

**Evaluating the Quality of
Radiotherapy Treatment Plans
with Uncertainty using Data
Envelopment Analysis**

Emma Stubington, B.Sc.(Hons.), M.Res



Submitted for the degree of Doctor of
Philosophy at Lancaster University.

14th February, 2020

Abstract

External beam radiation therapy is a common treatment method for cancer. Radiotherapy is planned with the aim of achieving conflicting goals: while a sufficiently high dose of radiation is necessary for tumour control, a low dose of radiation is desirable to avoid complications in normal, healthy, tissue. This thesis aims to support the radiotherapy treatment planning process for prostate cancer by evaluating the quality of proposed treatment plans relative to previous plans.

We develop a variable selection technique, autoPCA, to select the most relevant variables for use in our Data Envelopment Analysis (DEA) models. This allows us to evaluate how well plans perform in terms of achieving the conflicting goals of radiotherapy. We develop the uncertain DEA problem (uDEA) for the case of box uncertainty and show that for small problems this can be solved exactly. This study of uncertainty is motivated by the inherently uncertain nature of the treatment process. Robust DEA, uDEA and simulation are applied to prostate cancer treatment plans to investigate this uncertainty. We identify plans that have the potential to be improved, which clinicians then replan for us. Small improvements were seen and we discuss the potential difference this could make to planning cases that are more complex. To aid this, we develop a prototype software, EvaluatePlan, that assesses the efficiency of a plan compared to past treatment plans.

Acknowledgements

In loving memory of David Everest Baldock my number one fan and mathematical inspiration.

I am most grateful to my supervisors Matthias Ehrgott and Omid Nohadani. In particular to Matthias for his continual support and patience as I have written and developed this thesis. I am very grateful for the support from the staff at Rosemere Cancer Centre, in particular Glyn Shentall, Rachel Norris and Michael Biddlecombe, their expertise have been essential for this thesis.

I would like to thank everyone both past and present of the STOR-i Centre for Doctoral Training. In particular, my year group for all their support in helping me develop as a researcher. For everything, from love actually, dog walking and suits to crafting and everyday life. Thank you Jonathan Tawn and Kirsty Hagan for believing in me and supporting me from day one.

Thank you to “my long-suffering father who waded through every chapter not understanding a word of it, but loving it”. Without him, there would be far fewer commas. This thesis would have not been possible without my family, related and chosen, in particular my parents and silly Grandma. Thank you.

Most of all thank you Tom. For always making me smile and never letting me give up.

Declaration

I declare that the work in this thesis has been done by myself and has not been submitted elsewhere for the award of any other degree.

Emma Stubington

The word count is approximately 72,000 words.

Contents

Abstract	I
Acknowledgements	II
Declaration	III
Contents	IV
List of Figures	VIII
List of Tables	X
List of Abbreviations	XIII
List of Symbols	XVI
1 Motivation	1
1.1 Why prostate cancer?	2
2 Introduction	3
2.1 Radiotherapy	3
2.1.1 Prostate cancer	12
2.1.2 Treatment planning: Rosemere	13
2.2 Plan quality evaluation	16

2.2.1	Existing treatment plan evaluation methods	19
2.2.2	Introduction to data envelopment analysis and its application to radiotherapy treatment planning	24
2.3	Uncertainty	27
2.3.1	Sources of uncertainty in radiotherapy	28
2.3.2	Handling uncertainty in radiotherapy	31
2.3.3	What we regard as uncertainty	33
2.4	So what is this thesis about?	35
3	Mathematical background	38
3.1	Data Envelopment Analysis	38
3.1.1	Basic concept of efficiency	40
3.1.2	The mathematical models	43
3.1.3	Economies of scale	51
3.1.4	Model orientation	56
3.1.5	Some extensions to DEA models	58
3.2	Robust optimisation	61
3.2.1	Method	63
3.2.2	Robust DEA model	70
3.3	Sensitivity analysis	72
3.3.1	Sensitivity analysis in DEA	72
3.4	Uncertain DEA	81
3.4.1	The uDEA model	82
3.5	Principal Component Analysis and other variable reduction techniques	85
3.5.1	Principal Component Analysis	85
3.5.2	Partial Covariance	91

4 Treating uncertainty	93
4.1 Preliminary results	99
4.2 Fixed Uncertainty	110
4.2.1 Single changes in DMUs	111
4.2.2 Multiple changes in inputs and outputs	115
4.2.3 Inefficient and efficient DMUs' data can change	118
4.2.4 Overall best improvement in efficiency score	121
4.3 Changing value of u	123
4.3.1 Value of u for $N + M$ dimensions	124
4.4 Future research directions	150
 5 DEA in radiotherapy	 152
5.1 Introduction	152
5.2 Variable selection	153
5.2.1 Previous study	154
5.2.2 Rosemere data exploration	156
5.2.3 Applying PCA to prostate cancer data	164
5.2.4 Automated PCA Variable Selection	173
5.2.5 Other variable selection techniques	188
5.2.6 Conclusion	191
5.3 Nominal DEA	195
5.3.1 Bladder volume as an environmental variable	201
 6 Uncertainty in radiotherapy	 208
6.1 Simulation	210
6.2 Fixed uncertainty: Robust DEA	215
6.3 Comparison	216
6.4 Replanning	222

<i>CONTENTS</i>	VII
7 Software	229
7.1 Interfacing with Rosemere	230
7.2 Software functionality	230
7.3 How to use EvaluatePlan	231
7.4 Future software improvements	235
8 Conclusion and further work	236
Appendix A	242
Bibliography	243

List of Figures

2.1	The treatment process for radiotherapy.	6
2.2	Example CT scan.	8
2.3	MLC leaf pair schematic.	9
2.4	Treatment couch.	11
2.5	CT scan for a prostate cancer patient at Rosemere.	14
3.1	CRS and VRS efficient frontier, Example 1.	53
3.2	Input and output oriented model, efficient frontier, Example 1.	58
4.1	The lines that define the PPS, Example 3.	104
4.2	Value of uncertainty for facets defined by a single DMU.	129
4.3	Calculating the uncertainty geometrically, Example 5.	132
4.4	The effect of a decreasing gradient of efficient line segments.	137
4.5	The facet of the efficient frontier inefficient DMUs are compared with.	140
4.6	Possible facets of the efficient frontier DMUs can be compared to.	142
4.7	The hyperplane Π_{ABC} intersecting with the hyperplane $y=y_D$	144
4.8	The hyperplane Π_{ABC} and Π_{ABC}^u	146
4.9	The hyperplane $\Pi_{BC(x_1)}$, $\Pi_{BC(x_1)}^u$ and Π_{ABC}	150
5.1	gEUD rectum versus D_{95} , Dataset 1.	155
5.2	V_x input data, Dataset 2a.	157
5.3	D_x output data, Dataset 2a.	160

5.4	Volume data, Dataset 2a.	161
5.5	Correlation between the input variables, Dataset 2a	162
5.6	Correlation between the output variables, Dataset 2a	163
5.7	First two PCs for the output variables, Dataset 2a.	165
5.8	First two PCs for the outputs according to session number, Dataset 2a.	166
5.9	First two PCs for the input variables, Dataset 2a.	166
5.10	First two PCs for the inputs according to session number, Dataset 2a.	167
5.11	Distribution of the inputs PCs' standard deviation, Dataset 2a. . . .	179
5.12	Distribution of the outputs PCs' standard deviation, Dataset 2a. . . .	180
6.1	Plan 85 DVH comparing original plan and after replanning.	223
6.2	Plan 53 DVH comparing original plan and after replanning.	224
7.1	Launching and selecting the plan to evaluate in EvaluatePlan.	232
7.2	Results for plan 99 from EvaluatePlan.	233
7.3	Results for plan 94 from EvaluatePlan.	234

List of Tables

3.1	Hospital departments data, Example 1.	43
3.2	Hospital departments CRS efficiency scores, Example 1.	52
3.3	Hospital departments BCC efficiency scores, Example 1.	54
3.4	Hospital departments BCC slack variable values, Example 1.	55
3.5	Hospital departments BCC peers, Example 1.	55
4.1	Data and efficiency scores, Example 2.	97
4.2	Uncertainty u_{x_1}, u_{x_2} and u_y , Example 2.	98
4.3	Data with scaled uncertainty, Example 2.	98
4.4	Uncertainty U , Example 2.	99
4.5	Nominal data, Example 3.	103
4.6	DEA Distance for DMUs E and F , Example 3.	103
4.7	Uncertain data, Example 3.	105
4.8	DEA Distance for DMUs E^u and F^u , Example 3.	105
4.9	Minimum amount of uncertainty for DMUs E and F , Example 3. . .	105
4.10	The maximum possible changes to the efficiency score of DMU k . . .	122
4.11	Nominal data, Example 7.	141
4.12	Value of u for different facets of the efficient frontier, Example 7. . .	143
4.13	Data, efficiency scores and peers: two inputs, one output, Example 10.	148
4.14	The planes that intersect with the PPS, Example 10	149

5.1	Variables obtained from the clinical protocol.	158
5.2	Total volume in cm^3 of structures from the clinical protocol.	159
5.3	First four PCs for the input variables, Dataset 2a.	169
5.4	First four PCs for the output variables Dataset 2a.	169
5.5	First three PCs for the output variables, Dataset 2a.	171
5.6	Correlation between the input variables and the PCs, Dataset 2a. . .	172
5.7	Correlation between the output variables and the PCs, Dataset 2a. .	172
5.8	Input PC metrics, Dataset 2a.	174
5.9	Output PCs metrics, Dataset 2a.	174
5.10	Input variables selected by autoPCA, Dataset 2a.	185
5.11	Output variables selected by autoPCA $\gamma \in \{70, 80, 90\}$, Dataset 2a. .	186
5.12	Output variables selected by autoPCA $\gamma = 70$, Dataset 2a.	187
5.13	Order to remove variables, Dataset 2b.	189
5.14	Partial covariance method for the inputs, Dataset 2a.	190
5.15	Two variables with $> 76\%$ of the variation of the inputs, Dataset 2b.	191
5.16	Partial covariance method for the outputs, Dataset 2a	192
5.17	Comparing variable selection techniques, Dataset 2b.	193
5.18	Order to remove output variables, Dataset 2b.	194
5.19	Variables selected for our DEA models.	195
5.20	Nominal DEA results, Dataset 2b.	200
5.21	Nominal DEA results, peer frequency, Dataset 2b.	202
5.22	Nominal DEA results including Vol^{Bla} , Dataset 2b.	206
5.23	Nominal DEA results including Vol^{Bla} , peer frequency, Dataset 2b. .	207
6.1	Simulation results, efficiency score, $r = 10,000$	213
6.2	Number of efficient DMUs for increasing r	215
6.3	Robust DEA results, efficiency score.	217
6.4	Robust DEA results, value of u	218

6.5	Data, Example 11.	219
6.6	Simulation results, efficiency score $r \in \{10, 100, 1,000\}$, Example 11.	220
6.7	Robust DEA, efficiency score, Example 11.	220
6.8	Simulation results, efficiency score, $u = 0.461$, increasing r , Example 11.	221
6.9	Efficiency score for simulation, robust DEA and uDEA, Example 11.	222
6.10	Difference in variables for replan and original variables.	225
6.11	Comparing replans efficiency scores.	226
1	Dataset 1.	242

List of abbreviations

Medical

AG	Agility
CTV	Clinical Target Volume
BM	Beam Modulator
cDVH	cumulative DVH
CT	Computerised Tomography
DC	Dice Coefficient
D_x	Dose received in Gy by x% of the volume of a PTV
DNA	Deoxyribonucleic Acid
DVH	Dose Volume Histogram
EUD	Equivalent Uniform Dose
gEUD	generalised Equivalent Uniform Dose
GTV	Gross Tumour Volume
Gy	Unit of ionising radiation dose, gray
HT	Healthy Tissue
IMAT	Intensity Modulated Arc Therapy
IMRT	Intensity Modulated Radiation Therapy
MLC	MultiLeaf Collimator
MRI	Magnetic Resonance Imaging
NT	Normal Tissue

NTCP	Normal Tissue Complication Probability
OAR	Organ At Risk
OVH	Overlap Volume Histogram
PET	Positron Emission Tomography
PTV	Planning Target Volume
QC	Quality Control
V_x	Percentage of the volume of an OAR receiving x Gy or more of radiation dose

Maths

AR-IDEA	Assurance Region IDEA
BCC	Banker, Charnes, Cooper
CCR	Charnes, Cooper, Rhodes
CI s	Conformity Indices
CRS	Constant Returns to Scale
DEA	Data Envelopment Analysis
DMU	Decision Making Unit
DMU 0_{eff}	Efficient DMU 0
DMU 0_{ex}	Extreme efficient DMU 0
DMU 0_{ineff}	Inefficient DMU 0
DO	Decision Objective
IDEA	Imprecise DEA
LHS	Left Hand Side
LO	Linear Objective
LP	Linear Programme
NLP	Non-Linear Programme
PC	Principal Component

PCA	Principal Component Analysis
PPS	Production Possibility Set
RHS	Right Hand Side
s.t.	such that
sup	supremum
uDEA	uncertain DEA
u-LO	uncertain Linear Objective
VRS	Variable Returns to Scale

List of Symbols

α	Constant for defining a hyperplane, (4.5)
α	Level for quasi-efficient, efficiency score, Section 3.3
α_n	Maximum coefficient of a PC, Section 3.5
β	Constant for defining a hyperplane, (4.5)
γ	Constant for defining a hyperplane, (4.5)
Δ_{mi}	Uncertainty for DMU i output m
Δ_{ni}	Uncertainty for DMU i input n
ϵ	Infinitesimal constant
ζ	Perturbation vector
η	Column vector of decision variables
η_i^*	Values the variables η take at an optimal solution
θ	Efficiency score (variable)
θ^*	The efficiency score variable value at an optimal solution
$\bar{\theta}^*$	The efficiency score variable value at an optimal solution with uncertainty
κ	Positive constant
λ	Lagrange constant, Section 3.5
λ	Decision variable of length I , weights of the DMUs
λ^*	Values λ take at an optimal solution
$\bar{\lambda}_i^*$	Values λ take at an optimal solution with uncertainty
μ	Constant, $0 < \mu < 1$

ν	Optimal multipliers for the inputs
Ξ_0	Efficiency score for DMU ₀ in the CCR _{Mult} model, (3.3)
ξ_0	Efficiency score of DMU 0 in the CCR _{Frac} model
Π	Hyperplane defined by, (4.5)
π	Line, where two hyperplanes intersect
ρ	Positive constant
T_i^{jk}	Target for DMU i on the plane Π_{jk}
τ	Positive constant
v	Optimal multipliers for the outputs
Φ	N+M
ϕ	Number of efficient DMUs
Ψ	The collection of hyperplanes that define efficient facets
Ω	Universe of possible collections of uncertainty
(A, b, c, d)	Data of the problem
a	Region dependent parameter, Section 2.1
a_j	j^{th} vector of coefficients for PC j , Section 3.5
$a_j^{(L')}$	The matrix formed row-wise from each basic shift for constraint j
C_j	Delivered dose, Section 2.1
C	Correlation matrix, Section 3.5
c	Vector in the robust DEA model of $I + 1$ zeros followed by a 1
c^0	Nominal data
\mathfrak{D}	DEA distance
d	Euclidean distance
d_j^R	Dose distribution for voxel j in region R , Section 2.1
$\mathcal{E}_i(\mathcal{U})$	Robust DEA efficiency score
E_i	Solution to the robust DEA model, (3.17)
e	Vector of ones of length I

F	Matrix defining a polyhedral uncertainty set, (3.13)
g	Vector defining a polyhedral uncertainty set, (3.13)
g	Gradient
h	Non-binding input constraint
I	Number of DMUs
j	Index for constraints in the DEA models
j	Objective number, Section 2.1
k	Inefficient DMU under consideration
L	Number of affine transformations
l	Inefficient DMU not under consideration
l_1	L1 Norm
l_∞	Infinity norm
M	Number of output variables
M_j	Planned dose
m	Mapping
m^R	Total number of voxels in region R , Section 2.1
N	Number of input variables
\vec{n}	Normal vector to the plane
O	Number of environmental variables
p	Output constraint
p	Number of variable in PCA, Section 3.5
P	Point, $P = (x_{10}, \dots, x_{N0}, y_{10}, \dots, y_{M0})$
q	Index for a binding input constraint
q	Number of variable in PCA, Section 3.5
Q	Set of all binding input constraints
R	Region, OAR, PTV or NT, Section 2.1
r	Number of repetitions for simulations, Chapter 6

r	All vectors in a \mathbb{R} that define a hyperplane, (4.6)
S	Quality score, Section 2.1
S	Sample variance-covariance matrix
s_m^-	Slack for the m^{th} output constraint
s_n^+	Slack for the n^{th} input constraint
T	The PPS
$Tr(S)$	Trace of matrix S
t	Variable introduced in, (3.11)
\mathcal{U}_j	Uncertainty set for a single input/output j
\mathcal{U}_i^*	Collection of uncertainty with the minimum amount of uncertainty
\hat{u}	$u + \epsilon$
u_x	Minimum of x_{ni} and u
u_y	Minimum of y_{mi} and u
w	Dual variable for polyhedral uncertainty in (3.15)
X	Input data matrix
x	Data, Section 3.5
x'	Transformed input data
x_i	Vector of inputs for the i^{th} DMU
Y	Output data matrix
y_i	Vector of outputs for the i^{th} DMU
y'	Transformed output data
\mathcal{Z}	Perturbation set
Z	Environmental variables data matrix
z_i	Vector of environmental variables for the i^{th} DMU

Chapter 1

Motivation

One in every seven deaths worldwide is as a result of cancer; this is more than HIV, tuberculosis, and malaria combined. Moreover, the global cancer burden continues to increase with 21.7 million new cancer cases predicted by 2030, resulting in 13 million cancer deaths, largely due to an ageing population (Siegel et al., 2015). With the ever-increasing number of cancer cases, improvement in treatment techniques is essential.

With 60% of cancer deaths occurring in low/middle-income countries, lacking the medical resources and health systems to support the disease burden (Siegel et al., 2015), it is important that research into cancer treatment improvement reflects the needs of the larger population.

This thesis addresses the ongoing need for a multidisciplinary approach to cancer treatment and the potential contribution mathematical and statistical techniques can have. Optimisation techniques are required throughout the treatment process to aid in the identification of the treatment objectives and their relationship with clinical criteria. Improved understanding of this relationship will aid the detection of potential improvements, or confirm that past performance indicates possible improvements are unlikely. Data Envelopment Analysis (DEA) techniques are explored and the effect

uncertainty has on the treatment process is discussed and accounted for. Data from prostate cancer patients at Rosemere Cancer Centre at the Royal Preston Hospital are analysed and an initial implementation of clinical software is developed to improve the utility of the outcomes.

1.1 Why prostate cancer?

In 2012 the world health organisation led an extensive project called GLOBOCAN, with the aim of providing estimates of the prevalence of major types of cancer throughout the world (Ferlay et al., 2013). This highlighted the increasing number of cases of many cancers in both developed and third world countries. The study concluded that prostate cancer is the second most common cancer worldwide for males, accounting for 15% of all cancer cases and the fourth most common cancer overall, accounting for 7.9% of cases (Ferlay et al., 2013). The study also highlighted that two thirds of prostate cancer diagnoses are from more developed regions of the world. Some of this will be due to more specialised diagnostic equipment and increase in those seeking medical care. Therefore, these world statistics could in truth be closer to the European data where prostate cancer is the most common type of cancer in men, making up 25% of all cases in men. With an increase in cases of prostate cancer in Europe of 16% between 2000 and 2011 (Cancer-Research-UK, 2014) it is an important research area. Prostate cancer treatment is typically a more standard treatment than other cancers. This is largely due to it having only two main Organs At Risk (OARs), the bladder and the rectum, and the recurrence of similar shaped and positioned tumours within the prostate. This research uses prostate cancer as a motivating example with a view to extending the methods to more complicated cancer sites later.

Chapter 2

Introduction

2.1 Radiotherapy

Alongside surgery and chemotherapy, external beam radiotherapy is one of the major forms of cancer treatment and about two thirds of all cancer patients undergo a course of radiotherapy. Radiotherapy exploits a therapeutic advantage whereby cancer cells' are unable to recover as well as healthy cells from radiation damage. Moreover, radiotherapy is capable of delivering near conformal dose distributions to tumours with complex geometries. Further details on the medical physics of radiotherapy can be found in Bortfeld (2006).

Radiotherapy is a local treatment that uses radiation, either internally or externally to kill cancer cells. Typically internal radiotherapy treatment is only given to patients with early-stage cancer whereas external radiation can be used for both advanced and early-stage cancer. The focus of this thesis is on assessing the treatment planning process for Intensity Modulated Radiotherapy Treatment (IMRT) patients, an external radiotherapy treatment.

Radiotherapy can be used as a standalone treatment or in combination with other treatments. Used before or during other procedures radiation shrinks the tumour to

make surgery or chemotherapy more effective and used after other treatments it kills any cancer cells that might remain.

Radiotherapy uses the damaging effects of radiation on tumour cells. As the radiation passes through tissues it produces ionisation, causing physical and chemical changes in the cells and damage to their intracellular mechanisms. Cells are most sensitive to this effect immediately before mitosis, the point in the cell cycle when the genetic material duplicates prior to cell division. Ionising radiation damages the deoxyribonucleic acid (DNA) of the cells. Different cells have different length cell cycles which contributes to the cells' sensitivity to radiotherapy. Typically cancer cells replicate and divide a lot faster than healthy cells resulting in increased tumour growth compared to surrounding normal tissues. For external radiotherapy methods high energy x-ray beams are used. Currently photon beams, in IMRT, are the most common type of external radiation but other types of radiation such as particle beams of protons/electrons or heavy-charged particles can be used. This will usually be delivered over a number of outpatient sessions during several weeks. In this way the total dose is divided into smaller doses called fractions. By fractionating the total dose of radiation this allows healthy cells to recover in between treatment sessions. The aim is that more cancer cells than normal cells will be irradiated when they are most sensitive, resulting in selective tumour cell death with relative sparing of the healthy tissue (HT).

The treatment process

While radiotherapy is generally regarded as a targeted, local therapy, it is not possible to irradiate only the tumour. Therefore, the challenge in treatment planning is to deliver a high dose of radiation to the tumour while sparing surrounding organs. The treatment process for IMRT is illustrated schematically in Figure 2.1, each section of the process will now be explained.

Clinical protocol When a patient requires radiotherapy as part of their treatment a large multidisciplinary team including clinical oncologists, planning dosimetrists, radiographers and specialist nurses are required. To ensure each patient is treated appropriately for each cancer type a clinical protocol is used. This documents the treatment process from the referral process to patient follow up. It includes information on patient preparation, required imaging, structures to outline, machine set up and delivery and prescribed doses. Each hospital has their own clinical protocol for each cancer type although many hospitals will have similar clinical protocols as they are derived from clinical trials and medical guidelines. Clinical protocols are regularly reviewed and updated as techniques improve and treatment processes develop. The use of a clinical protocol ensures each patient is treated to the same standard regardless of the clinical team involved. For this research the use of a standardised treatment method allows us to compare patients as we know they are clinically comparable due to them following the same clinical protocol.

Imaging The first step in the treatment process is image acquisition. This is primarily done using a computerised tomography (CT) scan. This uses x-rays to create a detailed image of the structures inside the patient. CT imaging produces a more detailed higher resolution 3-dimensional image than standard x-rays which is why CT is normally used despite the higher risk to HT from increased ionising capabilities of the CT. CT scans provide a detailed image of the inside of the body so that the location, size and shape of the tumour and surrounding organs can be determined.

Other imaging techniques that can be used include Magnetic Resonance Imaging (MRI) and Positron Emission Tomography (PET). When multiple imaging techniques are used it leads to a series of data sets that must be correctly combined to create the required 3D data set. This can be particularly challenging as each imaging modality can result in different alignments in space, different pixel spacing and different displays

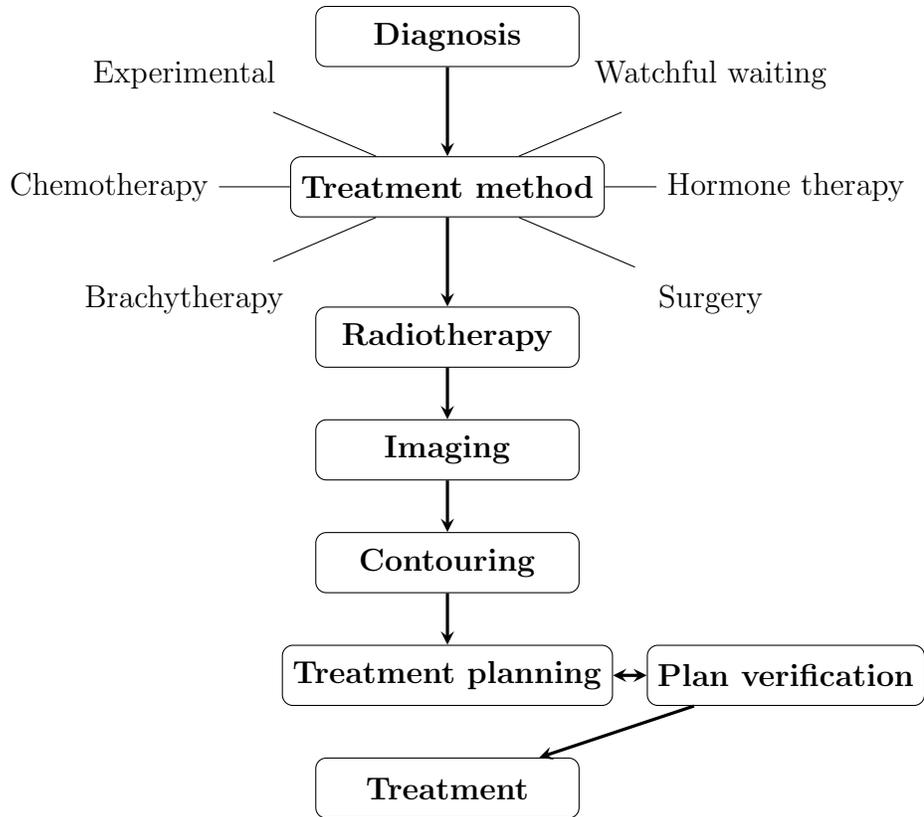


Figure 2.1: The treatment process for radiotherapy.

of anatomical structures. To overlay the different images correctly image registration and data fusion is used. This is the process of transforming each set of image data into the same co-ordinate system.

During image acquisition it is important the patient remains as still as possible to reduce the motion artefact in the images. Internal organ motion, breathing and patient fullness can cause blurring, streaks and changes in boundaries on the images. For prostate cancer patients small external tattoos are drawn on to the patient before treatment as a marker for patient alignment. Alternatively, metal artefacts or stakes, can be used as these will be present in each image and can be used to ensure correct orientation from every angle. For head and neck cancer patients plastic mesh masks are made to keep the patient still. Moulds can also be used to keep legs and arms still. It is important to correctly align the patient when the initial images are taken so

that the target points lie accurately at the isocenter¹ of the irradiation device during both planning and delivery.

Contouring After imaging a clinical oncologist will contour the important organs. This is the process where the organs are outlined on each of the individual CT scans. In particular they will identify where the cancer is, the Gross Tumour Volume (GTV) and the Organs at Risk (OARs). Figure 2.2 shows a CT image from a prostate cancer patient with the prostate, the GTV and the bladder and rectum, the OARs outlined. The schematic diagram on the right shows the relative location of the volumes. Here, CTV stands for Clinical Target Volume, which consists of the GTV and the estimate of microscopic tumour spread that cannot be detected in the CT scan. To account for uncertainties, planning target volumes (PTVs), i.e. volumes which are considered in treatment planning as targets for delivery of radiation dose, are then defined with additional margins around the CTV. These margins are detailed in the clinical protocol. Figure 2.2 illustrates that the ability to deliver a high dose to the PTVs is negatively affected by the OARs, here the bladder and rectum, which are immediately adjacent or even overlap the PTVs.

Creating the treatment plan The next stage in the treatment process is to create the treatment plan. Optimisation techniques are used to: 1. select the beam angles required to deliver treatment, 2. calculate the fluence map and 3. select the delivery sequence for the MLC. These three optimisation problems are often referred to as the geometry, intensity and realisation problems. All three of these optimisation problems cannot be dealt with at once, therefore they must be optimised separately and their results combined. Some methods incorporate the geometry and intensity or intensity and realisation problems into a combined optimisation problem. A brief explanation of each is given below; more detail can be found in Ehr Gott et al. (2010).

¹The isocenter is the point through which the central beam of radiation passes.

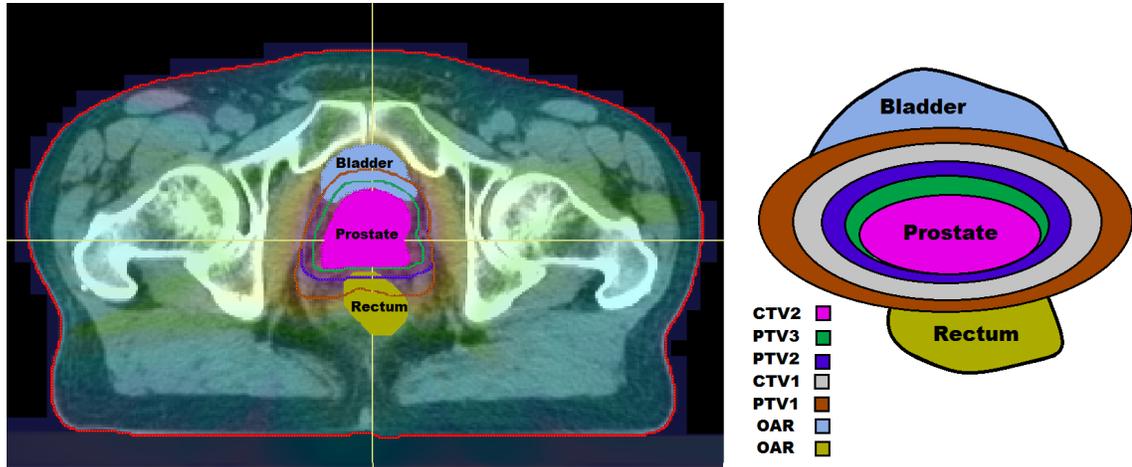


Figure 2.2: A CT scan with organs contoured and a schematic overview of relevant volumes for radiotherapy treatment planning.

Geometry: Beam angle selection Here the number of beams and the angles at which to deliver the beams are decided. The angles are determined by the position of the couch and gantry. This is particularly difficult in patients with complicated geometry such as large areas of overlapping tumour and OARs.

Intensity: Inverse treatment planning The correct dose of radiation to administer to the patient must be calculated. The goal is to find the intensity that results in a desirable dose distribution. This is regarded as an inverse planning problem as the clinicians know the desired dose of radiation to the PTV but need to calculate the corresponding intensity. The calculated intensity is used as the variable in a single-objective optimisation process to provide a single solution of the dose distribution that can then be used in further stages of the treatment planning process.

Radiation is measured in Gray² (Gy). The dose distribution is defined by the prescription to each area and limits that must be adhered to, e.g. for the prostate, treat the PTV with 74 Gy and limit the dose to the rectum to 50 Gy. The patient's image is divided into voxels and the amount of radiation to be deposited into each

²One Gray is defined to be the absorption of one joule of radiation energy per kg matter.

voxel is calculated. For details on how the dose deposition matrix is calculated see Holder (2005).

Realisation: Calculation of leaf positions/trajectories After solutions for the geometry and intensity problems have been found the leaf positions of the MLC must be calculated. A MLC is a large metal device comprising individually moving leaves that modulate the intensity to different areas, Figure 2.3. During treatment these leaves change position to block different areas of the radiation field. This shapes the radiation beam to the targeted area, creating a conformal shaping of the beams.

The radiation can be delivered via the step and shoot or dynamic method. In the step and shoot method the radiation beams are turned off while the MLC leaves change position. In the dynamic technique the dose is delivered continuously and the leaf pairs change position during the dose delivery. This results in decreased treatment time and hence reduces patient discomfort and therefore reduces the risk of patient movement.

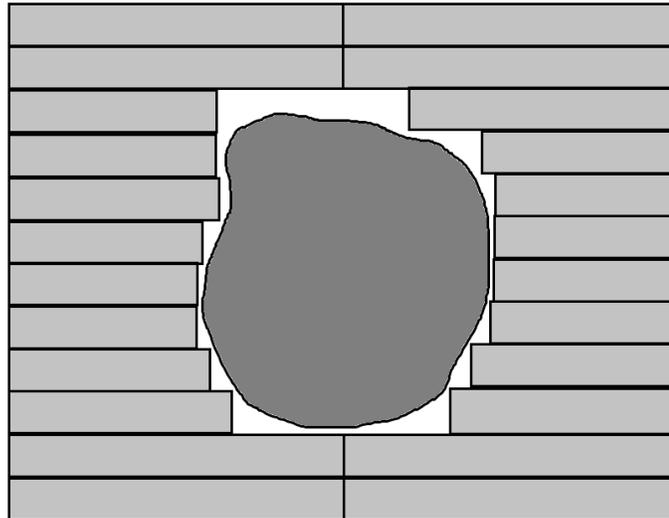


Figure 2.3: A schematic of a MLC showing the pairs of leaves shaped around the target area.

Treatment plan optimisation Once an initial treatment plan has been made planners will try and improve it to ensure prescribed doses are met and OAR tolerances are not exceeded. It is very difficult to design an optimal treatment plan due to the large number of parameters involved. For example, Ehrgott et al. (2010) note that for a five beam coplanar treatment there are 4.9×10^{10} possible beam arrangements. A MLC can have 8000 intensity parameters for each possible beam arrangement leading to billions of possible MLC leaf pair configurations. Therefore it is currently not feasible to plan all of the above three phases together.

A plan can be evaluated both through visual inspection and quantitative measures to check if it meets the required prescribed doses and restrictions on dose to the OARs. For complex cases it is unlikely to be an acceptable plan the first time round. As a result, the planner must then change parameters and re-plan until a satisfactory plan is found. For more details on this process see Schlegel and Mahr (2002). These methods however, are inefficient and are subject to human error. The trial and error aspect means the most desirable plan that could be achieved may not be found.

Treatment plan verification It is important that each plan is verified to ensure that it is feasible and it meets all the clinical criteria. Continuous evaluation of the treatment planning process is required due to the changes in weight of the patient, muscle composition and growth of the tumour leading to images becoming outdated. This makes treatment methods that take a long time to plan and perform unsuitable. Time, expertise and the cost of a treatment must be considered throughout the treatment process.

To aid plan verification a phantom can be used. A phantom is a model of the body, or a specific part of it that can be used to simulate the effect of radiation on tissue. In radiotherapy treatment a water phantom is often used because tissue consists mainly of water. Therefore, by calculating the radiation in the same volume

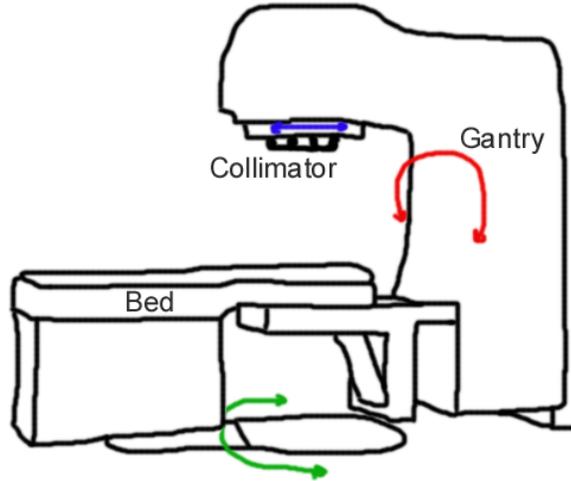


Figure 2.4: Treatment couch (bed) with linear accelerator and rotating gantry. Wikipedia and Narayane (2011)

of water reasonable estimates of the absorbed dose-distribution can be obtained. The use of a phantom also means the leaf pair sequence of the MLC can be checked i.e. the MLC can physically deliver the planned treatment.

Treatment During treatment the patient lies on a rotating couch. Above the couch a linear accelerator is fixed to a gantry that can rotate around its central axis, Figure 2.4. A linear accelerator forms the beams of radiation which then pass through the rotating gantry allowing the radiation to be directed at the patient from any angle. The head of the gantry contains a multileaf collimator (MLC) that shapes the radiation beams via pairs of moving leaves. The rotation of the couch and gantry means the radiation can be focused onto the patient from nearly every angle. The linear accelerator accelerates electrons by microwave technology to produce high energy x-rays. These x-rays then pass into the head of the linear accelerator where they will be shaped by a MLC before being emitted as the gantry rotates around the patient (Shepard et al., 1999).

With external treatment the risk to HT in the surrounding areas, particularly tissues within the path of radiation that have a relatively high cell division rate (e.g.

bowel and scalp hair follicles) is higher than that of internal radiotherapy. This can cause toxicity: diarrhoea and hair loss in these examples. Radiation oncologists endeavour to limit the radiation dose to non-tumour areas to limit toxicity.

Throughout the treatment planning process continual evaluation of the process itself is required to ensure the best treatment is received by the individual. This is particularly difficult in the optimisation processes involved in treatment planning due to the trade-offs between treatment of the tumour and sparing of the OARs. Treating the tumour can be achieved by using very high radiation doses. However, this will also effect healthy cells and potentially subject the patient to lethal doses of radiation. Conversely, by not irradiating at all, the patient may die of cancer. Neither is desirable. Therefore, a suitable balance must be found between treatment and sparing.

2.1.1 Prostate cancer

Prostate cancer is the second most common form of cancer in male patients and it accounts for a large number of radiotherapy patients. The normal prostate is a walnut sized gland present in males, that sits below the bladder above the penis. It is located in front of the rectum and the prostatic urethra runs through the centre of the prostate connecting the bladder and the penile urethra. Prostate cancer is the development of cancer cells in this gland.

Prostate cancer can be treated by a variety of methods including expectant management (watchful waiting), surgery, radiation therapy, cryotherapy, hormone therapy, chemotherapy and occasionally vaccine treatment. In 2012 the prostate cancer study group performed and published a large comparative effectiveness study to compare the success rates of all prostate cancer treatment options. They concluded that radiotherapy external beam treatments proved suitable for approximately 80% of the patients (PCRSF, 2013).

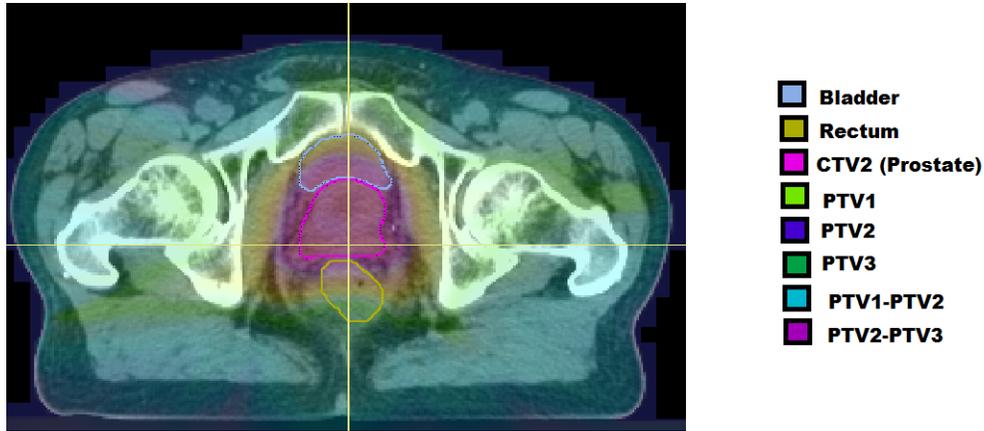
2.1.2 Treatment planning: Rosemere

Throughout this thesis we use data from Rosemere Cancer Centre, Royal Preston Hospital, Lancashire (from now on we refer to it as Rosemere) to motivate the real life applications of this research. Rosemere is a specialist centre based at the Royal Preston Hospital that provides care and treatment to patients in Lancashire and South Cumbria diagnosed with cancer. The radiotherapy department sees over 200 patients each day, treating patients with a wide range of cancer via treatment methods including chemotherapy, radiotherapy and surgery (Lancashire-Teaching-Hospitals, 2017). Here we focus on radiotherapy at Rosemere.

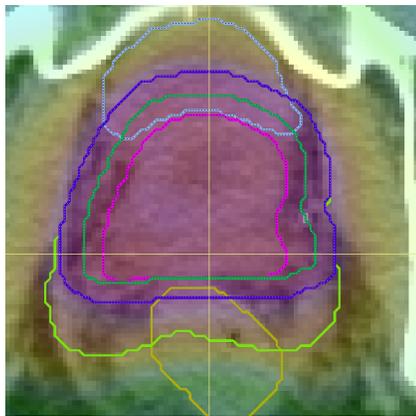
Rosemere's clinical protocol for prostate radiotherapy is based upon the CHHiP clinical trial and guidelines from the Royal College of Radiologists (RCR, 2018), Clinical Oncology Information Network (Mason et al., 1999) and British Association of Urological surgeons (BAUS, 2019). Throughout Chapters 5 to 7 we use variables that are extracted from Rosemere's clinical protocol, i.e. we convert the recommended prescribed doses to the PTVs and the constraints on the OARs to variables we can analyse.

In addition to the clinical protocol Rosemere documents guidance on how to plan IMRT treatment plans using the Pinnacle treatment planning system (Philips, 2009). This is called the work instruction. This is used to guide medical physicists, clinical oncologists and planning operators throughout plan creation. The working instruction details how plans should be made on the treatment planning system. This includes the structures to be defined based on the PTVs and OARs outlined by the clinicians. The work instruction also includes advice on how to improve plans and the IMRT objectives used in the planning software.

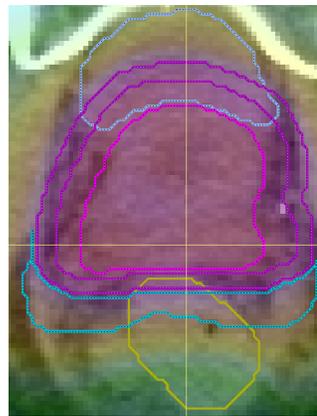
Rosemere treats prostate cancer patients with a three-dose level regime. To do this after the CTVs have been contoured the outlined structures are grown to define three new structures: PTV 1, PTV 2 and PTV 3, see Figure 2.5b. PTV 1 includes



(a) Bladder, rectum and prostate (CTV 2).



(b) PTV 1, PTV 2 and PTV 3.



(c) PTV 1-2 and PTV 2-3.

Figure 2.5: CT scan showing examples of outlining for a prostate cancer patient at Rosemere. Image using data from Rosemere with CERR (Deasy et al., 2003).

PTV 2 and 3 inside its defined area and PTV 2 includes PTV 3. The PTVs define three structures that include the CTVs and additional margins around the prostate (and seminal vesicles). The prescribed doses are for the PTV 1-2, PTV 2-3 and PTV 3 where PTV 1-2/PTV 2-3 are the structures PTV 1 not in PTV 2 or PTV 3 and PTV 2 not in PTV 3, see Figure 2.5c.

At Rosemere two schedules for Prostate cancer treatment are used. They are 60Gy in 20 fractions or 74Gy in 37 fractions. These are considered to be clinically equivalent treatments as a result of the CHHiP clinical trial concluding “using 60 Gy in

20 fractions is non-inferior to conventional fractionation using 74 Gy in 37 fractions and is recommended as a new standard of care for external-beam radiotherapy of localised prostate cancer” (Dearnaley et al., 2016).

If after an initial plan has been created the prescribed doses to the PTVs are not met or the dose to the OARs are too high, planners must replan. At Rosemere the work instruction documents the standard methods to employ if the desired prescribed doses and mandatory OAR tolerances are not met. One method to improve a plan is to reduce the PTV 1 posterior border to aid rectum sparing. Although this is common practice, these improved plans cannot be included in our analysis as they are no longer clinically comparable due to the redefinition of the PTV 1. Other methods Rosemere employ to improve a plan include adding additional constraints to the rectum or bowel inside the PTV 1 or adding a skin protector. If plans have these modifications, they can still be used in our analysis as they are still clinically comparable as no structure redefinition occurs.

To aid plan evaluation Rosemere examine a Dose Volume Histogram (DVH) (see Section 2.2.1) analysis spreadsheet to assess the DVH metrics. Plans will be replanned until all mandatory constraints are met and the clinician is satisfied with the dose deposition in the PTVs. Plan evaluation and finalisation can also include assessing the MLC shapes via a video of the leaf positions. This is to check that there are no large jumps between MLC positions. All treatment plans must be checked and signed off by the clinical oncologist in charge before treatment can commence.

At Rosemere two different MLC treatment machines are used. The Elekta Beam Modulator (BM) and the Elekta Agility (AG). A summary of the differences in the machine specifications can be found in Ruschin et al. (2016) and a more detailed overview of the MLCs can be found in Patel et al. (2005) and Thompson et al. (2014) respectively. The main difference between these machines are that the BM has 40 4mm leaf pairs and the AG has 80 5mm leaf pairs. In November 2015 when data

collection started, the majority of the treatments were delivered on the BM machine. However, now Rosemere only treat patients on the AG machine. Rosemere consider the two MLC machines to provide clinically comparable treatments. Literature to support this conclusion include Ruschin et al. (2016); Zhang et al. (2017) and Kantz et al. (2015).

