 27th of December (i.e., five days before the shock). Once information about the cold snap becomes available (i.e., weather forecasts), an increase in the SAP can be seen for the cold week case. This is because demand, and hence the cost of meeting demand, has increased. As in Figure 7.22, Figure 7.27 shows that the rise in the SAP lasts much longer than the first week in January. As before, the reason for this prolonged increase is the reduced levels of gas available from storage resulting from the cold week in January. The reduced amount of gas in storage means that more expensive sources of supply are needed to meet demand in the months following the cold snap.

7.5 Probability of supply being unable to meet demand

The final application of ROM considered in this chapter is to examine the probability $p$ that there is insufficient supply to meet demand in the UK gas market for at least one day in a year. As in all the previous sections of this chapter, this was done by applying ROM for the parameters for the year beginning on the 1st of April 2012. The probability of system failure is obtained by running one roll (or optimisation) of the model, with a 365 day horizon, with 1 scenario and repeating this either 100 or 1000 times. As before, each roll of the model corresponds to solving the optimisation

1Throughout this section the probability of system failure is defined as the probability of supply being unable to meet demand.
model once. The probability of supply being unable to meet demand is defined by dividing the total number of infeasible solutions (i.e., those where when supply < demand on at least one day in the year) by the total number of experiments. For each of the experiments there is a different stochastic demand scenario, generated as in Section 3.4.

There are a number of factors affecting the probability considered in this section. These are increasing levels of demand, the capacity of the different sources of supply, the time of year, and the initial amount of gas in storage. For a 0% increase in demand the scenarios were developed as in Section 7.1. For a \( \alpha \)% increase in demand the demand scenarios were developed in a manner similar to Section 7.2 using demand time series equal to SND multiplied by \((100 + \alpha)\%\). The two different sources of supply chosen to vary in this section are LNG and Norwegian supplies. These two sources are foreign sources of supply to the UK, which means that their availability are, at times, subject to high levels of uncertainty. This is because of factors such as, for example, demand from other regions and political crises in the Middle East. Imports through the BBL and IUK pipelines are also foreign sources of supply to the UK. However, as these two sources of supply are relatively small compared to LNG and Norway (see Chapter 2), their loss is not considered.

For Norway, the reduction was achieved in the model by decreasing the capacities of each of the Norwegian tranches (Table 7.1) to \( \beta \)%%. The values presented in Table 7.1 correspond to 100% Norwegian capacity. As described in Section 5.3 the level of LNG capacity presented in Table 7.2 is assumed at 70% of its true maximum capacity. LNG capacity of 100% corresponds to the unlikely situation where LNG ships are constantly coming in and out of the UK. The assumption of 70% is in line with the level assumed by the UK National Grid, [56, 57]. When LNG Capacity was set to \( \beta \)% in this section, the values in Table 7.1 are multiplied by \( \frac{\beta}{70} \). All other parameters used in this section are as described in Section 7.1.

Throughout this section the probability (\( p \)) of supply being unable to meet demand is defined to be highly unlikely if \( p \leq 0.001 \), possible if \( p \geq 0.001 \), and likely if \( p > 0.5 \). Despite this, it must be acknowledged that, system failure is always possible in the UK gas market even if the results in this section indicate the probability of it happening to be \( p \leq 0.001 \).

### 7.5.1 Varying demand, LNG capacity and Norwegian capacity

The Rolling Optimisation Model is first used to examine the probability of supply being unable to meet demand for varying levels of demand and LNG capacity. Table 7.4 shows the proportion of times the model was found to be infeasible out of 1000 experiments for increasing levels of demand and adjusted LNG capacities, while Figure 7.28 shows the corresponding heat map. All other parameters are as described in Section 7.1.

\[ \text{1} \] Throughout this chapter heat maps provide a graphical representation of the probability of system failure in the UK gas market for the various levels of demand and supply.
Table 7.4 and Figure 7.28 show that under current assumptions for SND (i.e., 0% increase) it is highly unlikely ($p \leq 0.001$) that supply would be unable to meet demand. This holds regardless of the level of LNG capacity. At the current assumed levels of LNG capacity (i.e., 70%), the probability of supply being unable to meet demand does not become possible until SND increases by 30% ($p = 0.06$). When SND increases by 50% the event of system failure becomes likely, at current assumed levels of LNG, as $p$ increases to 0.64.
Table 7.5: Fraction of infeasible solutions for increased SND and/or changes in Norwegian capacity.

Figure 7.29: Fraction of infeasible solutions for increased SND and/or changes in Norwegian capacity.

Table 7.5 and Figure 7.29 show a similar analysis, but with adjusted Norway capacities instead of LNG. All other parameters are as described in Section 7.1, which means that LNG is assumed to be at 70% of its true capacity. The analysis shows that under current assumptions for both SND (0% increase) and Norwegian capacity (100%), it is highly unlikely ($p < 0.001$) that supply would be unable to meet demand. This holds as long as Norwegian capacity is at least at 50% of its maximum. Once the capacity of Norwegian supplies goes below this level, the probability of system failure starts to increase from $p = 0.002$ to $p = 0.08$. This suggests that system failure is possible, but unlikely, under current assumptions for SND, if the Norwegian capacity is reduced.
significantly. When SND increases to 30%, system failure is also only likely once Norwegian capacity reduces significantly. Under current assumptions for Norwegian capacities (i.e., 100%), the likelihood of supply being insufficient for demand is possible \( (p \geq 0.001) \) once SND increases by 20% and likely \( (p \geq 0.5) \) once it increases by 50%.

<table>
<thead>
<tr>
<th>Norway capacity</th>
<th>LNG capacity</th>
<th>0%</th>
<th>10%</th>
<th>20%</th>
<th>30%</th>
<th>40%</th>
<th>50%</th>
<th>60%</th>
<th>70%</th>
<th>80%</th>
<th>90%</th>
<th>100%</th>
</tr>
</thead>
<tbody>
<tr>
<td>100%</td>
<td>0.001</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>90%</td>
<td>0.003</td>
<td>0.003</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>80%</td>
<td>0.01</td>
<td>0.001</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>70%</td>
<td>0.03</td>
<td>0.01</td>
<td>0.01</td>
<td>0.002</td>
<td>0.001</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>60%</td>
<td>0.05</td>
<td>0.001</td>
<td>0.01</td>
<td>0.003</td>
<td>0.001</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>50%</td>
<td>0.10</td>
<td>0.03</td>
<td>0.01</td>
<td>0.01</td>
<td>0.003</td>
<td>0.001</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>40%</td>
<td>0.20</td>
<td>0.06</td>
<td>0.03</td>
<td>0.02</td>
<td>0.01</td>
<td>0.01</td>
<td>0.002</td>
<td>0.001</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>30%</td>
<td>0.35</td>
<td>0.12</td>
<td>0.07</td>
<td>0.04</td>
<td>0.02</td>
<td>0.01</td>
<td>0.01</td>
<td>0.003</td>
<td>0.001</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20%</td>
<td>0.59</td>
<td>0.23</td>
<td>0.15</td>
<td>0.08</td>
<td>0.05</td>
<td>0.03</td>
<td>0.01</td>
<td>0.01</td>
<td>0.003</td>
<td>0.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10%</td>
<td>0.84</td>
<td>0.39</td>
<td>0.26</td>
<td>0.17</td>
<td>0.10</td>
<td>0.06</td>
<td>0.03</td>
<td>0.02</td>
<td>0.01</td>
<td>0.003</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0%</td>
<td>0.96</td>
<td>0.65</td>
<td>0.45</td>
<td>0.29</td>
<td>0.20</td>
<td>0.11</td>
<td>0.07</td>
<td>0.04</td>
<td>0.02</td>
<td>0.01</td>
<td>0.01</td>
<td></td>
</tr>
</tbody>
</table>

Table 7.6: Fraction of infeasible solutions with various changes to LNG and Norwegian capacity.

![Figure 7.30: Fraction of infeasible solutions with various changes to LNG and Norwegian capacity.](image)

Table 7.6 and Figure 7.30 show a similar analysis, but for adjusted levels of LNG and Norwegian capacity. For demand, no increase in SND is assumed, while all other parameters are
as described in Section 7.1. The results show that under the current assumed level of LNG capacity of 70%, it is highly unlikely \((p < 0.001)\) that supply will be unable to meet demand as long as Norwegian capacity is at least at 60% of its maximum. Once the capacity of Norwegian supplies goes below this level, system failure becomes possible, but still with small probability \((0.002 \leq p \leq 0.08)\). Similarly, under current assumptions regarding Norwegian capacity (100%), the probability of supply being unable to meet demand is extremely low \((p < 0.001)\) as long as LNG capacity is at least 30%. Once LNG capacity goes below 30%, system failure becomes possible with small probability as \(0.001 \leq p \leq 0.02\).

When LNG is completely removed from the UK gas market (i.e., 0% capacity), system failure is possible \((p \geq 0.001)\) regardless of the level of Norwegian capacity and likely \((p \geq 0.5)\) once the capacity of Norwegian supplies reduces to 20%. Similarly, when Norway supplies are removed from the market, system failure is possible \((p \geq 0.001)\) regardless of the level of LNG, and likely \((p \geq 0.5)\) once LNG reduces to 30%.

### 7.5.2 Varying the start date

As described in the previous section, the results presented in Tables 7.4 - 7.6 were developed using a SND time series that starts on the 1st of April. The following analysis investigates whether this choice of start date affects the results. Figure 7.31 shows the number of infeasible solutions found out of 100 experiments for a year’s worth of different start dates, for two different methods. The demand scenarios for each of these experiments were generated as in Section 7.1. For both methods, the capacities for LNG and Norway were held at 30% and 10% respectively. The difference between the methods is in the way the initial amount of gas in storage is derived, while all other parameters are the same as those used in Section 7.1.