When treating prostate cancer, the variability of the volume of the bladder can be reduced by patients emptying their bladder and then filling it with a set volume of fluid. Not only does this help with the reproducibility of the treatment but it helps reduce the small bowel and bladder toxicity (Waddle et al., 2018). However, for some patients this can be very uncomfortable or not possible. Preston are currently interested in the affect the volume of the bladder may have on the treatment.

2.2 Plan quality evaluation

During IMRT treatment planning, suboptimal plans are generated consistently. This has encouraged objective measures of plan quality and the development of software that can accurately predict if a plan can be improved or not (Moore et al., 2011). Such decisions occur daily in most radiotherapy clinics, yet there are few, if any, quantitative metrics to base decisions on when evaluating treatment plans. From a set of acceptable plans how do you select the best plan for an individual patient? How do you define best? The best plan can be chosen based on different attributes. For example, the plan that most effectively protects HT, the least expensive plan or the shortest delivery time. These different evaluation criteria result in a need for an additional tool that can aid professionals in their decision making; a tool that includes both past experience and data from previous successful treatments.

In Section 2.1 the three separate optimisation processes involved in creating a treatment plan were introduced and the parameters that are involved in it. Many

of the parameters interact but the nature of this interaction is unclear and it is hard to verify whether changing one parameter in the selected plan will result in a better plan. When radiotherapy treatment was first introduced plans were designed manually and hence the assessment of the plans could be done by the physicians. The limited capabilities of the radiotherapy equipment meant plans were relatively straightforward. As medical technology and computer capability has improved so has the complexity of planning, delivery and evaluation of the treatment. Shepard et al. (1999) highlighted the need for mathematical optimisation tools to aid treatment planning and assessment and it has been an active field in operational research since.

This challenge has been addressed in recent years by the development of mathematical models and algorithms that are based on multi-objective optimisation, which explicitly deals with the conflicting nature of the goals of radiotherapy and enables the exploration of trade-offs between them. Breedveld et al. (2009); Falkinger et al. (2012); Jee et al. (2007) and Wilkens et al. (2007) apply multi-objective goal programming, where a sequence of optimisations is conducted to achieve prioritised and rank ordered treatment goals pre-defined by the planner. The approaches of Bokrantz and Forsgren (2013); Craft et al. (2006); Lin et al. (2016) and Shao and Ehrgott (2016) compute a representative set of efficient plans that capture potential trade-offs in optimisation objectives. These methods are generally complemented by a so-called navigation method that assists the planner in selecting the best plan for the particular patient from the representative set. Craft and Monz (2010); Craft and Richter (2013); Ehrgott and Winz (2008); Lin and Ehrgott (2017) and Monz et al. (2008) are examples of such methods. Ruotsalainen (2009) proposes an interactive approach, where the planner iteratively adjusts preferences based on the knowledge learned from the generated plans and an optimisation run is repeated after each interaction with the planner.

In practice, however, the vast majority of cancer centres worldwide do not have

access to multi-objective optimisation based treatment planning systems. Hence, radiotherapy treatments are usually designed with a planning system that is based on single objective optimisation methods, often a weighted sum of the conflicting objectives mentioned before. Ehrgott et al. (2008) provide a review of optimisation problems in radiotherapy and Romeijn et al. (2004) investigate relationships between multi-objective and single-objective approaches.

The difficulty surrounding selection of the best treatment arises as human planners cannot be assumed to be moving toward some optimal solution. A more experienced planner will have an idea of what constitutes a good solution but they will not be able to determine whether or not it is truly the best plan that could be achieved in the current circumstances. It is difficult to compare individual patients against one another. The achievable doses to PTVs and the sparing of the OARs depends on many unique factors including the individual's anatomical geometry, tolerance to radiation and cancer progression. Although there are IMRT planning objectives which have been determined from population-based data to give physicians a suggested target, they are very general and, consequently, unsuitable for more irregular cases such as when a large overlap between the OAR and PTV exists. Without widespread IMRT optimisation engines being developed physicians run the risk of "inadvertently implementing treatment plans in which clinically relevant improvement would have been possible" (Moore et al., 2011). It would be beneficial to reduce this dependency on experience and patient variability.

In an ideal world a simple quantitative score could be produced that integrates all data and objectives, incorporates the relationship between irradiated tumour tissue and HT and outputs the best overall plan. However with so many parameters, variables and objectives this is not necessarily realistic. This is an area of the treatment planning that could offer a great improvement to the quality of patient care and increase the effectiveness and efficiency of the planning and treatment stages. It is for

these reasons that this area of research has been the subject of many studies and why it remains a valuable area into which to invest research time.

2.2.1 Existing treatment plan evaluation methods

A common method for comparing treatment plans is via the Dose Volume Histogram, DVH. This is a histogram that graphically summarises the predicted dose distribution for regions of interest. DVHs are constructed by binning the dose values for each voxel. The histogram provides details of the frequency of each dose throughout the volumes of interest. More commonly the cumulative distribution volume histogram (cDVH) is used and is referred to as DVH interchangeably. This is calculated by summing the DVH starting at the dose of interest up to the maximum dose. Both the volume (vertical) and dose (horizontal) axes can be displayed in terms of gray or in relative % volume or % dose, the latter making it easy to see if the minimum dose requirement for a particular PTV has been met by a particular plan. An example of a DVH can be seen in Chapter 6.

DVHs have the benefit of clearly presenting the uniformity of dose in the PTV. However, they are not suitable as a stand-alone decision making test. This is because of the loss of positional information in the volume(s) (Drzymala et al., 1991). DVHs are also subject to insensitivity in terms of hot and cold spots; they cannot detect small areas within the PTV(OAR) where the dose is significantly too low(high). This can have repercussions if it is a fast growing tumour or if an OAR is a serial organ³. Due to the binning process that creates the histogram care must be taken to pick suitable sized bins especially when analysing large volumes with DVHs. DVHs are used as the basis for many treatment plan metrics and assessment methods. For example many hospitals use V_x and D_x metrics. V_x denotes the percentage of the volume of an OAR

³Serial organs can tolerate low doses of radiation to the whole organ but cannot tolerate high doses to even a small volume.

receiving x Gy or more of radiation dose and D_x is the dose received in Gy by $x\%$ of the volume of a PTV. Both of these can be found using the DVH curves for the associated organ and offer a way of checking clinical criteria are met.

Conformity Indices (CIs) were introduced by the Radiation Therapy Oncology Group (RTOG) in Shaw et al. (1993) as a method of comparing the degree of conformity between DVHs. CIs assign a score to each treatment plan to allow comparison between different patients or multiple treatment plans from the same patient. There are many different proposed indices with varying degrees of interpretability. An overview can be found in Feuvret et al. (2006).

Normal Tissue Complication Probability (NTCP) models estimate the likelihood of complications to the NT during treatment. They aim to reduce complicated dosimetry and anatomic information to a single risk measure (Marks et al., 2010). There are a selection of NTCP models that can be used including DVH reduction, tissue architecture and multiple-metric models (Marks et al., 2010). NTCP models take into account the different organ architecture of a patient. In Schlegel and Mahr (2002) statistical models for NTCP are detailed. However they are rarely used clinically.

A common method of comparing treatment plans is to use decision objectives (DOs). DOs can be objectives from the inverse planning stages, a mathematical representation of one of the constraints or a surrogate objective that represents a criterion that is too hard to represent mathematically. There are a wide range of parameters that can be used as DOs including standard deviation of the PTV, mean DVH or Equivalent Uniform Dose (EUD) of the OARs. Plans are ranked according to how many of the prespecified DOs are met and the highest ranked plan is considered the best plan. However there is normally a disparity between the inverse planning objective space and the decision objective space; the objectives used to create the plan are not the same as those to assess the plan. This makes it difficult to relate the two stages of the planning process together. Phillips and Holdsworth (2011)

highlight this by comparing the different objectives used for the two stages in seven different publications in the International Journal of Radiation Oncology. None of the publications used the same set of objectives for the inverse planning and the plan assessment.

A similar concept to DVHs is the overlap volume histogram (OVH), a metric proposed by Wu et al. (2009) to describe the “spatial configuration of an OAR with respect to the target”. The OVH distance is signed such that an overlap is represented by a negative distance. The OVH represents the percentage of the OAR volume that overlaps the PTV. This produces a curve that characterises the relative spatial configuration of the two objects. OVHs were also used in Yang et al. (2013) to compare a new treatment (insertion of hydrogel before radiotherapy) to those plans without.

Bohsung et al. (2005) use a quality score to compare plans from 11 different centres in Europe. Six objectives based on over and under-dosage to the PTV, over-dosage to the OAR and over-dosage to the NT were compared by summing the difference between the planned dose M_j and the delivered dose, C_j for the j^{th} objective. The quality score was then calculated by

$$S = \sum_j \begin{cases} |M_j - C_j| & \text{if objective is violated} \\ 0 & \text{otherwise.} \end{cases}$$

In this way a plan that met all six dose objectives would have a score of 0 and plans could be ranked whereby increasing value in S represents increasing inferiority of plans.

In Niemierko (1997) the concept of equivalent uniform dose (EUD) was introduced for tumours to try to combat the “oversimplification made when the dosimetric aspects of a complex three-dimensional treatment plan are reduced in the patient’s records to a dose or a few doses at the reference point or points”. Two dose distributions were considered equivalent if they resulted in the same radiobiological response. In this way the EUD was defined as “the dose in Gy, which, when distributed uniformly across

the target volume, causes the survival of the same number of clonogens⁴ for any dose distribution (Niemierko, 1997). In Niemierko (1999) the concept was developed further to include NT. This was termed the generalised Equivalent Uniform Dose (gEUD). The gEUD penalises over-irradiation of the OAR and under-irradiation of the PTV, it enables comparison of different fluence maps and therefore can aid treatment plan selection.

To calculate the gEUD, the total dose deposited into each voxel in each region must be calculated. The regions refer to whether the voxel is in an OAR, PTV or NT and therefore require different doses of radiation. To do this, it is assumed that the total number of beams and the beam angles have already been decided upon. The dose distribution d_j^R for voxel j in region R is calculated and the gEUD for each region can be found,

$$gEUD_a^R(\mathbf{x}) = \left(\frac{1}{m^R} \sum_{j=1}^{m^R} (d_j^R)^a \right)^{\frac{1}{a}}.$$

Here m^R is the total number of voxels in region R and a is a region-dependent parameter. By setting $a < 0$ for the PTV and $a > 1$ for the OAR the model correctly penalises under(over)-irradiation in the PTV (OAR). Furthermore by increasing the size of a , $|a|$, the sensitivity of the gEUD to less(more) irradiated voxels in the PTV(OARs) increases. The difference in a also affects the optimisation problem, when $a \geq 1$ $gEUD_a^R(\mathbf{x})$ is a convex function whereas when $a < 0$ it is a concave function.

gEUD measures can be used in a multi-objective fluence map optimisation (Cabrera et al., 2014). The optimisation problem aims to minimise the gEUD of some OARs while still delivering the prescribed gEUD to the PTV and ensuring the gEUD of all the other OAR are below a set bound.

Lin et al. (2013) also use the gEUD to assess treatment plans. They use a DEA model to evaluate treatment plans for prostate cancer at Auckland Radiation Oncol-

⁴Clonogens are cells that can form clones of themselves, often but not limited to tumourous cells.

ogy in New Zealand. They compared 37 plans via the gEUD for the rectum, D_{95} for the prostate and overlap of the percentage volume of rectum overlapping the PTV. After replanning 5 of the plans that they believed could improve they concluded that “the results confirm that DEA is capable of identifying plan improvement potential and predicting the best attainable plan in terms of the inputs and outputs”. We will discuss this study more in Section 5.2.1 where we apply the method to data from Rosemere.

Another technique for treatment plan evaluation is to use a database of past successful treatment plans to help inform decisions for new treatment plans. In Wu et al. (2009) when a new plan is created the geometric configurations of the OAR and PTV are compared to those in the database via their OVH’s. In Moore et al. (2014) automatic treatment planning is used to provide planners with a good initial plan that can be used as the basis for their trial and error optimisation approach. This is done by comparing the doses to the OARs of the patient to a database of existing ones via their OVH’s. The system highlights those plans which fail to reach the objectives met by prior similar patients. The database exists as a separate entity outside the planning system to achieve optimal interfacing of the software with the different treatment planning systems. This means centres that use different software can still benefit from the shared knowledge of previous plans. This is particularly important for future radiotherapy treatment clinics being set up and new software being trialled as they will have no prior plans to assist them.

Rapidplan[®] software by Varian use a knowledge based system (KBS). Previous plans are saved in a database to provide an initial starting point for future treatment plans. This provides a prototype clinical strategy for the treatment of a tumour in a particular location in the body. By separating knowledge data from the optimisation programme it allows the knowledge to be adapted to new requirements. As treatment plans improve the knowledge can be updated further.

In Moore et al. (2011) a Quality Control (QC) tool was implemented into pre-existing IMRT planning software. Treatment plans are compared by looking at the ratio of the mean principal OAR dose (\mathbf{D}_{mean}^{OAR}) to the PTV prescription dose (\mathbf{D}_{pres}^{PTV}) and comparing this value to the fraction of the OAR overlapping the PTV, $\frac{V_{overlap}}{V_{OAR}}$. 42 IMRT plans from prostate and head & neck tumours, that have been previously implemented were used to create a model to predict the achievable mean doses. The $\frac{\mathbf{D}_{mean}^{OAR}}{\mathbf{D}_{pres}^{PTV}}$ and $\frac{V_{overlap}}{V_{OAR}}$ were plotted against each other resulting in a function with a lower-bound, representing the best organ sparing as a function of overlapping volume. The model's efficiency was then tested by integrating it directly into the treatment planning work-flow. The clinical prostate and head & neck IMRT plans generated in the 3 months prior to the introduction of the OC tool were compared with those generated in the following 4 months to determine the effect of the proposed QC technique. The study reports significant differences between the mean dose and the prescribed dose suggesting this is a suitable tool to increase normal tissue sparing and aid quality control of IMRT plans.

In Chapter 5 we will use some of the evaluation techniques discussed here to develop a plan evaluation method based on DEA.

2.2.2 Introduction to data envelopment analysis and its application to radiotherapy treatment planning

Data envelopment analysis (DEA) is a management science technique to perform relative performance analysis among a group of comparable “decision-making units” (DMUs). In the context of the evaluation of radiotherapy treatment plans, DMUs are the treatment plans and DEA assesses how well the plans perform in “transforming inputs into outputs”, *i.e.* delivering the prescribed dose to the tumour while limiting the dose delivered to OARs. The resulting efficiency score is relative to the set of plans considered in the study. A mathematical introduction to DEA can be found in

Chapter 3.1, and a description of its application in radiotherapy treatment planning is presented in Lin et al. (2013).

Data in the form of inputs and outputs is used to form an efficient frontier. This comprises of a subset of DMUs that represent the standard of performance that all DMUs should try to achieve. Only DMUs with inputs (outputs) that cannot be improved without worsening any other input or output will be on the efficient frontier. In this way the efficient frontier always represents the current best attainable result as determined by the dataset.

Efficiency is assessed by comparing a particular DMU to all other DMUs in the set and searching for possible improvements. To do this the relationship between the inputs and outputs must be considered. If it is believed that the outputs increase by the same proportional change as the inputs a constant returns to scale (CRS) model should be selected. Otherwise if it is believed the change in inputs does not result in a proportional change in outputs a variable returns to scale (VRS) model should be used. The chosen economy of scale affects restrictions put on the feasible set. The orientation of the DEA model must also be considered. This should be done based on which of the inputs/outputs the decision maker has control over and the desired objectives.

If we consider the radiation dose to OARs as the “cost” or input necessary to achieve the “profit” or output of radiation dose to the PTV, we can view the problem as one of input-output or performance analysis in an economic setting. This approach was developed by Lin et al. (2013), who use DEA for the evaluation of radiotherapy treatment plans for prostate cancer at Auckland Radiation Oncology in New Zealand.

In a radiotherapy setting the intensities, which are set by the oncologists can be thought of as inputs to the model and the dose distribution, which ultimately determines the medical outcome of the treatment, as the output. In this way a good treatment plan would minimise the dose to the HT and maximise the dose

to the PTVs. This lends itself to an input orientation method. DEA aims to assess treatment plans by comparing each new plan to previous treatment plans to provide a peer assessed performance measure. Each individual DMU receives an efficiency score which is an indication of how well it performs compared to other DMUs. Those DMUs deemed efficient create an efficient frontier that can offer further information on the relationship between the inputs and outputs and give guidance on realistic targets for future plans. This provides an enhanced understanding of the set of DMUs as individuals, instead of relating them to a perceived average.

DEA requires no assumptions to be made about the functional form of the efficient frontier or a priori weights for the inputs (outputs). Instead the efficient frontier is based on the best attainable results in a dataset of historical treatment plans. This ensures that the suggested improvements to plans are attainable with the particular hospital's limitations. However, this does mean that the efficient frontier may not include the best plan if it has not yet been achieved. This will be partially overcome by the addition of new plans to the database causing the efficient frontier to move closer to the true optimal frontier. The suboptimal plans are then expressed as changes that need to be made to the inputs to improve the plan.

Additional benefits to applying DEA to assess treatment plans is DEA's ability to handle multiple inputs and outputs. This is particularly important due to the conflicting nature of treatment objectives. However, care must be taken not to include too many variables as this can cause the DEA to lose discrimination power on the performance of the DMUs. DEA can adapt to exogenous variables *i.e.* variables that come from outside the model and are unexplained within the model. For example the overlap between the OAR and PTV or the volume of the patients' bladder.

In Chapter 3 we introduce the mathematical DEA models we will use and in Chapters 5 to 7 we apply them to data from Rosemere.

2.3 Uncertainty

The International Atomic Energy Agency (IAEA) published a report on the accuracy requirements and uncertainties in radiotherapy (Kurokawa, 2018). Their motivation for this report was that “there is no measurement (or procedure) in the radiation treatment process that can be performed perfectly; each step has a corresponding uncertainty”. They defined this uncertainty to quantitatively be “a parameter that characterizes the dispersion of values that can be obtained for a particular measurement when it is performed repeatedly”. This report is supported by over 150 countries including the United Kingdom and New Zealand. A review of the report is given in van der Merwe et al. (2017). They summarise the report in nine general recommendations of which three are applicable to this thesis.

1. “Each institution should determine uncertainties for their treatment procedures. Sample data are tabulated for typical clinical scenarios with estimates of the levels of accuracy that are practically achievable and suggested action levels”.
2. “Comprehensive quality assurance programs should be in place”.
3. “For reporting purposes, uncertainties should be presented”.

Recommendation 1 highlights the need for estimates of the uncertainty present throughout the treatment planning process. We use published estimates to advise us on suitable uncertainty values for our models in Chapter 5. Recommendation 2 expresses a desire for quality assurance programs. Our methods provide a quantitative evaluation of how well a plan compares to previous treatment plans. Recommendation 3 highlights the continual need to research and report findings on uncertainty. This will aid in further development of our models as more data on the uncertainties involved in the treatment process will allow us to adapt our models. This could be by calculating suitable allowable uncertainty levels for our simulation in Section 6.1 or

redefining our uncertainty sets in the robust optimisation model in Section 6.2. The report Kurokawa (2018) highlights that uncertainty in radiotherapy treatment is still a vast and growing area of research for both monitoring and determining the amount of uncertainty, and how to handle it in the treatment planning models.

2.3.1 Sources of uncertainty in radiotherapy

It is well known in radiation oncology that the outcomes of radiotherapy differ from the plans, i.e. the doses delivered to structures are usually slightly different from those calculated during treatment planning. This uncertainty in predicting radiation dose delivered to PTVs and OARs has many sources. The first is the inherent uncertainty in the computational models for radiation dose. These are mathematical models of the physical interaction of radiation (photons) with biological tissue representing the deposition of dose in the body. This process cannot be described with 100% accuracy. Further factors are the physical design of the treatment machines delivering the radiation. Small amounts of radiation leakage cannot be prevented due to the need to modulate intensity of radiation across beams. Many assumptions have to be made to prevent the dose-computation calculation becoming too complicated and time consuming. For example it is assumed that adjacent CT scan slices are identical in tissue composition to the slice in which the dose is being calculated; the lateral scatter from adjacent slices is negligible because a small enough field is being used (Khan, 2003) and when calculating the average number of surviving clonogenic cells the average can be found using a negative-exponential function of the dose (Bentzen, 2009).

Apart from these mathematical and engineering factors, a compounding issue arises due to the delivery of the treatment over several fractions. Since the patient returns for treatment over a number of days (37 or 20 in our Prostate case study) it is inevitable that the positioning of the relevant structures differs slightly from day

to day. Roeske et al. (1995) found that the prostate volume can change by $\pm 10\%$ and the rectal and bladder volumes can change by up to $\pm 30\%$ and Moiseenko et al. (2007) saw changes in bladder volume from 419cm^3 to 90cm^3 during treatment. In Booth and Zavgorodni (2005) they report that large variations in rectal dose (> 15 Gy) are possible when they evaluate the range of inter-fraction rectal dose variations for prostate cancer patients. Antolak et al. (1998) investigated changes in prostate volume throughout treatment and found that averaged over all patients and not including set-up errors, the mean displacements were slightly less than 0.5 cm. However for individuals the range of displacement was 0.03 cm to more than 1.5 cm. Das et al. (2008) observed high variability among planners and institutions, reporting that the median dose to the tumour can vary by $\pm 10\%$ of the prescribed dose across 96% of the patient population.

When defining the PTVs and OARs there is a large variability in human contouring accuracy (Moore et al., 2012). This is often dependent on the experience of the planner, the technology and the software used in the hospital. It has been found that even in organs that have high definition, from high boundary contrast, a CT scan contouring reproducibility was limited to 2mm (Moore, 2009). Other inter-institutional studies report even larger variation than this (Moore et al., 2012). Published expert gold standards are available for common tumours. However, if a tumour is irregular or there is a large overlap the variation can increase further. To improve accuracy the same structures can be repeatedly contoured but this increases the duration of an already time consuming process. If repeated contouring is used a metric called the dice coefficient/measure (DC) can be used to compare results (Allozi et al., 2010). The DC normalises the intersection of two sets, here the area of a contour, with the sum of constituents. This measures how closely two sets of contours are related to each other via the following metric

$$DC = \frac{2|A \cap B|}{|A| + |B|}.$$

The DC is available in many commercial treatment planning systems. However, this metric only looks at the differences in the volumes not the surface variation, a mathematically more complex task.

Throughout the treatment process images are required, initially as a diagnostic tool, then to develop a treatment plan and finally in treatment administration to ensure correct alignment of machinery and the patient (Chowdhury et al., 2010). Imaging technique errors can take a variety of formats from misalignment of the patient during imaging to technical and human error in the form of inattention or inexperience. One significant error that occurs after the imaging process as a result of poor imaging is that the tumour volume is not clearly defined. This occurs because of “microscopic extensions, which may or may not exist,” appearing on an image (Holder, 2005). Treatment planners need to take the irregular appearance of the tumour shape into account to allow for these microscopic extensions (Kessler, 2006).

One major area of subjectivity in treatment planning is the dependency on the treatment centre and its corresponding equipment. To understand this further and to assist in draft guidelines for the verification of IMRT, the QUality ASSurance of Intensity-MODulated beams in radiation Oncology (QUASIMODO) group set up an investigation between 11 of their centres throughout Europe (Bohsung et al., 2005). Each centre was provided with phantoms and pre-contoured CT scans (to remove the subjectivity of contouring) and detailed instructions on how to perform the treatment planning. The PTV was chosen to test the limitations of the techniques and methods of each centre. Specific objectives were set but there were no constraints on how to perform the treatment planning.

After each centre had completed the treatment planning the results were extracted, collected and then compared to the original dose objectives. Of the 11 plans only one plan fulfilled all the objectives. Of the remaining plans 6 failed some objectives but were still clinically acceptable and the remaining 4 were deemed clinically unaccept-

able. The plan that fulfilled all the objectives was from the only centre that used Intensity Modulated Arc Therapy (IMAT) as their treatment technique and IMAT is recognised to be perfectly adapted to the specific geometric arrangement of the PTV and OAR that occurred in this setting (Bohsung et al., 2005). The results highlighted the considerable differences between the DVHs of the PTV and OAR for different plans. Different centres make different compromises between target coverage, dose outside the PTV, number of beams and the number of monitor units. Although the plans did differ the study found that if the planning goals are clearly defined prior to treatment, comparable dose distributions can be achieved between different planning programmes.

From this section we conclude that there are many uncertainties associated with the radiotherapy treatment process. From the mathematical modelling of the dose distribution and the physical interaction of radiation with the biological tissues, to the variability in human contouring and treatment equipment alignment. We now consider how these uncertainties can be modelled.

2.3.2 Handling uncertainty in radiotherapy

We turn our attention to methods for modelling the uncertainties introduced in Section 2.3.1. Although we are ultimately concerned with modelling the overall uncertainty in a treatment plan, we need an appreciation of the way uncertainty is handled throughout the process. Therefore, here we give a brief overview of some of the methods used to model the uncertainty in radiotherapy treatment planning.

To model the uncertainty arising from the physical interaction of radiation with biological tissues Kozlovska et al. (2017) develop a physical-mathematical model of the distribution of optical radiation in biological tissues. This allows them to determine the change of intensity of optical radiation and investigate the effect of parameters including installation angle of the sensor, biological tissue thickness and the wavelength.

McMahon and Prise (2019) reviews the current mechanistic modelling techniques for radiation responses. They highlight that “due to the lack of distinction between different modelling assumptions there is very large uncertainty in the underlying biological parameters”, suggesting that methodology for handling these uncertainties is still an active area of research due to the complexity of the processes they are modelling.

During the planning stages and calculation of the dose distribution, model specification and assumptions on the dose calculation lead to many uncertainties. Lian and Xing (2004) investigate these uncertainties by modelling the inverse planning parameters as probability density functions to determine whether or not the final solution of the inverse planning is dependent on these parameters. They conclude that they are and highlight the need for accurate methods of modelling the uncertainty in the parameters. Budiarto et al. (2011) model the uncertainty in some of these parameters by using average values and variances of the dose distributions from patient-specific local probability distributions in the inverse treatment planning. They determine the probability distributions by modelling the uncertainties in the deformations of organs between treatment fractions.

Booth and Zavgorodni (1999) review the set-up error and organ motion uncertainty to collate the magnitudes of uncertainty for specific organ sizes and set-up techniques. However, they point out that often the methods used and the parameters measured to derive these values of uncertainty can differ greatly between studies. They conclude with a brief review of methods to account for these uncertainties. Their results indicate that, “when high confidence levels (e.g. 95%) are reached the total uncertainties in the dose location become the limiting factors for treatment accuracy”.

Other studies have used different methods to investigate the uncertainty in radiotherapy treatment planning. Booth and Zavgorodni (2001) study the effects of radiotherapy treatment uncertainties on the delivered dose distribution and tumour control probability. Whereas, Holloway et al. (2017) develop a method for acquiring

random range uncertainty probability distributions in proton therapy. Karger (2006) consider an alternative source of uncertainty when they investigate the biological models in treatment planning and their associated uncertainties. Additionally, there are many other methods for modelling these investigated uncertainties. For example, Smith et al. (2016) develop a multiobjective optimisation method to model the decisions made under uncertainty. This allows them to develop personalised treatment planning for IMRT prostate cancer patients.

2.3.3 What we regard as uncertainty

This discussion of the numerous sources of uncertainty may cast doubt on the DEA-based evaluation of treatment plans in Section 5.3. We assume the planning data to be exact and classify treatment plans as efficient or inefficient based on these data. However, it is likely that the values for the variables used in the analysis are imprecise. Hence, it is also possible that an inefficient plan does actually perform well in practice. In Chapter 6 we take uncertainty into account when evaluating the quality of treatment plans. We explore this opportunity to leverage uncertainty in order to identify treatment plans that are only considered inefficient due to the precise computation of plan data but that would perform well when considering the uncertainty.

We assume that the variables are in fact realisations from a range, called an uncertainty set, this will be introduced in Section 3.2. We consider how uncertainties in treatment planning affect treatment quality, as measured by the efficiency score from DEA.

In this thesis we assume that each OAR and PTV variable v can take any value in the interval $[v - \epsilon, v + \epsilon]$, where ϵ is uniformly distributed in $[0, u]$. Then we can sample from these distributions and simulate uncertainty in the treatment plans. The standard assumption is that uncertainty is proportional to the dose. The International

Commission on Radiation Units and measurements (Andreo et al., 2004) conclude that the available evidence for certain types of tumour points to the need for accuracy $\pm 5\%$. Combining the standard uncertainty value for dose determination and the uncertainty associated with Pinnacle for multileaf collimators, Henríquez and Castrillón (2008) suggest an uncertainty of 3.6% should be used.

The mathematical models we use to handle the uncertainty are drawn from the robust optimisation, sensitivity analysis and uDEA literature (Section 3.2, 3.3 and 3.4). These are then developed in Chapter 4 where we discuss how we handle uncertainty in the mathematical models and develop the uDEA model for box uncertainty, as motivated by the radiotherapy uncertainty literature.

Many of the models involved in treatment plan creation involve solving linear optimisation problems (see Section 2.1). It is often assumed that a small data uncertainty ($\leq 1\%$) can be ignored and the resulting solutions can be regarded as exact (Ben-Tal et al., 2009). However, in Ben-Tal and Nemirovski (2000) 94 NETLIB⁵ collections were investigated under random 0.01% perturbations of the data to investigate the true effect uncertainty can have. It was found that 13 of the 94 NETLIB problems became severely infeasible (with a non-negligible probability some of the constraints were violated by at least 50%). This suggests that even small amounts of uncertainty can have a large effect and that a technique is required that accounts for the uncertainty in the solution.

We conclude that the nature of uncertainty and its effects on modelling require further research from both a mathematical and radiotherapy point of view.

⁵Netlib is a collection of mathematical software, papers, and databases.

2.4 So what is this thesis about?

In this thesis, we determine: ‘how multi-criteria decision making (MCDM) methods can help support treatment planners, given that the plans are created with commercial treatment planning systems?’ We evaluate the quality of treatment plans generated by an existing system and provide recommendations to planners on possible improvements. Due to the conflicting nature of the goals of radiotherapy (tumour control and sparing of HT) MCDM methods are appropriate tools to deal with this question and we therefore develop further MCDM models here.

In Chapter 3 we introduce the mathematical background required for this Thesis. DEA (Section 3.1), a management science technique to perform relative performance analysis. Robust optimisation (Section 3.2), a method for modelling uncertainty in optimisation problems. Sensitivity analysis in DEA (Section 3.3), a technique used to investigate the effect that small local changes in a dataset have on an optimal solution and hence the decisions being made. uDEA (Section 3.4), a method to handle uncertainty in DEA and finally, variable selection methods (Section 3.5) with a focus on principal component analysis are discussed to aid application of our other methods to our data from Rosemere.

In Chapter 4, unlike other disciplines, we leverage the uncertainty present in the data such that inefficient DMUs may improve. In Section 4.1 we define preliminary results and concepts to be used. We consider fixed amounts of uncertainty in Section 4.2 to maximise the possible increase in efficiency score for an inefficient DMU. Building upon Ehrgott et al. (2018) in Section 4.3 we refine the concept of uDEA for the specific case of box uncertainty. We explore the relationship between the amount of uncertainty required for an inefficient DMU to be deemed efficient and how this efficiency is achieved. We address the question ‘when uncertainty is introduced, is it beneficial for an inefficient DMU to compare itself to a facet of the efficient frontier

not defined by its nominal peers?’ We determine the minimum amount of uncertainty required for a DMU to be deemed efficient. For small problems, this value can be easily calculated but for larger instances, this becomes computationally intensive as it requires the facets of the Production Possibility Set (PPS) to be known. The theory from this Chapter is then applied to the data from Rosemere in Chapters 5-7.

In Chapter 5 we present details of our case study which we introduced in Section 2.1.2. In Chapter 5.1 we outline the data we have obtained from Rosemere and describe the steps we followed in conducting this case study. In Section 5.2 we deal with the issue of data selection: which of the vast amount of data available for every single treatment plan is necessary and sufficient for quality evaluation? To do this we introduce a new method, autoPCA. autoPCA addresses some of the weaknesses of PCA with respect to the sample being used and the variables being selected. In Section 5.3 we apply DEA for the evaluation of treatment plan quality. In this way, we classify plans from Rosemere as efficient or inefficient based on the achieved values of other comparable plans.

After concluding in Chapter 5 that DEA is a suitable method for treatment plan quality evaluation, we address the need to account for uncertainty in our data in Chapter 6. In Section 6.1 we apply simulation techniques to deal with this issue and show how uncertainty affects the results of DEA. In Section 6.2 we apply robust DEA to the same dataset so that we can compare the two methods in Section 6.3. For a small subset of the data, our comparison includes results for the uDEA model. From these results we select a small subset of plans that we believe could benefit from replanning. Section 6.4 details the results of replanning by Rosemere.

In Chapter 7 we present a proof of concept for a piece of software, EvaluatePlan, to be trialled at Rosemere. EvaluatePlan is in the form of an R shiny app (Chang et al., 2019) and provides assessment of treatment plans to determine whether they are efficient. EvaluatePlan allows the user to select an amount (or range) of uncertainty

they wish to see if a plan will be considered efficient for.

In Chapter 8 we summarise our findings and suggest potential future research directions.

Chapter 3

Mathematical background

We now introduce the mathematical background required for this Thesis. The methods introduced are Data Envelopment Analysis (DEA) (Section 3.1), robust optimisation (Section 3.2), Sensitivity Analysis in DEA (Section 3.3), Uncertain DEA (Section 3.4) and finally variable selection methods (Section 3.5).

3.1 Data Envelopment Analysis

Charnes, Cooper and Rhodes (Charnes et al., 1978) first introduced Data Envelopment Analysis, DEA, in 1978. The literature and techniques developed quickly with 400 publications between 1978 and 1992 (Charnes et al., 1994). By 2016 the number had risen to over 10,300 publications and they continue to rise to this day (Emrouznejad and Yang, 2018). DEA has become popular because it combines theory and practice; practical applications are motivated by theoretical developments in the field, while the applications highlight downfalls in the theory that researchers need to address. Operations analysts, management scientists, industrial engineers and many others have used DEA throughout many different industries.

The original motivation behind DEA was to develop a method for assessing not-for-profit organisations such as schools and hospitals, corporations that could not

be assessed by the traditional method of comparing profit or stock market indices. However, since profit is not always a good indication of the potential improvement a firm can achieve it has also been used to compare profit-making organisations.

DEA calculates a performance measure for each Decision Making Unit, DMU, relative to all the other DMUs in the observed population.

Definition 3.1.1. *A Decision Making Unit, DMU, is defined to be the entity responsible for converting inputs to outputs.*

The term DMU was first introduced by Charnes et al. (1978) to highlight the distinction of DEA from more standard profit based efficiency measures. They introduced the efficiency measure for public programs and hence referred to the term ‘program’ as a collection of DMUs with common inputs and outputs. By introducing a somewhat vague definition of a DMU, it was hoped that the potential flexibility of the method would be realised. In this way, DMUs can include departments of organisations, banks, manufacturing units, universities, students exam results, hospital wards and many other programs from a wide range of applications.

For a collection of I DMUs, DEA consists of I mathematical programs, one for each of the individual DMUs. A performance measure is then calculated for each DMU relative to all the other DMUs in the observed population. This creates an efficient frontier with each DMU lying on or below it. In this way, DMUs are classified as efficient or inefficient where we use the Pareto-efficiency (sometimes called Pareto-koopmans-efficiency) definition from Charnes et al. (1981).

Definition 3.1.2. *A DMU is Pareto-efficient if and only if none of its inputs can be decreased without either (i) decreasing some of its outputs, or (ii) increasing some of its other inputs, and none of its outputs can be increased without either (i) increasing one or more of its inputs or (ii) decreasing some of its other outputs.*

In this way efficient DMUs will lie on the efficient frontier and inefficient DMUs

will lie beneath it. Any DMU that falls below the efficient frontier can be scaled against a linear combination of the DMUs on the frontier. DEA produces relative efficiency measures from the observed data and hence provides a realistic target for improvement. DEA focuses on the individual observations of each DMU instead of an overall average that can easily be distorted by outliers and environmental factors. DEA is an attractive method because it characterises each DMU by a single summary relative efficiency score and as Charnes et al. (1994) conclude, “specifying abstract statistical models and making inferences based on residual and parameter coefficient analysis is no longer required”.

First, in Section 3.1.1 we discuss the basic concept of efficiency and introduce Example 1: comparing hospital departments. In Section 3.1.2, we introduce some important definitions and common DEA models. In Sections 3.1.3 and 3.1.4, we look at economies of scale and model orientation before concluding in Section 3.1.5 with some further extensions that can be made to the DEA models. Throughout Section 3.1 we use Example 1 to demonstrate the concepts and methodology introduced.

3.1.1 Basic concept of efficiency

The Oxford English Dictionary defines efficiency as “the ratio of the useful work performed by a machine or in a process to the total energy expended or heat taken in”. This definition has been used in physics and engineering for many years and a natural extension to this definition is the ratio of outputs to inputs. In this way, efficiency can be thought of as the ratio of total outputs to total inputs in a system,

$$\text{efficiency} = \frac{\text{sum of outputs}}{\text{sum of inputs}}. \quad (3.1)$$

We wish to consider the efficiency in terms of possible alternatives in the system and hence we look at the relative efficiency. Considering all DMUs, which has the best ratio of inputs to outputs? These DMUs are assigned an efficiency score of 1, or

100%. The method ranks a set of DMUs and indicates which are efficient and those that have the potential to improve.

Charnes, Cooper and Rhodes coined the phrase DEA and introduced the Charnes Cooper Rhodes, CCR model, Charnes et al. (1978). Their paper was originally titled “Exposition, interpretations and extensions of the Farrell efficiency measure” highlighting the earlier version of DEA by Farrell (1957). Farrell was motivated by the need to produce a quantitative measure of “how far a given industry can be expected to increase its output by simply increasing the efficiency, without absorbing further resources”. His paper on “The measurement of productive efficiency” was used to compare agricultural production methods in the United States. In Section 3.1.4, we see that Farrell’s measure is what is now referred to as output-oriented DEA.

Mr Sturrock, a reviewer of Farrell’s paper, comments on his concern for a “system which depends on joining up the best results”. He worried that it would result in “hanging the carrot too high and the donkey will be discouraged” (Winsten, 1957). However, in DEA the goals are set by the data, rather than arbitrarily picked by expert opinion or ideal targets. As a result, we argue that DEA produces an attainable, realistic goal for other DMUs as long as the inputs and outputs are suitably chosen.

Here we define inputs and outputs for a DEA model in the following way.

Definition 3.1.3. *An input to a DEA model is a variable such that an increase will come at a cost either in a financial sense or in terms of resource/time allocation. It is any variable used to produce something of value for a DMU. Hence, it is a variable we wish to minimise to improve efficiency.*

Definition 3.1.4. *An output to a DEA model is any variable which describes the outcome financially or otherwise, from the resources required and processed in a system. Therefore, these variables should be increased as much as possible to increase the overall efficiency.*

We introduce a simple example with a single input and output in Example 1. Here we note that this is a simplified version of judging a hospital department efficiency as in practice there would be many more variables that would need to be considered. For a more in-depth hospital efficiency application of DEA see Jacobs (2001), where hospitals are compared against one another using variables such as the proportion of patients over 60, average available beds and percentage of total revenue spent on research. Example 1 is provided here to aid understanding of the DEA models presented in the remainder of this chapter not as a conclusive measure of hospital departments' efficiency.

Example 1: One input, one output: Comparing hospital departments

Consider eight different hospital departments, $A - H$, and data on the number of patients seen each week and the number of staff employed by the department, Table 3.1. From Definitions 3.1.3 and 3.1.4 the number of staff in each hospital department is the input (the cost to the system) and the number of patients is the output. The more patients that can be seen the more efficient the department is. Hence, we wish to assess which departments are efficient in converting staff time and expertise into patients attended to. Therefore, the departments $A - H$ are the DMUs.

From (3.1) we can compare the hospital departments by calculating the ratio of inputs and outputs. How much input is required to produce the desired output? I.e. the ratio of patients per staff member.

Here a fundamental assumption is that if one department is capable of attending to y patients with x staff then all other departments should also be able to achieve this too. For example, Department C can improve its efficiency by reducing its staff numbers to 15 for the same number of patients or increase the number of patients to 90 for the same number of staff. Here we note that Department H has an efficiency of zero. This is because it is the Human Resources Department of the hospital and hence sees no patients. Department H highlights the need to ensure DMUs are comparable.

Department	Output (No. of patients in tens of thousands)	Input (No. of staff in thousands)	Ratio $(\frac{\text{Output}}{\text{Input}})$	Relative efficiency $(100 * \frac{\text{Ratio}}{3.3})$
A	10	3	3.3	100
B	40	12	3.3	100
C	50	27	1.9	55.6
D	50	36	1.4	41.7
E	5	2	2.5	75
F	15	21	0.7	21.4
G	2	2	1	30
H	0	36	0	0

Table 3.1: Hospital departments data, Example 1.

They must be capable of performing the same task, here converting hospital staff expertise into attending to patients. This suggests Department H should not be included in the analysis as it is not performing the same tasks and does not have the same objective.

Department H will not be included in further analysis of this example. This results in a dataset that is strictly positive. Note, that the end column, Relative efficiency will be discussed later on in this section when we revisit Example 1.

3.1.2 The mathematical models

We now introduce some important definitions and common DEA models. Throughout the following notation will be used.

Consider I DMUs, each with M non-negative, non-zero outputs and N non-negative, non-zero inputs with $I > M + N$. The inputs and outputs of the i^{th} DMU are represented by the vectors $x_i \in \mathbb{R}^N$ and $y_i \in \mathbb{R}^M$. The input and output data for

all I DMUs can be arranged in $(N \times I)$ and $(M \times I)$ matrices respectively,

$$X = \begin{pmatrix} x_{11} & x_{12} & \dots & x_{1I} \\ x_{21} & x_{22} & \dots & x_{2I} \\ \vdots & \vdots & \ddots & \vdots \\ x_{N1} & x_{N2} & \dots & x_{NI} \end{pmatrix},$$

$$Y = \begin{pmatrix} y_{11} & y_{12} & \dots & y_{1I} \\ y_{21} & y_{22} & \dots & y_{2I} \\ \vdots & \vdots & \ddots & \vdots \\ y_{M1} & y_{M2} & \dots & y_{MI} \end{pmatrix}.$$

In this way each column of $X(Y)$ contains the input(output) data for DMU i . DEA models aim to evaluate the combination of inputs and outputs to maximise the efficiency for each individual DMU. Let DMU 0 be the DMU under consideration. Following from (3.1), the efficiency score ξ_0 of DMU 0 can be defined as,

$$\xi_0 = \frac{\text{weighted sum of outputs}}{\text{weighted sum of inputs}} = \frac{\sum_{m=1}^M v_m y_{m0}}{\sum_{n=1}^N \nu_n x_{n0}} = \frac{v^T y_0}{\nu^T x_0},$$

where $v = (v_1, v_2, \dots, v_M,)$ and $\nu = (\nu_1, \nu_2, \dots, \nu_N,)$ are the output and input weights for DMU 0. To determine the output and input weights that maximise the efficiency for DMU 0 the following fractional programming problem (3.2) must be solved

$$\mathbf{CCR}_{\text{Frac}} \quad \max_{v, \nu} \quad \xi_0 = \frac{v^T y_0}{\nu^T x_0} \quad (3.2a)$$

$$s.t \quad \frac{v^T Y}{\nu^T X} \leq 1 \quad (3.2b)$$

$$v, \nu \geq \epsilon, \quad (3.2c)$$

where ϵ is an infinitesimal constant. We note that here ϵ is not a number and hence in principal can never be approximated or written down. In practice when solving (3.2) via LP software ϵ is required to be defined by a very small number predefined in the solver. When the $\mathbf{CCR}_{\text{Frac}}$ model was first introduced in Charnes et al. (1978), (3.2c)

was modelled as the non-negativity constraints $v, \nu \geq 0$ for the weights. However, they modified this in Charnes et al. (1979) to (3.2c) to avoid any input or output being ignored when calculating the efficiency score. Here constraint (3.2b) ensures that the efficiency score for all I DMUs is less than or equal to one.

The system of equations (3.2) is the Charnes, Cooper and Rhodes, CCR model introduced in Charnes et al. (1978). This is a non-convex, nonlinear programming problem, however, it can be converted to a linear program as follows.

When maximising (3.2a) it is the relative magnitudes of the numerator and denominator that are of interest, not their individual values. Hence, choose weights ν such that $\nu^T x_0 = 1$. The inputs are strictly positive and following the procedure first introduced in Dinkelbach (1967), (3.2) can be rewritten,

$$\mathbf{CCR}_{\text{Mult}} \quad \max_{v, \nu} \quad \Xi_0 = v^T y_0 \quad (3.3a)$$

$$s.t. \quad \nu^T x_0 = 1 \quad (3.3b)$$

$$v^T Y - \nu^T X \leq 0 \quad (3.3c)$$

$$v, \nu \geq \epsilon. \quad (3.3d)$$

Here constraint (3.3b) constrains the weighted sum of the inputs of the DMU under consideration to be one. The objective in (3.2a) can then be replaced by the linear objective function in (3.3a). The objective (3.3a) is to maximise the weighted sum of the outputs and hence (3.3) is an output maximisation DEA program. If we instead set the weighted sum of outputs to equal one, (3.2) would become an input minimisation DEA program where the objective is to minimise the weighted sum of the inputs. The optimal objective function value of (3.3a), the output maximisation program is the reciprocal of the input minimisation DEA program. See Section 3.1.4 for more details on the input/output orientation of the DEA models.