For the first method in Figure 7.31, the initial amount of gas in storage was obtained from the results of the previous day. For example, the initial amount of gas in storage for the 2nd of April, for the first experiment, was obtained from the injections and withdrawals on the scenario independent day, \(d_1\), from the first experiment run on the 1st of April. For the 1st of April, the initial amount of gas in storage was at the levels reported by the UK National Grid’s website, [3], for the 1st of April 2012. This method corresponds with the method used to update ROM as described in Section 5.2.1.

As a large number of the experiments were found to be infeasible, the results of this method are questionable.\(^1\) As a result, the initial amount of gas in storage for each day for the second method in Figure 7.31 was obtained from actual data from the UK National Grid’s website, [3], starting on the 1st of April 2011 and ending on the 31st of March 2012.

\(^1\) Despite GAMS finding the problem infeasible, it still provides results for the outputs of the model and flags the unsatisfied constraints, [49]. In this case, these results are used to update the model.
Figure 7.31: Number of infeasible solutions with for different start days, 30% LNG capacity and 10% Norwegian capacity.

Figure 7.31 shows that, outside of winter, the number of the solutions found to be infeasible was relatively constant at around 30 (out of 100) for both methods. In winter however, the number of infeasible solutions found differs for both methods. For the first method, the probability of system failure starts to increase in November, while for the second method the probability stays around the 0.3 level. As both methods have different ways of determining the initial amount of gas in storage, this suggests that the initial amount of gas in storage affects the number of infeasible solutions found in winter months.

7.5.3 Varying the initial amount of gas in storage

The analysis in the previous section suggests that the initial amount of gas in storage influences the probability of supply being unable to meet demand. Figure 7.32 shows the number of infeasible solutions found out of 100 experiments with varying levels of the initial amount of gas in storage. This is done for two different start dates, the 1st of April and the 1st of January. For the 1st of April case, the demand scenarios associated with each experiment were generated in the same manner as Section 7.1. For the 1st of January case, the demand scenarios were generated in a similar manner using the process described in Section 3.4 but with the SND time series starting on the 1st of January 2010. As in the previous section, LNG capacity and Norwegian capacity are held at 30%
and 10% respectively. The initial amount of gas in each storage facility was varied as follows:

\[ IntStor_{so} = MinStor_{so} + \rho (MaxStor_{so} - MinStor_{so}), \]  

(7.1)

where \( \rho \) takes values in the set \( \{ 0, 0.05, 0.1, 0.15, 0.2, ..., 1 \} \). The variables \( IntStor_{so}, MinStor_{so} \) and \( MaxStor_{so} \) are as defined in Section 7.1 and represent the initial, minimum and maximum amount of gas in storage for storage facility \( so \). All other parameters are as presented in Section 7.1.

Figure 7.32: Fraction of infeasible solutions starting on 2 different days with varying initial amounts of gas in storage; 10% LNG capacity and 20% Norwegian capacity.

Figure 7.32 shows that when starting on the 1st of April, the initial amount of gas in storage does not influence the probability of system failure. This probability is constant, at 0.35. For the 1st January, the probability of supply being unable to meet demand varies greatly as the initial amount of gas in storage (i.e., \( \rho \)) changes.

This analysis confirms that the initial amount of gas in storage heavily affects the probability of supply being unable to meet demand, provided the initial day is in winter. If the start date is outside of winter it does not affect it as much. These results are not surprising, since it is clear that the UK gas market is most likely to fail to meet demand in winter. Thus, if storage levels are low outside of winter, there is plenty of time for injections before the high demand season comes. If, however, storage levels are low in winter, there is no time for injections, as storage facilities are required to withdraw gas in order to meet the high demand.
7.5.4 Start date in January and effect of a particularly cold week

In Section 7.5.2, it was shown that the probability of supply being unable to meet demand in the UK gas market, as determined by ROM, is dependent on the start date of the model. The analysis in Section 7.5.1 used demand scenarios that only started on the 1st of April. In contrast, the analysis in Section 7.5.2 used varying start dates for the demand scenarios, but only for certain reduced levels of LNG and Norwegian capacity. In this section, the probability of there being insufficient supply to meet demand is examined for a start date in January and for varying levels of LNG and Norwegian capacity. The effect of adding the additional stress of a particularly cold week in January, which brings about increased demand, is also examined. This was done by considering two different sets of demand scenarios developed using the process described in Section 3.4. The first were developed using SND for the year beginning on the 1st of January 2010, while the second set were developed from a time series that starts with cold SND for one week and finishes with SND for the rest of the year. Figure 7.33 shows the two time series used to develop the two different sets of stochastic demand scenarios. The initial levels of gas in storage, for the three different types of storage facilities, were all taken from actual levels supplied by the UK National Grid, [3], for the 1st of January 2012. All other parameters are as described in Section 7.1.

Figure 7.33: Time series with cold SND for the first week and SND thereafter and time series of SND.

The time series with a dashed line in Figure 7.33 represents a situation where the UK natural gas market is experiencing a high demand period at the start of January. In this period, the market knows that demand is high and that it is highly to stay like that for a couple days. Thereafter
however, the market does not know what demand will be, and can only assume SND. The other

time series in Figure 7.33 represents a situation where demand is assumed to be at SND levels. The two
time series differ only in the first week of January.

Table 7.7 and Figure 7.34 show the probabilities of supply being unable to meet demand for

scenarios based around SND and for those assuming a cold week. In contrast, Table 7.8 and Figure

7.35 show the probabilities for scenarios developed using the time series with the dashed line in

Figure 7.33.

<table>
<thead>
<tr>
<th>LNG capacity</th>
<th>0%</th>
<th>10%</th>
<th>20%</th>
<th>30%</th>
<th>40%</th>
<th>50%</th>
<th>60%</th>
<th>70%</th>
<th>80%</th>
<th>90%</th>
<th>100%</th>
</tr>
</thead>
<tbody>
<tr>
<td>100%</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>90%</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>80%</td>
<td>0.002 0.001</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>70%</td>
<td>0.01 0.003 0.001</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>60%</td>
<td>0.03 0.01 0.004 0.001</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>50%</td>
<td>0.04 0.03 0.02 0.01 0.002 0.001</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>40%</td>
<td>0.09 0.05 0.03 0.02 0.01 0.003 0.001</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>30%</td>
<td>0.18 0.10 0.06 0.04 0.03 0.01 0.004 0.001</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>20%</td>
<td>0.34 0.21 0.12 0.08 0.04 0.03 0.02 0.01 0.001</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>10%</td>
<td>0.57 0.39 0.25 0.15 0.09 0.05 0.03 0.02 0.01 0.003 0.001</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0%</td>
<td>1</td>
<td>0.63</td>
<td>0.45</td>
<td>0.28</td>
<td>0.17</td>
<td>0.10</td>
<td>0.06</td>
<td>0.04</td>
<td>0.03</td>
<td>0.01</td>
<td>0.004</td>
</tr>
</tbody>
</table>

Table 7.7: Fraction of infeasible solutions for changing LNG and Norwegian capacity for a day starting in January based on SND.

<table>
<thead>
<tr>
<th>LNG capacity</th>
<th>0%</th>
<th>10%</th>
<th>20%</th>
<th>30%</th>
<th>40%</th>
<th>50%</th>
<th>60%</th>
<th>70%</th>
<th>80%</th>
<th>90%</th>
<th>100%</th>
</tr>
</thead>
<tbody>
<tr>
<td>100%</td>
<td>0.004 0.003 0.002 0.002 0.001</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>90%</td>
<td>0.01 0.01 0.004 0.002 0.002 0.001 0.001</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>80%</td>
<td>0.02 0.01 0.004 0.004 0.002 0.002 0.001 0.001</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>70%</td>
<td>0.07 0.03 0.01 0.01 0.004 0.003 0.002 0.002 0.002 0.001</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>60%</td>
<td>0.16 0.08 0.04 0.02 0.01 0.01 0.003 0.002 0.002 0.001</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>50%</td>
<td>0.35 0.20 0.10 0.05 0.02 0.01 0.004 0.002 0.002 0.001</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>40%</td>
<td>0.62 0.41 0.24 0.12 0.07 0.03 0.01 0.01 0.004 0.003 0.002</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>30%</td>
<td>1   0.69 0.46 0.28 0.15 0.07 0.04 0.01 0.01 0.01 0.003</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>20%</td>
<td>0.99 1 0.76 0.53 0.34 0.19 0.09 0.05 0.02 0.01 0.01</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>10%</td>
<td>0.999 0.99 1 0.85 0.62 0.40 0.23 0.12 0.06 0.03 0.01</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0%</td>
<td>1   0.999 0.997</td>
<td>1</td>
<td>0.89</td>
<td>0.68</td>
<td>0.45</td>
<td>0.27</td>
<td>0.15</td>
<td>0.07</td>
<td>0.04</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 7.8: Fraction of infeasible solutions for changing LNG and Norwegian capacity for a day starting in January based on cold SND.
Figure 7.34: Fraction of infeasible solutions for changing levels of LNG capacity and Norwegian capacity for a day starting in January based on SND.

Figure 7.35: Fraction of infeasible solutions for changing levels of LNG capacity and Norwegian capacity for a day starting in January based on cold SND.
7.6 Summary

In this chapter, ROM was used to investigate the effect of various potential stresses on the UK gas market. This was done by applying the model to parameters for the year starting on the 1st of April 2012. The first stress test examined increasing demand scenarios in Section 7.2. This analysis showed that as yearly demand increases the supplies from LNG, Norway and through the BBL and IUK pipelines also increase. As demand increases, the amount of gas in LRS and MRS remains relatively constant, while the SAPs also increase. A similar analysis in Section 7.3 indicates that supplies from each source of supply and withdrawals from MRS all decrease as demand decreases. In terms of SAPs, the results also show that prices decrease in correlation with demand. Figures 7.12 and 7.14 indicate, however, that when SND decreases by 15% or more, SAPs and the amount of gas in MRS both become highly unrealistic. This suggests that if there was a large decrease in demand (i.e., 15% or more), then the model should be re-parameterised.

Section 7.4 details how ROM was used to examine three different shocks that may happen to the UK gas market: a sudden drop in LNG supplies in the middle of July, a sudden drop in LNG supplies at the start of January, and a cold week in January. In the first of these cases it was found that MRS and UKCS supplies make up the shortfall. In the case of a sudden drop in LNG in January, MRS is again the main source of supply that makes up the shortfall. In the case of a cold week in January, similar results again are seen. In contrast to the earlier cases however, the market
is able to anticipate the increased demand by increasing the amount of gas in MRS just before
the cold snap. In each of these cases, SAPs sharply increase once the shock to the market occurs.
However, once the shock is over, the SAPs remain relatively high due to the decreased amount of
gas in MRS.

The final analysis of this chapter regards the probability of supply being unable to meet demand
in the UK gas market, as determined by ROM. This analysis showed that this probability depends
on a number of factors such as capacity levels, demand levels, the time of year and the initial
amount of gas in storage. Under current assumptions, it was found that the event of system failure
is highly unlikely. The analysis also shows that the system remains robust until demand is increased
substantially and/or there is a large decrease in the capacities of some of the sources of supply.
Chapter 8

Summary

To complete this thesis, a summary is now presented and possible future work is discussed. The aim of this project was to model the flows of gas in the UK natural gas market whilst incorporating the stochastic nature of demand. Previously, a mathematical model of this market had not appeared in the literature. In Chapter 3 a stochastic process modelling demand was developed. It captures the daily information on demand that is known by the players in the UK gas market. On any given day, demand for that day is exactly known. For the next five days ahead demand is known only with increasing uncertainty, while after five days ahead all that is known is Seasonal Normal Demand (SND). The stochastic process for demand was informed by the statistical analysis of data from the 1st of October 2007 to the 31st of March 2012. The process is based on the statistical relationships established in Section 3.3 between historic actual demand and historic predicted demand. This analysis found that the error in predicting UK gas demand for one to five days ahead is correlated with the previous day’s error. Beyond the fifth day ahead, and for the rest of the process, the stochastic process for demand is based on the statistical relationship established between actual demand and Seasonal Normal Demand (SND) in Section 3.2. This analysis found that the difference between the log of actual demand and the log of SND follows an auto-regressive process of order one.

Before a detailed model of the UK gas market incorporating this stochastic process for demand was introduced, the different methodologies used to model gas markets were studied in Chapter 4. There are two main methods used to model gas markets namely, complementarity-based equilibrium models and cost minimisation models. Multiple objective functions can be included in complementarity-based models. This allows each player in the market to maximise their own profits through their own individual objective functions. This allows for market power (i.e., the power for one or more players in the model to influence the level of demand) in the model. In cost minimisation models, instead of each player maximising their own profits separately, there is an imaginary central planner that chooses how the different players in the market meet demand at least cost to the overall system. This is done using a single linear program and hence has only one objective function. As a result, market power cannot be incorporated into these types of models.
In Section 4.2.1 a complementarity-based equilibrium model is introduced. It is a simplified version of Gabriel et al’s model [11] and assumes that there is no market power. Market power is excluded as it is not appropriate, in most reasonable cases, in the context of the UK market, [12]. In Section 4.3.1 a corresponding cost minimisation model is presented. In Sections 4.4 and 4.5 these two models are proven to be equivalent. In Section 4.4 this is done by showing that the KKT conditions for optimality under both models reduce to the same set of equations, while in Section 4.5 it is shown using the principle of symmetry [28]. A detailed analysis of this equivalence had not been previously seen.

Cost minimisation, with its imaginary central planner, may be seen as an unintuitive approach to modelling natural gas markets. Complementarity-based equilibrium models are more intuitive as each player acts separately in order to maximise their own profits. They are, however, algebraically and structurally more complicated. This result shows that while cost minimisation may be unintuitive for natural gas markets, it can be used instead of the more intuitive, but complex, complementarity-based equilibrium approach when modelling markets without market power.

Using the results from Chapters 3 and 4 a detailed model of the UK gas market is introduced in Chapter 5. This model is known as the Rolling Optimisation Model (ROM) and constitutes the main result of the thesis. It is an extension of the deterministic cost minimisation model described in Section 4.3.1 and takes as an input demand scenarios simulated from the stochastic process for demand developed in Chapter 3. ROM replicates the daily decisions made in the UK gas market under uncertain information about future demand. The model decides the amount of gas to be injected to, or withdrawn from, the different storage facilities considered in the model, as well as how the different sources of supply meet the exactly known demand on the first day of the stochastic process. This is done at minimum cost whilst also ensuring all possible future demands are also met at minimum expected cost. Once these decisions are made the model moves to a new day where the same decisions are made using updated demand scenarios. Once these decisions are made, the model moves forward again in a similar manner and so on (see Figure 5.1). In Section 5.3 the parameters of ROM were chosen so as to best fit data from the UK gas market for the year beginning on the 1st of April 2010. The results of this analysis found that flows of gas simulated by ROM were qualitatively similar to actual flows. These flows included how the different sources of supply met demand, as well as the amount of gas in the different types of storage facilities. ROM was also used to simulate System Average Prices (SAPs) in the UK. Using similar parameters, ROM was then tested in Section 5.4 for data from the UK gas market for the year starting on the 1st of April 2011. This analysis again showed that results obtained from ROM fitted reasonably well to actual data. In Section 5.6 the effect of increasing the number of demand scenarios used in the model was examined. This analysis found that as the number of scenarios used in ROM increased, the amount of gas being supplied by the different sources of supply varied.

Ideally ROM should be run with as many demand scenarios as possible. However, Section 5.6 also noted the increasing computational cost associated with ROM as the number of scenarios increases. This provided the motivation for the introduction of scenario reduction techniques in
Chapter 6. The aim of this analysis was to use demand scenarios that accurately represent larger scenario sets, whilst at the same time maintaining a relatively low computational time for ROM. Three different scenario reduction techniques, based on the works described in [13, 14], were introduced in Chapter 6. These heuristic algorithms are known as backward reduction, simultaneous reduction and fast forward selection.

For the first time, these three algorithms were applied to gas demand scenarios developed using the stochastic process for demand of Chapter 3. It was found that the algorithms are all highly accurate and stable with respect to small perturbations in the data. The computational times associated with the three algorithms were also found to agree with the theoretical results derived in [14]. In Section 6.7, the effects of using these three algorithms on ROM were compared. The results showed minimal differences between the three algorithms. In Section 6.8, the benefit of using scenario reduction on the model was analysed. The results produced from ROM indicated that using a small number of scenarios generated from the fast forward selection algorithm were different to the results produced using a small number of scenarios generated without any scenario reduction technique. This showed the benefit of using the scenarios obtained from scenario reduction over using the same (small) number of randomly generated scenarios.

Following the scenario reduction chapter, ROM is used to investigate the effect of various stresses on the UK gas market in Chapter 7. The first stress test examined was for increased demand scenarios. This analysis showed that as yearly demand increases the supplies from Norway and through the BBL and IUK pipelines also increase. This analysis also showed that as demand increases, the amount of gas in storage remained relatively consistent, while the SAPs increase in correlation with demand. A similar analysis for decreasing demand scenarios showed that as demand decreased the amount of gas supplied from each of the different sources of supply decreased also. This analysis also showed that when demand decreased substantially the amount of gas in Medium-Range-Storage (MRS) became highly unrealistic suggesting that ROM should be re-parameterised if large changes occur in demand.

Following this analysis, ROM was then used to examine three shocks that may occur in the UK gas market. These shocks were a drop in LNG supplies in both July and January, as well as an extreme cold week in winter. Each of these shocks found that MRS was the main source of supply that made up the required short fall. In each of the three cases, SAPs were also found to sharply increase once the shock to the market has occurred. However, once the shock is over, the SAPs remained relatively high due to the decreased amount of gas in storage.

The final analysis of Chapter 7 was on the probability of supply being unable to meet demand in the UK gas market as determined by ROM. The analysis showed that this probability depends on a number of factors such as capacity levels, demand levels, the time of year and the initial amount of gas in storage. Under current assumptions on system parameters, it was found that system failure is highly unlikely. For this to become likely, there would have to be a substantial increase in demand and/or a large decrease in the capacities of some of the important sources of supply.
8. SUMMARY

8.1 Future work

As explained above, ROM provides a reasonably fit good to actual data. However, there are a number of areas where the model could be refined for usage by industrial partners. In this section some of these areas are discussed as potential future work as follows.

1. In Chapter 5 it is shown that the System Average Prices (SAPs) produced by ROM captured the seasonal varying effect of demand in actual SAPs in the UK gas market. However, as stated throughout this thesis, there are many other factors that affect gas prices in the UK, such as oil prices and international events (e.g., the Russian-Ukraine dispute in 2010). One area of future work for the project could be to incorporate these factors into the SAPs produced by ROM.