The proof that (3.2) and (3.3) are equivalent can be found in Cooper et al. (2004). Model (3.3) is now a linear program so the efficiency score for DMU 0 can be found

via the simplex method. It is important to note that both (3.2) and (3.3) calculate the efficiency score for DMU 0 only. To find the efficiency score for all I DMUs, (3.3) must be solved I times, once for each DMU. In this way, the optimal weights can be unique for each DMU¹. The weights are chosen by the data instead of being restricted by fixing them in advance. In model (3.3) the weights are often called multipliers and hence (3.3) is referred to as the CCR multiplier model, $\mathbf{CCR}_{\text{Mult}}$.

Although (3.3) can be solved by the simplex method, it is often easier to solve the dual form. Taking the dual of (3.3) gives the CCR envelopment model, $\mathbf{CCR}_{\text{Env}}$, (3.4).

$$\mathbf{CCR}_{\text{Env}} \quad \min_{\lambda, \theta} \theta_0 \quad (3.4a)$$

$$s.t \quad Y\lambda - y_0 \geq 0 \quad (3.4b)$$

$$X\lambda - \theta_0 x_0 \leq 0 \quad (3.4c)$$

$$\lambda \geq 0, \quad (3.4d)$$

where θ_0 is a scalar decision variable which represents the efficiency score for DMU 0. θ_0 is the dual variable associated with the constraint (3.3b) that normalises the weighted sum of the inputs and λ is a decision variable of length I .

Since a dual feasible solution is an upper bound on the optimal primal solution, if we have primal and dual feasible solutions that have equal objective functions then they must be optimal solutions. This follows from the weak duality theorem and gives us the strong duality theorem: if the primal(dual) problem has a finite optimal solution then the dual(primal) also has a finite optimal solution and these two values are equal. (For more details on the weak/strong duality properties in linear programming see any introductory linear programming course e.g. Dantzig and Thapa (1997)). Therefore, solving (3.4) or (3.3) gives an optimal solution for DMU 0. Hence, we can solve either to find the efficiency score for DMU 0.

¹To ease notation we use v_m and ν_n instead of v_{m0} and ν_{n0} throughout.

When solving an LP it is normally the number of constraints as opposed to the number of variables that affects the computational efficiency. In most DEA models there are significantly more DMUs being considered than variables and hence solving the dual problem (3.4) can be beneficial.

Definition 3.1.5. *An optimal solution to (3.4) for DMU 0 is the vector (λ^*, θ_0^*) .*

The primal model in (3.3) provides weights v and ν , which are the optimal multipliers for the inputs and outputs. In contrast, the dual model (3.4) provides weights λ for the DMUs. The linear optimisation problem (3.4) attempts to identify $(X\lambda, Y\lambda)$ such that $(X\lambda, Y\lambda)^T = (x_0, y_0)$ in the set spanned by the input and output data given by existing DMUs and defined by the constraints (3.4b) to (3.4d) that has outputs greater than or equal to DMU 0, and at the same time inputs less than or equal to those of DMU 0 multiplied by θ_0 . Among all such data points, the objective function (3.4a) alongside the constraint (3.4c) ensures that the point with the smallest input data is selected.

From (3.4) it can be seen that the solution $\theta_0^* = 1$, $\lambda_0^* = 1$, $\lambda_i^* = 0$ ($\forall i \neq 0$) is always feasible. Hence the optimal θ_0 is less than or equal to one. From the assumption that the data are greater than or equal to 0 and at least one variable is larger than 0, (3.4c) ensures θ_0 is greater than zero, hence $0 < \theta_0 \leq 1$. If the optimal value θ_0^* is less than one, DMU 0 is considered inefficient and the optimal solution of the DEA model provides evidence that it should be possible to scale down the input vector x_0 to $\theta_0^*x_0$ while maintaining the same output level. Furthermore, the nonzero entries in an optimal vector λ^* indicate which DMUs this suggestion is derived from. We note that for all inefficient DMUs, the corresponding zeroth entry in λ^* must be zero, $\lambda_0^* = 0$. Hence, DMUs with nonzero entries in λ^* are termed peers in DEA.

Definition 3.1.6. *For inefficient DMU 0, a peer is a DMU i such that $\lambda_i^* > 0$ in the optimal solution (λ^*, θ_0^*) to the DEA problem (3.4).*

Peers are the DMUs an inefficient DMU should be compared to in order to make suitable improvements. If $\theta_0^* = 1$, then the data set does not contain evidence that DMU 0 could be improved and DMU 0 is considered efficient.

Consider the CCR envelopment model in (3.4). Let $s_n^+ \geq 0$ ($s_m^- \geq 0$) be the slack in the n^{th} (m^{th}) input (output) constraint in (3.4) such that $s^+ = (s_1^+, s_2^+, \dots, s_n^+)$, $s^- = (s_1^-, s_2^-, \dots, s_m^-)$. This results in the following model,

$$\text{CCR}_{\text{Slack}} \quad \min_{\lambda, \theta, s^+, s^-} \theta_0 \quad (3.5a)$$

$$s.t. \quad Y\lambda - y_0 - s^- = 0 \quad (3.5b)$$

$$X\lambda - \theta_0 x_0 + s^+ = 0 \quad (3.5c)$$

$$\lambda, s^+, s^- \geq 0. \quad (3.5d)$$

In this manner, all the inequality constraints in (3.4) are now equalities and solving (3.5) gives the efficiency score of DMU 0 and the associated slack variables.

Definition 3.1.7. *DMU 0 is efficient if and only if $\theta_0^* = 1$ and $s^+, s^- = 0$ in (3.5).*

Definition 3.1.8. *DMU 0 is inefficient if and only if $\theta_0^* < 1$ in (3.5).*

From this we define a binding constraint and categorise efficient DMUs as weakly or strongly efficient.

Definition 3.1.9. *A binding constraint is a constraint for the m^{th} (n^{th}) output(input) such that $s_m^-(s_n^+) = 0$.*

Definition 3.1.10. *A DMU 0 is weakly efficient if and only if $\theta_0^* = 1$ and at least one slack variable, s_m^-, s_n^+ , is greater than 0 in (3.5).*

Definition 3.1.11. *A DMU 0 is strongly efficient if and only if $\theta_0^* = 1$ and all slack variables, s_m^-, s_n^+ , are equal 0 in (3.5).*

These definitions of weakly and strongly efficient DMUs can be understood by considering (3.5). Model (3.5) aims to minimise θ_0 . From constraint (3.5c) to minimise

θ_0 , $X\lambda$ must be as small as possible, this occurs when $X\lambda - \theta_0 x_0 = 0$ and $s_n^+ = 0$. Similarly, in constraint (3.5b) an increase in $Y\lambda$ will result in $X\lambda$ increasing or staying the same. Hence, from Definition 3.1.2, θ_0 will increase or stay the same. Therefore, at an optimal solution $Y\lambda$ is as small as possible, this occurs when $Y\lambda - y_0 = 0$ and $s_m^- = 0$. This is explored in more detail in Chapter 4 but this brief insight is given here to give intuition into the role of the slack variables in the definitions of weakly and strongly efficient DMUs.

Definition 3.1.12. *The efficient frontier is the convex hull of the data. It envelops all the data.*

All efficient DMUs will lie on the efficient frontier. We note here that the efficient frontier is based purely on the current data and therefore DMUs that are deemed efficient may in fact be able to improve further by increasing their outputs or decreasing their inputs. There are no assumptions made on what may be possible for the efficient DMUs. In principle, this means when we refer to efficiency, we are in fact talking about relative efficiency not absolute efficiency.

The efficient frontier forms an envelope around the observed data under the assumption that between feasible input-output pairs interpolation is valid. In this way, the envelope will be the smallest set containing all input-output relationships observed from the DMUs being assessed. For more details see Thanassoulis (2001). Any DMUs falling below the envelope are inefficient. Cooper et al. (1996) define efficiency as stated in Definition 3.1.13.

Definition 3.1.13. *The performance of DMU 0 is to be considered fully (100%) efficient if and only if the performance of other DMUs does not provide evidence that some of the inputs or outputs of DMU 0 could have been improved without worsening some of its other inputs or outputs.*

Definition 3.1.14. *The production possibility set, PPS, represents all possible input-output instances. It is enveloped from above by the efficient frontier and is the smallest convex set containing the observed data points. Formally, the PPS is defined to be*

$$T = \{(X, Y) | Y \geq 0 \text{ can be produced from } X \geq 0\}.$$

Definition 3.1.15. *The efficient frontier is the subset of points of set T satisfying the efficiency condition from Definition 3.1.7. It represents the standard of performance that all DMUs should try to achieve.*

Definition 3.1.16. *$(x_0, y_0) \in T$ is an extreme point of T if and only if*

$$\nexists (x_i, y_i), (x_j, y_j) \in T : x_0 = \lambda_i x_i + \lambda_j x_j, y_0 = \lambda_i y_i + \lambda_j y_j, \lambda_i + \lambda_j = 1.$$

Here we assume the following general axioms of the PPS as defined in Banker et al. (1984).

- Axiom 1- Convexity: If $(x_i, y_i) \in T$, $i = 1, \dots, I$ and $\lambda_i \geq 0$ are non-negative scalars such that $\sum_{i=1}^I \lambda_i = 1$ then $\left(\sum_{i=1}^I \lambda_i x_i, \sum_{i=1}^I \lambda_i y_i\right) \in T$
- Axiom 2- Inefficiency: If $(X, Y) \in T$ and $\bar{X} \geq X$ then $(\bar{X}, Y) \in T$. Similarly, if $\bar{Y} \leq Y$ then $(X, \bar{Y}) \in T$.

In this way T is the intersection of all \hat{T} satisfying axiom 1 and 2 above subject to $(x_i, y_i) \in \hat{T}$, $i = 1, \dots, I$ and hence is the smallest set to fulfil the axioms and is a polyhedral set.

We note that axiom 2 is sometimes called the free disposability of inputs and outputs axiom and that here we have omitted the ray unboundedness axiom, “if $(X, Y) \in T$ then $(\kappa X, \kappa Y) \in T$ for any $\kappa \geq 0$ ” as this only applies for constant returns to scales models and we wish to keep our definition general.

From Definitions 3.1.11, 3.1.14 and 3.1.16 we can define an anchor DMU. Anchor DMUs were first introduced in Thanassoulis and Allen (1998) to extend the DEA efficient frontier and were later named in Allen and Thanassoulis (2004).

Definition 3.1.17. *An anchor DMU is a strongly efficient DMU for which an input(output) can be increased(decreased) without entering the interior of the production possibility set.*

An anchor point is therefore an extreme point of the PPS that defines the transition from the section of the efficient frontier where DMUs are strongly efficient to the section where they are weakly efficient. For more details on anchor points in DEA see Bournol and Dulá (2009).

Example 1 Continued

We now calculate the efficiency scores for the seven hospital departments in Table 3.1 using the DEA model in (3.4). We see that Departments A and B are efficient with $\theta_A^* = \theta_B^* = 1$ and the remaining departments have $\theta^* < 1$. Returning to the results in Table 3.1 we see that the efficiency scores from solving (3.4) are consistent with the previous results from calculating the ratio of outputs to inputs. Departments A and B have the highest ratio of 3.3 and are the departments we find to be efficient. In Table 3.1 after calculating the ratio of outputs to inputs we can calculate the relative efficiency of each department. This is done by dividing each department ratio by the maximum achieved ratio, 3.3. This gives the final column in Table 3.1. Here Departments A and B are 100% efficient and Department F is the least efficient with relative efficiency 21.4%. We see that these are the same results we obtain in Table 3.2 from solving (3.4). Departments $C - G$ should be able to improve by comparing themselves to Department A , their peer.

3.1.3 Economies of scale

Both the CCR multiplier model (3.3) and the CCR envelopment model (3.4) act under the assumption of constant returns to scale, CRS. It is assumed the input and output

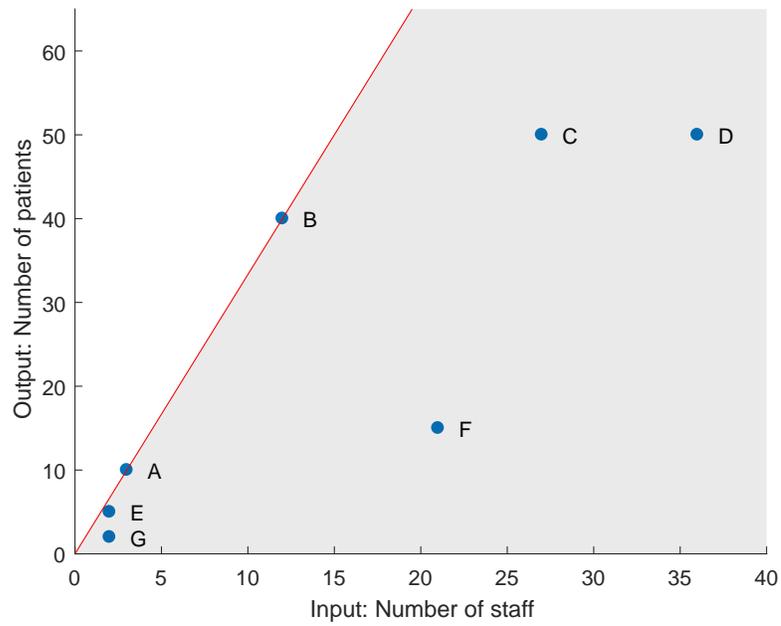
Department	CRS Efficiency score
A	1
B	1
C	0.56
D	0.42
E	0.75
F	0.21
G	0.30

Table 3.2: Hospital departments CRS efficiency scores, Example 1.

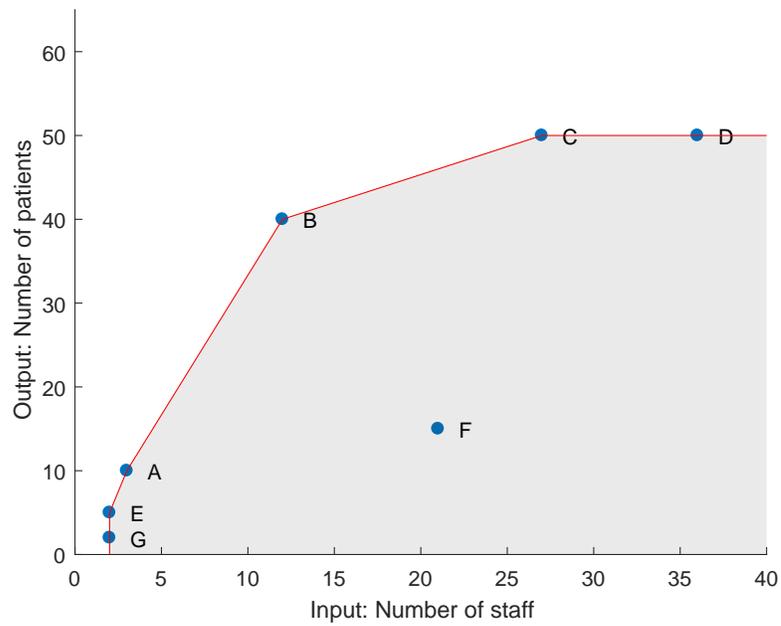
variables change proportionally. A single increase of one unit in an input will lead to an increase of one unit in output and doubling the inputs will double the outputs.

However, CRS is not always a suitable assumption, this was one of the main reasons many economists failed to adopt the methodology when it was first introduced (Ramanathan, 2003). If the inputs and outputs do not proportionally change it is more appropriate to assume variable returns to scale, VRS. Here the inputs and outputs are assumed to change at a variable rate (this includes the special cases of increasing and decreasing returns to scale). This changes both the efficient frontier and the production possibility set. This can be seen in Figure 3.1 where the efficient frontier and PPS are plotted for the Hospital Department data from Example 1. Figure 3.1a shows the CRS model and 3.1b the VRS model. The PPS is shown by the grey area on both figures. The VRS PPS is a subset of the CRS PPS. The efficient frontiers are shown by the red lines. The line segment A-B is part of the efficient frontier for both CRS and VRS otherwise they are different.

To model the VRS assumption the CRS DEA model must be adapted. This was first done by Banker, Charnes and Cooper and hence is named the BCC-DEA



(a) Constant returns to scale.



(b) Variable returns to scale.

Figure 3.1: CRS and VRS efficient frontier (red) and PPS (grey shaded region), Example 1.

Department	Efficiency
<i>A</i>	1
<i>B</i>	1
<i>C</i>	1
<i>D</i>	0.75
<i>E</i>	1
<i>F</i>	0.21
<i>G</i>	1

Table 3.3: Hospital departments BCC efficiency scores, Example 1.

model, (Banker et al., 1984). To modify (3.4) to exhibit VRS the following convexity constraint is added,

$$e^T \lambda = 1, \quad (3.6)$$

where e is a vector of ones. DEA models that include (3.6) are called BCC-DEA models or VRS-DEA models.

We demonstrate the differences between CRS and VRS with the hospital department Example 1.

Example 1 Continued

We first calculate the efficiency scores for the seven hospital departments with the BCC-DEA model. The results are in Table 3.3. Solving the BCC-DEA model for the hospital departments in Table 3.1 results in five departments being evaluated as efficient, Departments *A*, *B*, *C*, *E* and *G*.

Table 3.4 shows the corresponding slack variables for these five departments. From the slack variables, we can classify the departments as weakly or strongly efficient.

The input and output data from CCR-DEA Table 3.1 can be plotted, Figure 3.1a.

Efficient DMU	Slack variables		Efficiency categorisation
	s^+	s^-	
<i>A</i>	0	0	strong
<i>B</i>	0	0	strong
<i>C</i>	0	0	strong
<i>E</i>	0	0	strong
<i>G</i>	0	3	weak

Table 3.4: Hospital departments BCC slack variable values, Example 1.

Inefficient DMU	Peers					Efficiency Score
	<i>A</i>	<i>B</i>	<i>C</i>	<i>E</i>	<i>G</i>	
<i>D</i>	–	–	1	–	–	0.75
<i>F</i>	0.83	0.17	–	–	–	0.22

Table 3.5: Hospital departments BCC peers, Example 1.

We assume that if we have zero inputs we can produce zero outputs and if $(X, Y) \in T$ then $(\kappa X, \kappa Y) \in T$ for any $\kappa \geq 0$ where T is the PPS. From Table 3.1, Department *A* is efficient. Therefore, joining Department *A* with a straight line through the origin we obtain the CRS efficient frontier. This is shown in Figure 3.1a by the red line. The efficient frontier forms a ray from the origin through efficient Departments *A* and *B* and extends with a constant slope beyond Department *B*. All other departments not on the ray are inefficient. The shaded grey area in Figure 3.1a shows the PPS assuming CRS.

Similarly, from Table 3.3 Departments *A*, *B*, *C*, *E* and *G* are efficient under the BCC-DEA model. Joining these with a series of straight lines, we obtain the VRS efficient frontier shown in red in Figure 3.1b. The line through *A*, *B*, *C*, *E* and *G* envelopes all the observed input-output relations and is the efficient frontier for

the VRS model. The shaded grey area shows the PPS assuming VRS. Here we see that Department E is an anchor point (see Definition 3.1.17). Department E has the minimum input value of 2 and is strongly efficient whereas, Department G has the minimum input value of 2 but is weakly efficient so is not an anchor point.

We note here that any departments that are efficient for CRS, Departments A and B , are also efficient for VRS. However, the opposite is not true; Departments C , E and G are not efficient under CRS but are under the VRS assumption. The PPS under VRS is a subset of the CRS PPS.

For the remaining sections, we focus on VRS-DEA models.

3.1.4 Model orientation

The model orientation is dependent on whether we are interested in input conservation or output expansion as the indication of efficiency. The DEA model (3.3) introduced in Section 3.1.2 is an input minimisation model. Here we compare the different orientations of DEA models, input-oriented, output-oriented and briefly radial DEA. The orientation of a DEA model does not affect the classification of a DMU as inefficient or efficient. However, it does affect the DMUs peers.

The orientation of a DEA model should be chosen according to the application. For example, consider the following two applications:

- Assessing Schools
Inputs: Number of pupils and total available funding
Outputs: Exam results

- Assessing car manufacturing factories
Inputs: Raw materials, manufacturing cost
Outputs: Number of cars made.

When assessing schools the number of pupils and the funding available are very hard

to influence, one cannot easily reduce the number of students at a particular school. Therefore, it is reasonable to treat the inputs as fixed variables and seek to improve the outputs, that is increase the exam results. Hence, an output-oriented DEA model is suitable here.

In contrast, when manufacturing cars the factory is likely to have a fixed demand/target number of cars to produce and hence the outputs can be treated as fixed. We wish to reduce the raw materials and manufacturing costs and hence, an input-oriented model is suitable here.

Alternatively, to consider both input reduction and output enlargement the additive DEA model (also called slack based or radial DEA) can be used, see Charnes et al. (1985) for more details.

We demonstrate the difference between input and output oriented DEA models using the Hospital example introduced in Section 3.1.1 with the BCC-DEA model results from Table 3.3.

Example 1 Continued

To demonstrate the difference between input and output oriented DEA models we turn our attention to Department F in Table 3.1. Department F is inefficient as it falls below the efficient frontier in Figure 3.1b and solving (3.4) with the additional convexity constraint (3.6) gives an efficiency score of 0.21.

Department F is inefficient so we know that it has the potential to improve its efficiency score by either increasing its outputs while the inputs remain fixed (output-oriented DEA model) or decreasing its inputs for fixed outputs (input-oriented DEA model). This is equivalent to Department F being projected to the point J (21, 46) or I (4.5, 15) in Figure 3.2. For input-oriented DEA Department F is being projected to the efficient frontier section AB which is also part of the CRS efficient frontier. Whereas, for output-oriented DEA, Department F is projected to the efficient frontier section BC . From this we see that depending on the model orientation the

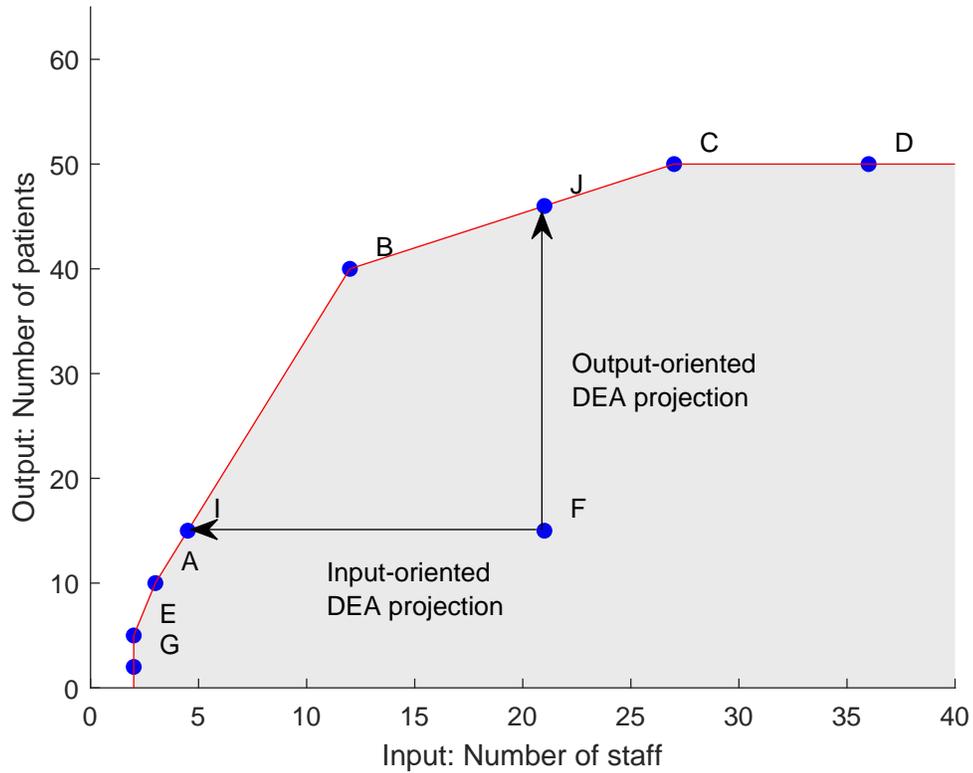


Figure 3.2: Difference between input and output oriented models for Department F under VRS-DEA model, Example 1.

Department will have different peers. Departments A and B for the input model and Departments B and C for the output model.

3.1.5 Some extensions to DEA models

Here we mention a few extensions that can be made to the basic DEA models to increase the modelling capability.

Environmental variables Sometimes there will be variables that cannot be influenced but must be considered when assessing the DMUs. For example, when assessing schools in Section 3.1.4 it may be that there are schools that require an entrance exam and those that are located in a socially deprived area. It would be wrong to compare

these two schools against each other as they are operating in very different conditions. To make fair comparisons additional variables could be included in the model such as a social deprivation index. However, this cannot be classified as an input or an output as we cannot change it, as the schools themselves have no control over the area they are located in. One way to introduce such variables is to use environmental variables (also called non-discretionary variables).

Let there be O environmental variables we wish to consider, where the environmental inputs for the i^{th} DMU are represented by the vector $z_i \in \mathbb{R}^O$. To modify (3.4) to include environmental variables an additional constraint is required for each environmental variable.

$$\sum_{i=1}^I \lambda_i z_{oi} - z_{o0} \leq 0 \quad o = 1, \dots, O \quad (3.7)$$

Here the main difference between (3.7) and the input constraints in (3.4c) is that the θ_0 is omitted. This is so that the environmental inputs play no direct role in the determination of the optimal value of θ_0 .

In the schools assessment example discussed above this would mean that if the social deprivation index was included as an input, only schools with the same or lower social deprivation index would be compared against one another to prevent unrealistic targets being set. Analysis has shown that accounting for these non-discretionary variables produces more meaningful and applicable results (Charnes et al., 1994).

Super-efficiency Another extension of DEA is the concept of super efficiency introduced by Andersen and Petersen (1993) to provide a framework to rank efficient DMUs based on parametric methods for the DEA CRS model. Their motivation stemmed from the weakness of DEA that in some DEA models, (especially if the number of variables is large relative to the number of observations), a large proportion of DMUs can be classified as efficient. This can be due to the performance in a single variable and hence may not reflect the true efficient DMUs. One way to over-

come this as discussed earlier in this Section is to examine the peers of the efficient DMUs, super-efficiency is an alternative method to provide a ranking of the efficient DMUs themselves.

Super-efficiency works by removing the efficient DMU under consideration, DMU 0, from the data and comparing DMU 0 to linear combinations of all the remaining DMUs. In this way DMU 0 is projected to the efficient frontier constructed from the remaining DMUs. This results in an efficient DMU being able to achieve an efficiency score greater than one. Seiford and Zhu (1999) extend the super-efficiency model to include VRS DEA models. However, they note that the super-efficiency model may become infeasible when applied to VRS DEA models due to the convexity constraint. The super-efficiency DEA model for VRS as introduced in Seiford and Zhu (1999) is

$$\begin{aligned}
 \min_{\lambda, \theta} \quad & \theta_0 \\
 \text{s.t.} \quad & \sum_{\substack{i=1 \\ i \neq 0}}^I \lambda_i y_{mi} - y_{m0} \geq 0 & m = 1, \dots, M \\
 & \sum_{\substack{i=1 \\ i \neq 0}}^I \lambda_i x_{ni} - \theta_0 x_{n0} \leq 0 & n = 1, \dots, N \\
 & \sum_{\substack{i=1 \\ i \neq 0}}^I \lambda_i = 1 \\
 & \lambda_i \geq 0 & i = 1, \dots, I, i \neq 0.
 \end{aligned}$$

Seiford and Zhu (1999) show that (3.8) becomes infeasible when at least one output of DMU 0 is larger than any convex combination of the remaining DMUs. As a result many of the super-efficiency studies have been focused on how to address this infeasibility issue in the VRS model. For a detailed introduction to super-efficiency and a discussion on the VRS infeasibility issue see Chen and Du (2015). We introduce the basic super-efficiency methodology here in preparation for Section 3.3 where the concept is used for sensitivity analysis in DEA.

Other extensions to the DEA models include the use of categorical inputs and outputs into the DEA model (Banker and Morey, 1986); incorporating judgement functions and a priori knowledge (Thompson et al., 1986; Brockett et al., 1997) or time-series analysis when observations for DMUs are available over multiple time periods (Sueyoshi and Goto, 2001). These extensions to DEA are by no means an exhaustive set. They are detailed here to demonstrate the increase in the range of situations that DEA can be applied to by incorporating these techniques.

From a mathematical standpoint, DEA solves a sequence of simple LP's. The solutions of these LP's provide important information which can help quantify inefficiencies and provide advice on suitable improvements that can be made.

We conclude this section with the complete input-oriented BCC-DEA envelopment model from (3.4) with constraint (3.6). Solving this I times gives the efficiency score for each DMU. In the following sections this model is referred to as the nominal DEA model.

$$\begin{aligned}
 \min_{\lambda, \theta} \quad & \theta_0 \\
 \text{s.t.} \quad & Y\lambda - y_0 \geq 0 \\
 & X\lambda - \theta_0 x_0 \leq 0 \\
 & e^T \lambda = 1, \\
 & \lambda, \theta_0 \geq 0.
 \end{aligned}$$

3.2 Robust optimisation

When modelling real-life optimisation problems it is unlikely that the data will be known exactly, there will most likely be numerous sources of uncertainty in the data. This can be from rounding errors, estimates made, lack of knowledge of a parameter, implementation errors or simply trying to model a complex real-world problem with

a small subset of parameters. As a result, methods that can account for the inherent uncertainty in data are required.

Robust optimisation grew out of earlier work on worst-case analysis and Abraham Wald's maximin model where the least worst outcome was sought, (Wald, 1945). However, this school of thought often results in overly pessimistic results. Soyster (1973) introduced the notion of robust optimisation under the name of inexact linear programming to solve linear optimisation problems for all data in convex sets. In contrast to the earlier maximin model, the methodology aimed to balance suboptimal solutions in the nominal certain data instance to avoid infeasible solutions when the data are uncertain. However, Soyster's approach was often still too pessimistic and too much of the optimality was sacrificed to ensure the problem was feasible. In the late 1990's this over-conservatism was addressed independently by Ben-Tal and Nemirovski (1998, 1999, 2000), El-Ghaoui and Lebret (1997) and El Ghaoui et al. (1998). In these papers ellipsoidal uncertainties were introduced and the terminology of a robust counterpart was first used. After these papers the robust optimisation literature grew quickly and it is still a growing active area of research.

Robust optimisation favours flexibility; unlike many models the estimated cost structure and probability distribution are not required. Instead, robust optimisation assumes the uncertain data resides in an uncertainty set. The uncertainty sets are chosen based on qualitative information or expert opinion and allow flexibility in the model. Robust models can be used to reduce the chance of over-optimising and to try and avoid a non-robust solution that inappropriately exaggerates the weaknesses of the estimated or sampled uncertainty. Unlike in sensitivity analysis, where uncertainty is accounted for post optimisation, in robust optimisation the uncertainty is incorporated into the initial model so that variations in the data are accounted for throughout the modelling process. In sensitivity analysis, the interest lies in how much the solution to the nominal problem can differ to that of the optimal solution to the perturbed

problem. Whereas, in robust optimisation the interest is in how much the nominal problem's optimal solution can violate the constraints of the perturbed problem, (Ben-Tal and Nemirovski, 2000).

In Section 3.2.1 we introduce the basic robust optimisation methodology and some common uncertainty sets used throughout the robust optimisation literature. In Section 3.2.2 we define the robust optimisation DEA problem which will be developed further in Chapter 4.

We note here that DEA and robust optimisation have different notation conventions. The main focus of this thesis is the application of DEA and hence, we follow more closely the DEA conventions. In this way some of the derivations for robust optimisation may not be the conventional robust optimisation textbook derivations but careful comparison will convince the reader that it is simply different notation. For example, in robust optimisation ' x ' usually represents the vector of variables, whereas, in DEA this is used to represent the input data, therefore here we use ' η ' to represent the variable vector.

3.2.1 Method

A linear optimisation (LO) problem can be written in the form

$$\text{LO : } \left\{ \min_{\eta} c^T \eta + d : A\eta \leq b, \eta \geq 0 \right\},$$

where $\eta \in \mathbb{R}^q$ is a column vector of decision variables and (c, d, A, b) represents the data of the problem arranged in a $(p + 1) \times (q + 1)$ data matrix

$$\left(\begin{array}{c|c} c^T & d \\ \hline A & b \end{array} \right).$$

Here the assumption is the data are exactly known. However, in real-life applications this is seldom true. If there is uncertainty arising in the data matrix then an

uncertain linear optimisation (u-LO) problem may be a more suitable model. Under this framework instead of assuming exact data we assume the objective and the constraint functions come from user defined uncertainty sets. As a result, by considering the uncertainty, the decisions made based on the u-LO results should be suitable for all constraints that can occur rather than a single data instance.

An u-LO problem is of the form

$$\text{u-LO : } \left\{ \min_{\eta} c^T \eta + d : A\eta \leq b, \eta \geq 0, \forall (c, d, A, b) \in \mathcal{U} \right\}, \quad (3.8)$$

with a given uncertainty set $\mathcal{U} \in \mathbb{R}^{(p+1) \times (q+1)}$. The problem in (3.8) is a collection of LO problems with data varying in a given uncertainty set \mathcal{U} , i.e. (3.8) is a family of problems, not a single problem. Equivalently (3.8) can be written as

$$\left\{ \min_{\eta} c^T \eta + d : \sup_{(c,d,A,b) \in \mathcal{U}} a_j \eta \leq b_j, \eta \geq 0, j = 1, \dots, p, \forall (c, d, A, b) \in \mathcal{U} \right\}, \quad (3.9)$$

where a_j is the j^{th} row from A and b_j is the corresponding RHS for constraint j . Here we use the supremum instead of a maximisation to ensure the problem is well defined. We model uncertainty by affine transformations, i.e. small perturbations are added that can vary over the uncertainty set thus accounting for the variability and uncertainty in the data. In this way the uncertainty set can be defined by,

$$\mathcal{U} = \left\{ \left[\begin{array}{c|c} c^T & d \\ \hline A & b \end{array} \right] = \underbrace{\left[\begin{array}{c|c} (c^0)^T & d^0 \\ \hline A^0 & b^0 \end{array} \right]}_{\text{Nominal data } D_0} + \sum_{l=1}^L \zeta_l \underbrace{\left[\begin{array}{c|c} (c^l)^T & d^l \\ \hline A^l & b^l \end{array} \right]}_{\text{Basic shifts } D_l} : \zeta \in \mathcal{Z} \subset \mathbb{R}^L \right\}, \quad (3.10)$$

where ζ is the perturbation vector varying in a given perturbation set \mathcal{Z} , $\left[\begin{array}{c|c} (c^0)^T & d^0 \\ \hline A^0 & b^0 \end{array} \right]$ is

the nominal data and $\sum_{l=1}^L \zeta_l \left[\begin{array}{c|c} (c^0)^T & d^l \\ \hline A^l & b^l \end{array} \right]$ models the L affine transformations used to model the uncertainty, here we will refer to the sum of this matrix as the basic shift matrix.

Example 1 Continued

If Departments A and B 's patient numbers are only known to within 5% accuracy the number of patients is actually in the region of 9.5 to 10.5 and 38 to 42 patients respectively. Then to assess the efficiency of Department C compared to Departments A and B the nominal data is

$$D_0 = \left[\begin{array}{c|c} (c^0)^T & d^0 \\ \hline A^0 & b^0 \end{array} \right] = \left[\begin{array}{cccc|c} 0 & 0 & 0 & 1 & 0 \\ \hline -10 & -40 & -15 & 0 & -15 \\ 3 & 12 & 21 & -21 & 0 \\ 1 & 1 & 1 & 0 & 1 \end{array} \right]$$

An example of basic shift matrices that can be applied to the nominal data to model the 5% uncertainty in Department A and B 's patient numbers is

$$D_A = \left[\begin{array}{cccc|c} 0 & 0 & 0 & 0 & 0 \\ \hline 0.5 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right] \quad D_B = \left[\begin{array}{cccc|c} 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right]$$

where the perturbation set is $\mathcal{Z} = \{\zeta \in \mathbb{R}^2 : -1 \leq \zeta_1, \zeta_2 \leq 1\}$. Here this models the scenario where the only uncertainty present is for the number of patients for Departments A and B .

The following assumptions will be useful as they allow us to simplify the model presented in (3.10). W.l.o.g. we briefly consider the main intuition behind the following assumptions:

1. The objective is certain
2. The constraints RHS' are certain
3. \mathcal{U} is convex and compact
4. The uncertainty is constraint-wise

For a detailed proof see Ben-Tal et al. (2009).

1. With an uncertain objective

$$\min_{\eta} c^T \eta + d \quad \forall (c, d) \in \mathcal{U}$$

we can introduce a new variable t such that (3.8) becomes

$$\left\{ \min_{\eta, t} t : c^T \eta + t \leq -d, A\eta \leq b, \eta \geq 0 \quad \forall (c, d, A, b) \in \mathcal{U} \right\}. \quad (3.11)$$

Therefore, we can always rewrite an uncertain linear objective as an inequality constraint.

2. We can assume the RHS of each constraint is equal to 0 and hence certain by introducing an additional variable for the RHS coefficient. If the objective is uncertain we can reformulate the objective as above and then introduce a new variable for d . Hence, from now on we remove the $+d$ in the objective term.
3. Ben-Tal et al. (2009) show that a robust feasible solution to (3.8) remains feasible when the uncertainty set is extended to its convex hull.
4. From (3.9), because every constraint $a_j \eta \leq b_j$ must be satisfied for all u we can consider each constraint individually.

We can now rewrite (3.8) as

$$\min_{\eta} c^T \eta \quad (3.12a)$$

$$s.t. \quad a_j \eta \leq b_j \quad \forall a_j \in \mathcal{U}_j \quad j = 1, \dots, p \quad (3.12b)$$

$$\eta \geq 0, \quad (3.12c)$$

where \mathcal{U}_j is the given uncertainty set, and hence, is a set of $p \times q$ real matrices. From now on we assume uncertainty occurs only in the A matrix.

This now allows us to form the robust counterpart of the u-LO problem.

Definition 3.2.1. *The Robust counterpart of (3.8) is the robust reformulation such that the robust value of the objective is minimised over all feasible solutions of the uncertain problem.*

To form the robust counterpart and hence, solve (3.8) we

1. Preserve the original certain constraint
2. Replace each of the original constraints $a_j\eta \leq b_j$ with its robust counterpart $a_j\eta \leq b_j \forall a_j \in \mathcal{U}_j$.

The robust counterpart is formed constraint wise. We denote a single constraint j from the basic shift matrix in (3.10) by $a_j^{(l)}$ where the bracketed superscript denotes the shifts $l = 1, \dots, L$ and $a_j^{(o)}$ represents the j^{th} constraint from the nominal data. Then the sum of the basic shifts for a single constraint j can be formed by multiplying ζ by the matrix $a_j^{(L')}$ where $a_j^{(L')} \in \mathbb{R}^{L \times q}$ is the matrix formed row-wise from each basic shift for constraint j , $a_j^{(L')} = (a_j^{(1)}; a_j^{(2)}; \dots; a_j^{(L)})$. We require solutions that are feasible for any data instance in our uncertainty set. Therefore, to solve 3.8 we apply robust reformulation techniques to remove the for all, (\forall) quantifier in each constraint (3.12b). This will not always be possible, hence, some robust counterparts are not tractable and other techniques such as simulation and heuristics are required. When a tractable reformulation of the uncertain constraint (3.12b) is possible it is derived by

1. Formulating the worst case reformulation of the constraint,

$$a_j^{(o)}\eta + \max_{a_j \in \mathcal{U}} \zeta a_j^{(L')} \eta \leq b_j.$$

2. Taking the dual of the problem, $\sup_{(c,d,A,b) \in \mathcal{U}} a_j\eta \leq b_j$, in (3.9).
3. Reformulating the original uncertain constraint to incorporate the dual from step 2.

We now demonstrate this process with a common uncertainty set, polyhedral uncertainty and a specific case of this, box uncertainty. There are many different uncertainty sets that can be used and should be selected according to the data. For more details on selecting the uncertainty set see Ben-Tal et al. (2009).

A polyhedral uncertainty set can be defined as:

$$\mathcal{Z} = \{\zeta : F\zeta^T + g \geq 0\}, \quad (3.13)$$

where $F \in \mathbb{R}^{p \times L}$, $\zeta \in \mathbb{R}^L$ and $g \in \mathbb{R}^p$ and \mathcal{Z} is defined in (3.10). A single constraint from the u-LO problem, (3.8) is,

$$(a_j^{(0)} + \zeta a_j^{(L')})\eta \leq b_j \quad \forall \zeta \in \mathcal{Z}.$$

The robust counterpart seeks the optimal solution to the worst-case scenario occurring. The inner maximisation problem of the uncertain LP is:

$$\max_{\zeta} \quad \zeta a_j^{(L')}\eta \quad (3.14a)$$

$$s.t. \quad F\zeta^T + g \geq 0. \quad (3.14b)$$

We now take the dual of the inner maximisation problem to obtain,

$$\min_w \quad g^T w \quad (3.15a)$$

$$s.t. \quad F^T w = -a_j^{(L')}\eta \quad (3.15b)$$

$$w \geq 0.$$

By strong duality, the optimal value of the primal inner maximisation problem (3.14) is the same as the optimal solution to the dual minimisation problem (3.15). Therefore, we replace the inner maximisation with its dual (3.15). This gives the following robust counterpart for polyhedral uncertainty,

$$\begin{aligned}
& \min_{\eta, w} c^T \eta \\
& s.t. \quad a_j^{(0)} \eta + g^T w \geq b_j & j = 1, \dots, p \\
& \quad \quad F^T w = -a_j^{(L')} \eta & j = 1, \dots, p \\
& \quad \quad w \geq 0 \\
& \quad \quad \eta \geq 0.
\end{aligned}$$

A special case of polyhedral uncertainty is box uncertainty. The unit box uncertainty set can be defined as:

$$\mathcal{Z} = \{\zeta : \|\zeta\|_\infty \leq 1\}.$$

This gives the inner maximisation problem:

$$\begin{aligned}
& \max_{\zeta} \zeta a_j^{(L')} \eta \\
& s.t. \quad \|\zeta\|_\infty \leq 1.
\end{aligned}$$

Taking the dual of the inner maximisation problem and inserting it into (3.12b) gives the following robust counterpart for box uncertainty,

$$\begin{aligned}
& \min_{\eta} c^T \eta \\
& s.t. \quad a_j^{(0)} \eta + \|a_j^{(L')} \eta\|_1 \geq b_j & j = 1, \dots, p \\
& \quad \quad \eta \geq 0.
\end{aligned}$$

Throughout this thesis we focus on box uncertainties. They are one of the simplest uncertainty sets to work with and their robust counterpart is tractable.

Choosing a suitable uncertainty set is a trade-off between the robustness of each individual realisation of the uncertain parameter and the size of the uncertainty set

(which should be as small as possible). A box constraint guarantees a constraint is never violated. However, there is only a small chance that all uncertain parameters take their worst case scenario so this can often be over-pessimistic. Not all robust optimisation uncertainty sets result in tractable robust counterparts. Polyhedral uncertainty sets result in LP robust counterparts, whereas ellipsoidal uncertainty sets result in a second order cone program.

3.2.2 Robust DEA model

Here we introduce the robust DEA model which will be the basis for the uncertain DEA, uDEA model introduced in Section 3.4.

In Section 3.1 we introduced the input oriented VRS DEA model which can be written as,

$$E_i = \min_{\theta, \lambda} \{ \theta_i : Y\lambda - y_i \geq 0, X\lambda - \theta_i x_i \leq 0, e^T \lambda = 1, \theta_i, \lambda \geq 0 \}, \quad (3.16)$$

where e is a vector of ones. Throughout this section we will develop the robust DEA model for the nominal DEA model, (3.16) but note that the methods discussed are not confined to the BCC input-oriented DEA model.

Following the reformulation in Ehrgott et al. (2018) we reorganise our input and output data in the following manner to ease development of the uDEA model.

$$A_i = \left[\begin{array}{c|c|c} -Y & y_i & 0 \\ \hline X & 0 & -x_i \end{array} \right] \text{ and } B = \left[\begin{array}{c|c|c} e^T & 0 & 0 \\ \hline 0 & 1 & 0 \end{array} \right].$$

Here the subscript on the A matrix is to show we are considering the data for the i^{th} DMU. Hence, (3.16) can be rewritten as,

$$E_i = \min_{\eta} \{ c^T \eta : A_i \eta \leq 0, B\eta = e, \eta \geq 0 \}, \quad (3.17)$$

where $c = (0, 0, \dots, 0, 1)^T$ and $\eta = (\lambda, 1, \theta_i)^T$. In the notation of Section 3.2 this is equivalent to $p = M + N$ and $q = I + 2$, where the two extra variables arise from

additional variables to model the original RHS of the constraints.

Let $(a_j)_i$ be the j^{th} row of A_i , $j = 1, \dots, M + N$, to ease notation from now on we will drop the subscript i on the $(a_j)_i$. The uncertainty arising in the data X , Y can then be introduced into the model by replacing each $a_j \eta \leq 0$ constraint with the set of constraints,

$$a_j \eta \leq 0 \quad \forall a_j \in \mathcal{U}_j, \quad (3.18)$$

where \mathcal{U}_j is the uncertainty set for a single input/output j . In this way each input/output can have an individual uncertainty set resulting in maximum model flexibility.

To ensure consistency with DEA methodology two restrictions must be placed on (3.18) which are not required in more general robust optimisation settings. First we must assume that all $a_j \in \mathcal{U}_j$ are positive so we must ensure the introduction of uncertainty does not introduce negative data. Secondly, we assume that the m^{th} (n^{th}) element of $y_i(x_i)$ agrees with $Y_{mi}(X_{ni})$. This ensures a DMU's data is consistent throughout the model. Without this restriction a DMU's input/output value could be different in the two places of the model it appears. This would not agree with DEA methodology.

Then we define the robust efficiency score for DMU i , $\mathcal{E}_i(\mathcal{U})$, to be the optimal value of the robust DEA model, $\mathcal{E}_i(\mathcal{U}) := \min_{\eta} \{c^T \eta : a_j \eta \leq 0, \forall a_j \in \mathcal{U}_j, \forall j, B\eta = e, \eta \geq 0\}$.

Unfortunately, the introduction of uncertainty in (3.18) generally results in a non-linear non-convex problem and hence cannot be solved using standard LP techniques. However, when we have box uncertainty the problem is tractable. This model will be developed further in Section 3.4 and Chapter 4.

3.3 Sensitivity analysis

Sensitivity analysis is used to investigate the effect small local changes in a dataset have on an optimal solution and hence the decisions being made. Charnes et al. (1984) regard sensitivity analysis as the mathematical programming equivalent of statistical significance testing as both consider the allowable range of data variation. In linear programming the interest is in the allowable changes in the data without a change in the set of vectors for the optimal basis whereas, in statistical analysis the interest may be in the hypothesised statistical distribution.

Traditionally, sensitivity analysis is used in linear programming to answer questions such as: What is the effect of an increase of ϵ in the RHS? Or if the cost of a variable y is reduced by ϵ should more y be produced? It is used to assess the sensitivity of an optimal solution to a linear program to certain changes in the dataset. This is done by changing a single objective function coefficient or RHS coefficient of the linear program at a time and observing the change this has on an optimal solution. The aim is to find ranges for which the data can change without the optimal solution value changing. For an example of sensitivity analysis in linear programmes see Dantzig and Thapa (1997). However, for DEA sensitivity analysis we are interested in changes in the coefficients of the data matrix and how this affects an optimal solution. As a result, modified sensitivity analysis techniques are required for DEA sensitivity analysis.

In the remainder of this section we introduce techniques for sensitivity analysis in DEA that will be built upon in Chapter 4.

3.3.1 Sensitivity analysis in DEA

In DEA sensitivity analysis we are interested in the changes that occur in the data, the coefficients of X and Y in (3.8). Changes in the data can change the value of

(λ, θ_0) in an optimal solution for DMU 0. We consider the effect changes to multiple data points have on an optimal solution.

In DEA the efficient frontier is determined by the extreme points in the dataset, see Definition 3.1.16. As a result, if the extreme points are themselves outliers wrong conclusions may be drawn. We confine our attention to the effects occurring from variations in the data itself as opposed to other variations such as model variations or variations in the number of DMUs. For an example of this type of variation see Simar and Wilson (1998) who consider variations in the sampling distributions.

The majority of sensitivity analysis for DEA focuses on the efficient DMUs and what changes in the nominal data will result in existing efficient DMUs remaining efficient. There are significantly fewer papers looking at the inefficient DMUs and how discrepancies in the data may affect them. This was highlighted by Gholam Abri et al. (2009) and Jahanshahloo et al. (2011) when they noted that “while the sensitivity analysis of an efficient unit’s classification has been extensively studied, the issue of an inefficient units estimation and classification seems to be ignored”. Here we provide an overview of some of the DEA sensitivity analysis literature that will contribute to the motivation and techniques to be used in the following chapters.

We first define some common terminology that is used throughout the DEA sensitivity analysis literature.

Definition 3.3.1. *A DMU’s stability region is defined to be the area of the PPS in which a DMU’s efficiency classification does not alter.*

In some DEA sensitivity analysis papers this is referred to as the radius of stability. We use these names interchangeably to keep consistency with the paper being discussed.

Definition 3.3.2. *A DMU 0 is extreme efficient if and only if:*

i. it is efficient, $\theta_0^ = 1$ and*

ii. it is the only DMU in its reference set, i.e. $\lambda_0^ = 1, \lambda_j = 0 \forall j \neq 0$.*

We note here that DMUs with $\theta_0^* > 1$ in the super-efficiency model (3.8) are the set of extreme efficient DMUs.

Throughout this section we denote the DMU under consideration by DMU 0 and the efficiency score from solving the nominal DEA problem by θ_0^* . We use the subscript ineff/eff/ex to show the paper considers an inefficient/efficient/extreme efficient DMU (see Definitions 3.1.8, 3.1.7 and 3.3.2) or no subscript when the paper considers both inefficient and efficient DMUs. Note that the subscript eff does not distinguish between efficient and extreme efficient DMUs.

Single changes

A large body of the DEA sensitivity analysis literature focuses on the effect changes have on a single DMU. This can be useful if there is a specific DMU for which more information is required. One of the earliest papers to discuss this was Charnes et al. (1984). By considering changes to a single output of DMU 0_{eff} they characterised the efficient DMUs by the size, large or small, of admissible data changes that result in the DMU remaining efficient.

Charnes and Neralic (1991) extended this idea by considering proportionate changes of all inputs and outputs of a single DMU 0_{eff} while the remaining DMUs' data are fixed. They derived sufficient conditions for DMU 0_{eff} to preserve its efficiency score, i.e. finding its radius of stability. Charnes et al. (1996) showed that when a single DMU is considered at a time, (all remaining DMUs' data are fixed), the radius of stability can be calculated via linear programming techniques for both l_1 and l_∞ norms. The method aimed to provide a “measure of the classification’s stability, especially

with respect to errors in the data". They applied their method to the input-oriented ratio model but the methodology can easily be extended to all DEA models. By changing a single input at a time they computed the largest ball surrounding DMU 0 such that the efficiency classification remained the same. Similar work was done by Seiford and Zhu (1998b) who sought a stability region for DMU 0_{eff} where a change in a single input or output does not stop DMU 0_{eff} being classified as efficient. Neralić and Wendell (2015, 2019) built upon this idea to find the largest hypercube such that DMU 0's efficiency status remains the same while the remaining DMUs' data are fixed. This is equivalent to perturbations within a fixed radius of DMU 0's nominal data under the l_∞ norm. They calculate the percentage change in DMU 0's data such that the efficiency calculation remains the same. This is done via an iterative algorithm that gradually increases the size of the symmetric hyperbox until the efficiency classification changes.

The previously discussed methods first require the nominal DEA problem to be solved to identify the efficient DMUs. Jahanshahloo et al. (2005b) developed a method to determine the radius of stability alongside the nominal efficiency score. The method is based on super-efficiency (see Section 3.1.5) and calculates a range for which the classification of DMU 0 holds when the DMU's input and output data vary and the other DMUs' data remain fixed. In this way, only a single linear program is required for each DMU instead of one for the nominal efficiency score and another to determine the range of stability. Their sensitivity analysis results agree with previous sensitivity analysis results but the computation time is significantly reduced.

Jahanshahloo et al. (2005a) also use super-efficiency techniques to investigate the changes that can be made to DMU 0_{eff} while the remaining DMUs' data remain fixed. However, in contrast to previous studies they determine the largest stability region by considering the three cases i) increase inputs and outputs, ii) increase inputs and decrease outputs and iii) decrease inputs and outputs; they then select the largest

of these three. This is done by removing DMU 0_{eff} from the observation set and forming a new efficient frontier via supporting hyperplanes that are binding at DMU 0_{eff} . The area between the nominal and new frontier is the stability region for DMU 0_{eff} . However, for large problems (in both input and output number and number of DMUs) this method is computationally intensive as it relies on the method from Huang et al. (1997) to determine all efficient surfaces of the model passing through DMU 0_{eff} . This method is not suitable for weakly efficient DMUs as removing them from the PPS does not change the efficient frontier. This led to Daneshvar et al. (2014) extending the methodology to include weakly efficient DMUs and anchor DMUs.

Boljunčić (2006) considers how the change in DMUs' data can affect extreme efficient DMUs. He notes that the larger the stability region for an extreme DMU the more reliable the obtained result is as the DMU is less affected by data misspecification or possible data errors. He assumes simultaneous changes of all inputs and outputs of DMU 0_{ex} but only allows a decrease of outputs and increase of inputs in the form of a small constant to each data point. In this way the region of stability is smaller than that derived in Jahanshahloo et al. (2005a). Boljunčić (2006) determines a region where DMU 0_{ex} remains efficient using an iterative parametric programming process to move to adjacent facets of the efficient frontier. Similar to this is the approach taken by Ghadimia and Ahadzadeh Namin (2009). However, here the inputs/outputs are multiplied by a small constant instead. They again calculate the region of efficiency for DMU 0_{ex} . DMU 0_{ex} is removed from the reference set and a linear programming model is used to determine the projection of DMU 0_{ex} to the facets of the new efficient frontier. Again the approach moves systematically from one facet to an adjacent one. They derive sufficient and necessary conditions for DMU 0_{ex} to remain efficient when iteratively changing each input/output of DMU 0_{ex} . They note that further research is required into the infeasibility of the linear programming model for all inputs and outputs.

Gholam Abri et al. (2009) observe that sometimes the behaviour of inefficient DMUs is very similar to the efficient DMUs and that although they are classified as inefficient they are in-fact operating satisfactorily. To overcome this they define a new efficiency category, ‘quasi-efficient’. Standard DEA models are used to classify each DMU as efficient when $\theta_0^* = 1$, quasi-efficient when $1 > \theta_0^* > \alpha$ and inefficient when $\theta_0^* \leq \alpha$, where α is determined by the conditions of the situation. The paper uses $\alpha = 0.7$ in the example but does not provide guidance on how to select this value. They consider both efficient and quasi-efficient DMUs to belong to the efficient category (despite not all of the DMUs having an efficiency score of one). The paper then proceeds to adapt the methodology from Charnes et al. (1992, 1996) to determine the stability radius for each DMU such that DMU 0’s classification changes from efficient or quasi-efficient to inefficient or vice versa.

Jahanshahloo et al. (2011) focuses on the inefficient DMUs. The paper considers methods such that DMU 0_{ineff} can improve to obtain an efficiency score of α as defined by the manager where α is close to one. This is done by finding the ‘necessary change region’, such that after these changes are made DMU 0_{ineff} will have an efficiency score of α . To do this the set of extreme efficient DMUs is first identified, the inputs of these DMUs are then multiplied by $\frac{1}{\alpha}$. This creates a new efficient frontier along which the transformed DMUs have efficiency score of α . The changes in inefficient DMUs to reach this new frontier are then found. DMU 0_{ineff} can improve by increasing outputs, decreasing inputs or a combination. In the two variable example given DMU 0_{ineff} follows a perpendicular line, the shortest distance from the nominal data to the transformed efficient frontier. The projection of DMU 0_{ineff} to the transformed efficient frontier provides the necessary change required for DMU 0_{ineff} ’s efficiency score to increase to α . DMU 0_{ineff} may move towards the new frontier via different strategies, this flexibility in the model allows a wide range of applications to be modelled.

Multiple changes

The previous DEA sensitivity analysis methods discuss changes to the data of a single DMU at a time. However, in many applications it is more realistic that all the DMUs' data suffer from uncertainty and this can affect the classification of the other DMUs.

One of the first DEA sensitivity analysis papers to consider simultaneous changes was Thompson et al. (1994). Motivated by data from Illinois coal mining and Kansas farming they believed that only small data changes were likely. Therefore, they considered changes to the solutions when all the data changed simultaneously in a small defined region. To implement this they increased(decreased) the efficient DMUs' inputs(outputs) in five percent increments and did the opposite to the inefficient DMUs. In this way they determined when a change in the efficiency ranking occurred.

Seiford and Zhu (1998a) also consider simultaneous data changes as they believe that "in reality possible data errors may occur for any DMU". Here they look at fixed percentage and absolute changes in a subset of inputs and outputs. Their method is a worst-case scenario approach where DMU 0_{eff} 's data changes unfavourably and all other DMUs' data changes favourably. Changes to DMU 0_{eff} are made that decrease the efficiency score (increase inputs and decrease outputs) and the efficiency score for the remaining DMUs increase (decrease inputs and increase outputs). A stability range for DMU 0_{eff} is calculated to determine how much change in the data can occur before it is reclassified as inefficient. They derive bounds on the possible changes for DMU 0_{eff} . Zhu (2001) continues this work by relaxing the assumption that the percentage changes of the data are the same for inputs and outputs. The method is based on super-efficiency, see Section 3.1.5 or Seiford and Zhu (1999). Here a subset of the inputs and outputs are changed for all DMUs first individually and then simultaneously.

An alternative approach is taken by Cooper et al. (1999) who consider bounded data, where the real data are known only within specified bounds or satisfy an ordinal

relationship. In this way the data are then variables whose exact value are unknown in advance and therefore must be determined. The bounds on the data form sets in which the true data vectors reside, similar to the idea of a box uncertainty set (see Section 3.2). However, this results in a nonlinear, nonconvex problem which is complicated further by the addition of ordinal relations between DMUs. To address this they develop a method they term Imprecise DEA, IDEA. Here the problem is transformed to a linear program equivalent by introducing new variables for products of data with multiplier variables (v, ν in (3.3)). These new variables are introduced such that the ordinal relations are preserved. In this way a linear program can be solved that allows for exact, ordinal and bounded data. The IDEA model is then extended to Assurance Region IDEA, AR-IDEA to include assurance regions on the multiplicative weights. An assurance region defines weights on the inputs and outputs rather than the data and is known within prescribed lower and/or upper bounds. The AR-IDEA model is solved by introducing new variables to translate the bounds on the multiplicative weights into sets of constraints.

Despotis and Smirlis (2002) also consider bounded data and provide an upper and lower bound on the efficiency score by considering best and worst case scenarios for DMU 0. The DMUs are classified into three categories: always efficient, efficient at the upper bound only and never efficient. When data are known to be imprecise but a bound on the data are not known they consider a single change to a single input of DMU 0_{ineff} to determine the level by which the input would need to reduce for DMU 0_{ineff} to be deemed efficient. This requires the addition of a non-linear constraint to the LP. A two-stage process using a bisection search is described that can be used to solve the nonlinear program in standard linear programming software. This is built on in He et al. (2016) who again consider bounded uncertainty of the DMUs data but allow changes in both the DMU under consideration and the remaining DMUs. They divide the data into sets of inputs and outputs they are interested in and consider

percentage data perturbations to DMU 0_{eff} and the remaining DMUs in order to calculate the stability region of each DMU and categorise them as in Despotis and Smirlis (2002).

Liu and Lai (2006) also consider subsets of inputs and outputs which they simultaneously deteriorate to determine the radius of stability for DMU 0_{eff} . The iterative deterioration is applied to a subset of DMUs including DMU 0_{eff} . This setting is motivated by determining the efficiency of Taiwan hospitals. Here the deterioration may only occur in the private hospitals in the north of Taiwan. They assume that the non-linear programmes which they use to model the data changes are feasible. However they conclude that this is not necessarily true. They assume that the minimum uncertainty for particular inputs and outputs can be found but note that this may not always be the case.

Other DEA sensitivity analysis techniques include the addition of a new DMU to the set and the effect this has. For example, Zamani and Borzouei (2016) use the method presented in Jahanshahloo et al. (2007) for finding strong defining hyperplanes of the production possibility set to determine the consequences of a new DMU being added to the solution set. The proposed method determines the stability region such that the new DMU is efficient and the originally efficient DMUs remain efficient.

From single changes in a single output in Charnes et al. (1984) to simultaneous data changes in all DMUs in all inputs and outputs in Thompson et al. (1994) there are many sensitivity analyses in DEA papers we have not mentioned here. Other similar related studies in DEA that contribute to this field of research include determining the least distance in DEA, (for an overview see Aparicio (2016)), and computational studies to determine all supporting hyperplanes of the DEA efficient frontier (Huang et al., 1997). We will build upon some of the methodology introduced in this section in Chapter 4 where, unlike much of the sensitivity analysis literature, the interest will be in how inefficient DMUs can change their efficiency status.

3.4 Uncertain DEA

In Section 3.1 we introduced DEA, a non-parametric method for assessing the efficiency of DMUs based on the dataset. Traditional DEA models assume that the data are known precisely. However, in many real-world applications the data are inherently uncertain. This could lead to a DMU being misclassified as inefficient when it is in-fact the imperfect data that is causing the DMU to be evaluated as inefficient. In this section we consider uncertain data to be the result of imperfect data due to a lack of accurate data where the ‘true’ amount of uncertainty can be any(or all) of the following: vague, unknown, obscure and arising from different sources. Ehrgott et al. (2018) introduced the concept of uncertain DEA, (uDEA) which aims to address such data instances. The reliability of the conclusions from a DEA model are intrinsically linked to the quality of the data used. The DMUs are compared against each other and assessed on the relationship between their inputs and outputs. Consequently, if not correctly accounted for uncertainties in the data can lead to the wrong conclusions.

uDEA is closely linked to robust optimisation methods and concepts introduced in Section 3.2 will be drawn upon. However, unlike in robust DEA, uDEA seeks to exploit the uncertain nature of the data to express each individual DMU under consideration in the most favourable way. uDEA aims to find the minimum amount of uncertainty required for an inefficient DMU to be rendered efficient, i.e. under what data instances would an inefficient DMU be categorised as efficient.

Here we introduce the general uDEA model, key definitions and configurations of uncertainty that will then be built upon in Chapter 4. We use similar notation to Ehrgott et al. (2018) and follow their definitions of uDEA.

3.4.1 The uDEA model

After solving the nominal DEA problem (3.8) a DMU is classified as efficient or inefficient. We now consider what effect introducing uncertainty has on this classification system and our perspective is altered as in Ehrgott et al. (2018) where we now classify DMUs as capable or incapable.

In Section 3.2.2 the robust DEA model was defined in (3.19) to be,

$$\mathcal{E}_i(\mathcal{U}) := \min_{\eta} \{c^T \eta : a_j \eta \leq 0, \forall a_j \in \mathcal{U}_j, \forall j, B\eta = e, \eta \geq 0\}. \quad (3.20)$$

We now extend (3.20) further by considering the effect of different uncertainty sets \mathcal{U}_j on the robust efficiency score $\mathcal{E}_i(\mathcal{U})$. To aid this we have the following definitions from Ehrgott et al. (2018).

Definition 3.4.1. *The uncertain inputs and outputs are:*

1. \mathcal{U}_j is an uncertainty set that models the possible values of the data a_j . Hence each $a_j \in \mathcal{U}_j$ is a possible row vector of input/output data for the j^{th} input/output.
2. $\mathcal{U} = \{\mathcal{U}_j : j = 1, \dots, N + M\}$ is a collection of uncertainty sets, or more succinctly, a collection of uncertainty. Hence \mathcal{U} contains the totality of uncertainty across all inputs and outputs.

Definition 3.4.2. Ω is the universe of possible collections of uncertainty.

Definition 3.4.3. An amount of uncertainty is a mapping

$$m : \Omega \rightarrow \mathbb{R}_+ : \mathcal{U} \mapsto m(\mathcal{U})$$

such that

- i. there is zero uncertainty if and only if there is no uncertainty i.e. $m(\mathcal{U}) = 0$ if and only if $|\mathcal{U}_j| = 1$ for $j = 1, \dots, M + N$, and
- ii. $m(\mathcal{U})$ is monotonic, i.e. $m(\mathcal{U}') \leq m(\mathcal{U}'')$ if $\mathcal{U}'_j \subseteq \mathcal{U}''_j$, $j = 1, \dots, M + N$.

This allows comparative evaluations between uncertainty sets as we have a numerical association with \mathcal{U} and hence can order the amount of uncertainty in increasing size. We do not explore this in detail as we focus mainly on box uncertainty where the notion of increasing uncertainty can easily be seen to be associated with increasing box size. Ehrgott et al. (2018) showed that as the uncertainty increases the efficiency of a DMU increases. A specific example of this for box uncertainty is explored in Section 4.3. As a result, the initial data provides a lower bound for the efficiency score a DMU can achieve.

In uDEA we search among all robust DEA solutions to find the minimum amount of uncertainty required to achieve the maximum possible efficiency score. We require a distinction between the original variables, $\theta_k, \lambda_i, i = 1, \dots, I$, of the DEA model (4.1) and optimal values, $\theta_k^*, \lambda_i^*, i = 1, \dots, I$, the variables take when (4.1) has been solved for the nominal data where the $*$ indicates that these are now realisations of a variable. In uDEA, we consider how the variable values of θ_k and λ change when uncertainty is introduced to (4.1). To distinguish between an initial optimal solution (λ^*, θ_k^*) and an optimal solution when uncertainty is introduced we use the notation $(\bar{\lambda}^*, \bar{\theta}_k^*)$ to show we are considering the values the variables take at an optimal solution to the uDEA problem.

The uDEA problem for the i^{th} DMU is:

$$\text{uDEA : } \quad \bar{\theta}_i^* = \sup_{0 \leq \theta_i \leq 1} \{ \theta_i : \min_{\mathcal{U} \in \Omega} \{ m(\mathcal{U}) : \mathcal{E}_i(\mathcal{U}) \geq \theta_i \} \} \quad (3.21a)$$

$$= \sup_{0 \leq \theta_i \leq 1} \{ \theta_i : \min_{\mathcal{U} \in \Omega} \{ m(\mathcal{U}) : \min_{\eta \geq 0} \{ c^T \eta : a_j \eta \leq 0, \forall a_j \in \mathcal{U}_j, \\ j = 1, 2, \dots, M + N, B \eta = e \} \geq \theta_i \} \}, \quad (3.21b)$$

where we again omit the dependence of i on the a terms to ease notation. Solving (3.21) for DMU i gives an optimal solution $(\bar{\eta}_i^*, \bar{\mathcal{U}}_i^*) = (\bar{\lambda}^*, \bar{\theta}_i^*, \bar{\mathcal{U}}_i^*)$, where $\bar{\theta}_i^*$ is the maximum efficiency score DMU i can achieve within the confines of the permissible uncertainty Ω ; $\bar{\mathcal{U}}_i^*$ is the collection of uncertainty with the minimum amount $m(\mathcal{U})$

of uncertainty required to achieve $\bar{\theta}_i^*$ and $\bar{\eta}_i^*$ contains the values of $\bar{\lambda}_i^*$ at the optimal solution and hence the peers for DMU i .

Solving (3.21) equates to finding the supremum, here used in case the maximum does not exist, of all θ_i such that the uncertainty is as small as possible, where the amount of uncertainty is defined in Definition 3.4.3, i.e., the robust efficiency score must be greater than or equal to θ_i . In this way by solving the uDEA problem we find the maximum possible efficiency score that can be obtained with the minimum amount of uncertainty. Here, as in classical DEA methodology, each DMU has the possibility of maximising its efficiency score. In addition, due to the uncertainty, each DMU can also choose the most favourable data instance, the best possible collection of inputs and outputs. Here the objectives are sequential: increase θ_i and then minimise $m(\mathcal{U})$. This is to follow DEA's underlying methodology of improvement for each individual DMU.

We note that if the nominal problem (3.8) has $\theta_i^* = 1$ then $\bar{\theta}_i^* = 1$, $\bar{\mathcal{U}}_i^* = \mathcal{U}_i^0$, $\bar{\eta}_i^* = (0, \dots, 1, \dots, 1, 1)$ where the first one is in the i^{th} position and $m(\mathcal{U}) = 0$ is the optimal solution. That is a DMU that is efficient in the nominal DEA problem will remain efficient in the u-DEA problem.

We now wish to distinguish between those DMUs that remain inefficient in the presence of uncertainty and those that can achieve efficiency. To do this we use the definition from Ehrgott et al. (2018) of capable and incapable DMUs.

Definition 3.4.4. *A DMU i under Ω is:*

- i. capable if $\bar{\theta}_i^* = \mathcal{E}_i(\mathcal{U}) = 1$ for some $\mathcal{U} \in \Omega$.*
- ii. weakly incapable if $\bar{\theta}_i^* = 1$ but $\mathcal{E}_i(\mathcal{U}) < 1$ for all $\mathcal{U} \in \Omega$.*
- iii. strongly incapable if $\bar{\theta}_i^* < 1$.*
- iv. incapable if it is either strongly or weakly incapable.*

In this way a DMU is incapable if it is inefficient for all uncertain data instances, i.e. even with the ability to select the most favourable data from Ω it has no claim on efficiency. The only way for an incapable DMU to become capable is if a change in Ω occurs.

3.5 Principal Component Analysis and other variable reduction techniques

In Section 2.1.2 we introduced Rosemere and the clinical protocol they use. To be able to use the data from Rosemere we need to be able to select suitable variables that represent the treatment plans. To do this we require variable selection techniques. Here we introduce Principal component analysis (PCA), and other techniques that we apply to the data from Rosemere in Section 5.2.

3.5.1 Principal Component Analysis

PCA is a statistical technique used for dimension reduction. PCA seeks to identify a basis that allows us to express the data set in the most meaningful fashion, i.e., the key information becomes readily apparent. The noise in the data is filtered out with this new basis and hidden structures are revealed. To do this, the variables are transformed to a new set of variables called the principal components (PCs). Each PC is a linear combination of the original variables. The first PC accounts for the largest proportion of variability in the data, i.e. the direction where there is the most variance and hence where the data are most spread out. Each following component will be linearly independent (uncorrelated) from the previous components. PCA is broadly used for data analysis and its popularity is driven by its simplicity. It is a non-parametric tool and can efficiently extract important information from complex data sets without major assumptions on the data.

We introduce the mathematical background, method and standard techniques for analysing and using PCA. In section 5.2 this will be used to aid variable selection for our DEA method. We note here that we do not discuss PCA-DEA, a method that uses the PCs as variables in the DEA model directly. This is because for our chosen application a weighted linear combination of variables does not make sense. For an overview of PCA-DEA see Adler and Golany (2007).

Consider a data set that consists of a sample of n observations on a vector of q variables $x = (x_1, x_2, \dots, x_q)$. We note here that a, i, j, k, q, z, α and λ are used here but bear no relation to the use of $a, i, j, k, q, z, \alpha, \lambda$ notation in other chapters. Then x is a random vector with $q \times q$ sample variance-covariance matrix S whose $(i, j)^{\text{th}}$ element, $s_{i,j}$, is the sample covariance between the i^{th} and j^{th} element of x when $i \neq j$ and the sample variance of the j^{th} element, s_j , when $i = j$. Note here we use the sample covariance matrix instead of the population covariance matrix as it is more realistic that we only have a sample of the whole dataset.

The first PC of this data is given by

$$z_1 = a_1^T x = \sum_{i=1}^q a_{i1} x_i,$$

where the vector of coefficients $a_1 = (a_{11}, a_{21}, \dots, a_{q1})$ is obtained by maximising the variance of z_1 subject to $a_1^T a_1 = 1$. This constraint on the maximisation ensures the answer is unique. This can be solved by using the Lagrange multiplier, λ ,

$$\max_{a_1} [a_1^T S a_1 - \lambda(a_1^T a_1 - 1)]. \quad (3.22)$$

Differentiation of (3.22) reveals that a_1 is an eigenvector of S with the corresponding eigenvalue λ_1 . Since this λ_1 maximises the variance of z_1 , it is the largest eigenvalue of S as

$$\text{var}(z_1) = a_1^T S a_1 = a_1^T \lambda_1 a_1 = \lambda_1.$$

In other words, the first PC z_1 retains the greatest amount of data variation, i.e. information on sample variability.

Similarly, the k^{th} PC of x is given by the transformation

$$z_k = a_k^T x \quad \forall k = 1, \dots, q,$$

where the vector of coefficients $a_k = (a_{1k}, a_{2k}, \dots, a_{qk})$ is given by

$$\text{var}(z_k) = a_k^T S a_k = \lambda_k.$$

This means that the k^{th} largest eigenvalue of S is the variance of the k^{th} PC z_k , which retains the k^{th} greatest fraction of the variation in the data set.

The resulting eigenvalues of the PCs' can be displayed graphically to highlight their relative importance. In this fashion, the first component reveals the direction where the data are most spread out.

PCA works on the basis that there is shared variance amongst variables, i.e., across the whole sample there is a dependence structure, (Hair Jr et al., 1995). In this way we must make the following assumptions:

- i. there is a linear relationship between the variables,
- ii. the mean and variance define the probability distribution of the data and
- iii. we have a sufficiently large dataset that it is representative of the true population.

Assumption (i) can be checked by calculating the Pearson correlation coefficient or plotting the variables pairwise if there are few variables. Assumption (ii) means a large variance in the dataset is meaningful, whereas a small variance is assumed to be caused by noise in the data. Assumption (iii) can be thought of as requiring sampling adequacy, i.e. the sample we have is a fair representation of the true population. For both assumption (ii) and (iii) the larger the sample size the less concerned by these assumptions we can be due to the central limit theorem.

In practice it is common to use the correlation matrix C instead of the sample variance matrix S where,

$$C(x_i, x_j) = \frac{s_{i,j}}{s_i s_j},$$

$s_{i,j}$ is the covariance between x_i and x_j and s_i, s_j is the variance of x_i and x_j . This overcomes the sensitivity of PCs to units of measurements which is particularly important if there are variables with largely different scales of magnitude or units. This ensures the variables with large variances do not dominate the first few PCs simply because they have a larger measurement scale. In this way, each variable is given equal weighting in the analysis.

After applying PCA the results must be interpreted. One problem with PCA is that the raw PCs are hard to interpret and evaluate. From the PC's coefficients it is often hard to identify the underlying patterns and relationships between the original variables and the new coefficients. Jolliffe (2002) note that, 'when we interpret PCs it is usually only the general pattern of the coefficients that is really of interest, not the values to several decimal places which may give a false impression of precision'. As a result in Jolliffe (2002) a simple +, - notation is used to look at the patterns occurring in the PCs. For each PC the absolute maximum coefficient is determined. Then for each PC $k = 1, \dots, q$, let

$$\alpha_k = \max_{i=1, \dots, I} |a_{ik}|,$$

where the a_k are the eigenvectors of S . Then, the following representation is used:

$$\text{If} \quad a_{ik} \geq \frac{1}{2}\alpha_k \quad + \quad (3.23a)$$

$$\frac{1}{2}\alpha_k > a_{ik} \geq \frac{1}{4}\alpha_k \quad (+) \quad (3.23b)$$

$$a_{ik} \leq -\frac{1}{2}\alpha_k \quad - \quad (3.23c)$$

$$\frac{1}{2}\alpha_k < a_{ik} \leq -\frac{1}{4}\alpha_k \quad (-) \quad (3.23d)$$

$$\text{else} \quad a_{ik} \quad \text{insignificant.} \quad (3.23e)$$

In this way patterns in the variables can easily be spotted. For example if the first PC has a + representation for all coefficients this suggest the first PC is a measure of overall size. It also shows if there are large positive and negative contributions in the PCs.

We now discuss some popular criteria that can be used to interpret the PCA results further. This may be in the form of choosing a number of PCs for further analysis or making decisions based on the first few PCs.

One popular method is to choose a predetermined percentage of variance the components must explain. This is an arbitrary number and is normally in the range of 80-95% variance depending on the data. Although this is an ad-hoc rule of thumb and there is limited formal justification, it is a popular method due to it being intuitively plausible and because in practice it often works well. An example of this method can be found in Jolliffe et al. (1982). They classify elderly individuals according to similar characteristics to provide insight into particular requirements the elderly may need from social services and other health care providers. From 20 variables they reduced the sample data to ten PCs that accounted for over 80 percent of the variation.

Another commonly quoted method is Kaiser's "eigenvalues-greater-than-one" rule (Kaiser, 1960). Here PCs that have eigenvalues less than one are discarded as they have less information than using one of the original variables. The motivation behind this method came from Kuder-Richardson's reliability of factors formulas (Kuder and Richardson, 1937) and from observations of the meaningfulness of PCs in terms of data applications. He concluded that "the number of eigenvalues greater than one of the observed correlation matrix led to a number of factors corresponding almost invariably, in a great number of studies, to the number of factors which practicing psychologists were able to interpret".

Alternatively, a scree test criterion can be used. The PCs are plotted according to the percentage variance they explain. A levelling off on the graph is sought, i.e. a

point where the rate of percentage variance explained vs the number of variables in the analysis reduces.

Sometimes the next stage of analysis requires whole variables to be selected from the PCA because weighted linear combinations make no sense in the application. Simply selecting the variables that have the largest contribution in the first few PCs is not sufficient (Jolliffe, 1973, 1972). So a number of techniques have been suggested in the literature.

One method is to look at the least significant PC and find the variable which contributes to it most. This variable is then removed and PCA is repeated until the desired number of PCs is reached. The justification behind this method is that the end PCs will be dominated by variables that explain less in the first PCs and so in removing them little information is lost, see Jolliffe (2002) for more details. The aim is that this results in variables which explain more of the variation in the earlier PCs being kept.

Another method is to compute the correlation between PCs and original variables. keeping those variables that correlate highly with PCs with larger eigenvalues. An example of this can be found in Uva et al. (2009) where they use PCA to analyse human and dog gene expression data from tumour and normal tissue cells. They calculate the correlation between the first few PCs with the original variables. Variables that correlate highly with the first few PCs are then selected for further analysis.

One downfall of PCA is its sensitiveness to the data currently being used, i.e. in small samples outliers can have a large influence on the conclusion drawn. This is particularly problematic when whole variables are being selected as a few outliers can cause an insignificant variable to be chosen for the further analysis.

Although normalising the data can help remove bias in variables that have largely different scales it does not prevent data with underlying structures distorting the PCs. An example of this will be seen in Section 5.2.3 where the first PC divides into two

distinct groups due to the number of fractions distorting the PCs.

Many of these techniques are subjective and parameters such as the percentage variance explained should be based on the dataset at hand. Despite the subjectivity of how to apply and interpret PCA it remains a very popular technique across many different disciplines.

In Section 5.2.4 we develop a new variable selection method based on PCA that aims to overcome PCA's sensitiveness to the current sample and provides a technique for choosing whole variables from a larger collection.

3.5.2 Partial Covariance

An alternative method for variable selection is based on the proportion of the total variance retained when conditioning out a subset of variables. This is called the partial covariance method.

Let x be a q dimensional normally distributed random vector with mean 0 and sample covariance matrix S . The aim of the partial covariance method is to select variables that explain as much variation in the data as possible. Therefore, we reorder the q variables of the data into two groups $i = 1, \dots, p$ are the variables we keep and $i = p + 1, \dots, q$ are the variables we remove. The number of variables to keep should be as small as possible but large enough that it represents the full dataset. For small values of q it is possible to evaluate all combinations of $\binom{q}{p}$ variables.

First, we partition the covariance matrix such that

$$S = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix},$$

where S_{11} is the $p \times p$ covariance matrix of the $1, \dots, p$ variables to keep, S_{22} is the $(q-p) \times (q-p)$ covariance matrix of the $p+1, \dots, q$ variables to remove and $S_{12} = (S_{21})^T$ is the $(p \times q-p)$ covariance matrix between the subsets of variables to keep and remove.

Then the conditional covariance of the $i = p + 1, \dots, q$ variables given the $i = 1, \dots, p$ variables can be calculated. We denote this conditional covariance by $S_{2|1}$ and calculate it via

$$S_{2|1} = S_{22} - S_{21}S_{11}^{-1}S_{12},$$

(Morrison, 1990). This can be repeated for all combinations of variables. The trace² of $S_{2|1}$ ($Tr(S_{2|1})$) represents the size of the remaining variance in the variables $i = p + 1, \dots, q$ which have not been accounted for in the $i = 1, \dots, p$ variables that are kept. By selecting the variable set with the smallest value of $Tr(S_{2|1})$ for any given p we retain the most information. By selecting a level of variance we wish to retain, (Adler and Yazhensky (2010) suggest 76%) we can find the smallest number of variables required to explain the set level.

To aid the variable selection process McCabe (1984) suggests calculating the percentage of variance explained:

$$\%Explained = 1 - \frac{\text{sum of the eigenvalues of } S_{2|1}}{\text{sum of the eigenvalues of } S}. \quad (3.24)$$

If using normalised data the calculation of (3.24) is simplified to

$$\%Explained = 1 - \frac{Tr(S_{2|1})}{q}.$$

This method will be applied to data from Rosemere in Section 5.2.5.

²The trace of a square matrix is defined to be the sum of the diagonal elements.

Chapter 4

Treating uncertainty

Throughout this Chapter we investigate the relationship between the amount of uncertainty required for an inefficient DMU to be deemed efficient and the DEA distance of the DMU under consideration. The DEA distance is defined in Definition 4.1.6 and can be thought of as the Euclidean distance from a DMU to its projection onto a facet of the efficient frontier. We consider which facet of the efficient frontier should be the target of this projection. When uncertainty is introduced is it beneficial for an inefficient DMU to compare itself to a facet of the efficient frontier not defined by its nominal peers? We leverage the uncertainty present in the data such that inefficient DMUs may improve. We wish to determine the minimum amount of uncertainty required for an inefficient DMU to be evaluated as efficient.

In Section 3.1 we introduced the input-oriented VRS DEA model in matrix notation. We repeat the model here in summation notation to aid development of the following theorems.

Consider I DMUs, each with M non-negative, non-zero outputs and N non-negative, non-zero inputs. We assume that the PPS has dimension $N + M$ and that $I > M + N$. The inputs and outputs of the i^{th} DMU are represented by the column vectors $x_i \in \mathbb{R}^N$ and $y_i \in \mathbb{R}^M$. The efficiency score θ_k^* for DMU k can be

found by solving the linear optimisation problem

$$\min_{\lambda, \theta} \theta_k \quad (4.1a)$$

$$s.t. \quad \sum_{i=1}^I \lambda_i y_{mi} - y_{mk} \geq 0 \quad m = 1, \dots, M \quad (4.1b)$$

$$\sum_{i=1}^I \lambda_i x_{ni} - \theta_k x_{nk} \leq 0 \quad n = 1, \dots, N \quad (4.1c)$$

$$\sum_{i=1}^I \lambda_i = 1 \quad (4.1d)$$

$$\lambda_i \geq 0 \quad i = 1, \dots, I. \quad (4.1e)$$

Here y_{mi} is the m^{th} output and x_{ni} is the n^{th} input value for DMU i . $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_I) \in \mathbb{R}^I$ is a decision variable. From Definition 3.1.5 the solution $(\lambda, \theta_k) = (0, 0, \dots, 1, \dots, 0, 1)$, where the first one is in the k^{th} position is feasible. The feasibility of this solution ensures the maximum value that θ_k can take is 1, i.e. an optimal value θ_k^* of model (4.1) is less than or equal to one. In this way, we denote an optimal solution for DMU k by (λ^*, θ_k^*) . $\lambda^* \in \mathbb{R}^I$ is a vector of weights (not necessarily unique) at an optimal solution. In this chapter we use the following definition of efficiency.

Definition 4.0.1. *DMU k is efficient if and only if $\theta_k^* = 1$.*

Therefore, if $\theta_k^* < 1$ DMU k is inefficient. To find the efficiency scores for the I DMUs the linear program (4.1) must be solved I times, once for each of the I DMUs.

Model (4.1) assumes that the data x_{ni} , y_{mi} are exact and DMUs are classified as inefficient or efficient based on these data. However, the data of many real-world applications are inherently uncertain. Hence, it is possible that an inefficient DMU performs well in practice, i.e. it is the uncertainty in the data that stops it being classified as efficient. In this chapter, we study the uncertain nature of the data and the effect this has on an individual DMU's efficiency score. Building upon Ehrgott et al. (2018), (see Section 3.4) we refine the concept of uDEA for the specific case where

we have box uncertainty. This refinement is motivated by the real-world radiotherapy application as introduced in Section 2.1.

In DEA sensitivity analysis (see Section 3.3), the focus is on how an efficient DMU can remain efficient. We are instead interested in the inefficient DMUs and how they can become efficient, i.e. we wish to leverage the uncertainty present in the data such that inefficient DMUs may improve and determine the minimum amount of uncertainty required for an inefficient DMU k to be evaluated as efficient.

In Section 4.1 we introduce some important definitions and results to be used throughout the chapter. In Section 4.2 we start by considering a fixed amount of uncertainty. In Section 4.3 we examine the effect of a variable amount of uncertainty. We derive the minimum amount of uncertainty required for an inefficient DMU k to become efficient both geometrically and via the DEA model (4.1). The section concludes with a short discussion of future directions to solving the problem.

Throughout this chapter, we borrow the terminology and methodology of robust optimisation, DEA sensitivity analysis and uDEA from Sections 3.2, 3.3 and 3.4.

As we wish to solve the uncertain model (3.21) for inefficient DMU k (uDEA) as introduced in Section 3.4, we repeat the uDEA model here to highlight the differences between the models we use in the following sections.

$$\text{uDEA : } \quad \bar{\theta}_i^* = \sup_{0 \leq \theta_i \leq 1} \{ \theta_i : \min_{\mathcal{U} \in \Omega} \{ m(\mathcal{U}) : \mathcal{E}_i(\mathcal{U}) \geq \theta_i \} \}. \quad (4.2)$$

As in Section 3.4 we use the notation $(\bar{\lambda}^*, \bar{\theta}_i^*, \bar{\mathcal{U}}_i^*)$, to denote the values the variables take at an optimal solution to the uDEA problem (4.2).

To aid the development of the next section, we consider a simplified version of the uDEA problem (4.2) by focusing on the constraints for inefficient DMU k . In Section 4.2, we assume a fixed uncertainty amount, therefore, we omit $\min_{\mathcal{U} \in \Omega}$ in (4.2). In this way, for fixed u , we obtain a linear programme which is equivalent to solving the robust DEA model (3.19). We assume that $\bar{\theta}_i^* = 1 \forall i \neq k$ holds throughout, for justification see Theorem 4.1.16 and Corollary 4.1.17. This gives the following

model, where Δ_{mi} , Δ_{ni} are the uncertainty associated with the efficient DMUs and Δ_{mk} , Δ_{nk} the uncertainty associated with the inefficient DMU.

$$\min_{\theta_k, \lambda, \Delta} \theta_k \quad (4.3a)$$

$$s.t. \sum_{i=1}^I \lambda_i (y_{mi} + \Delta_{mi}) - (y_{mk} + \Delta_{mk}) \geq 0 \quad m = 1, \dots, M \quad (4.3b)$$

$$\sum_{i=1}^I \lambda_i (x_{ni} + \Delta_{ni}) - \theta_k (x_{nk} + \Delta_{nk}) \leq 0 \quad n = 1, \dots, N \quad (4.3c)$$

$$\sum_{i=1}^I \lambda_i = 1 \quad (4.3d)$$

$$\lambda_i \geq 0 \quad i = 1, \dots, I \quad (4.3e)$$

$$\Delta_{mi}, \Delta_{ni} \in [-u, +u] \quad m = 1, \dots, M \quad n = 1, \dots, N \quad i = 1, \dots, I. \quad (4.3f)$$

Definition 4.0.2. An optimal solution to (4.3) for DMU k is the vector

$$(\bar{\lambda}^*, \bar{\theta}_k^*, \Delta_{mi}^*, \Delta_{ni}^*), \quad m = 1, \dots, M, \quad n = 1, \dots, N, \quad i = 1, \dots, I.$$

This gives the following definition.

Definition 4.0.3. Box uncertainty is defined to be

$$\{x_{ni}^u \in [x_{ni}^-u, x_{ni}^+u], y_{mi}^u \in [y_{mi}^-u, y_{mi}^+u], \quad m = 1, \dots, M, \quad n = 1, \dots, N, \quad i = 1, \dots, I\}. \quad (4.4)$$

Where x_{ni} and y_{mi} are the nominal data and x_{ni}^u and y_{mi}^u are the values the data takes in the uncertainty set. In this way, the true data for each DMU falls in an interval of length $2u$ for each input and output. W.l.o.g. we assume that the possible deviation from the nominal data value is the same for each input/output. This assumption can be made without restricting the uncertainty sets of each variable to be equal, in the following way.