2. While daily SAPs were studied in this project, there are other important types of gas prices in the UK gas market, e.g., forward prices. Gas forward prices are the prices one pays to have natural gas delivered at some specified time in the future. These types of prices are not taken into account in this project and as such form the basis for another area of potential future work.

3. In the UK gas market, most supply (e.g., UKCS, Norway) arrives in the UK through pipelines. The only exception is LNG, which arrives in the UK in large shipments from around the world. In ROM however, LNG is modelled in the same manner as all the other sources of supply and the model does not take into account the arrival of large shipments of LNG into the UK. ROM assumes that LNG is available at a constant rate. As a result, another area of future work for this project could be the modelling of large shipments of LNG into UK.

4. In Chapter 6, the various scenario reduction techniques developed in [13, 14] are applied to gas demand scenarios. The analyses in [13, 14] also detail how these reduction techniques can be used in the creation of scenario trees. Another area of future work could be to use these scenario trees on the gas demand scenarios developed in this project (see, for example, electricity market models such as WILMAR [55]). Using scenario trees should lead to simplified scenarios with, for example, one high-demand scenario, one low-demand scenario and an average-demand scenario. This is in contrast to the demand scenarios used throughout this thesis, which vary throughout the year and thus represent periods of both high and low demand.

5. In Chapter 3 a stochastic process for UK gas demand is developed. This process models demand as a whole. Chapter 2 explains how demand in the UK is made of four main sectors: domestic demand, industrial & commercial demand, demand for power generation, and exports. Instead of modelling UK gas demand as a whole, an area of interest to industrial partners could be to model each of these different areas of demand individually.
Bibliography


Appendix A

Statistical analysis of actual demand and predicted demand

This appendix contains figures for the statistical analysis of the difference between the natural log of actual demand and the natural log of two- to five-day ahead predictions (see Section 3.3.1).

A.1 Two-day ahead predictions

Figure A.1: Time series of actual demand and two-day ahead predictions.
A. STATISTICAL ANALYSIS OF ACTUAL DEMAND AND PREDICTED DEMAND

Figure A.2: The difference between the natural log of actual demand and the natural log of two-day ahead predictions.

Figure A.3: Auto-correlations for the difference between the natural log of actual demand and the natural log of two-day ahead predictions.
A.1 Two-day ahead predictions

Figure A.4: Partial auto-correlations for the difference between the natural log of actual demand and the natural log of two-day ahead predictions.

Figure A.5: Residuals from the AR(1) process defined in equation (3.16) for the difference between the natural log of actual demand and the natural log of two-day ahead predictions.
A. STATISTICAL ANALYSIS OF ACTUAL DEMAND AND PREDICTED DEMAND

![Sample Autocorrelation Function (ACF)](image)

Figure A.6: Auto-correlations for the residuals from the AR(1) process of equation (3.16).

![Histogram for the residuals from the AR(1) process of equation (3.16)](image)

Figure A.7: Histogram for the residuals from the AR(1) process of equation (3.16).
Figure A.8: The residual versus fitted value plot for the residuals from the AR(1) process of equation (3.16).
A.2 Three-day ahead predictions

Figure A.9: Time series of actual demand and three-day ahead predictions.
A.2 Three-day ahead predictions

Figure A.10: The difference between the natural log of actual demand and the natural log of three-day ahead predictions.

Figure A.11: Auto-correlations for the difference between the natural log of actual demand and the natural log of three-day ahead predictions.
Figure A.12: Partial auto-correlations for the difference between the natural log of actual demand and the natural log of three-day ahead predictions.

Figure A.13: Residuals from the AR(1) process defined in equation (3.17) for the difference between the natural log of actual demand and the natural log of three-day ahead predictions.
Figure A.14: Auto-correlations for the residuals from the AR(1) process of equation (3.17).

Figure A.15: Histogram for the residuals from the AR(1) process of equation (3.17).
Figure A.16: The residual versus fitted value plot for the residuals from the AR(1) process of equation (3.17).
A.3 Four-day ahead predictions

Figure A.17: Time series of actual demand and four-day ahead predictions.
A. STATISTICAL ANALYSIS OF ACTUAL DEMAND AND PREDICTED DEMAND

Figure A.18: The difference between the natural log of actual demand and the natural log of four-day ahead predictions.

Figure A.19: Auto-correlations for the difference between the natural log of actual demand and the natural log of four-day ahead predictions.
A.3 Four-day ahead predictions

Figure A.20: Partial auto-correlations for the difference between the natural log of actual demand and the natural log of four-day ahead predictions.

Figure A.21: Residuals from the AR(1) process defined in equation (3.18) for the difference between the natural log of actual demand and the natural log of four-day ahead predictions.
Figure A.22: Auto-correlations for the residuals from the AR(1) process of equation (3.18).

Figure A.23: Histogram for the residuals from the AR(1) process of equation (3.18).
Figure A.24: The residual versus fitted value plot for the residuals from the AR(1) process of equation (3.18).
A.4 Five-day ahead predictions

![Diagram showing time series of actual demand and five-day ahead predictions.](image)

Figure A.25: Time series of actual demand and five-day ahead predictions.
A.4 Five-day ahead predictions

Figure A.26: The difference between the natural log of actual demand and the natural log of five-day ahead predictions.

Figure A.27: Auto-correlations for the difference between the natural log of actual demand and the natural log of five-day ahead predictions.
Figure A.28: Partial auto-correlations for the difference between the natural log of actual demand and the natural log of five-day ahead predictions.

Figure A.29: Residuals from the AR(1) process defined in equation (3.19) for the difference between the natural log of actual demand and the natural log of five-day ahead predictions.
Figure A.30: Auto-correlations for the residuals from the AR(1) process of equation (3.19).

Figure A.31: Histogram for the residuals from the AR(1) process of equation (3.19).
Figure A.32: The residual versus fitted value plot for the residuals from the AR(1) process of equation (3.19).
Appendix B

Gas market models

B.1 Gabriel et al’s mixed complementarity based model

In this appendix, a detailed presentation of the problems of the different players in Gabriel et al’s model [11] is presented. This model is discussed in detail in Section 4.2.

B.1.1 Pipeline operator’s problem

The first problem considered in [11] is the pipeline operator’s. The pipeline operator controls \( A \) different pipelines between the \( N \) different nodes which is where the other players are located. The pipes leaving/entering at node \( n \) are denoted \( A(n) \). The pipeline operator attempts to maximise the flows of gas through its pipe as follows:

\[
\text{Max} \quad \sum_{y \in Y} \sum_{s=1}^{3} \text{days}_s \tau_{asy} f_{asy}
\]

subject to:

\[
0 \leq f_{asy} \leq \bar{f}_a \quad \forall s, y,
\]

where \( f_{asy} \) is the amount of gas that flows through pipeline \( a \) in season \( s \) and year \( y \) and \( \tau_{asy} \) is the price per unit that the pipeline operator receives. This price is exogenous to this problem and is determined by the overall equilibrium model. Equation (B.1) constraints the amount of gas flowing through pipeline \( a \) by providing upper and lower bounds. The amount of days in season \( s \) is \( \text{days}_s \). Gabriel et al explain that, as this problem is a linear program, the KKT conditions associated with it are both necessary and sufficient for optimality.

B.1.2 Producer’s problem

The next problem to be considered in Gabriel et al’s model is the producer’s problem. In the model there are \( C \) producers. In each of the seasons the producers sell gas to the marketers, while
in season 1 they also sell gas to the storage operators. Each producer $c$ at node $n$ attempts to maximise their profits as follows:

$$\text{Max} \sum_{y \in Y} \sum_{s=1}^{3} \text{days}_s \pi_{nsy} q_{c,sy} - \text{days}_s c_p^r(q_{c,sy})$$

subject to:

$$0 \leq q_{c,sy} \leq \bar{q}_c \quad \forall s, y, \quad (B.2)$$

$$\sum_{y \in Y} \sum_{s=1}^{3} \text{days}_s q_{c,sy} \leq \text{prod}_c, \quad (B.3)$$

where $q_{c,sy}$ is the amount of gas produced by producer $c$, in season $s$, in year $y$. The unit price that producers at node $n$ get for doing this is $\pi_{nsy}$ while $c_p^r(.)$ is producer $c$’s quadratic cost function. Equation (B.2) constrains the amount that gas producer $c$ can produce in a season by providing upper and lower bounds, while equation (B.3) constrains the total amount of gas produced by producer $c$. Gabriel et al model producers in [11] as price takers in an environment of perfect competition. This means that the price that they receive for producing, $\pi_{nsy}$, is exogenous to this problem and is determined by the overall equilibrium model. Gabriel et al explain that if the cost function, $c_p^r(.)$, is convex, then the KKT conditions associated with this problem are both necessary and sufficient for optimality.