We denote the $m(n)^{\text{th}}$ row of $Y(X)$ the output(input) data by $Y_m(X_n)$. Then for the M outputs and N inputs let $u_{y_1}, \dots, u_{y_M}, u_{x_1}, \dots, u_{x_N}$ be the corresponding uncertainty. Let $U = u_{y_1} \times u_{y_2} \times \dots \times u_{y_M} \times u_{x_1} \times \dots \times u_{x_N}$. Transform the original data

DMU	X_1	X_2	Y	θ^*
A	140	92	60	1
B	280	115	60	0.800
C	420	161	180	0.857
D	560	230	420	1

Table 4.1: Data and efficiency scores, Example 2.

such that, $Y'_1 = Y_1 * \frac{U}{u_{y_1}}, \dots, Y'_M = Y_M * \frac{U}{u_{y_M}}, X'_1 = X_1 * \frac{U}{u_{x_1}}, \dots, X'_N = X_N * \frac{U}{u_{x_N}}$. Then using the transformed data with uncertainty U is equivalent to using the original data with $u_{y_1}, \dots, u_{y_M}, u_{x_1}, \dots, u_{x_N}$. This is because DEA is not affected by scaling, Ali and Seiford (1990), therefore the inputs/outputs can be scaled so that the amount of uncertainty is the same. This is demonstrated in the following example.

Example 2: Scaling uncertainty

Consider the four DMUs $A - D$ with inputs X_1 and X_2 and output Y , the data for these are presented in Table 4.1. Applying model (4.1) identifies DMU A and D as efficient and DMUs B and C as inefficient. Their efficiency scores are given in the last column of Table 4.1. The uncertainty associated with each variable is $u_{x_1} = 3, u_{x_2} = 2, u_y = 4$. We add uncertainty to each data point as described in Table 4.2, this uncertainty configuration will be explained later in Section 4.2. Here they can be thought of as realisations from the interval of length $2u_j, j \in \{x_1, x_2, y\}$ around each data point. The efficiency scores for the uncertain data are given in Table 4.2.

We now scale the data in Table 4.1 such that $Y' = Y * \frac{U}{u_y}, X'_1 = X_1 * \frac{U}{u_{x_1}}$ and $X'_2 = X_2 * \frac{U}{u_{x_2}}$ where $U = u_{x_1} \times u_{x_2} \times u_y = 3 \times 2 \times 4 = 24$. The data for the scaled variables Y', X'_1 and X'_2 are given in Table 4.3 along with the efficiency scores. We note that as expected, the efficiency score for the scaled data is the same as for the nominal data in Table 4.1.

DMU	X_1^u	X_2^u	Y^u	$\bar{\theta}^*$
A	$140 + u_{x_1}$	$92 + u_{x_2}$	$60 - u_y$	1
B	$280 - u_{x_1}$	$115 - u_{x_2}$	$60 + u_y$	0.859
C	$420 - u_{x_1}$	$161 - u_{x_2}$	$180 + u_y$	0.900
D	$560 + u_{x_1}$	$230 + u_{x_2}$	$420 - u_y$	1

Table 4.2: Uncertainty u_{x_1}, u_{x_2} and u_y , Example 2.

DMU	X'_1	X'_2	Y'	θ^*
A	1120	1104	360	1
B	2240	1380	360	0.800
C	3360	1932	1080	0.857
D	4480	2760	2520	1

Table 4.3: Data with scaled uncertainty, Example 2.

Due to the scaling in Table 4.3 we can now use the value U for all the data. This is done in Table 4.4. Again we see that the efficiency scores in Table 4.2 and 4.4 are the same, thus we can use the scaled data to ease calculations. Examination of the peers for the DMUs in Table 4.2 and 4.4 are also the same with the same λ_i values.

Throughout the rest of the chapter we assume this scaling has already occurred. We drop the ' notation and use u to denote the amount of uncertainty applied to all the data. We assume the uncertainty is never larger than the data and hence the data remains strictly positive. Each data-point takes values in the range $x_{ni}^u \in [x_{ni} - u_x, x_{ni} + u]$ and $y_{mi}^u \in [y_{mi} - u_y, y_{mi} + u]$ $i = 1, \dots, I$, $m = 1, \dots, M$, $n = 1, \dots, N$ where $u_x = \min(x_{ni}, u)$ and $u_y = \min(y_{mi}, u)$. *W.l.o.g.* we will assume that $u < x_{ni}$ and $u < y_{mi} \forall i = 1, \dots, I$, $m = 1, \dots, M$, $n = 1, \dots, N$. Therefore, $x_{ni}^u \in [x_{ni} - u, x_{ni} + u]$ and $y_{mi}^u \in [y_{mi} - u, y_{mi} + u]$ and $x_{ni}, y_{mi} > 0$.

DMU	\mathbf{X}'_1^u	\mathbf{X}'_2^u	\mathbf{Y}'^u	$\bar{\theta}^*$
A	$140 + U$	$92 + U$	$60 - U$	1
B	$280 - U$	$115 - U$	$60 + U$	0.859
C	$420 - U$	$161 - U$	$180 + U$	0.900
D	$560 + U$	$230 + U$	$420 - U$	1

Table 4.4: Uncertainty U , Example 2.

4.1 Preliminary results

To start, we introduce some important definitions and results to be used throughout this chapter. Throughout this chapter we explore the effect uncertainty in the data has on the DEA model. This results in us referring to points in the PPS that have not been observed in the data. Here we will refer to them as artificial DMUs and use the notion from Thanassoulis and Allen (1998) that these are DMUs that have definite values for inputs and outputs, but have not been observed in practice, i.e. they are obtained based on the inputs and outputs of other DMUs. This allows us to refer to points on the PPS that are defined by uncertain data of existing DMUs. In particular we define the virtual DMU k^u .

Definition 4.1.1. *The virtual DMU i^u is defined to be the DMU with inputs and outputs $i^u = (x_i + u, y_i - u), \forall i \neq k, i^u = (x_i - u, y_i + u)$ if $i = k$.*

This allows us to consider projections from the point k^u to the efficient frontier.

As introduced in Section 3.1 the PPS of (4.1) is a polyhedron defined by the intersection of a finite number of half spaces. We have assumed that the PPS has dimension $N+M$ and hence the facet of the efficient frontier inefficient DMU k is being projected to can be defined by at most $\Phi = N + M$ DMUs. To aid the development of the following sections we define a $\Phi - 1$ dimensional hyperplane. A hyperplane in

Φ dimensions can be defined by the parametric equation

$$\alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_N x_N + \beta_1 y_1 + \beta_2 y_2 + \cdots + \beta_M y_M = \gamma. \quad (4.5)$$

Where $\alpha_1, \dots, \alpha_N, \beta_1, \dots, \beta_M$ and γ are constants.

In vector notation this is

$$\Pi_\Phi = \{r \in \mathbb{R}^d : r \bullet \vec{n}_1 = \gamma\} \quad (4.6)$$

where $\vec{n}_1 = (\alpha_1, \dots, \alpha_N, \beta_1, \dots, \beta_M)$ is a normal vector to the plane and the dot product of two vectors $a, b \in \mathbb{R}^d$ is defined to be $a \bullet b = a_1 b_1 + a_2 b_2 + \cdots + a_d b_d$.

Definition 4.1.2. Ψ is the collection of hyperplanes that define efficient facets.

Definition 4.1.3. An extreme point of the PPS formed by the intersection of several facets $\Pi \in \Psi$.

An extreme point will be an existing efficient DMU.

Definition 4.1.4. The shortest (Euclidean) distance from a point $P = (x_{10}, \dots, x_{N0}, y_{10}, \dots, y_{M0})$ to a hyperplane (4.5) is

$$d(P, \Pi_\Phi) = \frac{|\alpha_1 x_{10} + \cdots + \alpha_N x_{N0} + \beta_1 y_{10} + \cdots + \beta_M y_{M0} - \gamma|}{\sqrt{\alpha_1^2 + \cdots + \alpha_N^2 + \beta_1^2 + \cdots + \beta_M^2}}. \quad (4.7)$$

We now state a key result of this chapter, the proof of which will be given in Section 4.3. Here the results are given for the input oriented DEA model.

Result 4.1.5. The DEA distance, $\mathfrak{D}(i, \Pi_\Phi)$, from DMU i to the hyperplane, Π_Φ defined by (4.5) is

$$\mathfrak{D}(i, \Pi_\Phi) = \frac{|\alpha_1 x_{1i} + \cdots + \alpha_N x_{Ni} + \beta_1 y_{1i} + \cdots + \beta_M y_{Mi} - \gamma|}{\sqrt{\alpha_1^2 + \cdots + \alpha_N^2}}. \quad (4.8)$$

Definition 4.1.6. The DEA distance, $\mathfrak{D}(i, \Pi)$, is the Euclidean distance from DMU i to the plane Π given by (4.8).

Definition 4.1.7. *The DEA distance, $\mathfrak{D}(i)$, is the minimum DEA distance, $\mathfrak{D}(i, \Pi)$, for all hyperplanes $\Pi \in \Psi$.*

It follows from Definition 4.1.7 that DMU i is efficient in the nominal DEA problem if and only if $\mathfrak{D}(i) = 0$. We note that by the nature of the nominal DEA problem $\mathfrak{D}(i, \Pi_\iota)$, where ι is the hyperplane defined by the peers of DMU i will always give the minimum DEA distance $\mathfrak{D}(i)$.

Definition 4.1.8. Π^u is the plane formed from the translation $\Pi \rightarrow \Pi^u$ where the points x, y of Π are transformed to the points x', y' via

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} -u \\ u \end{bmatrix}.$$

This leads to the following definition of u , the amount of uncertainty.

Definition 4.1.9. u_k is the minimum amount of uncertainty required such that $\mathfrak{D}(k^u) = 0$.

Here u is the amount of uncertainty $m(\mathcal{U})$ for the specific case of box uncertainty. From now on we denote this by u . In this way u_k is the minimum amount of uncertainty required for DMU k to be deemed efficient. For any DMU i that is efficient in the nominal DEA problem $u_i = 0$.

In the standard DEA problem, here VRS input oriented model (3.8), each inefficient DMU k is projected to a point on the efficient frontier along a trajectory of fixed outputs. The point DMU k is projected to on the efficient frontier can be defined by a convex combination of its peers. We consider where DMU k would be projected to if it is projected to an alternative hyperplane of the efficient frontier. We call these target points and define them here.

Definition 4.1.10. *The target point $T(k, \Pi)$ is defined to be the point on the plane Π that DMU k is projected to.*

When $\Pi \in \Psi$ is a hyperplane such that $\mathfrak{D}(k^u, \Pi)$ is minimal, then $T(k, \Pi)$ is the projection of DMU k to the facet of the efficient frontier defined by the peers of DMU k in the nominal DEA problem i.e. it is the standard DEA projection for DMU k . This is seen in Example 3 where DMU E is projected to the point i on the efficient frontier. When Π is not the facet of the efficient frontier such that $\mathfrak{D}(k^u, \Pi)$ is minimal the projection will be to a virtual point not on the efficient frontier. That is if the facet of the efficient frontier is extended where would this intercept with DMU k be when DMU k 's outputs remain fixed but the inputs can vary. This can be seen in Example 3 where DMU E is projected to the points g and h which are not on the efficient frontier but are on hyperplanes which define facets of the efficient frontier. We note that in the standard DEA problem the projection to Π when Π is a facet of the efficient frontier such that $\mathfrak{D}(k, \Pi)$ is minimal is the only projection in the standard DEA model. This is because in the nominal DEA problem there is only one point a DMU is projected to, the point that renders the DMU efficient on the nominal efficient frontier. Hence, our use of projection in the definition of the target point is not a standard DEA term. To help understand these definitions we demonstrate the case of a single input and output with the example below.

Example 3: DEA Distance

Consider the six DMUs pictured in Figure 4.1 and whose nominal data are listed in Table 4.5. In Figure 4.1 the efficient frontier is shown by the red line sections between the efficient DMUs $A - D$. By extending the line sections going through pairs of efficient DMUs inputs and outputs (and through the DMU with the smallest(largest) $x(y)$ value) we can show the hyperplanes, (here lines as we are in \mathbb{R}^2) that intersect with the PPS. These are shown in Figure 4.1 by the dashed lines. This means $\Psi = \{\Pi_A, \Pi_{AB}, \Pi_{BC}, \Pi_{CD}, \Pi_D\}$. Where Π_{ij} is the hyperplane going through the points defined by the data of DMU $i, j \in \{A, B, C, D\}$, Π_A is the vertical line $x = x_A$ and Π_D is the horizontal line $y = y_D$.

DMU	In	Out	Efficiency score
A	1	1	1
B	3	4	1
C	7	7	1
D	10	8	1
E	8	5	0.542
F	6	2	0.278

Table 4.5: Nominal data, Example 3.

$\Pi \in \Psi$	$\mathbf{T}(\mathbf{E}, \Pi)$	$\mathbf{T}(\mathbf{F}, \Pi)$	$\mathfrak{D}(\mathbf{E}, \Pi)$	$\mathfrak{D}(\mathbf{F}, \Pi)$
Π_A	$g = (1, 5)$	$(1, 2)$	7	5
Π_{AB}	$h = (\frac{11}{3}, 5)$	$(\frac{5}{3}, 2)$	$\frac{13}{3}$	$\frac{13}{3}$
Π_{BC}	$i = (\frac{1}{3}, 5)$	$(\frac{1}{3}, 2)$	$\frac{11}{3}$	$\frac{17}{3}$
Π_{CD}	$g = (1, 5)$	$(-8, 2)$	7	14
Π_D	-	-	∞	∞

Table 4.6: DEA Distance for DMUs E and F , Example 3.

Therefore, for inefficient DMUs E and F there are five lines that intersect with the PPS that we can calculate the DEA distance for. These are shown in Figure 4.1. We can then calculate the DEA distances for the inefficient DMUs to their target points on each line in Ψ , these are given in Table 4.6. The target points for DMU E for each of the lines are shown in Figure 4.1 by the points $g - i$. We note here that in the nominal DEA DMUs E and F can never be projected to the line Π_D .

As expected we see that $\mathfrak{D}(E) = 3.66$ is obtained when DMU E is projected to Π_{BC} as DMU E has peers DMU B and C in the nominal DEA and similarly $\mathfrak{D}(F) = 4.33$ is obtained when DMU F is projected to Π_{AB} as DMU F has peers DMU A and B in the nominal DEA.

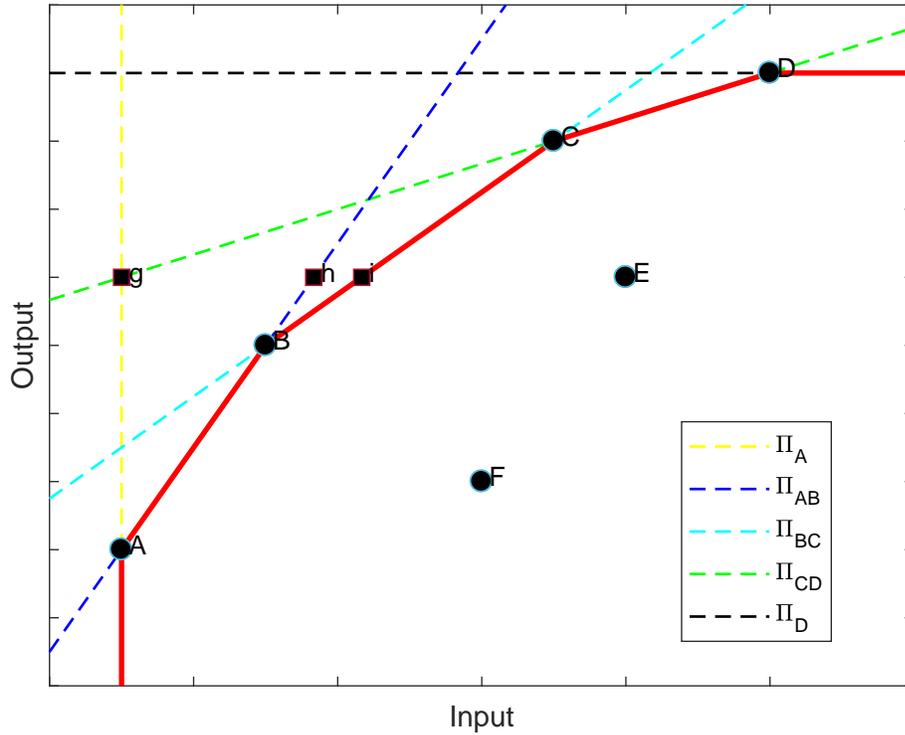


Figure 4.1: The lines that define the PPS, Example 3.

To calculate $\mathfrak{D}(k^u, \Pi^u)$, for inefficient DMUs E and F we transform the data according to Definitions 4.1.1 and 4.1.8. This gives the data in Table 4.7. The uncertain DEA distances can then be calculated as in Table 4.8. To find the minimum amount of uncertainty required for DMUs E and F to become efficient we find $\min\{u, s.t. \mathfrak{D}(k^u, \Pi^u) = 0, \Pi^u \in \Psi^u\}$. As a result, we calculate the value of u for each $\Pi^u \in \Psi^u$. These are shown in Table 4.9. We note that for Π_D in Table 4.8 we do not have a DEA distance but we do have a minimum amount of uncertainty for Π_D in Table 4.9. However, the amount of uncertainty for Π_D in Table 4.9 includes ϵ as we require $u_E > 1.5$ and $u_F > 3$. This will be explained later in Example 4.

Therefore, the minimum amount of uncertainty required for DMUs E and F to become efficient are $u_E = 0.79$ and $u_F = 1.21$. Both first become efficient on the line segment BC .

DMU	In	Out
A	1+u	1-u
B	3+u	4-u
C	7+u	7-u
D	10+u	8-u
E	8-u	5+u
F	6-u	2+u

Table 4.7: Uncertain data, Example 3.

$\Pi^u \in \Psi$	$\mathfrak{D}(E^u, \Pi^u)$	$\mathfrak{D}(F^u, \Pi^u)$
Π_A^u	$7-2u$	$5-2u$
Π_{AB}^u	$ \frac{13}{3} - \frac{10}{3}u $	$ \frac{13}{3} - \frac{10}{3}u $
Π_{BC}^u	$ \frac{11}{3} - \frac{14}{3}5u $	$ \frac{17}{3} - \frac{14}{3}u $
Π_{CD}^u	$ 7 - 8u $	$ 14 - 8u $
Π_D^u	-	-

Table 4.8: DEA Distance for DMUs E^u and F^u , Example 3.

$\Pi^u \in \Psi^u$	u_E	u_F
Π_A^u	3.50	2.50
Π_{AB}^u	1.30	1.30
Π_{BC}^u	0.79	1.21
Π_{CD}^u	0.88	1.75
Π_D^u	$1.50+\epsilon$	$3+\epsilon$

Table 4.9: Minimum amount of uncertainty of u such that DMUs E and F are efficient on Π^u , Example 3.

Let $s_m^- \geq 0$ ($s_n^+ \geq 0$) be the slack in the m^{th} (n^{th}) output (input) constraint in (4.1) so that (4.1b) and (4.1c) can be written:

$$\sum_{i=1}^I \lambda_i y_{mi} - y_{mk} - s_m^- = 0 \quad m = 1, \dots, M \quad (4.9a)$$

$$\sum_{i=1}^I \lambda_i x_{ni} - \theta_k x_{nk} + s_n^+ = 0 \quad n = 1, \dots, N. \quad (4.9b)$$

Proposition 4.1.11. *For an inefficient DMU k and for any optimal solution to (4.1) there exists at least one binding input constraint.*

Proof. Let (θ_k^*, λ^*) be an optimal solution for DMU k to (4.1). Then there is always at least one binding constraint. Assume this is an output constraint p and there are no binding input constraints. Then for each input n , $n = 1, \dots, N$, constraint (4.9b) can be written as:

$$\frac{\sum_{i=1}^I \lambda_i^* x_{ni}}{x_{nk}} + \frac{(s_n^+)^*}{x_{nk}} = \theta_k^*,$$

where all the $(s_n^+)^* > 0$. *W.l.o.g.* let input q have the largest value of $\frac{\sum_{i=1}^I \lambda_i^* x_{ni}}{x_{nk}}$.

Given that $x_{ni} > 0 \forall n, i$, then

$$\frac{\sum_{i=1}^I \lambda_i^* x_{qi}}{x_{qk}} \geq \frac{\sum_{i=1}^I \lambda_i^* x_{ni}}{x_{nk}} \quad \forall n \neq q \quad (4.10a)$$

$$\Leftrightarrow \frac{(s_q^+)^*}{x_{qk}} \leq \frac{(s_n^+)^*}{x_{nk}} \quad \forall n \neq q. \quad (4.10b)$$

That means we can reduce $(s_q^+)^*$ to 0 and reduce θ_k^* by $\frac{s_q^+}{x_{qk}}$. Then all remaining $\frac{(s_n^+)^*}{x_{nk}}$ must reduce by $\frac{(s_q^+)^*}{x_{qk}}$ which is possible from (4.10b). Because (4.1) is a minimisation problem, $(s_q^+)^* = 0$ is feasible, hence this is a contradiction and we must have at least one binding input constraint. \square

Therefore, from Proposition 4.1.11 there will always be at least one binding input constraint. From now on we define Q to be the set of all binding input constraints.

Select a binding input constraint q , then with uncertainty (4.9) becomes

$$\sum_{i=1}^I \lambda_i (y_{mi} + \Delta_{mi}) - (y_{mk} + \Delta_{mk}) - s_m^- = 0 \quad m = 1, \dots, M \quad (4.11a)$$

$$\sum_{i=1}^I \lambda_i (x_{qi} + \Delta_{ni}) - \theta_k (x_{qk} + \Delta_{nk}) = 0 \quad q \in Q \quad (4.11b)$$

$$\sum_{i=1}^I \lambda_i (x_{ni} + \Delta_{ni}) - \theta_k (x_{nk} + \Delta_{nk}) + s_n^+ = 0 \quad n = 1, \dots, N \notin Q.$$

Lemma 4.1.12. *For a binding input constraint q , if $\sum_{i=1}^I \lambda_i^* y_{mi} < \sum_{i=1}^I \bar{\lambda}_i^* y_{mi}$, then $\sum_{i=1}^I \lambda_i^* x_{qi} \leq \sum_{i=1}^I \bar{\lambda}_i^* x_{qi}$.*

Proof. Consider a decrease in a single output, m , of a single DMU r , $r \neq k$. If the output y_{mr} decreases by the maximum amount, $\Delta_{mr} = u$, at a new optimal solution we have

$$\sum_{\substack{i=1 \\ i \neq r}}^I \bar{\lambda}_i^* y_{mi} + \bar{\lambda}_r^* (y_{mr} - u) - y_{mk} + (s_m^-)^* = 0, \quad (4.12)$$

i.e. the LHS of (4.12) decreases by $\bar{\lambda}_r^* u$ compared to (4.1b).

If $\bar{\lambda}_r^* u \leq (s_m^-)^*$, then the slack in (4.12) decreases by $\bar{\lambda}_r^* u$ compared to (4.1b), so $(s_m^-)^* = (s_m^-)^* - \bar{\lambda}_r^* u$ and the decrease in output y_{mr} has no effect on the efficiency score. Hence, we concentrate on the case where $\bar{\lambda}_r^* u > (s_m^-)^*$.

If $\bar{\lambda}_r^* u > (s_m^-)^*$, then $(s_m^-)^* = 0$ and $\sum_{i=1}^I \lambda_i^* y_{mi}$ increases such that $\sum_{i=1}^I \bar{\lambda}_i^* y_{mi} = \sum_{i=1}^I \lambda_i^* y_{mi} + \bar{\lambda}_r^* u - (s_m^-)^*$. This means $\sum_{i=1}^I \bar{\lambda}_i^* y_{mi} > \sum_{i=1}^I \lambda_i^* y_{mi}$, i.e. the weighted sum of the outputs has increased compared to the nominal constraint in (4.1b).

Constraint q is binding, therefore any change in the value of λ will cause the weighted sum of the inputs to increase (or stay the same). In this way increasing the weighted sum of the outputs results in the weighted sum of the inputs increasing (or staying the same). \square

Furthermore, an increase in a single output of a single DMU can result in a decrease in efficiency score.

Lemma 4.1.13. *For a binding input constraint q , if $\sum_{i=1}^I \lambda_i^* y_{mi} > \sum_{i=1}^I \bar{\lambda}_i^* y_{mi}$ then $\sum_{i=1}^I \lambda_i^* x_{qi} \geq \sum_{i=1}^I \bar{\lambda}_i^* x_{qi}$.*

Proof. Analogous to Lemma 4.1.12. □

Theorem 4.1.14. *For box uncertainty defined in (4.4), there exists an optimal solution, $(\bar{\lambda}_k^*, \bar{\theta}_k^*, \Delta_{mi}^*)$, to (4.3) such that the data point the DMU picks is an extreme point of the uncertainty set, i.e.*

$$\Delta_{mi}^* = {}^+u, \quad m = 1, \dots, M$$

$$\Delta_{ni}^* = {}^+u, \quad n = 1, \dots, N.$$

Proof. This follows from well known results in linear and convex optimisation, (for example see Bertsimas and Tsitsiklis (1997)). The optimal solution of a linear function over a feasible region (convex polygon) will exist at an extreme point of the feasible region. The details of which can be found in Dantzig (1960) where an inductive proof of the simplex method is provided. This result is analogous for (3.19), that an optimal solution will exist at an extreme point of the uncertainty set. □

We use Theorem 4.1.14 to simplify the results in the rest of this chapter by focusing only on the extreme points of the box. Therefore, (4.3f) becomes $\Delta_{mi}, \Delta_{ni} \in \{-u, +u\}$ $m = 1, \dots, M, n = 1, \dots, N, i = 1, \dots, I$.

In the following theorem we use the definition of an incapable DMU as introduced in Ehrgott et al. (2018) and defined in Section 3.4, Definition 3.4.4.

Theorem 4.1.15. *For box uncertainty defined in (4.4) there can never be an incapable DMU when $\Omega \in \mathbb{R}^+$.*

Proof. Consider DMU k . If all the input constraints are binding then the DMU k is efficient and we are done. Otherwise select a non-binding input constraint h . Assume

DMU k is incapable i.e. there is no finite u such that k becomes efficient. Then,

$$\nexists u \text{ such that } \bar{\theta}_k^* = 1 \text{ and } \sum_{i=1}^I \lambda_i(x_{hi} + \Delta_{hi}) - \theta_k(x_{hk} + \Delta_{hk}) = 0, \Delta_{mi}, \Delta_{ni} \in \{-u, +u\}.$$

For DMU k to be efficient, we require $\theta_k = 1$. This occurs when

$$\begin{aligned} \sum_{i=1}^I \lambda_i(x_{hi} + \Delta_{hi}) &= x_{hk} + \Delta_{hk} \\ \Leftrightarrow x_{hk} - \sum_{i=1}^I \lambda_i x_{hi} &= \sum_{i=1}^I \lambda_i \Delta_{hi} - \Delta_{hk}. \end{aligned}$$

If $x_{hk} - \sum_{i=1}^I \lambda_i x_{hi} > 0$, set $u = \sum_{i=1}^I \lambda_i \Delta_{hi}$, $\Delta_{hk} = -u$ and if $x_{hk} - \sum_{i=1}^I \lambda_i x_{hi} < 0$, set $u = -\sum_{i=1}^I \lambda_i \Delta_{hi}$, $\Delta_{hk} = u$. Let $u = \frac{1}{2}(|x_{hk} - \sum_{i=1}^I \lambda_i x_{hi}|)$. Then $\bar{\theta}_k^* = 1$. Therefore, there exists u such that $\bar{\theta}_k^*$ equals one and $\sum_{i=1}^I \lambda_i(x_{hi} + \Delta_{hi}) - \theta_k(x_{hk} + \Delta_{hk}) = 0$ and we can never have an incapable DMU. \square

An alternative proof to this can be found by adapting the proof of Theorem 4 in Cooper et al. (2001), that there is always a finite optimum value that establishes a radius of stability.

Theorem 4.1.16. *For inefficient DMUs k and l with box uncertainty defined in (4.4), there exists an optimal solution to (4.2) for DMU k such that $\lambda_l = 0$.*

Proof. Let DMU k be the inefficient DMU under consideration. DMU l is inefficient with $\theta_l^* < 1$. Let u' be an amount of uncertainty such that DMU l becomes a peer to DMU k . Then DMU l must be efficient, $\bar{\theta}_l^* = 1$.

The value of θ_l can increase to $\bar{\theta}_l^* = 1$ by increasing its outputs and decreasing its inputs or by efficient DMUs increasing their inputs and decreasing their outputs. Then DMU l will become part of an existing facet of the efficient frontier or it will be a new extreme point of the efficient frontier.

If DMU l becomes part of an existing facet of the efficient frontier, it can be written as a convex combination of other efficient DMUs and hence DMU k 's efficiency

score does not improve. A new extreme point will lie outside the nominal DEA PPS. Therefore any facet defined by a new extreme point will be further away (or the same) DEA distance from DMU k than $\mathfrak{D}(k)$, the nominal DEA minimum DEA distance. Therefore, it is not beneficial for DMU k to be compared to the facets containing the new extreme point and solving the uDEA model for DMU k will select the original position for DMU l . Therefore, there is an optimal solution for DMU k such that $\lambda_l = 0$. \square

In this way, we only need to consider DMUs that are efficient in the nominal DEA problem (4.1).

Corollary 4.1.17. *For DMUs with $\theta_i^* = 1$, for any inefficient DMU k with box uncertainty defined in (4.4), there exists an optimal solution to (4.2) for DMU k such that $\bar{\theta}_i^* = 1$.*

Proof. This follows from Theorem 4.1.16. \square

Consequently, DMUs that are efficient in the nominal problem (4.1) will remain efficient at least until the point when DMU k becomes efficient when uncertainty is introduced.

4.2 Fixed Uncertainty

For any DMU k , solving (4.1) gives an optimal solution vector (θ_k^*, λ^*) . Where θ_k^* is the efficiency score for DMU k with the nominal data. We are interested in the change in efficiency score as small changes (i.e. uncertainty) are introduced into the dataset. Here we consider a fixed amount of uncertainty i.e. in (4.3) $\Delta_{mi}, \Delta_{ni} \in \{-u, +u\} \forall m, n$.

In Section 4.2.1, we consider the introduction of a fixed uncertainty amount to a single input or output. In Sections 4.2.2 and 4.2.3, simultaneous changes to all inputs

and outputs with fixed uncertainty are explored. The overall best improvement in efficiency score for fixed uncertainty is given in Section 4.2.4.

From Theorem 4.1.16, for inefficient DMUs k and l , there exists an optimal solution to (4.3) for DMU k such that $\lambda_l = 0$. Therefore, we can assume inefficient DMUs do not become peers to other inefficient DMUs. Consequently, in the following sections we assume that from the I DMUs, we have $I - 1$ efficient DMUs and a single inefficient DMU, DMU k . Then $\theta_k^* < 1$ and $\theta_i^* = 1 \forall i = 1, \dots, I, i \neq k$.

In the rest of this section we show that for fixed uncertainty u the following theorem holds.

Theorem 4.2.1. *To maximise the possible increase in efficiency score for inefficient DMU k , solving the linear program (4.3) will result in the following uncertainty being selected:*

$$\begin{aligned} \Delta_{mi} &= -u, \quad \Delta_{mk} = +u & i \neq k, \quad m = 1, \dots, M \\ \Delta_{ni} &= +u, \quad \Delta_{nk} = -u & i \neq k, \quad n = 1, \dots, N. \end{aligned}$$

This is similar to the idea introduced in DEA sensitivity analysis in Charnes et al. (1996) that for inefficient DMUs the northwest corner of a DMU's uncertainty box will "dominate every other member of the cell". We follow a similar process to earlier DEA sensitivity analysis papers where we first consider a single change to a DMU and then extend to simultaneous changes in the dataset.

4.2.1 Single changes in DMUs

Consider a single change in the data, i.e. a single x_{ni} or y_{mi} changes by $\pm u$ while the remaining data are fixed. Let DMU $r(k)$ be the efficient(inefficient) DMU under consideration. The change in efficiency score will be largest when we have a binding input constraint. This can be seen in Lemma 4.1.12 where the role of the slack variables and the possible changes in the weighted sum of the inputs and outputs are

demonstrated. So from now on we consider only binding input constraints. When choosing a binding input constraint q the possible scenarios are:

- Increase an input of an efficient DMU r by u ,

$$\Delta_{qr} = +u, \Delta_{ni} = \Delta_{mi} = 0, i \neq r, n \neq q, \forall m$$
- Decrease an input of an efficient DMU r by u ,

$$\Delta_{qr} = -u, \Delta_{ni} = \Delta_{mi} = 0, i \neq r, n \neq q, \forall m$$
- Increase an output of an efficient DMU r by u ,

$$\Delta_{m'r} = +u, \Delta_{ni} = \Delta_{mi} = 0, i \neq r, m \neq m', \forall n$$
- Decrease an output of an efficient DMU r by u ,

$$\Delta_{m'r} = -u, \Delta_{ni} = \Delta_{mi} = 0, i \neq r, m \neq m', \forall n$$
- Increase an input of the inefficient DMU k by u ,

$$\Delta_{qk} = +u, \Delta_{ni} = \Delta_{mi} = 0, i \neq k, n \neq q, \forall m$$
- Decrease an input of the inefficient DMU k by u ,

$$\Delta_{qk} = -u, \Delta_{ni} = \Delta_{mi} = 0, i \neq k, n \neq q, \forall m$$
- Increase an output of the inefficient DMU k by u ,

$$\Delta_{m'k} = +u, \Delta_{ni} = \Delta_{mi} = 0, i \neq k, m \neq m', \forall n$$
- Decrease an output of the inefficient DMU k by u ,

$$\Delta_{m'k} = -u, \Delta_{ni} = \Delta_{mi} = 0, i \neq k, m \neq m', \forall n$$

Change in inputs for efficient DMUs

Lemma 4.2.2. *For a single input of an efficient DMU r changed by u , the maximum increase in θ_k^* that can occur is $\frac{u}{x_{nk}}$.*

Proof. Consider an increase in an efficient DMU r 's input, q . Solving (4.1) gives an optimal solution (θ_k^*, λ^*) , when there is no uncertainty in the data.

For binding constraint q the LHS of (4.3c) increases by $\bar{\lambda}_r^* u$ compared to (4.1c) and $\sum_{i=1}^I \lambda_i^* x_{qi}$ must decrease or $\theta_k^* x_{nk}$ must increase. However, $\sum_{i=1}^I \lambda_i^* x_{qi}$ cannot decrease because the constraint for input q is binding so $\sum_{i=1}^I \lambda_i^* x_{qi}$ is already as small as possible. This means $\sum_{i=1}^I \bar{\lambda}_i^* x_{qi} = \sum_{i=1}^I \lambda_i^* x_{qi}$. DMU k 's data are fixed therefore, θ_k^* must increase by

$$\begin{aligned} \sum_{\substack{i=1 \\ i \neq r}}^I \bar{\lambda}_i^* x_{qi} + \bar{\lambda}_r^* (x_{qr} + u) - \bar{\theta}_k^* x_{qk} &= \sum_{i=1}^I \lambda_i^* x_{qi} - \theta_k^* x_{qk} \\ \bar{\theta}_k^* - \theta_k^* &= \frac{\sum_{i=1}^I \bar{\lambda}_i^* x_{qi} - \sum_{i=1}^I \lambda_i^* x_{qi} + \bar{\lambda}_r^* u}{x_{qk}} \\ &= \frac{\bar{\lambda}_r^* u}{x_{qk}}. \end{aligned}$$

Therefore, θ_k^* must increase by $\frac{\bar{\lambda}_r^* u}{x_{qk}}$. This is at its maximum when $\bar{\lambda}_r^* = 1$ and θ_k^* increases by $\frac{u}{x_{qk}}$. Similarly, if a single input r decreases by u , θ_k^* must stay the same or decrease by a maximum of $\frac{\bar{\lambda}_r^* u}{x_{qk}}$.

Therefore, the maximum increase in θ_k^* that can occur when a single input of an efficient DMU r is changed by u occurs when the input is increased by u . \square

Change in outputs for efficient DMUs

Consider a decrease in a single output m of an efficient DMU r . When we considered the change in single inputs it was clear the change could affect θ_k^* as all the input constraints (4.1c) involve the term $\theta_k x_{nk}$ in the nominal problem, (4.1). Although θ_k does not appear directly in the output constraints in (4.1b), changing them can still affect the value of θ_k^* .

Lemma 4.2.3. *For a single output of an efficient DMU r changed by u , the maximum increase in θ_k^* that can occur is*

$$\max_{q \in Q} \frac{\max_{i \neq k} x_{qi} - \min_{i \neq k} x_{qi}}{x_{qk}}. \quad (4.13)$$

Proof. This follows from Lemmas 4.1.12 and 4.1.13. In Lemma 4.1.12 an increase in the weighted sum of the outputs compared to the nominal constraint in (4.1b) results in the weighted sum of the inputs increasing (or staying the same). The maximum increase that can occur in the weighted sum of the inputs occurs when $\lambda_{\hat{i}} = 0$ for \hat{i} s.t. $x_{q\hat{i}} = \max_{i \neq k} x_{qi}$ in the nominal solution and $\lambda_{\hat{i}} = 1$ in the solution of the uncertain problem and $\lambda_{\bar{i}} = 1$ for \bar{i} s.t. $x_{q\bar{i}} = \lambda_i \min_{i \neq k} x_{qi}$ in the nominal and $\lambda_{\bar{i}=0}$ in the uncertain model. This results in a change in efficiency score given by (4.13). \square

Change in inputs for the inefficient DMU

We now assume that the efficient DMUs' data are fixed and consider what occurs if a single change is made to the inefficient DMU k 's data.

Lemma 4.2.4. *For a single input of the inefficient DMU k changed by u , the maximum increase in θ_k^* that can occur is $\frac{u}{x_{nk}}$.*

Proof. Choose a binding input constraint, q . If the input value for DMU k , x_{qk} is increased, $\Delta_{nk} = +u$, the LHS of (4.3c) will decrease by $\bar{\theta}_k^* u$ compared to the nominal value in (4.1c). If there are no other binding constraints $\bar{\theta}_k^*$ will reduce until a binding constraint occurs. If there is another binding constraint, input constraint q will cease to be binding and the slack for constraint q will be greater than zero, $(s_q^+)^* > 0$ and the value of the efficiency score will decrease, $\bar{\theta}_k^* \leq \theta_k^*$.

If the input value for DMU k , x_{qk} is decreased, $\Delta_{nk} = -u$, the LHS of (4.3c) will increase by $\bar{\theta}_k^* u$ compared to the nominal value in (4.1c). If there exists another binding constraint the optimal weights cannot change, $\sum_{i=1}^I \lambda_i^* x_{qi} = \sum_{i=1}^I \bar{\lambda}_i^* x_{qi}$, and the efficiency score for DMU k must increase,

$$\sum_{i=1}^I \lambda_i^* x_{qi} - \theta_k^* x_{qk} = \sum_{i=1}^I \bar{\lambda}_i^* x_{qi} - \bar{\theta}_k^* (x_{qk} - u)$$

$$\bar{\theta}_k^* - \theta_k^* = \frac{u \bar{\theta}_k^*}{x_{qk}}.$$

If input constraint q is the only binding constraint the change in efficiency score above is an upper bound. Therefore, the maximum change in efficiency score occurs when the input is reduced by u . This results in a change in efficiency score of $\frac{u\bar{\theta}_k^*}{x_{qk}} \leq \frac{u}{x_{qk}}$, because $\bar{\theta}_k^* \leq 1$. \square

Change in outputs for the inefficient DMU

As noted, when single changes in efficient DMUs are considered, although θ_k does not appear directly in the output constraints (4.1b), changing them can still affect the value of θ_k^* . For this case, we have the following two Lemmas.

Lemma 4.2.5. *If $\Delta_{mk} = +u$, $\bar{\theta}_k^* \geq \theta_k$.*

Lemma 4.2.6. *If $\Delta_{mk} = -u$, $\bar{\theta}_k^* \leq \theta_k$.*

The proofs of these are very similar to the proof of Lemma 4.1.12.

Therefore, from Lemmas 4.2.2, 4.2.3, 4.2.4 4.2.5 we conclude the following result:

Result 4.2.7. *An inefficient DMU k 's efficiency score will not decrease if any or all of the following occur: DMU k 's inputs decrease or outputs increase or an efficient DMU r 's inputs increase or outputs decrease.*

4.2.2 Multiple changes in inputs and outputs

Now consider multiple changes in the inputs and outputs of the efficient(inefficient) DMUs with the inefficient(efficient) DMU's data fixed. From Result 4.2.7 in Section 4.2.1 we consider only changes that increase(decrease) the efficient DMUs' inputs(outputs) and decrease(increase) the inefficient DMU's inputs(outputs).

Proposition 4.2.8. *The maximum increase in DMU k 's efficiency score when all efficient DMU r 's data can change is*

$$\max_{q \in Q} \frac{\max_{i \neq k} x_{qi} - \min_{i \neq k} x_{qi} + u}{x_{qk}}.$$

Proof. The largest changes occur when there is a binding input and output constraint. Assume the constraint for output p is binding. Then the LHS of (4.11a) decreases by u compared to (4.1b). Therefore,

$$\sum_{i=1}^I \bar{\lambda}_i^* y_{pi} \geq \sum_{i=1}^I \lambda_i^* y_{pi}, \quad (4.14)$$

and from Lemma 4.1.13 the weighted sum of the inputs can increase. The LHS of (4.11b) increases by u compared to (4.1c) so $\sum_{i=1}^I \bar{\lambda}_i^* y_{pi} < \sum_{i=1}^I \lambda_i^* y_{pi}$ or $\bar{\theta}_k^* > \theta_k^*$. From (4.14) $\sum_{i=1}^I \bar{\lambda}_i^* y_{pi} < \sum_{i=1}^I \lambda_i^* y_{pi}$ cannot occur. Hence the efficiency score must increase by

$$\frac{\sum_{i=1}^I \bar{\lambda}_i^* x_{qi} - \sum_{i=1}^I \lambda_i^* x_{qi} + u}{x_{qk}}.$$

The maximum value this can take occurs when $\sum_{i=1}^I \bar{\lambda}_i^* x_{qi}$ is as large as possible and $\sum_{i=1}^I \lambda_i^* x_{qi}$ is as small as possible, i.e. the maximum increase in efficiency score when changing efficient DMU r is

$$\max_{q \in Q} \frac{\max_{i \neq k} x_{qi} - \min_{i \neq k} x_{qi} + u}{x_{qk}}.$$

□

Changing all efficient DMU r 's data in Proposition 4.2.8, results in a maximum increase in DMU k 's efficiency score, which is larger than changing a single input of DMU r as in Lemma 4.2.2.

From Result 4.2.7, for the inefficient DMU we decrease the inputs and increase the outputs, $\Delta_{nk} = -u$, $\Delta_{mk} = u$ where Δ_{mi} , $\Delta_{ni} = 0$ $i \neq k$.

Proposition 4.2.9. *The maximum increase in efficiency score when all inefficient DMU k 's data changes is*

$$\max_{q \in Q} \frac{\max_{i \neq k} x_{qi} - \min_{i \neq k} x_{qi} + u}{x_{qk}}. \quad (4.15)$$

Proof. Consider the p^{th} output constraint

$$\sum_{i=1}^I \lambda_i y_{pi} - (y_{pk} + u) - s_m^- = 0. \quad (4.16)$$

If $u < s_m^-$ the LHS of (4.16) decreases by u compared to (4.11a), so $(s_m^-)^* = (s_m^-)^* + u$ and the increase in output y_{mk} has no effect on the efficiency score, $\bar{\theta}_k^* = \theta_k^*$.

If $u > s_m^-$ then $(s_m^-)^* = 0$ and $\sum_{i=1}^I \lambda_i^* y_{mi}$ must increase by $u - (s_m^-)^*$. As $s_m^- \leq 0$ the largest increase occurs when $(s_m^-)^* = 0$, i.e. output constraint p is binding. We seek the maximum increase in efficiency score, so, if it exists we select a binding output constraint p . If no binding output constraint exists (4.15) is an upper bound for the increase in efficiency score. From Lemma 4.1.12 this means $\sum_{i=1}^I \lambda_i^* x_{ni}$ must increase or stay the same, $\sum_{i=1}^I \bar{\lambda}_i^* x_{qi} \geq \sum_{i=1}^I \lambda_i^* x_{qi}$. Select a binding input constraint q . There will always be at least one binding input constraint from Proposition 4.1.11. Consider multiple changes with $s_p^- = 0$, $\Delta_{mk} = +u$ and $\Delta_{nk} = -u$ in (4.11). The LHS of (4.11b) increases by $\sum_{i=1}^I \bar{\lambda}_i^* x_{qi} - \sum_{i=1}^I \lambda_i^* x_{qi} + u\theta_k^*$ compared to (4.1c). Then we have,

$$\begin{aligned} \sum_{i=1}^I \bar{\lambda}_i^* x_{qi} - \bar{\theta}_k^* (x_{qk} - u) &= \sum_{i=1}^I \lambda_i^* x_{qi} - \theta_k^* x_{qk} \\ \bar{\theta}_k^* - \theta_k^* &= \frac{\sum_{i=1}^I \bar{\lambda}_i^* x_{qi} - \sum_{i=1}^I \lambda_i^* x_{qi} + u\bar{\theta}_k^*}{x_{qk}}. \end{aligned}$$

The maximum value this can take occurs when $\sum_{i=1}^I \bar{\lambda}_i^* x_{qi}$ is as large as possible and $\sum_{i=1}^I \lambda_i^* x_{qi}$ is as small as possible, i.e. the maximum increase in efficiency score when DMU k 's data changes is

$$\max_{q \in Q} \frac{\max_{i \neq k} x_{qi} - \min_{i \neq k} x_{qi} + u}{x_{qk}}.$$

□

Changing all inefficient DMU k 's data in Proposition 4.2.9, results in a maximum increase in DMU k 's efficiency score, which is not smaller than changing a single input of DMU k as in Lemma 4.2.4.

4.2.3 Inefficient and efficient DMUs' data can change

From Section 4.2.2, by changing only the efficient DMUs data the most improvement to DMU k 's efficiency score is achieved by increasing the inputs and decreasing the outputs. By changing only the inefficient DMU's data the most improvement to DMU k 's efficiency score is achieved by decreasing the inputs and increasing the outputs. We now consider the case where both inefficient and efficient DMUs data can change.

First restrict the changes to inputs or outputs only. The possible scenarios are:

1. Increase all inputs, $\Delta_{ni} = +u$
2. Increase all outputs, $\Delta_{mi} = +u$
3. Decrease all inputs, $\Delta_{ni} = -u$
4. Decrease all outputs, $\Delta_{mi} = -u$
5. Increase(decrease) efficient(inefficient) DMUs' inputs,
 $\Delta_{ni} = +u, \Delta_{nk} = -u, i \neq k$
6. Decrease(increase) efficient(inefficient) DMUs' inputs,
 $\Delta_{ni} = -u, \Delta_{nk} = +u, i \neq k$
7. Increase(decrease) efficient(inefficient) DMUs' outputs,
 $\Delta_{mi} = +u, \Delta_{mk} = -u, i \neq k$
8. Decrease(increase) efficient(inefficient) DMUs' outputs,
 $\Delta_{mi} = -u, \Delta_{mk} = +u, i \neq k$

However, from Result 4.2.7, we only want to consider scenarios that increase DMU k 's efficiency score. Therefore, we only need to consider scenario five and eight in the above list.

Proposition 4.2.10. *The maximum increase in efficiency score of an inefficient DMU k when changing the inputs of all DMUs is*

$$\max_{q \in Q} \frac{2u}{x_{qk}}.$$

Proof. From Result 4.2.7, consider the change in efficiency score if the efficient(inefficient) DMUs' inputs increase(decrease) i.e. $\Delta_{ni} = +u$, $\Delta_{nk} = -u$, $i \neq k$. At an optimal solution (4.1c) becomes,

$$\sum_{i=1}^I \bar{\lambda}_i^*(x_{qi} + u) - \bar{\theta}_k^*(x_{qk} - u) = 0. \quad (4.17)$$

The LHS of (4.17) has increased by $u(1 + \bar{\theta}_k^*)$ compared to (4.1c) so $\sum_{i=1}^I \bar{\lambda}_i^* x_{qi} - u(1 + \bar{\theta}_k^*) = \sum_{i=1}^I \lambda_i^* x_{qi}$ or $\bar{\theta}_k^* x_{qk} = \theta_k^* x_{qk} + u(1 + \bar{\theta}_k^*)$. The q^{th} input constraint is binding so $\sum_{i=1}^I \bar{\lambda}_i^* x_{qi} = \sum_{i=1}^I \lambda_i^* x_{qi}$. Then the change in efficiency score can be calculated,

$$\begin{aligned} \sum_{i=1}^I \bar{\lambda}_i^*(x_{qi} + u) - \bar{\theta}_k^*(x_{qk} - u) &= \sum_{i=1}^I \lambda_i^* x_{qi} - \theta_k^* x_{qk} \\ \bar{\theta}_k^* - \theta_k^* &= \frac{u(1 + \bar{\theta}_k^*)}{x_{qk}}. \end{aligned}$$

The maximum value $\bar{\theta}_k^*$ can take is one. Therefore the maximum increase in efficiency score of inefficient DMU k when changing the inputs of all DMUs is

$$\max_{q \in Q} \frac{2u}{x_{qk}}.$$

□

Similarly for the outputs we have the following proposition. The proof is similar to the proof of Proposition 4.2.10 and hence is omitted.

Proposition 4.2.11. *The maximum increase in efficiency score of inefficient DMU k when changing the outputs of all DMUs is*

$$\max_{q \in Q} \frac{\max_{i \neq k} x_{qi} - \min_{i \neq k} x_{qi}}{x_{qk}}.$$

We now consider the potential improvement in efficiency score if both the inputs and outputs can change.

Proposition 4.2.12. *The maximum increase in efficiency score when changing all DMUs is*

$$\max_{q \in Q} \frac{\max_{i \neq k} x_{qi} - \min_{i \neq k} x_{qi} + 2u}{x_{qk}}.$$

Proof. From Result 4.2.7 changing the outputs, $\Delta_{mi} = -u$, $\Delta_{mk} = +u$, $i \neq k$, results in the efficiency score increasing or remaining the same. When the inputs change, the largest increase in efficiency score occurs when $\Delta_{ni} = +u$, $\Delta_{nk} = -u$, $i \neq k$. From Proposition 4.1.11, at least one of the input constraints must be binding. Choose a binding input constraint, q . If a binding output constraint p exists, (4.11) becomes,

$$\sum_{\substack{i=1 \\ i \neq k}}^I \lambda_i (y_{pi} - u) - (y_{pk} + u) = 0 \quad (4.18a)$$

$$\sum_{\substack{i=1 \\ i \neq k}}^I \lambda_i (x_{qi} + u) - \theta_k (x_{qk} - u) = 0. \quad (4.18b)$$

At an optimal solution the LHS of (4.18a) decreases by $2u$ compared to (4.1b) so $\sum_{i=1}^I \bar{\lambda}_i^* y_{pi} = \sum_{i=1}^I \lambda_i^* y_{pi} + 2u$. This means $\sum_{i=1}^I \bar{\lambda}_i^* x_{qi} \geq \sum_{i=1}^I \lambda_i^* x_{qi}$. Additionally, the LHS of (4.18b) increases by $u(1 + \bar{\theta}_k^*)$ compared to (4.1c). Therefore, the overall change in efficiency score can be calculated,

$$\begin{aligned} \sum_{i=1}^I \bar{\lambda}_i^* (x_{qi} + u) - \bar{\theta}_k^* (x_{qk} - u) &= \sum_{i=1}^I \lambda_i^* x_{qi} - \theta_k^* x_{qk} \\ \bar{\theta}_k^* - \theta_k^* &= \frac{\sum_{i=1}^I \bar{\lambda}_i^* x_{qi} - \sum_{i=1}^I \lambda_i^* x_{qi} + u + \bar{\theta}_k^* u}{x_{qk}}. \end{aligned}$$

The maximum change in efficiency score occurs when $\sum_{i=1}^I \bar{\lambda}_i^* x_{qi}$ and $\bar{\theta}_k^*$ are as large as possible and $\sum_{i=1}^I \lambda_i^* x_{qi}$ is as small as possible, i.e. the maximum increase in efficiency score when changing all DMUs is

$$\max_{q \in Q} \frac{\max_{i \neq k} x_{qi} - \min_{i \neq k} x_{qi} + 2u}{x_{qk}}.$$

□

4.2.4 Overall best improvement in efficiency score

Finally, we examine which of the previous scenarios, individual and multiple changes for both inefficient and efficient DMUs, result in the most improvement in DMU k 's efficiency score. Table 4.10 summarises the results from Section 4.2. From Table 4.10 we conclude that the maximum improvement to DMU k 's efficiency score is achieved when the inputs and outputs of all the DMUs data can change with $\Delta_{i \neq k}^{ni} = +u$, $\Delta_{i \neq k}^{mi} = -u$, $\Delta_{nk} = -u$, $\Delta_{mk} = +u$. In particular, for DMU k 's efficiency score to improve (4.3) becomes,

$$\min_{\theta_k, \lambda} \theta_k \quad (4.19a)$$

$$s.t. \quad \sum_{i=1}^I \lambda_i (y_{mi} - u) - (y_{mk} + u) \geq 0 \quad m = 1, \dots, M \quad (4.19b)$$

$$\sum_{i=1}^I \lambda_i (x_{ni} + u) - \theta_k (x_{nk} - u) \leq 0 \quad n = 1, \dots, N \quad (4.19c)$$

$$\sum_{i=1}^I \lambda_i = 1 \quad (4.19d)$$

$$\lambda_i \geq 0 \quad i = 1, \dots, I \quad (4.19e)$$

$$u \geq 0. \quad (4.19f)$$

This completes the proof of Theorem 4.2.1 and gives the following result.

Result 4.2.13. *For fixed uncertainty there is always an optimal solution in which the inefficient DMU benefits from increased outputs and decreased inputs and the efficient DMUs' inputs increase and outputs decrease.*

Scenario	Maximum Increase $\bar{\theta}_k^* - \theta_k^*$	Corresponding result
Increase efficient DMU input by u	$\frac{u}{x_{nk}}$	Lemma 4.2.2
Reduce efficient DMU output by u	$\max_{q \in Q} \frac{\max_{i \neq k} x_{qi} - \min_{i \neq k} x_{qi}}{x_{qk}}$	Lemma 4.2.3
Reduce inefficient DMU input by u	$\frac{u}{x_{nk}}$	Lemma 4.2.4
Increase inefficient DMU output by u	$\max_{q \in Q} \frac{\max_{i \neq k} x_{qi} - \min_{i \neq k} x_{qi}}{x_{qk}}$	Similar to Lemma 4.2.3
All efficient DMUs data can change	$\max_{q \in Q} \frac{\max_{i \neq k} x_{qi} - \min_{i \neq k} x_{qi} + u}{x_{qk}}$	Proposition 4.2.8
All inefficient DMUs data can change	$\max_{q \in Q} \frac{\max_{i \neq k} x_{qi} - \min_{i \neq k} x_{qi} + u}{x_{qk}}$	Proposition 4.2.9
All inputs can change	$\max_{q \in Q} \frac{2u}{x_{qk}}$	Proposition 4.2.10
All outputs can change	$\max_{q \in Q} \frac{\max_{i \neq k} x_{qi} - \min_{i \neq k} x_{qi}}{x_{qk}}$	Proposition 4.2.11
All DMUs can change	$\max_{q \in Q} \frac{\max_{i \neq k} x_{qi} - \min_{i \neq k} x_{qi} + 2u}{x_{qk}}$	Proposition 4.2.12

Table 4.10: The maximum possible changes to the efficiency score of DMU k .

4.3 Changing value of u

From Theorem 4.2.1, to maximise the possible increase in efficiency score for inefficient DMU k , the linear program (4.3) becomes (4.19). We now consider the case where we have a changing value of u . What is the minimum amount of u required such that inefficient DMU k becomes efficient? Ehrgott et al. (2018) prove that as the amount of uncertainty increases the efficiency score must increase. In the case of box uncertainty this is very intuitive.

Theorem 4.3.1. *For inefficient DMU k with uncertainty defined by (4.4), increasing the value of u increases the efficiency score.*

Proof. Solving (4.19) for fixed u gives $\theta_k = \bar{\theta}_k^*$. Consider an increase in u such that $\hat{u} = u + \epsilon$. Let $\hat{\theta}_k$ be the new efficiency score when \hat{u} and $\hat{\lambda}$ are the corresponding weights. The constraints (4.19b) and (4.19c) become,

$$\sum_{i=1}^I \hat{\lambda}_i (y_{mi} - u - \epsilon) - (y_{mk} + u + \epsilon) \geq 0 \quad m = 1, \dots, M \quad (4.20a)$$

$$\sum_{i=1}^I \hat{\lambda}_i (x_{ni} + u + \epsilon) - \hat{\theta}_k (x_{nk} - u - \epsilon) \leq 0 \quad n = 1, \dots, N. \quad (4.20b)$$

Select a binding input constraint for input q . At optimality the LHS of (4.20b) has increased by $\epsilon(1 + \hat{\theta}_k)$ compared to (4.19c). Consequently, $\sum_{i=1}^I \bar{\lambda}_i^*(x_{qi} + u) > \sum_{i=1}^I \hat{\lambda}_i(x_{qi} + u)$ or $\bar{\theta}_k^*(x_{qk} - u) < \hat{\theta}_k(x_{qk} - u)$. However, $\sum_{i=1}^I \bar{\lambda}_i^*(x_{qi} + u)$ cannot decrease because the constraint for input q is binding so, $\sum_{i=1}^I \bar{\lambda}_i^*(x_{qi} + u)$ is already as small as possible and $\sum_{i=1}^I \bar{\lambda}_i^*(x_{qi} + u) = \sum_{i=1}^I \hat{\lambda}_i(x_{qi} + u)$. Then the introduction of $\hat{u} = u + \epsilon$ results in

$$\begin{aligned} \sum_{i=1}^I \bar{\lambda}_i^*(x_{qi} + u) - \bar{\theta}_k^*(x_{qk} - u) &= \sum_{i=1}^I \hat{\lambda}_i(x_{qi} + u + \epsilon) - \hat{\theta}_k(x_{qk} - u - \epsilon) \\ -\bar{\theta}_k^*(x_{qk} - u) &= \epsilon - \hat{\theta}_k(x_{qk} - u - \epsilon) \\ \hat{\theta}_k - \bar{\theta}_k^* &= \frac{\epsilon(1 + \hat{\theta}_k)}{x_{qk} - u}. \end{aligned}$$

Then, $x_{qk} - u > 0$ due to the assumption of strictly positive data. As a result $\frac{\epsilon(1+\hat{\theta}_k)}{x_{qk}-u} > 0$, $\hat{\theta}_k - \bar{\theta}_k^* > 0$ and the efficiency has increased. \square

In this way θ_k^* , the efficiency score when there is no uncertainty, provides a lower bound for $\bar{\theta}_k^*$, DMU k 's efficiency score when uncertainty is present.

4.3.1 Value of u for $N + M$ dimensions

When considering the improvement DMU k can achieve we wish to calculate the smallest amount of uncertainty to deem DMU k efficient. We must calculate the DEA distance from DMU k to the efficient frontier. We consider the input-oriented envelopment DEA model with VRS and hence the outputs are fixed. Therefore we must project the inefficient DMUs to the efficient frontier while the inefficient DMUs outputs' remain fixed. This distance will be greater than or equal to the Euclidean distance from the inefficient DMU to the efficient frontier.

Result 4.3.2. *The DEA distance, $\mathfrak{D}(k, \Pi_\Phi)$, from an inefficient DMU k to the efficient frontier facet, Π_Φ defined by (4.5) is*

$$\mathfrak{D}(k, \Pi_\Phi) = \frac{|\alpha_1 x_{1k} + \cdots + \alpha_N x_{Nk} + \beta_1 y_{1k} + \cdots + \beta_M y_{Mk} - \gamma|}{\sqrt{\alpha_1^2 + \cdots + \alpha_N^2}}. \quad (4.21)$$

Proof. Let the facet of the efficient frontier inefficient DMU k is projected to be Π_Φ . The inefficient DMU k has co-ordinates $(x_{1k}, \dots, x_{Nk}, y_{1k}, \dots, y_{Mk})$ and does not lie on the hyperplane Π_Φ . Consider these hyperplanes:

$$\Pi_{y1} : y_1 = y_{1k}$$

$$\Pi_{y2} : y_2 = y_{2k}$$

$$\vdots$$

$$\Pi_{yM} : y_M = y_{Mk}$$

Let Π_N be the plane where Π_Φ intersects with $\Pi_{y_1}, \Pi_{y_2}, \dots, \Pi_{y_M}$. Then

$$\Pi_N = \{r \in \mathbb{R}^N : r \bullet \vec{n}_2 = \gamma - \beta_1 y_{1k} - \dots - \beta_M y_{Mk}\},$$

where $\vec{n}_2 = (\alpha_1, \dots, \alpha_N, 0, \dots, 0)$. Then the DEA distance from DMU k to the hyperplane Π_Φ is equivalent to finding the Euclidean distance from DMU k to the plane Π_N . From (4.7), this is

$$\mathfrak{D}(k, \Pi_\Phi) = \frac{|\alpha_1 x_{1k} + \dots + \alpha_N x_{Nk} + \beta_1 y_{1k} + \dots + \beta_M y_{Mk} - \gamma|}{\sqrt{\alpha_1^2 + \dots + \alpha_N^2}}.$$

□

Theorem 4.3.3. *For an inefficient DMU k , the minimum amount of uncertainty required to be projected to the target point $T(k, \Pi_\Phi)$ on Π_Φ as defined in (4.5) is*

$$u = \frac{|\alpha_1 x_{1k} + \dots + \alpha_N x_{Nk} + \beta_1 y_{1k} + \dots + \beta_M y_{Mk} - \gamma|}{2|\alpha_1 - \dots - \alpha_N + \beta_1 + \dots + \beta_M|}. \quad (4.22)$$

Proof. The DEA distance from the point k to the hyperplane Π_Φ is given by (4.21).

From Theorem 4.2.1 we choose the following data instance $(x_{1k} - u, \dots, x_{Nk} - u, y_{1k} + u, \dots, y_{Mk} + u)$, and $(x_{1i} + u, \dots, x_{Ni} + u, y_{1i} - u, \dots, y_{Mi} - u)$ $i = 1, \dots, I$, $i \neq k$ and denote it k^u . Let Π_Φ^u be the hyperplane that goes through the virtual DMUs $i = 1, \dots, I$, $i \neq k$ with uncertainty defined by Theorem 4.2.1. Π_Φ^u is parallel to Π_Φ because all the inputs have increased by u and all the outputs have decreased by u . As a result the normal to the hyperplane will be the same.

$$\Pi_\Phi^u = \{r \in \mathbb{R}^\Phi : r \bullet \vec{n}_1 = \gamma^u\},$$

where $\gamma^u = \gamma + u(\alpha_1 + \alpha_2 + \dots + \alpha_N - \beta_1 - \beta_2 - \dots - \beta_M)$.

Consider these hyperplanes:

$$\begin{aligned}\Pi_{y_1}^u &: y_1 = y_{1k} + u \\ \Pi_{y_2}^u &: y_2 = y_{2k} + u \\ &\vdots \\ \Pi_{y_M}^u &: y_M = y_{Mk} + u.\end{aligned}$$

Let Π_N^u be the plane where Π_Φ^u intersects with $\Pi_{y_1}^u, \Pi_{y_2}^u, \dots, \Pi_{y_M}^u$ then

$$\Pi_N^u = \{r \in \mathbb{R}^N : r \bullet \vec{n}_2 = \gamma^u - \beta_1 y_{1k} - \dots - \beta_M y_{Mk}\},$$

and the DEA distance from the point k^u to the hyperplane Π_Φ^u is found by projecting the point k^u to Π_Φ^u along Π_N^u :

$$\begin{aligned}\mathfrak{D}(k^u, \Pi_\Phi^u) &= \frac{|\alpha_1(x_{1k} - u) + \dots + \alpha_N(x_{Nk} - u) + \beta_1(y_{1k} + u) + \dots + \beta_M(y_{Mk} + u) - \gamma|}{\sqrt{\alpha_1^2 + \dots + \alpha_N^2}} \\ &= \mathfrak{D}(k, \Pi_\Phi) + \frac{|2u(-\alpha_1 - \dots - \alpha_N + \beta_1 + \dots + \beta_M)|}{\sqrt{\alpha_1^2 + \dots + \alpha_N^2}}.\end{aligned}\quad (4.23)$$

Therefore, by introducing an uncertainty of u the proximity of the efficient frontier to the uncertain box around DMU k reduces by at most $\frac{|2u(-\alpha_1 - \dots - \alpha_N + \beta_1 + \dots + \beta_M)|}{\sqrt{\alpha_1^2 + \dots + \alpha_N^2}}$. For DMU k to be on the facet of the efficient frontier defined by Π_Φ we need to introduce uncertainty such that the DEA distance from the point k^u to the hyperplane Π_Φ^u is 0 i.e. $\mathfrak{D}(k^u, \Pi_\Phi^u) = 0$. From (4.23) this gives

$$\mathfrak{D}(k, \Pi_\Phi) = \frac{|2u(-\alpha_1 - \dots - \alpha_N + \beta_1 + \dots + \beta_M)|}{\sqrt{\alpha_1^2 + \dots + \alpha_N^2}}.\quad (4.24)$$

Substituting (4.21) into (4.24) and rearranging to make u the subject gives

$$u = \frac{|\alpha_1 x_{1k} + \dots + \alpha_N x_{Nk} + \beta_1 y_{1k} + \dots + \beta_M y_{Mk} - \gamma|}{2|-\alpha_1 - \dots - \alpha_N + \beta_1 + \dots + \beta_M|}.$$

□

Alternatively the value of u can be derived via the equations and inequalities in the nominal DEA problem (4.1) and the uncertain DEA problem in (4.19). The nominal

DEA problem for DMU k is

$$\min_{\theta_k, \lambda} \theta_k \quad (4.25a)$$

$$s.t. \quad \lambda_1 y_{11} + \lambda_2 y_{12} + \cdots + \lambda_\Phi y_{1\Phi} - y_{1k} \geq 0 \quad (4.25b)$$

$$\vdots$$

$$\lambda_1 y_{M1} + \lambda_2 y_{M2} + \cdots + \lambda_\Phi y_{M\Phi} - y_{Mk} \geq 0 \quad (4.25c)$$

$$\lambda_1 x_{11} + \lambda_2 x_{12} + \cdots + \lambda_\Phi x_{1\Phi} - \theta_k x_{1k} \leq 0 \quad (4.25d)$$

$$\vdots$$

$$\lambda_1 x_{N1} + \lambda_2 x_{N2} + \cdots + \lambda_\Phi x_{N\Phi} - \theta_k x_{Nk} \leq 0 \quad (4.25e)$$

$$\lambda_1 + \lambda_2 + \cdots + \lambda_\Phi = 1 \quad (4.25f)$$

$$\lambda_1, \lambda_2, \dots, \lambda_\Phi \geq 0. \quad (4.25g)$$

The equality in (4.25f) can be removed by setting $\lambda_1 = 1 - \lambda_2 - \cdots - \lambda_\Phi$. For DMU k to become efficient (4.25b:4.25e) must be binding because we are considering a hyperplane with $\Phi - 1$ dimensions. Hence, they can be replaced with equalities. Rearranging (4.25b) gives

$$\lambda_2 = \frac{y_{1k} - y_{11} + \lambda_3(y_{11} - y_{13}) + \cdots + \lambda_\Phi(y_{11} - y_{1\Phi})}{y_{12} - y_{11}}.$$

Then,

$$\begin{aligned} \lambda_3 = & \frac{(y_{2k} - y_{21})}{\zeta} + \frac{\lambda_4((y_{21} - y_{24})(y_{12} - y_{11}) - (y_{11} - y_{14})(y_{22} - y_{21}))}{\zeta} + \cdots \\ & + \frac{\lambda_\Phi((y_{21} - y_{2\Phi})(y_{12} - y_{11}) - (y_{11} - y_{1\Phi})(y_{22} - y_{21}))}{\zeta}, \end{aligned}$$

with $\zeta = (y_{23} - y_{21})(y_{12} - y_{11}) + (y_{11} - y_{13})(y_{22} - y_{21})$. Repeating this process results in (4.25e) giving a value for θ_k with no λ 's present. This process is repeated for the DEA model in (4.19) to find $\bar{\theta}_k^*$. We require the smallest u such that $\bar{\theta}_k^* = 1$. Therefore, we set the expression for $\bar{\theta}_k^*$ obtained from (4.19) equal to one. We then rearrange the

equation to make u the subject. This gives

$$u = \frac{|\alpha_1 x_{1k} + \cdots + \alpha_N x_{Nk} + \beta_1 y_{1k} + \cdots + \beta_M y_{Mk} - \gamma|}{2|\alpha_1 + \cdots + \alpha_N + \beta_1 + \cdots + \beta_M|},$$

where $\alpha_1, \dots, \alpha_N, \beta_1, \dots, \beta_M$ are defined as in (4.5). This is the same value for u as derived geometrically in (4.22). We demonstrate the complete process for $M = 1$, $N = 1$ and $M = 1$, $N = 2$ in Examples 6 and 9.

In this way for each hyperplane in Ψ , the uncertainty required for DMU k to be projected to its corresponding target point can be calculated. However, as the size of the problem, $N + M$ and I , increases the number of possible facets on the efficient frontier increases exponentially (Briec and Leleu, 2003). Therefore, as the problem size increases, explicitly calculating the value of u for each hyperplane and then selecting the minimum is not feasible. There are methods to calculate all the facets see for example, Briec and Leleu (2003); Frei and Harker (1999); Olesen and Petersen (2015) and Fukuyama and Sekitani (2012).

The following examples are provided to demonstrate the above theorems.

Single input and output With a single input and output the value of u in Theorem 4.3.3 can easily be found. Furthermore, the facet of the efficient frontier that requires the least uncertainty for an inefficient DMU to be projected to it can be found explicitly.

Lemma 4.3.4. *Consider a DEA problem (4.1) with three DMUs A, B, C and $M = 1$ and $N = 1$. For an inefficient DMU C with peers A and B , and $g = \frac{y_B - y_A}{x_B - x_A}$ the minimum amount of uncertainty required to be projected to the target point $T(C, \Pi_{AB})$ on Π_{AB} defined by the inputs and outputs of DMU A and B is*

$$u = \frac{g(x_C - x_A) - y_C + y_A}{2(1 + g)}. \quad (4.26)$$

Proof. This follows directly from Theorem 4.3.3 with $M = 1$, $N = 1$. □

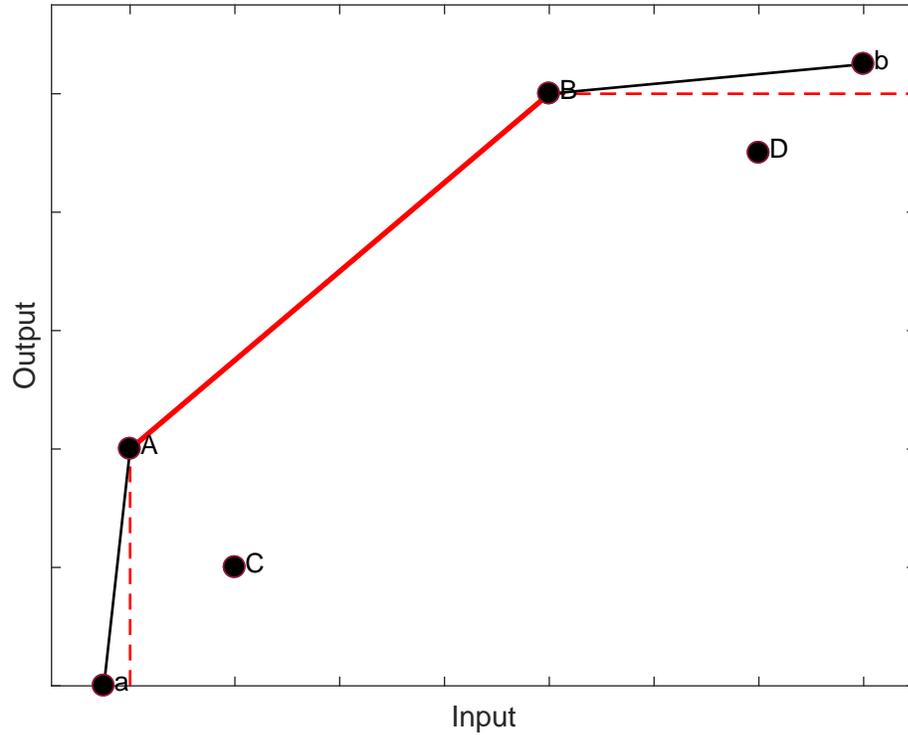


Figure 4.2: Calculating the value of uncertainty for facets defined by a single DMU.

Lemma 4.3.4 allows us to calculate the amount of uncertainty required for a DMU C to become efficient when $N = 1$, $M = 1$ and it is being compared to a hyperplane of the efficient frontier that is defined by two DMUs. However, inefficient DMUs can also be compared to hyperplanes of the efficient frontier defined by a single DMU, $x = \min(x_i)$ and $y = \max(y_i)$ in the following way.

Example 4: Calculating the value of uncertainty for facets of the efficient frontier defined by a single DMU

Let $x = x_A$ for $y \in \{0 : y_A\}$ be a facet of the efficient frontier, vertical red dashed line in Figure 4.2. Let DMU C have peer DMU A such that $x_C > x_A$, $y_C < y_A$ as in Figure 4.2. Define a virtual DMU a such that $x_a = x_A - \epsilon$ and $y_a = 0$ then DMU C is projected to the line segment aA . Then the gradient of the line aA is

$$g_{aA} = \frac{y_A - 0}{x_A - (x_A - \epsilon)} = \frac{y_A}{\epsilon}.$$

From (4.26),

$$\begin{aligned} u_{aA} &= \frac{\frac{y_A}{\epsilon}(x_C - x_A) - y_C + y_A}{2(1 + \frac{y_A}{\epsilon})} \\ 2u_{aA} \left(\frac{\epsilon + y_A}{\epsilon} \right) &= \frac{y_A(x_C - x_A) + \epsilon(y_A - y_C)}{\epsilon} \\ u_{aA} &= \frac{y_A(x_C - x_A) + \epsilon(y_A - y_C)}{2(\epsilon + y_A)} \end{aligned}$$

But ϵ is an infinitesimally small constant so we take the limit as $\epsilon \rightarrow 0$, then

$$u_{aA} \rightarrow \frac{y_A(x_C - x_A)}{2y_A} = \frac{x_C - x_A}{2}.$$

Similarly, to compare to the facet of the efficient frontier defined by $y = \max(y_i)$, define a virtual DMU b , $x_b = x_B + \kappa$, $y_b = y_B + \epsilon$, where κ is a positive constant such that $x_B + \kappa > x_D$. Then DMU D is projected to the line segment Bb , which has gradient,

$$g_{Bb} = \frac{y_B + \epsilon - y_B}{x_B + \kappa - x_B} = \frac{\epsilon}{\kappa}.$$

From (4.26),

$$\begin{aligned} u_{Bb} &= \frac{\frac{\epsilon}{\kappa}(x_D - x_B) - y_D + y_B}{2(1 + \frac{\epsilon}{\kappa})} \\ 2u_{Bb} \left(\frac{\kappa + \epsilon}{\kappa} \right) &= \frac{\epsilon(x_D - x_B) + \kappa(y_B - y_D)}{\kappa} \\ u_{Bb} &= \frac{\epsilon(x_D - x_B) + \kappa(y_B - y_D)}{2(\kappa + \epsilon)}. \end{aligned}$$

We take the limit as $\epsilon \rightarrow 0$, then

$$u_{Bb} \rightarrow \frac{\kappa(y_B - y_D)}{2\kappa} = \frac{y_B - y_D}{2}.$$

So for DMU D projected to the line segment $y = y_B$ any value of $u > \frac{y_B - y_D}{2}$ will result in DMU D being classified as efficient on the line segment $y = y_D$.

We now provide a detailed example of Theorem 4.3.3 for $M = 1$ and $N = 1$.

Example 5: Calculating the uncertainty geometrically, $M = 1$, $N = 1$

Consider the DEA problem introduced in Lemma 4.3.4. DMU A and B are peers to inefficient DMU C . Therefore, $x_A \leq x_C$ and $y_A \leq y_C \leq y_B$. These inequalities define the shaded region in Figure 4.3a and the efficient frontier is depicted by the red line AB . For input-oriented envelopment DEA model with VRS, DMU C is projected to the point D shown in Figure 4.3b. Under box uncertainty each data-point can take any value in the range

$\{x_i \in [x_i^-u], y_i \in [y_i^+u], i = A, B, C\}$. This is illustrated in Figure 4.3c by the blue boxes around the data of the DMUs and can be thought of as DMU i 's ($i = A, B, C$) data being a realisation from anywhere inside the square (including the boundary). We wish to find which points should be chosen from these boxes to increase the efficiency score for DMU C . From Theorem 4.2.1 we increase the outputs and decrease the inputs for DMU C and do the opposite to the remaining DMUs as in (4.19). This is shown by the points $A^u = (x_A + u, y_A - u)$, $B^u = (x_B + u, y_B - u)$ and $C^u = (x_C - u, y_C + u)$ in Figure 4.3d. The green line A^uB^u represents the new efficient frontier for the chosen data. To find the smallest value of u such that DMU C is considered efficient is equivalent to finding the smallest box size such that C^u first touches the line A^uB^u .

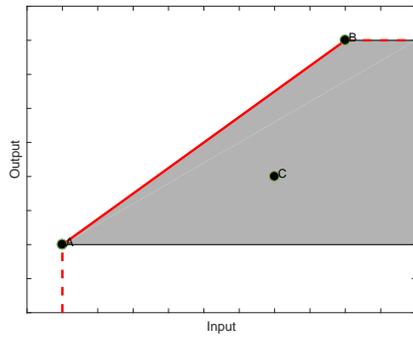
The lines AB and A^uB^u are parallel with the gradient

$$g = \frac{y_B - y_A}{x_B - x_A}.$$

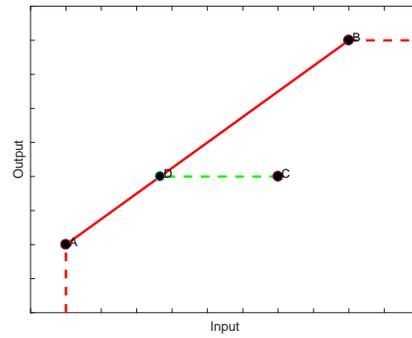
Define points D and E as in Figure 4.3e

$$D = \left(x_A + \frac{(y_C - y_A)}{g}, y_C \right)$$

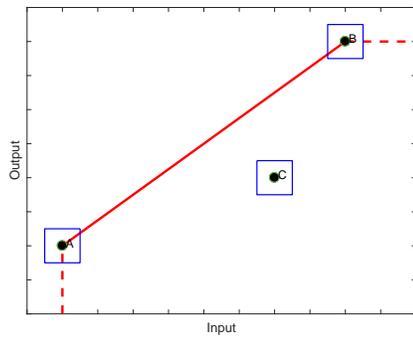
$$E = \left(x_A + u + \frac{(y_C + 2u - y_A)}{g}, y_C + u \right) = \left(x_A + u + \frac{(y_C + 2u - y_A)}{g}, y_C + u \right).$$



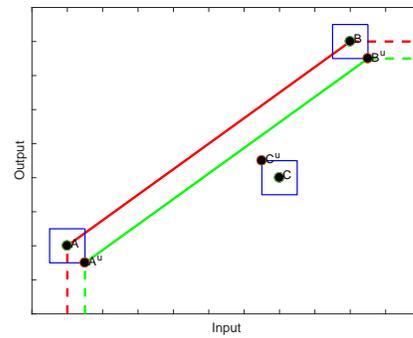
(a) Restricted area of interest.



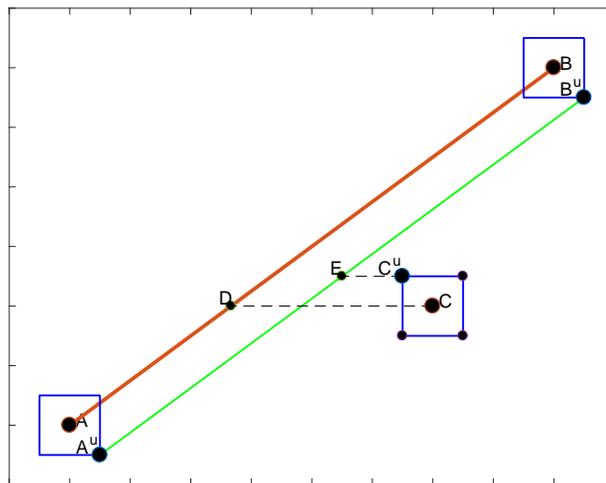
(b) DMU C is projected to the point D .



(c) Box uncertainty.



(d) Best position for DMU C .



(e) Projection of C and C^u to D and E .

Figure 4.3: Calculating the uncertainty geometrically, Example 5.

Before uncertainty is considered, the DEA distance from C to its corresponding projection onto AB is

$$\mathfrak{D}(C, AB) = x_C - x_D = x_C - x_A - \frac{y_C - y_A}{g}.$$

By decreasing DMU C 's input by u and increasing that of DMUs A and B by u the proximity of the efficient frontier $A^u B^u$ to C^u is:

$$\begin{aligned} \mathfrak{D}(C^u, A^u B^u) &= x_{C^u} - x_E \\ &= x_C - u - x_A - u - \frac{(y_C + 2u - y_A)}{g} \\ &= \mathfrak{D}(C, AB) - 2u \left(1 + \frac{1}{g}\right). \end{aligned}$$

When $\mathfrak{D}(C^u, A^u B^u) = 0$ C^u is on the efficient frontier $A^u B^u$. To find the minimum amount of uncertainty and hence where DMU C will be deemed efficient we solve $\mathfrak{D}(C^u, A^u B^u) = 0$ and rearrange to find the value of u . This gives,

$$\begin{aligned} 0 &= \mathfrak{D}(C^u, A^u B^u) \\ 0 &= \mathfrak{D}(C, AB) - 2u \left(1 + \frac{1}{g}\right) \\ 0 &= x_C - x_A - \frac{y_C - y_A}{g} - 2u \left(1 + \frac{1}{g}\right) \\ u &= \frac{g(x_C - x_A) - y_C + y_A}{2(1 + g)}. \end{aligned}$$

Therefore, DMU C will be deemed efficient when $A = (x_A + u, y_A - u)$, $B = (x_B + u, y_B - u)$ and $C(x_C - u, y_C + u)$ and $u = \frac{g(x_C - x_A) - y_C + y_A}{2(1 + g)}$.

Alternatively, the minimum amount of uncertainty can be derived from the DEA model as shown in Example 6.

Example 6: Calculating the uncertainty from the DEA model, $M = N = 1$

Recall the DEA model for a single input and output where inefficient DMU C has peers A and B .

$$\min_{\theta_C, \lambda} \theta_C \quad (4.27a)$$

$$s.t. \quad \lambda_A y_A + \lambda_B y_B - y_C \geq 0 \quad (4.27b)$$

$$\lambda_A x_A + \lambda_B x_B - \theta_C x_C \leq 0 \quad (4.27c)$$

$$\lambda_A + \lambda_B = 1 \quad (4.27d)$$

$$\lambda_A, \lambda_B \geq 0. \quad (4.27e)$$

The equality in (4.27d) can be removed by setting $\lambda_A = 1 - \lambda_B$. For DMU C to be efficient (4.27b) and (4.27c) must be binding, hence they can be replaced with equalities. Then from (4.27b) and (4.27c) λ_B and θ_C are,

$$\lambda_B = \frac{y_C - y_A}{y_B - y_A}$$

$$\theta_C = \frac{(1 - \lambda_B)x_A + \lambda_B x_B}{x_C}.$$

Select uncertainty as in Theorem 4.2.1. Consider a change in the data of u , then λ_B and θ_C become λ_B^u and θ_C^u ,

$$\lambda_B^u = \frac{(y_C + u) - (y_A - u)}{(y_B - u) - (y_A - u)} = \frac{y_C - y_A + 2u}{y_B - y_A} \quad (4.28a)$$

$$\theta_C^u = \frac{(1 - \lambda_B^u)(x_A + u) + \lambda_B^u(x_B + u)}{(x_C - u)} = \frac{\lambda_B^u(x_B - x_A) + x_A + u}{(x_C - u)}. \quad (4.28b)$$

We wish to calculate the smallest value of u such that $\theta_C^u = 1$. Therefore set (4.28b) equal to one, substitute in (4.28a) and rearrange to make u the subject,

$$1 = \frac{\lambda_B^u(x_B - x_A) + x_A + u}{(x_C - u)}$$

$$2u \left(\frac{1}{g} + g \right) = (x_C - x_A) - \frac{1}{g}(y_C - y_A)$$

$$u = \frac{g(x_C - x_A) - y_C + y_A}{2(1 + g)},$$

and $g = \frac{y_B - y_A}{x_B - x_A}$.

This is the same value for u as derived geometrically in (4.26).

In Example 5 and 6 the amount of uncertainty required for DMU C to become efficient on the efficient facet defined by DMU A and B is calculated. However, there may be additional efficient DMUs that are not peers to DMU C that would result in a smaller value of u if DMU C were compared to them. This leads to the following important question. When is it beneficial for an inefficient DMU to compare itself to different facets of the efficient frontier?

With a single input and output the efficient DMUs can be ordered by increasing input and output. Here each facet has dimension $N + M - 1 = 1$. Therefore, the facets are line segments and can be defined by two extreme points of the efficient frontier. Solving the nominal DEA problem (4.1) I times, once for each of the I DMUs, with $N = 1$, $M = 1$ determines all efficient DMUs. There are a finite number, ϕ , of efficient DMUs whose inputs and outputs represent an extreme point of the PPS. These ϕ points can be ordered $x_1 < x_2 < \dots < x_\phi$ and $y_1 < y_2 < \dots < y_\phi$. In this way each consecutive pair define a facet of the efficient frontier.

In Theorem 4.3.6 we show that for $M = N = 1$ we can determine the minimum amount of uncertainty required for a DMU to become efficient by determining the facet which requires the minimum amount of uncertainty for DMU k to be deemed efficient. To reach this conclusion we prove the following theorem which demonstrates that when $M = N = 1$ as the gradient of the line segments on the efficient frontier decrease, the amount of uncertainty required for a DMU to become efficient decreases.

Theorem 4.3.5. *Consider the nominal DEA problem (4.1) with $N = 1$ and $M = 1$. For inefficient DMUs a DEA distance τ from the efficient frontier, as the rate at which the outputs change compared to the inputs on the efficient frontier decreases, the required uncertainty for an inefficient DMU to become efficient when compared to that line segment decreases.*

Proof. Let A , B , C be efficient DMUs such that their inputs and outputs represent

an extreme point of the PPS, and $x_A < x_B < x_C$ and $y_A < y_B < y_C$. Then the efficient frontier contains the line segments AB and BC . The line segments have gradients g_1 and g_2 respectively,

$$g_1 = \frac{y_B - y_A}{x_B - x_A} \quad g_2 = \frac{y_C - y_B}{x_C - x_B}. \quad (4.29)$$

where $g_1 > g_2$. In other words, from the efficient DMU with the smallest input and output to the efficient DMU with the largest input and output the line segments have decreasing gradients. Let

$$\begin{aligned} D &= (x_A + \rho_1 + \tau, y_A + g_1\rho_1) & E &= (x_B + \rho_2 + \tau, y_B + g_2\rho_2) \\ F &= (x_A + \rho_1, y_A + g_1\rho_1) & G &= (x_B + \rho_2, y_B + g_2\rho_2), \end{aligned}$$

where ρ_1 , ρ_2 and τ are positive constants as in Figure 4.4. In this way the inefficient DMUs D and E have peers A, B and B, C , respectively. The point $F(G)$ is the point on the efficient frontier that DMU $D(E)$ is projected to and $\mathfrak{D}(D, AB) = \mathfrak{D}(E, BC) = \tau$. This is shown in Figure 4.4. From (4.26), the amount of uncertainty required for DMU D and E to be projected to the line segment AB , BC respectively is

$$u_D = \frac{g_1(x_D - x_A) - y_D + y_A}{2(1 + g_1)} = \frac{g_1\tau}{2(1 + g_1)} \quad (4.30a)$$

$$u_E = \frac{g_2(x_E - x_B) - y_E + y_B}{2(1 + g_2)} = \frac{g_2\tau}{2(1 + g_2)}. \quad (4.30b)$$

But, $g_1 > g_2$ so g_1 can be rewritten as $g_1 = g_2 + \epsilon$ ($\epsilon > 0$) and (4.30a) becomes

$$u_D = \frac{\tau(g_2 + \epsilon)}{2(1 + g_2 + \epsilon)} = \frac{\tau}{2} \left(\frac{g_2 + \epsilon}{1 + g_2 + \epsilon} \right).$$

Then we can show that $u_D > u_E$ via a contradiction. Assume $u_D < u_E$, then

$$\begin{aligned} \frac{g_2}{1 + g_2} &> \frac{g_2 + \epsilon}{1 + g_2 + \epsilon} \\ \Leftrightarrow g_2(1 + g_2 + \epsilon) &> (g_2 + \epsilon)(1 + g_2) \\ \Leftrightarrow g_2 + g_2^2 + g_2\epsilon &> g_2 + g_2^2 + \epsilon + g_2\epsilon \\ \Leftrightarrow 0 &> \epsilon. \end{aligned}$$

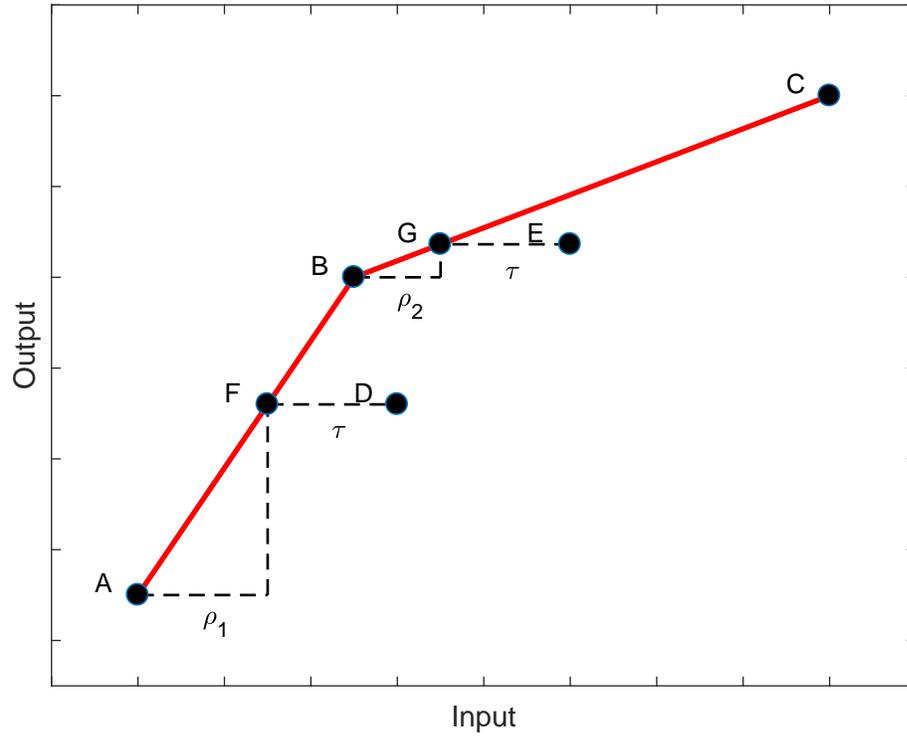


Figure 4.4: The effect of a decreasing gradient of efficient line segments on the value of u .