### B.1.3 Storage operator’s problem

The next problem to be considered in Gabriel et al’s model is the storage operator’s problem. In the model there are $R$ storage operators. Each of these buys gas from the producers in season 1 and then sell this gas to the marketers in seasons 2 and 3. Each storage operator $r$ at node $n$ attempts to maximise their profits as follows:

$$\text{Max} \sum_{y \in Y} \left[ \text{days}_2 \gamma_{n2y} x_{r2y} + \text{days}_3 \gamma_{n3y} x_{r3y} - \text{days}_1 c_{s}^{st}(\sum_{a \in A(n)} g_{ary}) \right.\left. - \sum_{a \in A(n)} \text{days}_1 (\tau_{a1y} + r_{a1y} + \pi_{n2(a)1y}) g_{ary} \right]$$

subject to:

$$\text{days}_2 x_{r2y} + \text{days}_3 x_{r3y} - \text{days}_1 \sum_{a \in A(n)} g_{ary} (1 - \text{loss}_a)(1 - \text{loss}_r) = 0 \quad \forall y \quad (B.4)$$

$$x_{rsy} \leq \bar{x}_r \quad s = 2, 3 \quad \forall y, \quad (B.5)$$
B.1 Gabriel et al’s mixed complementarity based model

\[ \sum_{a \in A(n)} g_{ary} \leq g \forall y, \quad (B.6) \]

\[ \sum_{s=2,3} days_s x_{rsy} \leq k \forall y, \quad (B.7) \]

\[ 0 \leq g_{ary} \forall a \in A(n), \forall y, \quad (B.8) \]

where \( x_{rsy} \) is the amount of gas storage operator \( r \) sells in season \( s \) and year \( y \) and \( g_{ary} \) is the amount of gas they buy in year \( y \) through pipeline \( a \). Note: as the storage operators only buy gas in season 1 there is no subscript \( s \) in \( g_{ary} \). The price per unit that storage operators at node \( n \) receive for selling their gas is \( \gamma_{nsy} \). The quadratic cost function for storage operator \( r \) is \( c_{st}^{sf}(.) \). It is a function of the total amount of gas the storage operator buys. The pipeline operator’s charge for using pipeline \( a \) is \( \tau_{a1y} \), while \( \tau_{a1y}^{reg} \) is a fixed pre-determined regulated charge that the storage operators must pay for using pipeline \( a \) in year \( y \). The price storage operators pay producers at node \( n \) associated with pipe \( a \) for their gas is \( \pi_{n2(a)1y} \).

Equation (B.4) ensures that the total amount of gas sold by the storage operator is equal to the total amount bought after taking into account losses from the pipelines and storage facilities, \( loss_a \) and \( loss \), respectively. Equation (B.5) provides an upper bound on the amount of gas being sold by storage operator \( r \) in season \( s \) and year \( y \), while equation (B.7) provides an upper bound on the total amount of gas sold. Equation (B.6) provides an upper bound on the total amount of gas being bought by storage operator \( r \), while equation (B.8) ensures that these values are non-negative.

Gabriel et al model storage operators as price takers in an environment of perfect competition. This means that the price they receive for selling gas, \( \gamma_{nsy} \), is exogenous to this problem and is determined by the overall equilibrium model. Gabriel et al explain that if the cost function, \( c_{st}^{sf}(.) \), is convex, then the KKT conditions associated with this problem are both necessary and sufficient for optimality.

B.1.4 Marketer’s problem

The next problem in [11] is the marketer’s problem. In the model there are \( M \) marketers who buy gas from the producers in every season, the storage operators in seasons 2 and 3 and the peak gas operator in season 3 only. The marketers then sell this gas to the different demand sectors. Gabriel et al model four different demand sectors, namely

1. Residential,
2. Commercial,
3. Industrial,
4. Electrical power.

Each of these are represented by a subscript, \( k = 1, 2, 3, 4 \), respectively. In contrast to the rest of the players in this model the marketers do not operate in an environment of perfect competition.
Gabriel et al assume Nash-Cournot competition amongst each of the marketers. As a result, each marketer has their own inverse demand function that determines the price at which they sell their gas. These functions determine, endogenously, the amount of gas demand in each sector. They do this by allowing the marketers to set their price of gas in order to maximise their profits. For example, if the marketers set a very high price, then demand will be low. Conversely, if they set a low price, then demand will be high. Each marketer $m$ at node $n$ attempts to maximise their profits as follows:

$$\text{max} \sum_{k \in K} \sum_{y \in Y} [days_1 \theta_{n1y}^k (h_{m1y}^k + h_{-m1y}^k) h_{m1y}^k$$

$$+days_2 \theta_{n2y}^k (h_{m2y}^k + u_{m2y}^k + h_{-m2y}^k + u_{-m2y}^k)(h_{m2y}^k + u_{m2y}^k)$$

$$+days_3 \theta_{n3y}^k (h_{m3y}^k + u_{m3y}^k + v_{my}^k + h_{-m3y}^k + u_{-m3y}^k + v_{-my}^k)(h_{m3y}^k + u_{m3y}^k + v_{my}^k)]$$

$$- \sum_{y \in Y} [\prod_{s=1}^3 \sum_{a \in A(n)} days_s (\tau_{asy}^r + \tau_{asy}^{\text{reg}} + \pi_{n2(a)sy}) h_{amsy}^k)$$

$$+days_2 \gamma_{n2y} u_{m2y} + +days_3 \gamma_{n3y} u_{m3y} + days_3 \beta_{ny} v_{my}]$$

subject to:

$$\sum_{k \in K} days_s h_{msy}^k - \sum_{a \in A(n)} days_s (1 - loss_a) h_{amsy} = 0, \ \forall s, y, \quad (B.9)$$

$$\sum_{k \in K} days_s u_{msy}^k - days_s u_{msy} = 0, \ \forall s = 2, 3, \ \forall y \quad (B.10)$$

$$\sum_{k \in K} days_3 v_{my}^k - days_3 v_{my} = 0, \ \forall y, \quad (B.11)$$

$$0 \leq h_{msy}^k \ \forall k, s, y, \quad (B.12)$$

$$0 \leq h_{amsy} \ \forall a \in A(n), s, y, \quad (B.13)$$

$$0 \leq u_{msy}^k \ \forall k, s = 2, 3, y, \quad (B.14)$$

$$0 \leq u_{msy} \ \forall s = 2, 3, y, \quad (B.15)$$

$$0 \leq v_{my}^k \ \forall y, \quad (B.16)$$

$$0 \leq v_{my} \ \forall y, \quad (B.17)$$
where \( \theta_{n_{sy}}^k(\cdot) \) is the inverse demand function for demand sector \( k \) of marketers at node \( n \) for season \( s \) and year \( y \). It takes as an input the total amount of gas used by the marketers at node \( n \) to supply sector \( k \) in season \( s \) and year \( y \). Table B.1 gives an explanation for the remaining undefined variables.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>( h_{m_{sy}}^k )</td>
<td>Flow rate of gas to demand sector ( k ) from marketer ( m ) in season ( s ) and year ( y )</td>
</tr>
<tr>
<td>( h_{m_{-sy}}^k )</td>
<td>Flow rate of gas to demand sector ( k ) from all marketers except marketer ( m ) in season ( s ) and year ( y )</td>
</tr>
<tr>
<td>( h_{a_{msy}} )</td>
<td>Flow rate of gas to marketer ( m ) from producers in season ( s ) through pipeline ( a )</td>
</tr>
<tr>
<td>( u_{m_{sy}}^k )</td>
<td>Rate of storage gas shipped to sector ( k ) from marketer ( m ) in season ( s ) and year ( y )</td>
</tr>
<tr>
<td>( u_{m_{-sy}}^k )</td>
<td>Rate of storage gas shipped to sector ( k ) from all marketers except marketer ( m ) in season ( s ) and year ( y )</td>
</tr>
<tr>
<td>( u_{m_{sy}} )</td>
<td>Rate of storage gas shipped to marketer ( m ) in season ( s ) and year ( y )</td>
</tr>
<tr>
<td>( v_{m_{sy}}^k )</td>
<td>Rate of peak gas to demand sector ( k ) from marketer ( m ) in year ( y )</td>
</tr>
<tr>
<td>( v_{m_{-sy}}^k )</td>
<td>Rate of peak gas to demand sector ( k ) from all marketers except marketer ( m ) in year ( y )</td>
</tr>
<tr>
<td>( v_{m_{-sy}} )</td>
<td>Rate of peak gas shipped to marketer ( m ) in year ( y )</td>
</tr>
</tbody>
</table>

Table B.1: Variables associated with the marketers problem of Gabriel et al’s model

Equation (B.9) ensures that the total amount of gas supplied by marketer \( m \) to the different demand sectors from producers is equal to the total amount of gas bought by marketer \( m \) from the producers after taking into account pipeline losses. Similarly, equation (B.10) ensures that the total amount of gas supplied by marketer \( m \) to the different demand sectors from the storage operators is equal to the total amount of gas bought by marketer \( m \) from the storage operators. Equation (B.11) ensures that the total amount of gas supplied by marketer \( m \) to the different demand sectors from the peak gas operators is equal to the total amount of gas bought by marketer \( m \) from the peak gas operators. Equations (B.12) - (B.17) ensure that each of these quantities are non-negative.

Gabriel et al explain in [11] that as long as the non-linear terms in the objective function are concave in their variables, then the KKT conditions associated with this problem are both necessary and sufficient for optimality.

### B.1.5 Peak gas operator’s problem

The final problem to be considered in Gabriel at al’s model is the peak gas operator’s problem. In the model there are \( P \) peak gas operators. Each of these sells gas to the marketers in season 3 only. Each peak gas operator \( p \) at node \( n \) attempts to maximise their profits as follows:

\[
Max \sum_{y \in Y} \text{days}_3(\beta_{m_{wy}}w_{py} - c_p^w(w_{py}))
\]
subject to:

\[ 0 \leq w_{py} \leq \overline{w}_p, \quad (B.18) \]

where \( w_{py} \) is the amount of gas that peak gas operator \( p \) sells in year \( y \) while \( \beta_{ny} \) is the price per unit those at node \( n \) receive for doing this. Note: as the peak operators only sell gas in season 3 there is no subscript \( s \) in \( w_{py} \) or \( \beta_{ny} \). The quadratic cost function for peak gas operator \( p \) is \( c^{pg}_p(\cdot) \). It is a function of the total amount of gas the peak gas operator sells. Equation (B.18) provides upper and lower bounds for the amount of gas a peak gas operator sells in year \( y \). Gabriel et al. model peak gas operators as price takers in an environment of perfect competition. This means that the price they receive for selling gas, \( \beta_{ny} \), is exogenous to this problem and is determined by the overall equilibrium model.