But this is a contradiction as $\epsilon > 0$ so

$$\frac{g_1\tau}{2(1+g_1)} > \frac{g_2\tau}{2(1+g_2)} \text{ and } u_D > u_E.$$

Therefore, for inefficient DMUs a DEA distance τ from the efficient frontier, as the gradient of the line segment on the efficient frontier decreases the required uncertainty for an inefficient DMU to become efficient when compared to that line segment decreases. □

Hence inefficient DMUs will never benefit from being compared to line segments of the efficient frontier with a larger gradient than the gradient of the peers from solving the nominal DEA problem (4.1). This gives rise to the following theorem,

Theorem 4.3.6. For $N = 1$, $M = 1$ and ϕ efficient DMUs, order and label the ϕ efficient DMUs by increasing input and output such that $x_1 < x_2 < \dots < x_{\phi-1} < x_\phi$ and $y_1 < y_2 < \dots < y_{\phi-1} < y_\phi$. Then, for any inefficient DMU k , the facet of the efficient frontier which requires the minimum amount of uncertainty for DMU k to be projected to its corresponding target point can be determined by:

$$\begin{aligned}
y_k + x_k &\leq y_1 + x_1 && \text{efficient frontier is } x = x_1 \\
y_1 + x_1 &\leq y_k + x_k \leq y_2 + x_2 && \text{efficient frontier is } EF_{1,2} \\
y_2 + x_2 &\leq y_k + x_k \leq y_3 + x_3 && \text{efficient frontier is } EF_{2,3} \\
y_3 + x_3 &\leq y_k + x_k \leq y_4 + x_4 && \text{efficient frontier is } EF_{3,4} \\
&\vdots && \vdots \\
y_{\phi-1} + x_{\phi-1} &\leq y_k + x_k \leq y_\phi + x_\phi && \text{efficient frontier is } EF_{\phi-1,\phi} \\
y_\phi + x_\phi &\leq y_k + x_k && \text{efficient frontier is } y = y_\phi.
\end{aligned}$$

Where $EF_{i,j}$ is the straight line formed between (x_i, y_i) and (x_j, y_j) .

Proof. Choose a subset of consecutive efficient DMUs, for example, A , B , $C < \phi$.

Let DMU k be inefficient with peers A and B where

$$(x_k, y_k) = (x_B + \mu(x_C - x_B), y_B - \epsilon)$$

$0 < \mu < 1$ and $0 < \epsilon < y_B - y_A$. This ensures DMU k has peers A and B and $x_A < x_B < x_k < x_C$.

Let the gradient of the line segment defined by AB and BC be g_1, g_2 as in (4.29). Let u_{AB} , u_{BC} be the uncertainty required for DMU k to be projected to the line segment AB , BC respectively. Then from (4.26),

$$\begin{aligned}
u_{AB} &= \frac{g_1(x_k - x_A) - y_k + y_A}{2(1 + g_1)} = \frac{g_1((x_B - x_A) + \mu(x_C - x_B)) - (y_B - y_A) + \epsilon}{2(1 + g_1)} \\
u_{BC} &= \frac{g_2(x_k - x_B) - y_k + y_B}{2(1 + g_2)} = \frac{g_2(\mu(x_C - x_B)) + \epsilon}{2(1 + g_2)}.
\end{aligned}$$

It is beneficial for DMU k to be compared to BC if $u_{AB} > u_{BC}$,

$$u_{AB} = \frac{g_1((x_B - x_A) + \mu(x_C - x_B)) - (y_B - y_A) + \epsilon}{2(1 + g_1)} > \frac{g_2(\mu(x_C - x_B)) + \epsilon}{2(1 + g_2)} = u_{BC}.$$

This occurs when

$$\mu = \frac{\epsilon}{x_C - x_B}.$$

Therefore, when

$$y_k + x_k < y_B + x_B \quad \text{the peers for DMU } k \text{ remain the same}$$

$$y_k + x_k > y_B + x_B \quad \text{the peers for DMU } k \text{ change.}$$

This is because less uncertainty is required for DMU k to be projected to its corresponding target point on the facet defined by BC then AB . Continued application of the above for increasing A,B,C gives the desired result. \square

We note the use of less than or equal signs in Theorem 4.3.6. These are used and overlap because if $x_k + y_k = x_i + y_i$ $i = 2, \dots, \phi$ then the value of uncertainty required for DMU k to be projected to its corresponding target point on $EF_{i-1,i}$ and $EF_{i,i+1}$ are equal. This can be seen in Figure 4.5 where DMU k is on one of the diagonal lines between colour blocks.

Consequently, it is sometimes beneficial for inefficient DMUs to be compared to facets of the efficient frontier which they are not compared to in the nominal DEA problem (4.1). In this way we can determine the minimum amount of uncertainty required for a DMU to become efficient by determining the facet which requires the minimum amount of uncertainty for DMU k to be projected to its corresponding target point for the case with $N = 1$ and $M = 1$. The areas in which the minimum uncertainty is achieved on each facet can be seen in Figure 4.5. Here we note that inefficient DMUs in the region $y_k + x_k < y_A + x_A$, depicted by the red region in Figure 4.5, will be compared to the facet of the efficient frontier given by $x = x_A$, shown by a vertical dotted line in Figure 4.5. If a DMU is on the line $x = x_A$, $y \in (0, y_A)$

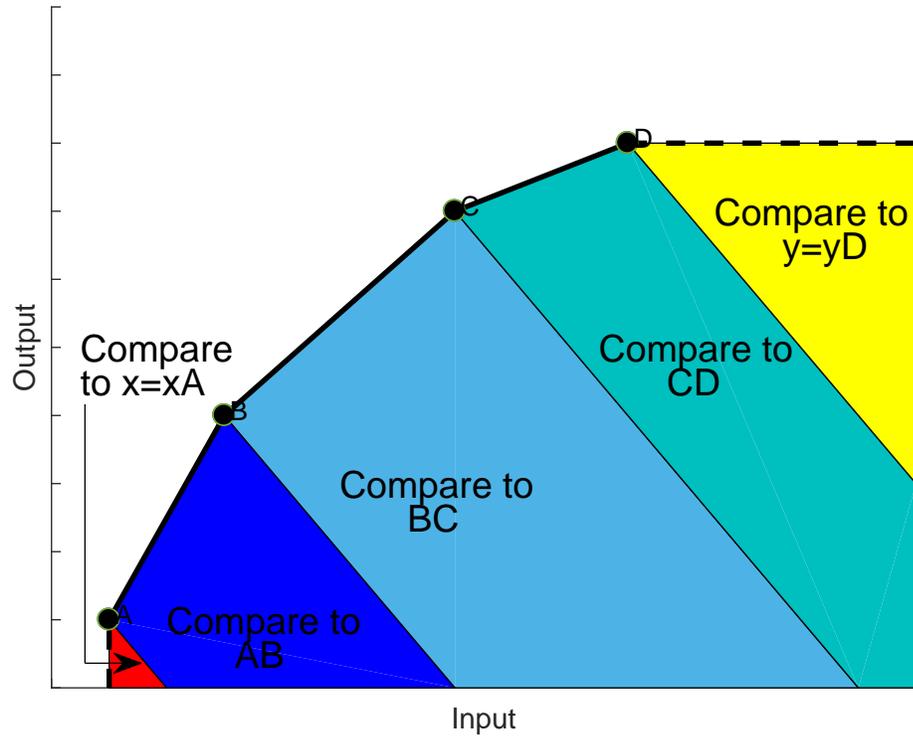


Figure 4.5: The facet of the efficient frontier inefficient DMUs are compared with.

it would be possible for DMUs to increase their outputs further without increasing the inputs. However, they are still deemed efficient as we are considering an input orientation where the outputs must remain fixed. When inefficient DMUs are in the region $y_\phi + x_\phi < y_k + x_k$, the yellow region in Figure 4.5, they will be compared to the facet of the efficient frontier given by $y = y_\phi$. Any DMUs on the line $y = y_\phi$ after solving the nominal DEA problem will require $u = \epsilon$ to become efficient. That is because on the dashed line $y = y_\phi$ any infinitesimally small value for u will result in $y_k > y_\phi$.

We demonstrate the case of a single input and output with Example 7.

DMU	In	Out	Efficiency Score
A	1	1	1
B	3	4	1
C	7	7	1
D	10	8	1
E	8	5	0.542
F	6	2	0.278

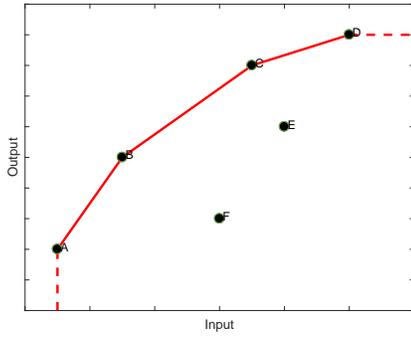
Table 4.11: Nominal data, Example 7.

Example 7: Numerical: one input, one output

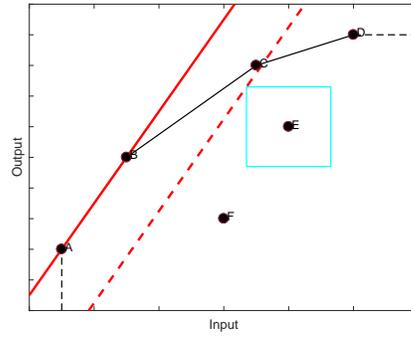
Consider the six DMUs pictured in Figure 4.6a and whose nominal data are listed in Table 4.11. Solving the nominal DEA problem (4.1) for all six DMUs gives the efficiency scores in Table 4.11 where DMUs $A - D$ are efficient and DMUs $E - F$ are inefficient. This can be seen in Figure 4.6a where the efficient DMUs are joined by the red line to depict the efficient frontier.

From Theorem 4.3.6 the facet of the efficient frontier the inefficient DMUs should be compared to can be calculated. We have $y_A + x_A = 2$, $y_B + x_B = 7$, $y_C + x_C = 14$, $y_D + x_D = 18$, $y_E + x_E = 13$ and $y_F + x_F = 8$ therefore both DMUs E and F should be compared to BC . The uncertainty required for DMUs E and F to be projected to its corresponding target point on each of the hyperplanes which intersect the efficient frontier can be calculated from (4.26) and is presented in Table 4.12. As expected, the smallest value for u_E and u_F occurs when DMUs E and F are compared to BC .

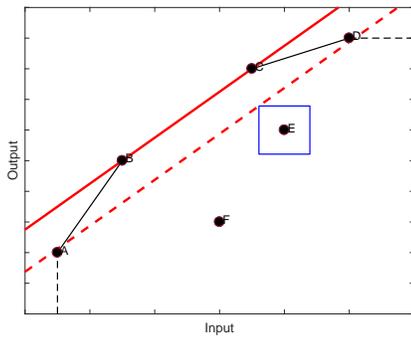
In this way, for a single input and output we can determine the hyperplane which requires the minimum amount of uncertainty for DMU k to be projected to its corre-



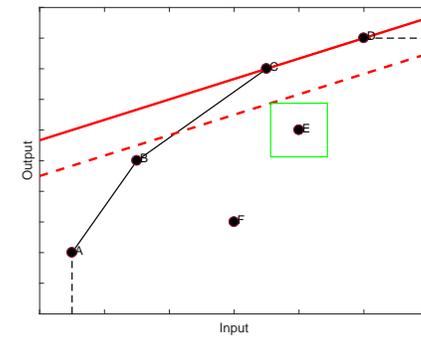
(a) The efficient frontier in Example 7.



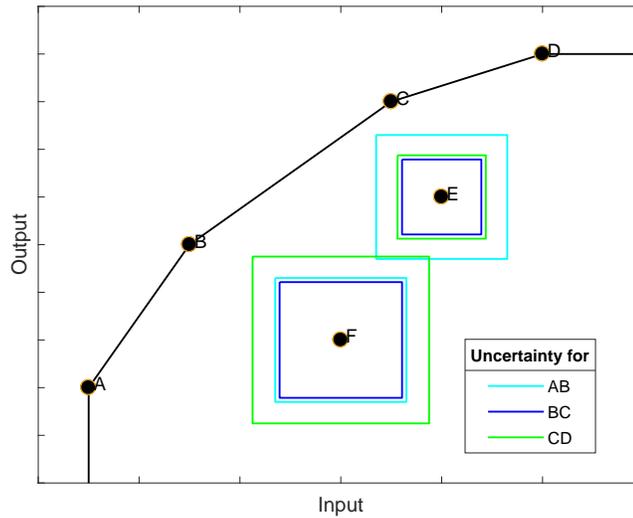
(b) The extended facet of AB.



(c) The extended facet of BC.



(d) The extended facet of CD.



(e) The required uncertainty for DMUs E and F to become efficient on different facets.

Figure 4.6: Diagrams to show the possible facets of the efficient frontier DMU E and F can be compared to.

Efficient frontier section	u_E	u_F
AB	1.3	1.3
BC	0.786	1.214
CD	0.875	1.75

Table 4.12: Value of u for different facets of the efficient frontier, Example 7.

sponding target point and hence the facet on which the DMU will be deemed efficient.

However, for higher dimensions it is much harder to explicitly compute the facets of the efficient frontier. With three variables the efficient frontier can be visualised for simple examples. In Examples 8 and 9 we again demonstrate that the value of u for each hyperplane that intersects the efficient frontier can be derived geometrically and via the DEA equations. Example 10 is a small example with two inputs and one output that considers the efficient facets for three efficient DMUs to demonstrate the increasing complexity.

Example 8: Calculating the uncertainty geometrically, $N = 2$, $M = 1$

Let DMUs $A - C$ be efficient and DMU D inefficient with peers $A - C$. We denote the two inputs x_{1i} , x_{2i} and the output y_i for $i \in \{A, B, C, D\}$.

When an inefficient DMU has three peers, the facet of the efficient frontier it is compared with can be calculated. Following the procedure in Section 4.3.1, the equation of the hyperplane with DMUs $A - C$, here denoted by Π_{ABC} , is

$$\Pi_{ABC} : \alpha_1 x_1 + \alpha_2 x_2 + \beta_1 y = \gamma,$$

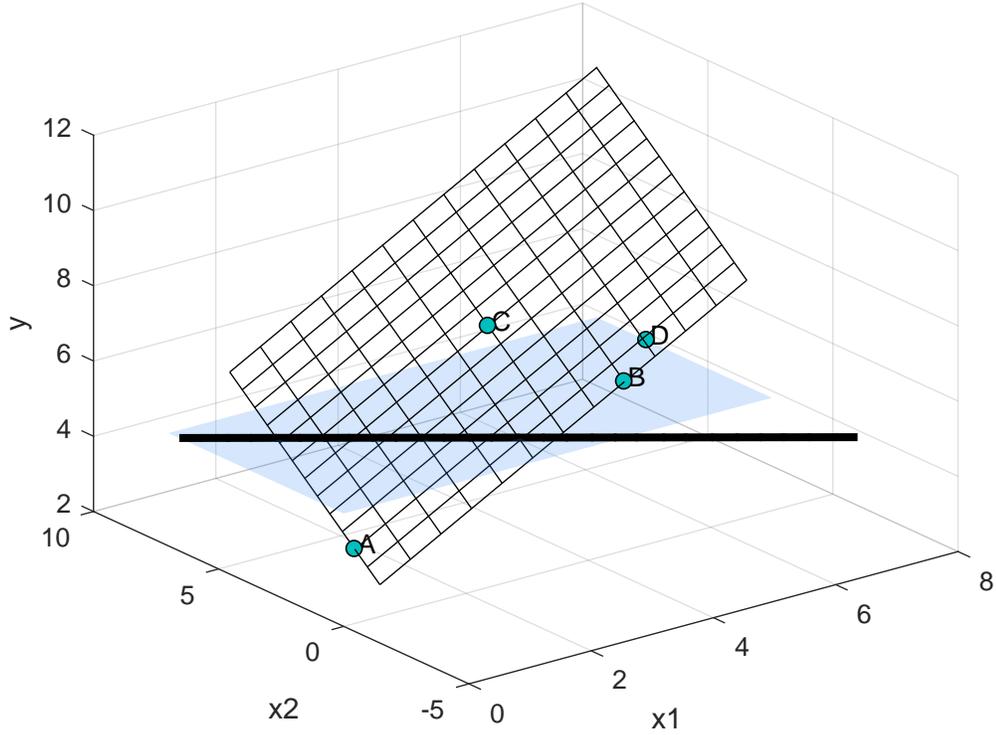


Figure 4.7: The hyperplane Π_{ABC} , defined by efficient DMUs A, B and C (cross-hatched), intersecting with the hyperplane $y=y_D$ (blue).

$$\alpha_1 = (x_{2B} - x_{2A})(y_C - y_A) - (x_{2C} - x_{2A})(y_B - y_A) \quad (4.31a)$$

$$\alpha_2 = -(x_{1B} - x_{1A})(y_C - y_A) + (x_{1C} - x_{1A})(y_B - y_A) \quad (4.31b)$$

$$\beta_1 = (x_{1B} - x_{1A})(x_{2C} - x_{2A}) - (x_{1C} - x_{1A})(x_{2B} - x_{2A}) \quad (4.31c)$$

$$\gamma = (\alpha_1 x_{1A} + \alpha_2 x_{2A} + \beta_1 y_A). \quad (4.31d)$$

In vector notation this is

$$\Pi_{ABC} = \{r \in \mathbb{R}^3 : r \bullet \vec{n}_1 = \gamma\},$$

where $\vec{n}_1 = (\alpha_1, \alpha_2, \beta_1)$ is a normal vector to the plane and γ is a constant. The inefficient DMU D has co-ordinates (x_{1D}, x_{2D}, y_D) . For input oriented DEA an improvement in the inputs is sought. Therefore, DMU D is projected to the efficient

frontier Π_{ABC} along the hyperplane $\Pi_{y_D} : y = y_D$. This can be seen in Figure 4.7, where Π_{ABC} , (cross-hatched) intersects with Π_{y_D} , (blue) along, a line. The hyperplane Π_{y_D} has normal vector $\vec{n}_2 = (0, 0, 1)$. The hyperplanes Π_{ABC} and Π_{y_D} intersect along the line

$$\pi = \{r \in \mathbb{R}^2 : r \bullet \vec{n}_2 = \gamma - \beta_1 y_D\},$$

and the DEA distance from DMU D to its projected point on the hyperplane Π_{ABC} is

$$\mathfrak{D}(D, \Pi_{ABC}) = \frac{|\alpha_1 x_{1D} + \alpha_2 x_{2D} + \beta_1 y_D - \gamma|}{\sqrt{\alpha_2^2 + \alpha_1^2}}.$$

Consider the addition of uncertainty and denote the new efficient frontier Π_{ABC}^u and the DEA distance from D^u to Π_{ABC}^u by $\mathfrak{D}(D^u, \Pi_{ABC}^u)$. From Theorem 4.2.1 the possible realisation of the uncertain data $(x_{1i} + u, x_{2i} + u, y_i - u)$ for $i \in \{A, B, C\}$ and $(x_{1D} - u, x_{2D} - u, y_D + u)$ is selected. We denote these new points A^u, B^u, C^u, D^u , and wish to calculate the smallest value of u required for DMU D to become efficient.

From (4.31a)-(4.31c) we note that the u 's will cancel each other out in the equations for $\alpha_1, \alpha_2, \beta_1$ so $\alpha_1^u = \alpha_1, \alpha_2^u = \alpha_2, \beta_1^u = \beta_1$. This is because the uncertain efficient frontier is parallel to the original efficient frontier, as seen in Figure 4.8. However, the γ value for Π_{ABC}^u will change.

$$\begin{aligned} \gamma^u &= -(\alpha_1(x_{1A} + u) + \alpha_2(x_{2A} + u) + \beta_1(y_A - u)) \\ &= \gamma - u(\alpha_1 + \alpha_2 - \beta_1). \end{aligned}$$

The hyperplanes Π_{ABC}^u and $\Pi_{y_D}^u$ intersect along the line π^u .

$$\begin{aligned} \pi^u &= \{r \in \mathbb{R}^2 : r \bullet \vec{n}_2 = \gamma^u - \beta_1 y_D\} \\ &= \{\gamma - \beta_1 y_D - u(\alpha_1 + \alpha_2 - \beta_1)\}, \end{aligned}$$

and the DEA distance from the point D^u to the hyperplane Π_{ABC}^u is:

$$\mathfrak{D}(D^u, \Pi_{ABC}^u) = \mathfrak{D}(D, \Pi_{ABC}) + \frac{2u(-\alpha_1 - \alpha_2 + \beta_1)}{\sqrt{\alpha_1^2 + \alpha_2^2}}.$$

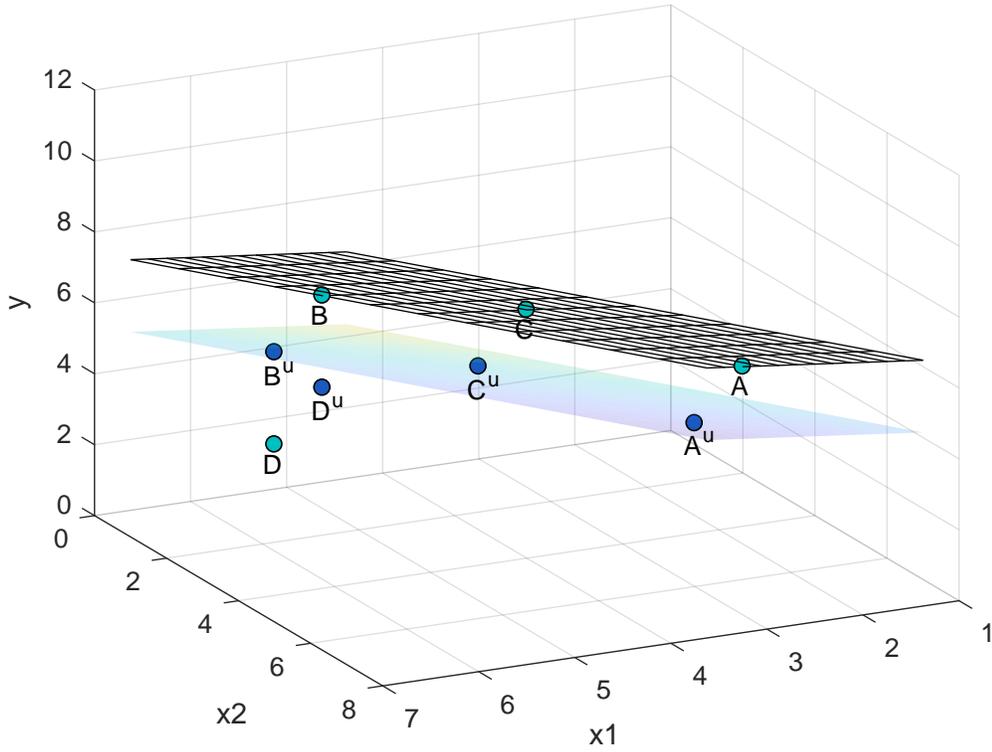


Figure 4.8: The hyperplane Π_{ABC} , defined by efficient DMUs A, B and C (cross-hatched), and the hyperplane Π_{ABC}^u defined by A^u , B^u and C^u (coloured).

By introducing an uncertainty of u , the proximity of the efficient frontier to the uncertain box around DMU D reduces by at most $\frac{2u(-\alpha_1 - \alpha_2 + \beta_1)}{\sqrt{\alpha_1^2 + \alpha_2^2}}$. Before uncertainty is considered, the required improvement to make DMU D efficient is $\mathfrak{D}(D, \Pi_{ABC})$. For DMU D to be considered efficient when we consider uncertainty we require $\mathfrak{D}(D^u, \Pi_{ABC}^u) = 0$, i.e.

$$\mathfrak{D}(D, \Pi_{ABC}) = \frac{2u(-\alpha_1 - \alpha_2 + \beta_1)}{\sqrt{\alpha_1^2 + \alpha_2^2}}. \quad (4.32)$$

Rearranging (4.32) gives,

$$u = \frac{|\alpha_{x1}x_{1D} + \alpha_{x2}x_{2D} + \alpha_y y_D + \gamma|}{2|-\alpha_{x1} - \alpha_{x2} + \alpha_y|}. \quad (4.33)$$

This is the same value for u as given by (4.22) with $N = 2$, $M = 1$.

The value of u with two inputs and one output can also be derived from the nominal DEA problem.

Example 9: Uncertainty from DEA model, two inputs, one output

The nominal DEA problem for DMU D is

$$\min_{\theta_D, \lambda} \theta_D \quad (4.34a)$$

$$s.t. \quad \lambda_A y_A + \lambda_B y_B + \lambda_C y_C - y_D \geq 0 \quad (4.34b)$$

$$\lambda_A x_{1A} + \lambda_B x_{1B} + \lambda_C x_{1C} - \theta_D x_{1D} \leq 0 \quad (4.34c)$$

$$\lambda_A x_{2A} + \lambda_B x_{2B} + \lambda_C x_{2C} - \theta_D x_{2D} \leq 0 \quad (4.34d)$$

$$\lambda_A + \lambda_B + \lambda_C = 1 \quad (4.34e)$$

$$\lambda_A, \lambda_B, \lambda_C \geq 0. \quad (4.34f)$$

The equality in (4.34e) can be removed by setting $\lambda_A = 1 - \lambda_B - \lambda_C$. For DMU D to be deemed efficient (4.34b, 4.34c and 4.34d) must be binding, hence they can be replaced with equalities. Rearranging (4.34b) gives

$$\lambda_B = \frac{y_D - \lambda_C y_C - (1 - \lambda_C) y_A}{y_B - y_A}. \quad (4.35)$$

Substituting (4.35) into (4.34c) gives,

$$\lambda_C = \frac{(y_B - y_A)(\theta_D x_{1D} - x_{1A}) + (y_A - y_D)(x_{1B} - x_{1A})}{(y_B - y_A)(x_{1C} - x_{1A}) + (y_A - y_C)(x_{1B} - x_{1A})}. \quad (4.36)$$

Then substituting (4.36) into (4.34d) gives

$$\theta_D = \frac{(y_B - y_D)(\alpha_1 x_{1A} + \alpha_2 x_{2A}) + (y_D - y_A)(\alpha_1 x_{1B} + \alpha_2 x_{2B})}{(y_B - y_A)(\alpha_1 x_{1D} + \alpha_2 x_{2D})},$$

where $\alpha_1 = (x_{2B} - x_{2A})(y_C - y_A) - (x_{2C} - x_{2A})(y_B - y_A)$ and $\alpha_2 = -(x_{1B} - x_{1A})(y_C - y_A) + (x_{1C} - x_{1A})(y_B - y_A)$. Repeating this with the uDEA model (4.19), gives

$$\bar{\theta}_D^* = \frac{(y_B - y_D - 2u)(\alpha_1(x_{1A} + u) + \alpha_2(x_{2A} + u))}{(y_B - y_A)(\alpha_1(x_{1D} - u) + \alpha_2(x_{2D} - u))} + \frac{(y_D - y_A + 2u)(\alpha_1(x_{1B} + u) + \alpha_2(x_{2B} + u))}{(y_B - y_A)(\alpha_1(x_{1D} - u) + \alpha_2(x_{2D} - u))}. \quad (4.37)$$

DMU	x_1	x_2	y	Efficiency score	$\lambda_i \ i \in \{A, B, C, D\}$
A	1	2	3	1	0.419
B	5	1	6	1	0.324
C	4	4	7	1	0.257
D	7	5	5	0.438	0

Table 4.13: Data, efficiency scores and peers: two inputs, one output, Example 10.

We require the smallest u such that $\bar{\theta}_D^* = 1$. Setting (4.37) equal to one and rearranging gives,

$$u = \frac{|\alpha_1 x_{1D} + \alpha_2 x_{2D} + \beta_1 y_k + \gamma|}{2|-\alpha_1 - \alpha_2 + \beta_1|},$$

where $\beta_1 = ((x_{1B} - x_{1A})(x_{2C} - x_{2A}) - (x_{1C} - x_{1A})(x_{2B} - x_{2A}))$ and $\gamma = -(\alpha_1 x_{1A} + \alpha_2 x_{2A} + \beta_1 y_A)$. This is the same value for u as derived geometrically in (4.33).

We now consider a numerical example to demonstrate the different facets of the efficient frontier DMU D can be compared with.

Example 10: Two inputs, one output

Consider the four DMUs pictured in Figure 4.7 whose nominal data are listed in Table 4.13. Solving the nominal DEA problem (4.1) gives the efficiency scores in Table 4.13. DMUs $A - C$ are efficient and DMU D is inefficient with peers A, B and C . This can be seen in Figure 4.7 where the efficient DMUs are on the hyperplane $\Pi_{ABC} : -10x_1 - 7x_2 + 11y - 9 = 0$. From (4.33), the minimum amount of uncertainty required for DMU D to be considered efficient on Π_{ABC} is $u = \frac{59}{56} \approx 1.05$.

Let Π_{ijk} denote the plane going through the points defined by the inputs and outputs of DMU i, j and k , Π_z be the plane where the $\min(\max) \ x(y)$ value is fixed and $\Pi_{ij(z)}$ denote the plane going through the points defined by the inputs and outputs of DMU i and j and taking any value in the z direction. Consider the different hyperplanes, $\Pi \in \Psi$, that lie on the efficient frontier as in Table 4.14.

Plane	Equation of plane	Value of U
Π_{ABC}	$-10x_1 - 7x_2 + 11y - 9 = 0$	1.05
$\Pi_{BC(x_1)}$	$-1x_2 + 3y - 17 = 0$	0.875
Π_{x_1}	$-10x_1 + 10 = 0$	3
$\Pi_{AC(x_2)}$	$4x_1 - 3y + 5 = 0$	1.286
Π_{x_2}	$-7x_2 + 7 = 0$	2
$\Pi_{AB(y)}$	$1x_1 + 4x_2 - 9 = 0$	1.8
Π_y	$11y - 77 = 0$	1

Table 4.14: The planes that intersect with the PPS of Example 10 that inefficient DMU D can be compared to.

In this way, we conclude that the smallest amount of uncertainty required for DMU D to become efficient is $u = 0.875$. This occurs when the line BC is moved freely on the x_1 axis and can be seen in Figure 4.9 where D^u is now on the hyperplane $\Pi_{BC(x_1)}^u$ (pale blue hyperplane in Figure 4.9). This agrees with the value calculated from (4.33).

Therefore, to solve the uDEA problem for box uncertainty, the extreme points of the PPS must be found and then the facets can be calculated. However, for increasing numbers of variables, it is not practical to geometrically consider each facet of the efficient frontier. As Olesen and Petersen (2003) conclude, it is inescapable that “all facet generating algorithms proposed so far are exponential clearly poses a serious problem. It is for this reason some times argued that it is simply not practical to investigate the facial decomposition of a polyhedral DEA production possibility set until the day when an efficient algorithm is made available”. As in Section 3.3, we conclude that for large DEA problems it is not currently feasible to explicitly calculate all possible facets of the efficient frontier as it cannot be done in polynomial time.

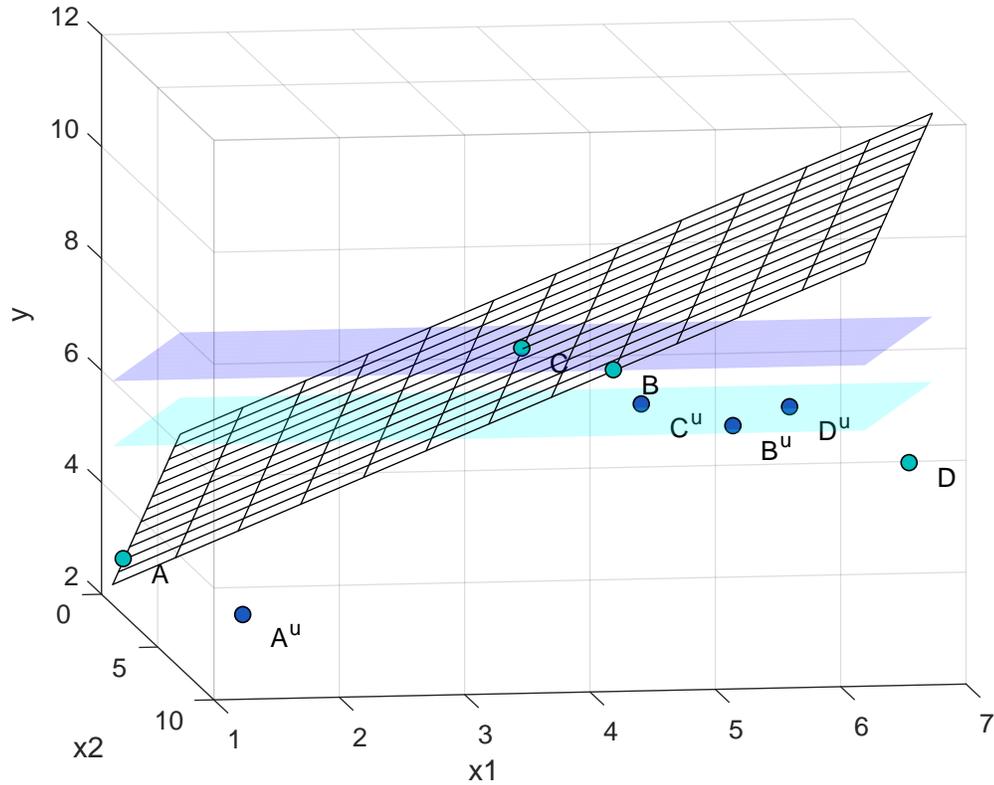


Figure 4.9: The hyperplane $\Pi_{BC(x_1)}$ (dark blue), $\Pi_{BC(x_1)}^u$ (light blue) and Π_{ABC} (cross-hatched). The point defined by DMU D 's uncertain data, D^u is on the hyperplane $\Pi_{BC(x_1)}^u$.

4.4 Future research directions

Here we develop a methodology for solving the uDEA model for box uncertainty. When the facets of the PPS are known the uDEA problem can be solved exactly but as $M + N$ and I increase this become computationally intensive. Therefore, methods that do not rely on the full specification of the efficient frontier's facets would be beneficial.

Ehrgott et al. (2018) provide a first order algorithm for solving the uDEA problem when the uncertainty is ellipsoidal. However, an exact method for all uDEA instances, if it exists, is required. These problems are non-linear and non-convex. Therefore,

further research into heuristic approaches to allow the problem to be solved in a sensible time frame will be beneficial.

In model (4.19), when the amount of uncertainty is not fixed nonlinear terms arise from the λ and θ_k terms. An alternative approach would be to determine a suitable simplified model, that captures the properties of the uDEA model, by introducing binary variables to linearise the nonlinear constraints. This is an area we need to explore further.

Chapter 5

DEA in radiotherapy

5.1 Introduction

The purpose of this chapter is to discuss how to evaluate treatment plans, i.e., we want to undertake a comparison among a set of treatment plans to determine how well they manage to deliver the prescribed dose of radiation to the PTV and how well they can spare OARs from radiation damage at the same time. If we consider the radiation dose to OARs as the “cost” or input necessary to achieve the “profit” or output of radiation dose to the PTV, we can view the problem as one of input-output or performance analysis in an economic setting. This approach has been pioneered by Lin et al. (2013), who use DEA for the evaluation of radiotherapy treatment plans for prostate cancer at Auckland Radiation Oncology, a private radiation therapy centre in Auckland, New Zealand. In the context of the evaluation of radiotherapy treatment plans, DMUs are the treatment plans and DEA assesses how well the plans perform in “transforming inputs into outputs,” i.e., delivering the prescribed dose to the tumour while limiting the dose delivered to OARs. The resulting efficiency score is relative to the set of plans considered in the study. The DEA models we use are introduced in Section 3.1.

Throughout this chapter, we use data from Rosemere. The data consists of prostate patient treatment plans from Preston hospital delivered to patients between November 2015 and July 2018. The plans were planned for delivery on either an AG or BM MLC (see Section 2.1). Due to the continual adaptation of the planning methods, machinery and treatment techniques as the research progressed, we learn that not all the data provided is suitable for inclusion in the analysis. As a result, we use three different datasets, Dataset 1, 2a and 2b as the analysis proceeds. Dataset 1 is included in the appendix, Section A. Dataset 1 comprises 51 plans, Dataset 2a contains 87 plans for patients 1-99 and Dataset 2b is a subset of Dataset 2a consisting of plans for patients 30-99. Not all the patients plans were suitable for inclusion in our data analysis hence we have 87 plans for patients 1-99. Similarly there are only 66 comparable plans for patients 30-99.

In Section 5.2 we investigate different methods for variable selection. In Section 5.3 we apply DEA to find the nominal efficiency scores for each treatment plan.

5.2 Variable selection

To be able to apply DEA to our data, and hence, calculate the treatment plans' efficiency scores, we must first identify suitable variables. As discussed in Section 3.1, DEA is subject to loss of discrimination and hence, we wish to use as small a subset of variables as possible. In Section 5.2.1 we first adapt the methodology of Lin et al. (2013) and conclude that the variables used in Lin et al. (2013) are not suitable for our analysis. As a result, in Section 5.2.2 we identify and explore a subset of variables extracted from Rosemere's clinical protocol. In Section 5.2.3 we apply PCA to Dataset 2a as a variable reduction technique. In Section 5.2.4 we develop an autoPCA variable selection method to allow the technique to be used for hospitals with different protocols. In Section 5.2.5 we discuss other variable selection techniques

before concluding in Section 5.2.6.

In Section 5.2.1 we aim to replicate the study from Lin et al. (2013). Here we use Dataset 1. The data is for the generalised equivalent uniform dose (gEUD) to the rectum, D_{95}^{PTV1} and D_{95}^{PTV3} only as these are the variables Lin et al. (2013) consider. As discussed in Section 5.2.1 it is clear that these variables are not suitable for our study. As a result, in Section 5.2.2 we identify ten inputs and nine outputs that clinicians believe influence treatment plan quality along with eight volumes. These plan variables are listed in Tables 5.1 and 5.2. We note here that when reporting the results the plan numbers are not sequential due to the removal of plans that we learn are not comparable due to missing data, incomplete treatments or a non-standard treatment plan. We use the resulting Dataset 2a for the initial data analysis in Section 5.2.2 and the development of the autoPCA in Section 5.2.4. However, after processing these data Rosemere changed their treatment protocol and as a result, all the BM machines are no longer used for treatment. Although AG and BM plans are considered clinically comparable, we feel that including BM plans in the dataset when they are no longer being used clinically is unsuitable. As a result, we remove the BM plans leaving a dataset of 66 plans that are for patients 30-99 in Dataset 2b. Each plan in this dataset has been used to treat a patient using an AG treatment machine, each plan adheres to the same clinical protocol with the same structures outlined and therefore is considered clinically comparable. As a result, this final dataset is used for the DEA analysis and further research in Chapter 6.

5.2.1 Previous study

As a first step, we try to adopt the model of Lin et al. (2013) and apply it to the initial data from Rosemere, Dataset 1, as done in Stubington et al. (2019). The single input in Lin et al. (2013) is the gEUD rectum, see Section 2.2.1. The input gEUD for the rectum is an averaging quantity that measures the homogeneity of the dose

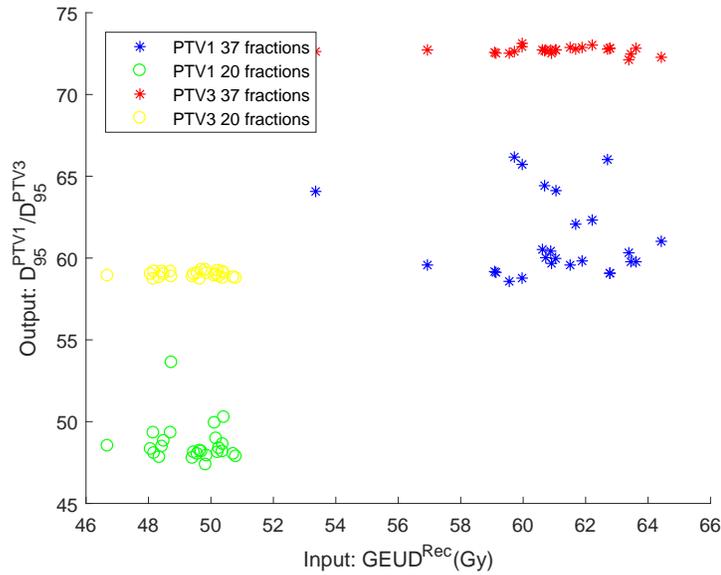


Figure 5.1: gEUD rectum versus D_{95} for PTV 1 and PTV 3 for the 51 treatment plans from Rosemere, Dataset 1.

delivered to the rectum. In their DEA model, the single output is D_{95} for the prostate. We note that the data in Lin et al. (2013) contain only prostate cases with a single PTV, in contrast to the data set from Rosemere Cancer Centre where three PTVs are present. We want to choose a definition of output, and therefore a structure for which to compute D_{95} that is as close as possible to what has been used in Lin et al. (2013). After consultation with clinical staff in charge of the data from Auckland and Rosemere, it was decided that the single PTV from Auckland plans can be considered equivalent to (and is similarly defined as) PTV 3 for the Rosemere data set.

In Figure 5.1 we plot the Rosemere data from Table 1 using the variables from the Auckland study, separately with D_{95} for either PTV 1 or PTV 3 as the single output, and gEUD for the rectum as single input to be analyzed by DEA.

From Figure 5.1 it is clear that the plans are not comparable using the data from Rosemere. There are four distinct clusters, distinguished by the number of fractions and whether PTV 1 or PTV 3 is considered as planning target volume in the D_{95}

measure. Hence, it is impossible to carry out DEA using the data shown in Figure 5.1. We must find a way to control for the two different treatment regimes, and we must consider which inputs and outputs to select. Here we note that neither gEUD for the rectum nor D_{95} for PTV 1 or PTV 3 are recorded according to the Rosemere clinical protocol. We had to extract this information from the plans specifically to produce Figure 5.1. This hinders application of the method in routine hospital processes, since extra effort is needed to extract data. Hence, we are interested in exploring which of the data that are already reported for the protocol (because constraints are placed on them) could and should be used for DEA. Here, we need to select OAR data as inputs and PTV data as outputs in DEA terminology.

5.2.2 Rosemere data exploration

In Section 5.2.1 we have concluded that we cannot perform DEA with the same inputs and outputs considered in Lin et al. (2013) for our data set. As a result, we extract a list of variables from the clinical protocol, see Section 2.1.2 for more details. In Table 5.1 we have identified 19 different variables, ten inputs and nine outputs that clinical staff consider relevant for quality assessment of the radiotherapy plans obtained. Here, V_x denotes the percentage of the volume of an organ at risk (either rectum or bladder) receiving x Gray (Gy) or more of radiation dose and D_x is the dose received in Gy by $x\%$ of the volume of a PTV. If we need to refer to these values for a particular structure, we will use the structure name as a superscript, shortening PTV1-PTV2 and PTV2-PTV3 to PTV1-2 and PTV2-3. In addition to these dose-volume values, we extracted data on the total volume in cm^3 of the rectum, bladder, prostate and PTVs, Table 5.2. Radiotherapy planners at Rosemere suggested that these might influence treatment plan quality. Since we only have 87 DMUs (plans), the use of 26 variables is considered excessive (see, e.g. Coelli et al. (2005) for more information) and will not lead to meaningful results from DEA. Hence, in Section 5.2.3 we attempt

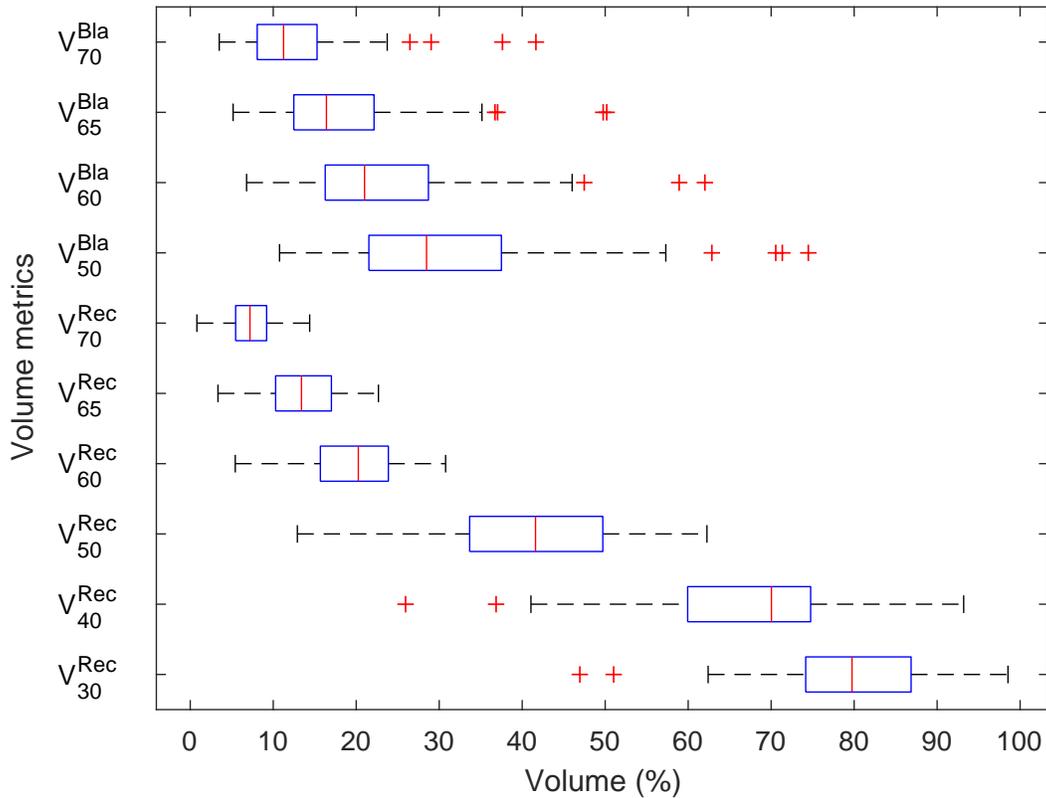


Figure 5.2: V_x input data, Dataset 2a.

to identify a subset of measures listed in Table 5.1 which account for most of the variability present in our data set and which allow a sound application of DEA. To begin with we first explore the variables identified from the clinical protocol listed in Table 5.1 and the volume variables in Table 5.2.