Gabriel et al explain that if the cost function \((c^{pg}_p(\cdot))\) is convex, then the KKT conditions associated with this problem are both necessary and sufficient for optimality.

### B.1.6 Market clearing conditions

In order for this mixed complementarity-based equilibrium model to be solved, the different problems must be connected. The following market clearing conditions do this.

#### B.1.7 Market clearing conditions for pipeline \( a \)

The first market clearing conditions considered are those for pipeline \( a \), as follows:

\[
days_1 f_{a1y} = \sum_{r \in R(n_1(a))} days_1 g_{ary} + \sum_{m \in M(n_1(a))} days_1 h_{amy}, \quad \forall y, \quad (B.19)\
\]

\[
days_s f_{asy} = \sum_{m \in M(n_s(a))} days_s h_{amsy}, \quad s = 2, 3 \quad \forall y. \quad (B.20)\
\]

Equation (B.19) ensures that in season 1 the total amount of gas flowing through pipeline \( a \) goes from the producers to either the marketers or storage operators. Equation (B.20) ensures that the total amount of gas flowing through pipeline \( a \) in season \( s \) (for \( s = 2, 3 \)) must go from the producers to the marketers.

#### B.1.8 Market clearing conditions at production node \( n \)

The next market clearing conditions are those associated with producers at node \( n \). They are:

\[
\sum_{c \in C(n)} days_1 q_{cly} = \sum_{a \in A(n)} (\sum_{r \in R(n_1(a))} days_1 g_{ary} + \sum_{m \in M(n_1(a))} days_1 h_{amy}), \quad \forall y, \quad (B.21)\
\]

\[
\sum_{c \in C(n)} days_s q_{cys} = \sum_{a \in A(n)} (\sum_{m \in M(n_1(a))} days_1 h_{amy}), \quad s = 2, 3, \quad \forall y. \quad (B.22)\
\]
Equation (B.21) ensures that in season 1 the total amount of gas produced by the producers goes to either the marketers or storage operators. Equation (B.22) ensures that in seasons 2 and 3 the total amount of gas produced by the producers goes to the marketers.

**B.1.9 Storage operator’s market clearing condition for node** $n = n_1(a)$

The next market clearing condition is that associated with the storage operators at node $n$, as follows:

$$
\sum_{r \in R(n)} days sx_{rasy} = \sum_{m \in M(n)} days sx_{masy}, \quad s = 2, 3, \quad \forall y. \quad \text{(B.23)}
$$

Equation (B.23) ensures that the amount of gas being sold by the storage operators, in seasons 2 and 3 in year $y$, equals the amount of gas being bought by the marketers from storage.

**B.1.10 Market clearing conditions for the peak gas operator**

The final market clearing condition is that associated with the peak operators as follows:

$$
\sum_{p \in P(n)} u_{py} = \sum_{m \in M(n)} v_{my}, \quad \forall y. \quad \text{(B.24)}
$$

Equation (B.24) ensures that for year $y$ the amount of gas being sold by the peak gas operators equals the amount of gas being bought by the marketers from peak gas operators.

**B.2 Matrix structure of complementarity-based model**

In this appendix the structures of the matrices and vectors, used in Section 4.5 are presented. Firstly consider the following matrix structures:

$$
Y^0_n = \begin{bmatrix}
1 \\
1 \\
\vdots \\
1
\end{bmatrix}, \quad \text{(B.25)}
$$

$$
Y^{n \times kn}_1 = \begin{bmatrix}
Y^0_n & Y^0_n & \cdots & Y^0_n
\end{bmatrix}. \quad \text{(B.26)}
$$

$$
Y^n_2 = \begin{bmatrix}
1 \\
1 \\
\vdots \\
1
\end{bmatrix}. \quad \text{(B.27)}
$$
Y_{3}^{k_{1}n \times k_{1}k_{2}n} = \begin{bmatrix}
Y_{2}^{n} & Y_{2}^{n} & \cdots & Y_{2}^{n} \\
Y_{2}^{n} & Y_{2}^{n} & \cdots & Y_{2}^{n} \\
\vdots & \vdots & \ddots & \vdots \\
Y_{2}^{n} & \cdots & \cdots & Y_{2}^{n}
\end{bmatrix},
\quad (B.28)

where the matrix $Y_{0}^{n}$ is the identity matrix of size $n$. The matrix $Y_{1}^{n \times kn}$ is a matrix with $n$ rows and $kn$ columns containing $k$ matrices that take the structure described by $Y_{0}^{n}$. The matrix $Y_{2}^{n}$ is a lower triangular matrix of size $n$ where each non-zero value is 1. The matrix $Y_{3}^{k_{1}n \times k_{1}k_{2}n}$ is a matrix with $k_{1}n$ rows and $k_{1}k_{2}n$ columns. Every $n$ rows contain $k_{2}$ matrices that take the structure described by the matrix $Y_{2}^{n}$.

Using these definitions, the matrices $D$ and $A$, used in Section 4.5, take the form

\begin{equation}
D = \begin{bmatrix}
-Y_{0}^{PD} & Y_{1}^{PD \times SOPD} & -Y_{0}^{SOD} & Y_{0}^{SOD} \\
Y_{1}^{PD \times SOPD} & Y_{0}^{PD} & -Y_{0}^{SOD} & -Y_{1}^{SOD} \\
-Y_{0}^{SOD} & -Y_{1}^{SOD} & Y_{0}^{SOD} & Y_{1}^{SOD}
\end{bmatrix},
\quad (B.29)
\end{equation}

and

\begin{equation}
A = \begin{bmatrix}
Y_{0}^{PD} & Y_{0}^{SOD} & -Y_{0}^{SOD} & -Y_{0}^{PD} \\
-Y_{0}^{PD} & Y_{0}^{SOD} & -Y_{0}^{SOD} & -Y_{0}^{PD} \\
Y_{0}^{SOD} & -Y_{0}^{SOD} & Y_{0}^{SOD} & Y_{0}^{SOD} \\
Y_{3}^{SO \cdot D \times SO \cdot P \cdot D} & -Y_{3}^{SO \cdot D \times SO \cdot 1 \cdot D} & -Y_{3}^{SO \cdot D \times SO \cdot P \cdot D} & Y_{3}^{SO \cdot D \times SO \cdot 1 \cdot D}
\end{bmatrix},
\quad (B.30)
\end{equation}

where $D$ is a $(PD + SOD + D) \times (2PD + SOPD + 2SOD)$ matrix and $A$ is a $(3PD + 2SOPD + 4SOD) \times (2PD + SOPD + 2SOD)$ matrix.
The vectors $x, v, u$ have the structures

$$x = \begin{bmatrix}
Q_{1,1} \\
\vdots \\
Q_{P,D} \\
I_{1,1,1} \\
\vdots \\
I_{SO,P,D} \\
W_{1,1} \\
\vdots \\
W_{SO,D} \\
H_{1,1} \\
\vdots \\
H_{P,D} \\
U_{1,1} \\
\vdots \\
U_{SO,D}
\end{bmatrix},$$  \hspace{1cm} (B.31)

$$v = \begin{bmatrix}
\pi_{1,1} \\
\vdots \\
\pi_{P,D} \\
\gamma_{1,1} \\
\vdots \\
\gamma_{SO,D} \\
\lambda_{Demand_1} \\
\vdots \\
\lambda_{Demand_D}
\end{bmatrix},$$  \hspace{1cm} (B.32)
where the vector $x$ is of length $2PD + SOPD + 2SOD$, the vector $v$ is of length $PD + SOD + D$ and the vector $u$ is of length $3PD + 2SOPD + 4SOD$. Similarly, the vectors $c$, $e$ and $b$ have the
structure

\[ c = \begin{bmatrix} -c_1 \\ \vdots \\ -c_P \\ -a_1 \\ \vdots \\ -a_{SO} \\ -b_1 \\ \vdots \\ -b_{SO} \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad (B.34) \]

\[ e = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ -Demand_1 \\ \vdots \\ -Demand_D \end{bmatrix}, \quad (B.35) \]
\[
\begin{bmatrix}
Q_1^{\text{max}} \\
\vdots \\
Q_P^{\text{max}} \\
0 \\
\vdots \\
0 \\
I_1^{\text{max}} \\
\vdots \\
I_{SO}^{\text{max}} \\
0 \\
\vdots \\
0 \\
W_1^{\text{max}} \\
\vdots \\
W_{SO}^{\text{max}} \\
0 \\
\vdots \\
0 \\
\text{MaxCap}_1 - \text{IntCap}_1 \\
\vdots \\
\text{MaxCap}_{SO} - \text{IntCap}_{SO} \\
-\text{MinCap}_1 + \text{IntCap}_1 \\
\vdots \\
-\text{MinCap}_{SO} + \text{IntCap}_{SO} \\
0 \\
\vdots \\
0
\end{bmatrix},
\]

where the vector \( c \) is of length \( 2PD + SOPD + 2SOD \), the vector \( e \) is of length \( PD + SOD + D \) and the vector \( b \) is of length \( 3PD + 2SOPD + 4SOD \).
Appendix C

Rolling Optimisation Model

This appendix contains extra detail associated with Chapter 5.

C.1 GAMS code associated with ROM for the year beginning in April 2010

In this appendix the GAMS code associated with the Rolling Optimisation Model for the year beginning in April 2010 is presented. See Section 5.3 for discussion.

```
SETS
P Producers / P1*P18 /
D Day / D2*D365 /
S Scenarios / S1*S3 /
K Var / K1*K2 /
R Rolls / R1*R365 /
SO Storage Operators / SO1*SO3 /

ALIAS (D,DD);

PARAMETERS
C(P) Cost per unit for producers
/P1 58.464
p2 57.611
p3 56.552
p4 69.985
p5 60.985
p6 64.322
p7 48.375
p8 58.637
```
C. ROLLING OPTIMISATION MODEL

MaxCap(P) Maximum capacities for producers

<table>
<thead>
<tr>
<th>p9</th>
<th>55.242</th>
</tr>
</thead>
<tbody>
<tr>
<td>p10</td>
<td>58.865</td>
</tr>
<tr>
<td>p11</td>
<td>71.458</td>
</tr>
<tr>
<td>p12</td>
<td>64.189</td>
</tr>
<tr>
<td>p13</td>
<td>55.084</td>
</tr>
<tr>
<td>p14</td>
<td>59.352</td>
</tr>
<tr>
<td>p15</td>
<td>61.626</td>
</tr>
<tr>
<td>p16</td>
<td>69.543</td>
</tr>
<tr>
<td>p17</td>
<td>70.977</td>
</tr>
<tr>
<td>p18</td>
<td>60.336</td>
</tr>
</tbody>
</table>

MinCap(P) Minimum capacities for producers

<table>
<thead>
<tr>
<th>p9</th>
<th>38</th>
</tr>
</thead>
<tbody>
<tr>
<td>p10</td>
<td>36</td>
</tr>
<tr>
<td>p11</td>
<td>12</td>
</tr>
<tr>
<td>p12</td>
<td>11</td>
</tr>
<tr>
<td>p13</td>
<td>16</td>
</tr>
<tr>
<td>p14</td>
<td>14</td>
</tr>
</tbody>
</table>

MaxW(SO) Maximum withdrawal rate for storage facilities

| SO1 | 43 |
| SO2 | 43 |
| SO3 | 35 |

MaxI(SO) Maximum injection rate for storage facilities
C.1 GAMS code associated with ROM for the year beginning in April 2010

/SO1 43
SO2 43
SO3 35
/

StorMax(SO) Maximum capacity of storage facilities
/SO1 3300
SO2 810
SO3 180
/

StorInt(SO) Initial amount of gas in storage facilities
/SO1 440
SO2 295
SO3 56
/

StorMin(SO) Minimum capacity of storage facilities
/SO1 440
SO2 169
SO3 39
/

a(D,SO) unit cost of injection
b(D,SO) unit cost of withdrawal
ai(SO) unit cost of injection on scenario independent day one
bi(SO) unit cost of withdrawal on scenario independent day one

Demand(D,S) Demand on day d in scenario S
prob(S) Probability associated with scenario S

b1(D,D) lower triangular matrix
;

SCALAR Demand1 Demand on scenario independent day one ;
MinCap(P)=0;

*Define lower triangular matrix
b1(D,DD)=1 $ (ORD(D) GE ORD(DD));

*Define lag for short-run cost of storage
SCALAR lag /26/;

*Define unit cost of injection to storage
a(D,’SO1’)=0.664$ (ORD(D) GT lag)+0.069$ (ORD(D) LE lag);
a(D,’SO2’)=0.664$ (ORD(D) GT lag)+0.069$ (ORD(D) LE lag);
a(D,’SO3’)=2;
122 *Define unit cost of withdrawal from storage

123 b(D,'SO1')=0.019$(ORD(D) GT lag)+0.064$(ORD(D) LE lag);

124 b(D,'SO2')=0.664$(ORD(D) GT lag)+0.019$(ORD(D) LE lag);

125 b(D,'SO3')=2;

128 ai(SO)=a('D2',SO);

129 bi(SO)=b('D2',SO);

132 *These variables are constrained to be non-negative

133 POSITIVE VARIABLE Inj,W,Inj1,W1;

135 VARIABLES

136 Q(P,D,S) Gas supplied by source P on day D in scenario s

137 Inj(SO,D,S) Gas injected by storage facility so on day d in scenario s

138 W(SO,D,S) Gas withdrawn by storage facility so on day d in scenario s

139 Q1(P) Gas supplied by source p for the first day

140 Inj1(SO) Gas injected by storage facility so on the first day

141 W1(SO) Gas withdrawn by storage facility so on the first day

142 Z Total cost i.e. objective function ;

144 EQUATIONS

145 COST Objective function

146 SUPPLY Demand-supply constraint

147 CapMax Maximum capacity constraint for producers

148 CapMin Minimum capacity constraint for producers

149 StorEqMax Maximum storage constraint

150 StorEqMin Minimum storage constraint

151 InjCap Injection constraint

152 WithCap Withdrawal constraint

153 SUPPLY1 Demand-supply constraint on day 1

154 CapMax1 Maximum capacity constraint for producers on day 1

155 CapMin1 Minimum capacity constraint for producers on day 1

156 StorEqMax1 Maximum storage constraint on day 1

157 StorEqMin1 Minimum storage constraint on day 1

158 InjCap1 Injection constraint on day 1

159 WithCap1 Withdrawal constraint on day 1

160 ;

162 COST.. Z =E= SUM(P,C(P)*Q1(P)) + SUM(SO,ai(SO)*Inj1(SO) + bi(SO)*W1(SO)) + SUM(S,Prob(S)*(+SUM((P,D), C(P)*Q(P,D,S)) +SUM((SO,D),b(D,SO)*W(SO,D,S)) +SUM((SO,D),a(D,SO)*Inj(SO,D,S))));

167 SUPPLY(D,S) .. SUM(P,Q(P,D,S))+SUM(SO,W(SO,D,S)-Inj(SO,D,S)) =E= Demand(D,S);
C.1 GAMS code associated with ROM for the year beginning in April 2010

```
170  CapMax(P,D,S).. Q(P,D,S) =L= MaxCap(P);
171  CapMin(P,D,S).. Q(P,D,S) =G= MinCap(P);

173  StorEqMax(SO,D,S).. Inj1(SO) - W1(SO) + \sum(DD, B1(D,DD)*Inj(SO,DD,S)) 
174      - \sum(DD, B1(D,DD)*W(SO,DD,S)) =E= StorMax(SO) - StorInt(SO);
175
177  StorEqMin(SO,D,S).. Inj1(SO) - W1(SO) + \sum(DD, B1(D,DD)*Inj(SO,DD,S)) 
178      - \sum(DD, B1(D,DD)*W(SO,DD,S)) =G= StorMin(SO) - StorInt(SO);
179
182  InjCap(SO,D,S) .. Inj(SO,D,S) =E= MaxI(SO);
183  WithCap(SO,D,S) .. W(SO,D,S) =E= MaxW(SO);
184  SUPPLY1.. \sum(P, Q1(P)) + \sum(SO, W1(SO) - Inj1(SO)) =E= Demand1;
185  CapMax1(P) .. Q1(P) =L= MaxCap(P);
186  CapMin1(P) .. Q1(P) =G= MinCap(P);
187  StorEqMax1(SO) .. Inj1(SO) - W1(SO) =E= StorMax(SO) - StorInt(SO);
188  StorEqMin1(SO) .. Inj1(SO) - W1(SO) =G= StorMin(SO) - StorInt(SO);
189  InjCap1(SO) .. Inj1(SO) =E= MaxI(SO);
190  WithCap1(SO) .. W1(SO) =E= MaxW(SO);
191
192  MODEL ROLLOPT /ALL/;
193
*Demands for scenario independent days taken from an .xls file
194  $CALL GDXXRW.EXE year_apr_10_1R.xls par=Act_demand
195  PARAMETER Act_demand(R,K);
196  $GDXIN year_apr_10_1R.gdx
197  $LOAD Act_demand
198  $GDXIN

*Demand on day d in scenario S associated with roll R taken from a .csv file
200  TABLE scen(R,D,S)
201  $ondelim
202  $include stochastic_demand_scenarios_n3_scen_red_N3_apr10.csv
203  $offdelim
204
208  *Demands are outputted from MATLAB as whole numbers
209  *Therefore, to ensure accuracy, they are all multiplied by 10^6 before being
210  *outputted from MATLAB. This division corrects that.
211  scen(R,D,S) = scen(R,D,S) / 1000000;
212
213  LOOP (R,
214  *Update demands
215      Demand1=Act_demand(R,’k1’);
216      Demand(D,S) = scen(R,D,S);
```
C. ROLLING OPTIMISATION MODEL

\[
\text{Prob}(S) = \frac{1}{\text{CARD}(S)};
\]

*Solve model

\[
\text{SOLVE} \ \text{ROLLOPT USING LP MINIMIZING } Z;
\]

*Update initial amount of gas in storage

\[
\text{StorInt}(SO) = \text{StorInt}(SO) + \text{Inj1.l}(SO) - W1.l(SO);
\]

C.2 Parameter Estimation

In this appendix the simulated annealing algorithm is presented, see Section 5.3.1 for discussion.

Simulated annealing algorithm

**Step 0:** \( \text{Current parameters} := \text{Initial parameters} \)

\( \text{Current error} := \text{Initial error} \)

\( \text{Temp}_0 := \text{Initial temperature} \)

**Step i:**

if \( i \mod 4 = 0 \)

\( \text{Temp}_i := 0.9 \text{Temp}_{i-1} \)

end if

Perturb parameters: \( \text{Test parameters} := \text{Current parameters} + \epsilon \)

Calculate \( \text{Test error} \) using \( \text{Test parameters} \) and equation (5.14)

if \( \text{Test error} < \text{Current error} \)

\( \text{Current parameters} := \text{Test parameters} \)

\( \text{Current error} := \text{Test error} \)

else if \( \exp(\frac{\text{Current error} - \text{Test error}}{\text{Temp}_i}) < \text{Uniform}[0, 1] \)

\( \text{Current parameters} := \text{Test parameters} \)

\( \text{Current error} := \text{Test error} \)

end if

if \( \text{Test error} < \text{Best error} \)

\( \text{Best parameters} := \text{Test parameters} \)

\( \text{Best error} := \text{Test error} \)

end if

C.3 ROM tested for the year beginning in October 2010

In this appendix, the parameters and results of ROM, when applied to data for the year starting on the 1st of October 2010, are presented. They are similar to those described in Section 5.3.
Table C.1: The costs and maximum capacities (in mcm) of the different tranches for each source of supply.

<table>
<thead>
<tr>
<th>Tranche</th>
<th>$c_p$</th>
<th>$Q_{p}^{max}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>UKCS₁</td>
<td>58.464</td>
<td>48</td>
</tr>
<tr>
<td>UKCS₂</td>
<td>57.611</td>
<td>61</td>
</tr>
<tr>
<td>UKCS₃</td>
<td>56.552</td>
<td>57</td>
</tr>
<tr>
<td>Norway₁</td>
<td>60.985</td>
<td>21</td>
</tr>
<tr>
<td>Norway₂</td>
<td>64.322</td>
<td>54</td>
</tr>
<tr>
<td>Norway₃</td>
<td>48.375</td>
<td>30</td>
</tr>
<tr>
<td>Norway₄</td>
<td>58.637</td>
<td>23</td>
</tr>
<tr>
<td>LNG₁</td>
<td>55.242</td>
<td>38</td>
</tr>
<tr>
<td>LNG₂</td>
<td>58.865</td>
<td>36</td>
</tr>
<tr>
<td>LNG₃</td>
<td>71.458</td>
<td>12</td>
</tr>
<tr>
<td>BBL₁</td>
<td>64.189</td>
<td>11</td>
</tr>
<tr>
<td>BBL₂</td>
<td>55.084</td>
<td>16</td>
</tr>
<tr>
<td>BBL₃</td>
<td>59.352</td>
<td>14</td>
</tr>
<tr>
<td>IUK₁</td>
<td>61.626</td>
<td>11</td>
</tr>
<tr>
<td>IUK₂</td>
<td>69.543</td>
<td>16</td>
</tr>
<tr>
<td>IUK₃</td>
<td>70.977</td>
<td>27</td>
</tr>
<tr>
<td>IUK₄</td>
<td>60.336</td>
<td>18</td>
</tr>
</tbody>
</table>

Table C.2: The total maximum capacity (in mcm) for each source of supply, \[1, 2\].

<table>
<thead>
<tr>
<th>Source of supply</th>
<th>Total maximum capacity</th>
</tr>
</thead>
<tbody>
<tr>
<td>UKCS</td>
<td>166</td>
</tr>
<tr>
<td>Norway</td>
<td>128</td>
</tr>
<tr>
<td>LNG</td>
<td>86</td>
</tr>
<tr>
<td>BBL</td>
<td>41</td>
</tr>
<tr>
<td>IUK</td>
<td>72</td>
</tr>
</tbody>
</table>

Table C.3: Parameters associated with long-, medium- and short-range storage (in mcm) for ROM, \[1, 2\].

<table>
<thead>
<tr>
<th></th>
<th>LRS</th>
<th>MRS</th>
<th>SRS</th>
</tr>
</thead>
<tbody>
<tr>
<td>$IntStor_{so}$</td>
<td>3291</td>
<td>633</td>
<td>45</td>
</tr>
<tr>
<td>$MaxStor_{so}$</td>
<td>3300</td>
<td>810</td>
<td>180</td>
</tr>
<tr>
<td>$MinStor_{so}$</td>
<td>440</td>
<td>169</td>
<td>45</td>
</tr>
<tr>
<td>$I_{so}^{max}$</td>
<td>43</td>
<td>43</td>
<td>35</td>
</tr>
<tr>
<td>$W_{so}^{max}$</td>
<td>43</td>
<td>43</td>
<td>35</td>
</tr>
</tbody>
</table>
C. ROLLING OPTIMISATION MODEL

Figure C.1: Actual demand profile from the UK gas market starting on the 1st of October 2010.

Figure C.2: Predicted demand profile from the UK gas market starting on the 1st of October 2010.
Figure C.3: Actual and calibrated predicted SAPs starting on the 1st of October 2010. These prices were calibrated using equation (5.13) with $\alpha = 1.461$ and $\beta = 31.481$.

Figure C.4: Actual and simulated amount of LRS starting on the 1st of October 2010.
Figure C.5: Actual and simulated amount of MRS starting on the 1st of October 2010.

Figure C.6: Actual and simulated amount of SRS starting on the 1st of October 2010.
C.4 Effect of increasing number of scenarios on Rolling Optimisation Model

The figures in this appendix shows mean yearly supply from the UKCS, Norway and BBL (with error bars) from ROM for $n = 1, \ldots, 6, 10, 20, 50$. It also shows the mean yearly injections to each storage facility combined. See Section 5.6 for details.

![Mean plus error bars for the total yearly UKCS supply from ROM for $n = 1, \ldots, 6, 10, 20, 50$.]
Figure C.8: Mean plus error bars for the total yearly Norwegian supply from ROM for \( n = 1, \ldots, 6, 10, 20, 50 \).

Figure C.9: Mean plus error bars for the total yearly injections to storage for ROM for \( n = 1, \ldots, 6, 10, 20, 50 \).
C.4 Effect of increasing number of scenarios on Rolling Optimisation Model

Figure C.10: Mean plus error bars for the total yearly BBL supply from ROM for $n = 1, \ldots, 6, 10, 20, 50$. 

Appendix D

Scenario reduction

Appendix D contains the appendices associated with Chapter 6.

D.1 Stochastic process for demand based on 2008 -’09 data.

This appendix details a stochastic process modelling natural gas demand in the UK. It is similar to that described in Section 3.4; however, it is only based on a subset of the data, from the 1st of October 2008 to 30th of September 2009. The process is defined as follows:

\[
\begin{align*}
\text{dem}_{t,t} &= \text{ActDem}_t, \\
\text{dem}^s_{t,t+1} &= \text{ActDem}_{t+1} + \sqrt{93}\epsilon_{t+1}, \\
\text{dem}^s_{t,t+2} &= \text{ActDem}_{t+2} + \sqrt{231}\epsilon_{t+2}, \\
\text{dem}^s_{t,t+3} &= \text{ActDem}_{t+3} + \sqrt{317}\epsilon_{t+3}, \\
\text{dem}^s_{t,t+4} &= \text{ActDem}_{t+4} + \sqrt{379}\epsilon_{t+4}, \\
\text{dem}^s_{t,t+5} &= \text{ActDem}_{t+5} + \sqrt{447}\epsilon_{t+5},
\end{align*}
\]

\[
\ln(\text{dem}^s_t) = \ln(SND_t) - 0.008 + 0.878(\ln(\text{dem}^s_{t-1}) - \ln(SND_{t-1})) + 0.017\text{Weekend}_t + \sqrt{0.002}\epsilon_t, \quad \forall \ s, \ t > t + 5.
\]

All variables are as described in Section 3.4. The variable Weekend\(_t\) is a binary variable that takes the value 1 if day \(t\) is a Saturday or Sunday and 0 otherwise.

D.2 Computational time associated with operations count

Figures D.1 and D.2 shows the numerical results for computational times associated with operations counts of Section 6.5.1 for the backward, simultaneous backward reduction and fast forward selection algorithms. In particular it shows these counts when \(n = \frac{N}{2}\) and \(n = N - 1\). When
$n = \frac{N}{2}$ equations (6.33) and (6.34) both become of $O(N^3)$. Similarly, when $n = N - 1$ equations (6.33) and (6.34) become of $O(N^2 \log N)$ and $O(N^3)$ respectively.

Figure D.1: Log-log plot of computational time versus the number of original scenarios ($N$) for $n = N/2$. 
D.3 Comparison of the effect of using scenario reduction algorithms on ROM

Figures D.3 - D.5 contain the mean plus error bars comparing the effect of backward reduction, simultaneous backward reduction and fast forward selection algorithms on the results obtained from ROM. This is done for the mean yearly UKCS supplies, as well as the total yearly injections and withdrawals to and from storage. See Section 6.7 for discussion.
D. SCENARIO REDUCTION

Figure D.3: Mean plus error bars for the total yearly supply from the UKCS for $n = 1, \ldots, 6, 10$ and $N = 1000$.

Figure D.4: Mean plus error bars for the total yearly injections into storage for $n = 1, \ldots, 6, 10$ and $N = 1000$. 
D.4 Benefit of using a scenario reduction algorithm with ROM

Figures D.6 - D.8 are associated with Section 6.8 where the benefit of using the fast forward selection algorithm on ROM is analysed.
D. SCENARIO REDUCTION

Figure D.6: Mean plus error bars for the total yearly supply from Norway using the fast forward selection algorithm for $N = 100, 1000$ and with no reduction.

Figure D.7: Mean plus error bars for the total yearly supply through BBL using the fast forward selection algorithm for $N = 100, 1000$ and with no reduction.
Figure D.8: Mean plus error bars for the total yearly supply from LNG using the fast forward selection algorithm for $N = 100, 1000$ and with no reduction.