We obtained 87 distinct treatment plans, 52 of which were 20 fraction treatments and 35 of which were 37 fraction treatments. We visualise the data in the boxplots of Figures 5.2 and 5.3. Notice that the output data in Figure 5.3 is dose per fraction to adjust for the number of treatment fractions, this will be explained further in Section 5.2.3.

Figure 5.2 shows that there is more variability within the lower rectal variables $V_{30} - V_{40}$ than the higher rectal constraints $V_{50} - V_{70}$ data suggesting Rosemere pri-

Structure	37 fractions	20 fractions
Rectum	V_{30}	$V_{24.6}$
	V_{40}	$V_{32.4}$
	V_{50}	$V_{40.8}$
	V_{60}	$V_{48.6}$
	V_{65}	$V_{52.8}$
	V_{70}	V_{57}
Bladder	V_{50}	$V_{40.8}$
	V_{60}	$V_{48.6}$
	V_{65}	$V_{52.7}$
	V_{70}	$V_{56.8}$
PTV 1-2	D_{99}	D_{99}
	D_{98}	D_{98}
	D_{50}	D_{50}
PTV 2-3	D_{99}	D_{99}
	D_{98}	D_{98}
	D_{50}	D_{50}
PTV 3	D_{99}	D_{99}
	D_{98}	D_{98}
	D_{50}	D_{50}

Table 5.1: Variables obtained from the clinical protocol.

Structure volume	Notation
Rectum	Vol^{Rec}
Bladder	Vol^{Bla}
Prostate	Vol^{Pros}
PTV 1-2	Vol^{PTV1-2}
PTV 2-3	Vol^{PTV2-3}
PTV 3	Vol^{PTV3}

Table 5.2: Total volume in cm^3 of structures from the clinical protocol.

oritise meeting the higher rectal constraints consistently. Figure 5.3 shows that there is more variability within the D_x^Y data for PTV1-2 than PTV 3, where there is little variability, suggesting that Rosemere is very good at meeting the higher dose to the internal PTV 3. The PTV1-2 variables are slightly positively skewed meaning that the median is typically less than the mean. The range of the PTV1-2 data is also larger than that of the remaining output variables.

In Figure 5.4 we observe that the largest variability in volume is in the bladder, this may suggest it is a suitable variable to consider as an environmental variable in our DEA models. All the prostate and associated PTV volumes are positively skewed meaning the distribution has a larger mean than median.

To explore the data further and to understand the underlying relationships between the variables we calculate the Pearson's correlation coefficient, see Section 3.5. We then plot the pairwise relationship between each input or output variable together with the correlation coefficient in Figures 5.5 and 5.6. The diagonal of the plot shows the distribution of each variable and the off diagonal shows the correlation between pairs of variables. To determine whether the correlation between variables is statistically significant we test the hypothesis at the 0.05 significance level that the correlation coefficient is significantly different from zero. The correlation coefficients

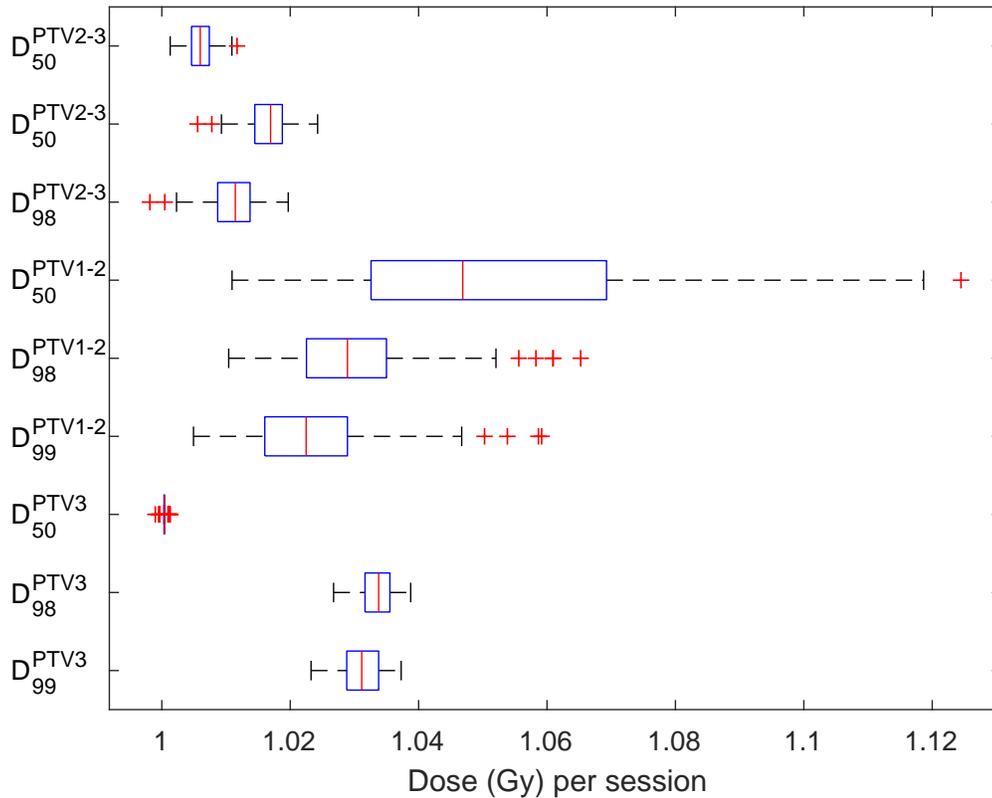


Figure 5.3: D_x output data, Dataset 2a.

in red in Figures 5.5 and 5.6 indicate the variable pairs that are statistically greater than zero.

In Figure 5.5 we note that as expected there are no negative correlations between the input variables. This means we do not have conflicting variables within the input set and they all behave according to Definition 3.1.3. All four bladder constraints are highly correlated with correlation coefficients between 0.94 and 0.99 for all pairs. However, the rectal constraints split into two groups of higher and lower rectal constraints: $V_{60} - V_{70}$ are strongly correlated with correlation coefficients in the range 0.92-0.97 and $V_{30} - V_{40}$ are strongly correlated with correlation coefficient 0.85. However, the correlation between these two groups is weaker, for example, the correlation between V_{30} and V_{70} is only 0.43.

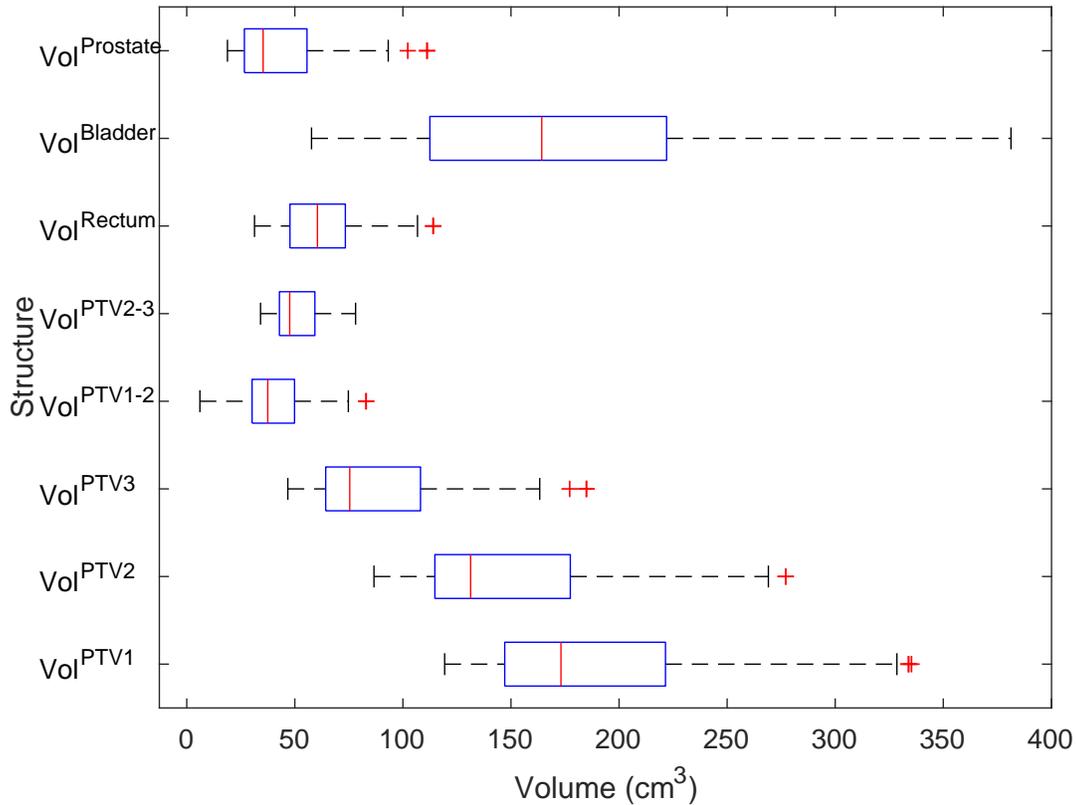


Figure 5.4: Volume data, Dataset 2a.

Looking at Figure 5.5 it would be natural to look at the correlation between V_{60}^{Rec} , V_{65}^{Rec} and V_{70}^{Rec} which have pairwise correlation of 0.92, 0.97 and 0.97 and deem them similar enough that only one of these variables is required. However Dyson et al. (2001) demonstrate that removing variables based purely on their correlation can lead to dramatic differences in results. Therefore, less adhoc methods are required and explored here.

In Figure 5.6 we look at the relationship between the output variables. Here we notice that for all three PTVs, the D_{99} and D_{98} are highly correlated with correlation coefficients between 0.98 and 0.99. The corresponding D_{50} for PTV1-2 and PTV2-3 with D_{99} and D_{98} are fairly high: 0.63, 0.64, 0.7 and 0.74. However, the correlation between D_{99}^{PTV3} and D_{98}^{PTV3} with D_{50}^{PTV3} is not significantly different from zero. There

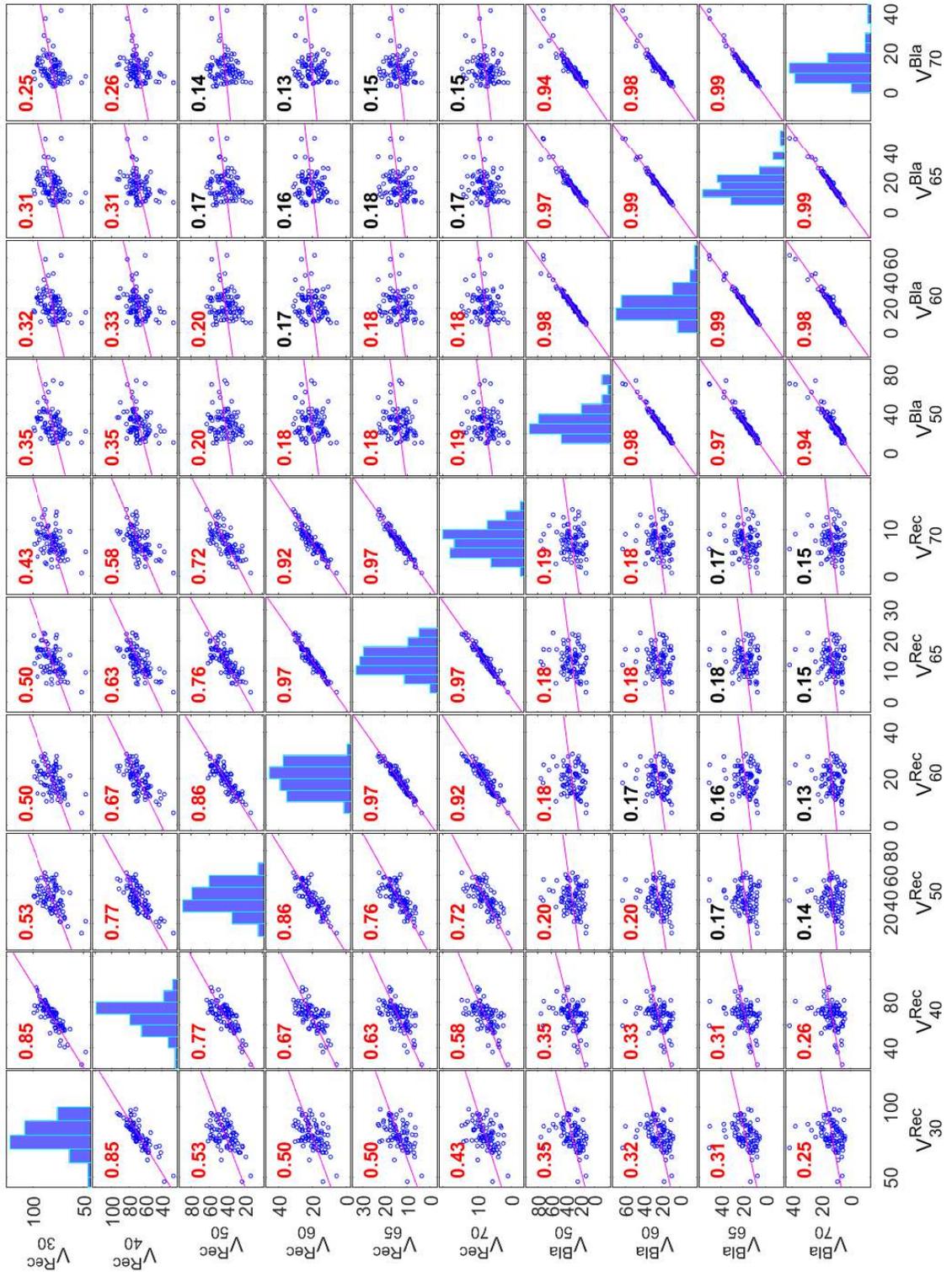


Figure 5.5: Correlation between the input variables, Dataset 2a

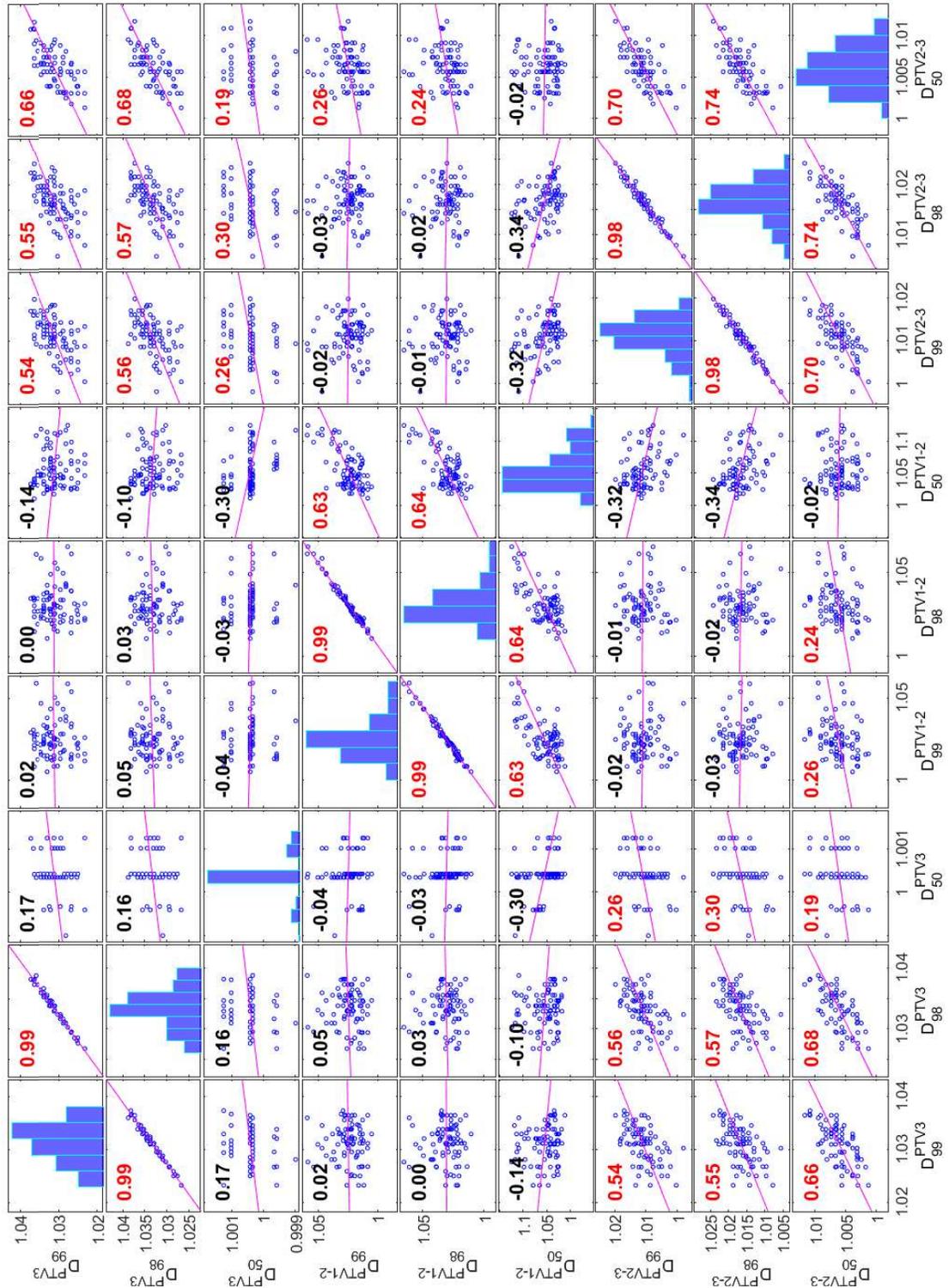


Figure 5.6: Correlation between the output variables, Dataset 2a

is also no correlation between PTV 3 and PTV1-2 for all three D_x metrics.

The diagonal plots in both the input and output distributions in Figures 5.5 and 5.6 appear to be fairly normally distributed. With more data, we would expect this distribution to be more pronounced.

5.2.3 Applying PCA to prostate cancer data

In Section 3.5, we introduced the general PCA model as a data reduction technique. Here we describe how we have applied it to the data set of plans obtained from Rosemere before describing the method used to select variables for DEA.

There are some studies that use the derived PCs directly in the DEA model (see Section 3.5.1). However, this is not suitable in a radiotherapy treatment plan context as taking a vector of treatment plan metrics does not transcribe to a clinical meaning. Hence, we wish to use it as a decision making aid to select variables to be used in our DEA models.

PCA is susceptible to the sample currently being used; depending on the sample the variables selected can be very different. It is important to make sure we have accounted for the effect the current sample has on variable selection. Since it is the OAR and PTV variables that define the quality of a treatment, we first perform PCA for the OAR and PTV variables and ignore the total volume variables. Throughout we use the correlation matrix instead of the sample variance matrix to overcome the sensitivity of PCs to units of measurements, this is explained in more detail in Section 3.5 where PCA is first introduced.

We apply PCA to the output variables and plot the first two PCs against each other. This is shown in Figure 5.7 where each point represents an observation, i.e., a treatment plan. The fact that two distinct clusters of plans form in Figure 5.7 suggests there is some underlying data structure. Looking into which plan corresponds to each point we realise that the group on the left of Figure 5.7 is all the 20 fraction plans

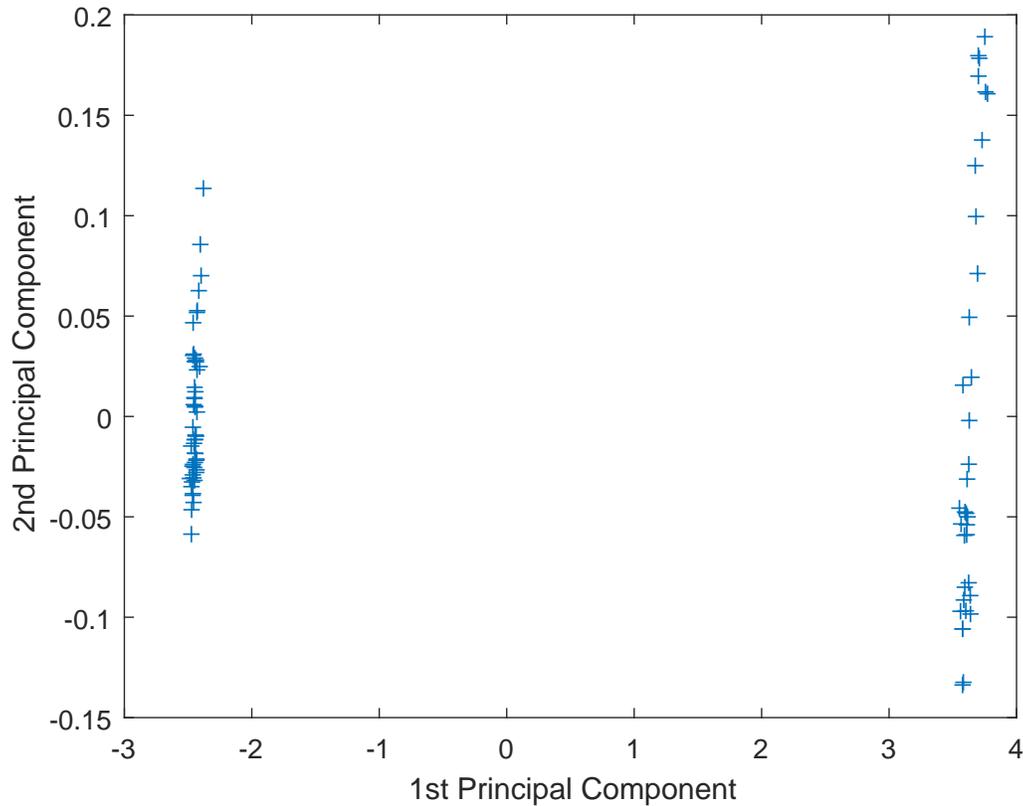
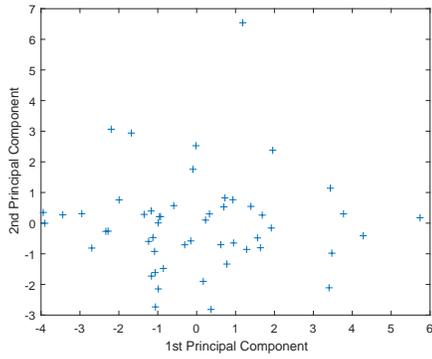


Figure 5.7: First two PCs for the output variables, Dataset 2a.

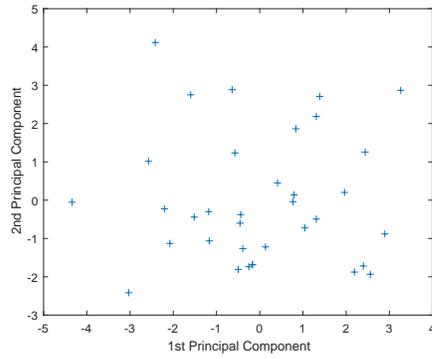
and the 37 fraction plans are on the right. This suggests that we need to account for the number of fractions in a treatment plan before we can apply our variable selection methods. We next reproduce the component plot to compare the effect of the number of fractions by plotting the first two PCs for the 20 and 37 fraction plans separately in Figure 5.8. The plot in Figure 5.8 shows that the two clear groups in the data do not exist when we split the data according to the number of fractions.

We repeat this for the input data in Figure 5.9 where the two distinct clusters do not form and we can see no discernible difference when we plot the 20 and 37 fractions separately in Figure 5.10.

Figures 5.7 to 5.10 suggest that it is variation within the output data, not the input data, that causes the two clusters to form in the PCA. This is caused by the



(a) 20 fraction plans outputs.



(b) 37 fraction plans outputs.

Figure 5.8: First two PCs for the outputs according to session number, Dataset 2a.

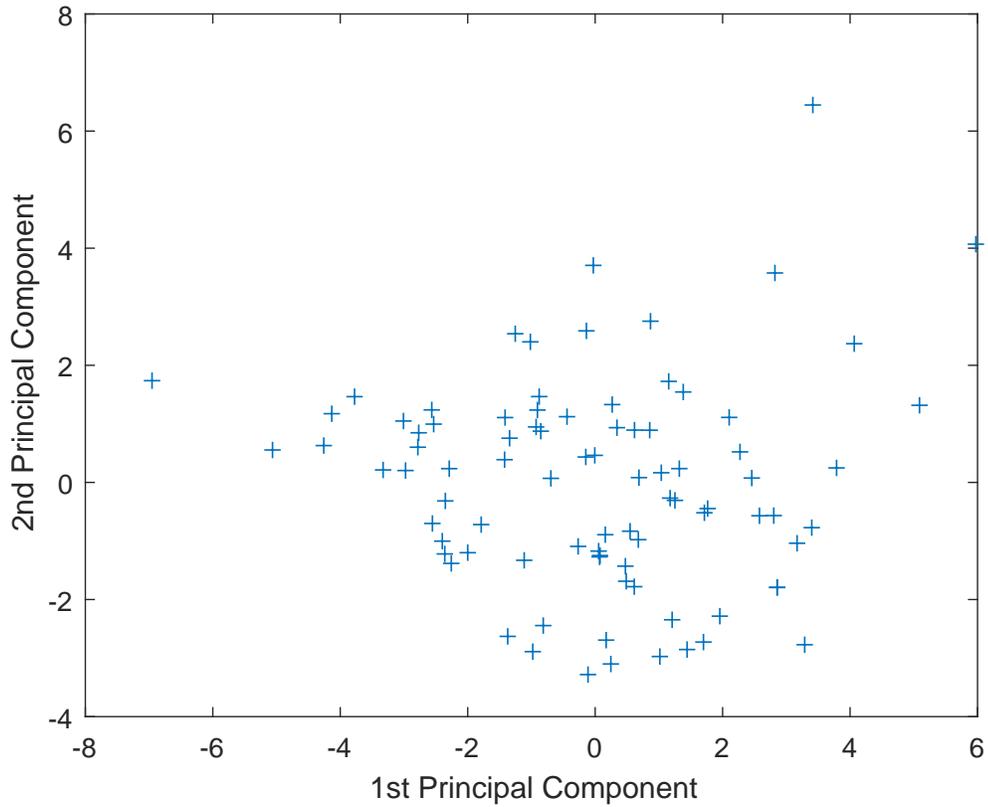
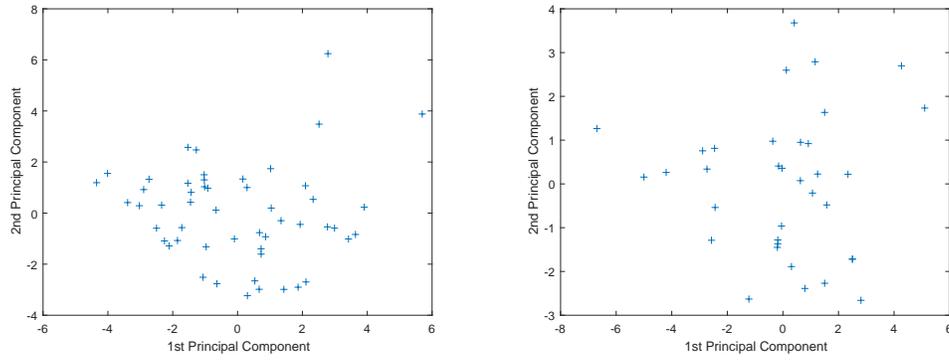


Figure 5.9: First two PCs for the input variables, Dataset 2a.



(a) 20 fraction plans inputs.

(b) 37 fraction plans inputs.

Figure 5.10: First two PCs for the inputs according to session number, Dataset 2a.

definition of the OAR and PTV data, i.e. the OAR data are V_x metrics, so that they are already taking into account the difference in the number of fractions as the percentage volumes receiving, for example, 50 Gy in 37 fractions and 40 Gy in 20 fractions. The PTV data are D_x metrics measured as the dose in Gy received by a certain percentage of the volume of a PTV. Due to the different prescribed dose for 37 and 20 fraction treatments, their values are very different for the two different treatment regimes.

Furthermore, Figure 5.8 suggests that performing PCA for the sets of plans following the two different treatment regimes separately removes the two clusters. Hence, by accounting for the difference between 20 and 37 fraction plans, the distinct clusters no longer form in PCA and we can proceed. To account for the difference between the treatment regimes, we note that the prescribed dose per fraction as well as the total prescribed dose is different for both treatment regimes. Hence, simply dividing by the number of fractions will not eliminate the clusters. We therefore divide PTV D_x data by the prescribed dose to normalise data between the two types of treatment. This introduces ratio variables which can be problematic for DEA methodology, this will be discussed in Section 5.3. From now on, we use the normalised PTV data throughout, however, we will still refer to them with the same variable names.

Here we consider variable selection techniques for the inputs and outputs separately as opposed to all at once. This is because we know that the inputs and outputs are closely linked, the higher the dose to a PTV the higher the dose to an OAR. This could cause us to get unmeaningful results if we apply PCA to the variables all together. This is in part due to us requiring at least one input and output variable for the DEA model in Section 5.3. Therefore, here we keep the variables separate to ensure we have considered the main contributor to both the inputs and outputs for our model.

Simplified + - notation

We now apply the simplified notation of Jolliffe (2002) to look at the patterns occurring in the PCs. As introduced in Section 3.5 the raw PCs are often hard to interpret and evaluate, therefore, the simplified notation can be a beneficial tool.

The first four PCs from applying PCA to the input data of Dataset 2a are on the left hand side of Table 5.3 and the simplified notation from Jolliffe (2002) gives the representation on the right hand side. This data simplification highlights some of the hidden patterns in the dataset. PC 1 has significant positive coefficients for all variables; this suggests that the first PC is a measure of overall dose deposition to the OAR. As discussed in Section 5.2.2 from the correlation matrix Figure 5.5 for the inputs we can see 3 main groups forming: V_{30} and V_{40} rectum, bladder constraints and the higher rectal constraints $V_{50} - V_{70}$. This is reflected in the PCs' scores. Here the second PC splits the data into bladder constraints and higher rectal constraints. From this, it would seem sensible to choose one bladder constraint and one higher rectal constraint to explain the inputs. The third PC eigenvalue is 0.89, which is less than one so contributes less than one of the original variables and so could justify excluding PC 3 according to Kaiser's rule (see Section 3.5.1).

Repeating the above PC simplification for the output data gives the first four PCs

Variable	PC 1	PC 2	PC 3	PC 4	PC 1	PC 2	PC 3	PC 4
V_{30}^{Rec}	0.30	-0.10	0.67	-0.43	+	(-)	+	(-)
V_{40}^{Rec}	0.35	-0.16	0.50	0.17	+	(-)	+	
V_{50}^{Rec}	0.32	-0.27	0.02	0.76	+	-		+
V_{60}^{Rec}	0.34	-0.31	-0.24	0.04	+	-	(-)	
V_{65}^{Rec}	0.34	-0.30	-0.29	-0.29	+	-	(-)	(-)
V_{70}^{Rec}	0.32	-0.28	-0.36	-0.33	+	-	-	(-)
V_{50}^{Bla}	0.30	0.38	-0.02	0.00	+	+		
V_{60}^{Bla}	0.30	0.39	-0.06	0.05	+	+		
V_{65}^{Bla}	0.30	0.40	-0.07	0.00	+	+		
V_{70}^{Bla}	0.28	0.41	-0.13	0.00	+	+		
EigenValues	5.35	3.22	0.89	0.33	5.35	3.22	0.89	0.33
% Explained	53.54	32.19	8.92	3.29	53.54	32.19	8.92	3.29

Table 5.3: First four PCs for the input variables, Dataset 2a.

Variable	PC 1	PC 2	PC 3	PC 4	PC 1	PC 2	PC 3	PC 4
D_{99}^{PTV3}	-0.42	0.05	0.40	-0.36	-		+	-
D_{98}^{PTV3}	-0.43	0.06	0.40	-0.34	-		+	-
D_{50}^{PTV3}	-0.18	-0.10	-0.69	-0.64	(-)		-	-
$D_{99}^{\text{PTV1-2}}$	-0.01	0.60	-0.18	-0.02		+	(-)	
$D_{98}^{\text{PTV1-2}}$	0.00	0.59	-0.20	0.00		+	(-)	
$D_{50}^{\text{PTV1-2}}$	0.15	0.49	0.20	-0.01	(-)	+	(-)	
$D_{99}^{\text{PTV2-3}}$	-0.44	-0.03	-0.20	0.42	-		(-)	+
$D_{98}^{\text{PTV2-3}}$	-0.45	-0.04	-0.22	0.38	-		(-)	+
$D_{50}^{\text{PTV2-3}}$	-0.43	0.18	0.00	0.13	-	(+)		
EigenValues	3.94	2.60	1.04	0.79	3.94	2.60	1.04	0.79
% Explained	43.82	28.93	11.57	8.83	43.82	28.93	11.57	8.83

Table 5.4: First four PCs for the output variables Dataset 2a.

in Table 5.4. Here the first two PCs explain 72.7%, the first three 84.3% and the first four 93.2% of the variation in the data. The simplification on the right hand side of Table 5.4 suggests the first two PCs distinguish between PTV 3 and PTV2-3 variables in the first PC and PTV1-2 in the second PC. From this and the correlation matrix Figure 5.6 there appears to be no discernible difference between the PC scores for D_{98} vs D_{99} variables. This may suggest we can use D_{99} or D_{98} but both are not required. To analyse this further we repeat the PCA excluding D_{99} and D_{98} values in turn. This gives the results in Tables 5.5. The same patterns are seen when using both D_{98} and D_{99} so it would suggest it is only necessary to use one of these metrics.

Correlation between variables and PCs

Another method that can be used to identify patterns in the dataset is to compare the correlation between the original variables and the resulting PCs. For Dataset 2a the input values are in Table 5.6. The first PC is highly correlated with all the 10 variables again suggesting the first PC is a measure of overall size. The second PC is positively correlated with the bladder metrics and negatively correlated with the higher rectal metrics, agreeing with the previous analysis.

This is repeated for the output variables in Table 5.7. Here the first PC is highly negatively correlated with D_{99}^{PTV3} and D_{98}^{PTV3} and all PTV2-3 metrics. The second PC is highly correlated with the D_{99}^{PTV1-2} and D_{98}^{PTV1-2} variables, suggesting it is the higher PTV constraints that are most significant. We note that there is very little correlation between the final three PCs and the original variables; the largest correlation is 0.09. In fact, from PC 3 downwards there is no correlation greater than 0.57, suggesting that most of the variation in the data has been accounted for.

The methods discussed in this section are somewhat subjective. We can identify similar patterns occurring across the different techniques, however, they do not provide a method for choosing suitable variables without user input. Therefore, we now

Variable	PC 1	PC 2	PC3	PC 1	PC 2	PC 3
D_{99}^{PTV3}	-0.51	0.12	-0.25	-		(-)
D_{50}^{PTV3}	-0.28	-0.19	0.90	-	(-)	+
D_{99}^{PTV1-2}	0.02	0.68	0.31		+	(+)
D_{50}^{PTV1-2}	0.25	0.64	0.04	(+)	+	
D_{99}^{PTV2-3}	-0.55	0.01	-0.11	-		
D_{50}^{PTV2-3}	-0.54	0.29	-0.10	-	(+)	
Eigenvalues	2.45	1.72	0.88	2.45	1.72	0.88
% Explained	40.76	28.64	14.60	40.76	28.64	14.60

(a) Excluding D_{98} variables.

Variable	PC 1	PC 2	PC3	PC 1	PC 2	PC 3
D_{98}^{PTV3}	-0.50	0.14	-0.27	-		(-)
D_{50}^{PTV3}	-0.28	-0.20	0.89	-	(-)	+
D_{98}^{PTV1-2}	0.02	0.67	0.33		+	(+)
D_{50}^{PTV1-2}	0.23	0.64	0.04	(+)	+	
D_{98}^{PTV2-3}	-0.57	-0.01	-0.08	-		
D_{50}^{PTV2-3}	-0.54	0.28	-0.11	-	(+)	
Eigenvalues	2.51	1.73	0.88	2.51	1.73	0.88
% Explained	41.78	28.77	14.60	41.78	28.77	14.60

(b) Excluding D_{99} variables.

Table 5.5: First three PCs for the output variables, Dataset 2a.

Variable	PC 1	PC 2	PC 3	PC 4	PC 5	PC 6	PC 7	PC 8	PC 9	PC10
V_{30}^{Rec}	0.69	-0.18	0.64	-0.25	-0.14	0.00	0.04	0.00	0.00	0.00
V_{40}^{Rec}	0.80	-0.29	0.47	0.09	0.21	-0.03	-0.06	0.00	0.00	0.00
V_{50}^{Rec}	0.75	-0.48	0.02	0.44	-0.08	0.01	0.07	0.02	0.00	0.00
V_{60}^{Rec}	0.78	-0.56	-0.23	0.03	-0.10	0.01	-0.11	-0.05	-0.01	0.00
V_{65}^{Rec}	0.78	-0.53	-0.28	-0.17	-0.01	-0.03	-0.05	0.07	0.01	0.00
V_{70}^{Rec}	0.75	-0.51	-0.34	-0.19	0.10	0.03	0.11	-0.03	0.00	0.00
V_{50}^{Bla}	0.70	0.68	-0.02	0.00	0.01	0.18	-0.02	0.02	-0.03	-0.01
V_{60}^{Bla}	0.70	0.71	-0.06	0.03	0.00	0.03	0.00	-0.01	0.04	0.03
V_{65}^{Bla}	0.69	0.72	-0.07	0.00	-0.01	-0.05	0.00	-0.01	0.03	-0.03
V_{70}^{Bla}	0.65	0.73	-0.12	0.00	-0.01	-0.15	0.02	0.00	-0.04	0.01

Table 5.6: Correlation between the original input variables and the PCs, Dataset 2a.

Variable	PC 1	PC 2	PC 3	PC 4	PC 5	PC 6	PC 7	PC 8	PC 9
D_{99}^{PTV3}	-0.84	0.07	0.41	0.33	-0.07	0.07	0.00	0.03	-0.04
D_{98}^{PTV3}	-0.85	0.10	0.40	0.30	-0.04	0.08	0.00	-0.03	0.04
D_{50}^{PTV3}	-0.35	-0.16	-0.71	0.57	0.14	0.02	0.00	0.00	0.00
$D_{99}^{\text{PTV1-2}}$	-0.01	0.96	-0.19	0.02	-0.20	0.01	0.00	-0.06	-0.03
$D_{98}^{\text{PTV1-2}}$	0.00	0.96	-0.20	0.00	-0.19	0.04	0.01	0.06	0.03
$D_{50}^{\text{PTV1-2}}$	0.31	0.80	0.20	0.01	0.47	0.11	0.00	0.00	0.00
$D_{99}^{\text{PTV2-3}}$	-0.88	-0.05	-0.20	-0.37	0.04	0.19	-0.08	0.00	0.00
$D_{98}^{\text{PTV2-3}}$	-0.90	-0.07	-0.22	-0.34	0.06	0.10	0.09	-0.01	0.00
$D_{50}^{\text{PTV2-3}}$	-0.84	0.29	0.00	-0.11	0.12	-0.42	-0.01	0.00	0.00

Table 5.7: Correlation between the original output variables and the PCs, Dataset 2a.

explore alternative PCA based methods in the following sections.

5.2.4 Automated PCA Variable Selection

We now develop a method of variable selection based on PCA to overcome some of the problems discussed in Section 3.5 and to strengthen the observations from Sections 5.2.2 and 5.2.3. We wish to create a process for variable selection that can be automated. This will allow us to tailor the method to different hospitals where a different set of variables may accurately explain the protocol. We wish to develop a method that selects variables as opposed to PCs to use in the DEA model.

In Table 5.3 we noted that the first PC is a measure of overall dose to the OARs. The coefficients all take values between 0.28 and 0.35 and are considered significant via the representation in (3.23) as introduced in Jolliffe (2002). We argue that these suggest that there is very little variability between the PCA coefficients and hence, selecting a variable based on this PC will not be beneficial. However, this recommendation is based on comparison with the other PCs which have a larger coefficient range. For example, PC 2 has coefficients between $[-0.23, 0.41]$ and PC 3's coefficients are between $[-0.36, 0.67]$. If these PCs also have very little variation in the coefficients we do not want to exclude PC 1. In this way, we wish to consider both the range of the coefficients and the variance between them. Table 5.8 gives the variance between a PC's coefficients and the range. To three decimal places PC 1 has variance 0 whereas, the remaining PCs have a variance of 0.111. This suggests that there is very little variability in the coefficients for the first PC compared to future PCs. Whereas, in Table 5.9 we see that the variance for PC 1 is less different to the other PCs. We need a measure that captures the notion of a small/large variance in the context of the current data. From this, we can determine whether a PC contains unique information about the variance in the variables.

Another consideration is the occurrence of $+ve$ and $-ve$ coefficients in the PCs. We

Principal component	Variance	Mean	Absolute mean	Max	Min
1	0.000	0.316	0.316	0.346	0.282
2	0.111	0.016	0.301	0.407	-0.310
3	0.111	0.001	0.237	0.673	-0.364
4	0.111	-0.003	0.209	0.763	-0.434
5	0.111	-0.007	0.221	0.697	-0.460
6	0.111	-0.003	0.208	0.719	-0.620
7	0.111	0.008	0.246	0.578	-0.567
8	0.111	0.000	0.221	0.728	-0.528
9	0.111	-0.003	0.228	0.539	-0.558
10	0.111	0.000	0.182	0.654	-0.718

Table 5.8: Input PC metrics, Dataset 2a.

Principal component	Variance	Mean	Absolute mean	Max	Min
1	0.058	-0.245	0.279	0.155	-0.453
2	0.080	0.200	0.238	0.595	-0.099
3	0.122	-0.054	0.277	0.403	-0.693
4	0.122	-0.051	0.256	0.417	-0.644
5	0.120	0.064	0.253	0.804	-0.338
6	0.123	0.041	0.229	0.382	-0.846
7	0.125	-0.001	0.188	0.660	-0.737
8	0.125	-0.001	0.231	0.602	-0.601
9	0.125	0.006	0.229	0.596	-0.610

Table 5.9: Output PCs metrics, Dataset 2a.

cannot simply select the largest *+ve* contribution as the sign of any PC is completely arbitrary. Similarly, considering only the absolute largest contribution could result in us excluding an equally significant variable. For example, in Table 5.4 PC 3 D_{50}^{PTV1-2} has coefficient 0.2 and D_{50}^{PTV2-3} has coefficient -0.2. This means they contribute equally to PC 3. The occurrence of both + and - in the same PC suggests we need to consider both as potential variables. In addition to this, we must decide on the overall percentage variance we wish to explain.

Algorithm 1 provides a framework for variable selection, we will refer to it as autoPCA. We define three parameters α , β and γ , which are dependent on the data and application of the problem. They can be thought of as measures of the significance of the variation in the PCs' in terms of the variance between PCs, the sign of the coefficients in an individual PC and the overall percentage variance we wish to explain.

To choose the variables for our DEA model we take samples of size k from our dataset. For each sample, we consider each PC in turn. For PC j we first determine whether the variance is statistically significant compared to the variance of the other PCs. If it is not, we move on to PC $j + 1$ without selecting a variable. If it is significant, we identify whether there are significant positive and negative coefficients. If there are, we choose two variables, the variable with the largest positive contribution and the variable with the largest negative contribution. Otherwise, we select a single variable with the highest absolute value. We then calculate the total percentage variance explained so far from the sum of the eigenvalues of the so far considered PCs. If the total desired percentage variance has been met then we stop the variable selection process, otherwise we continue to PC $j + 1$. This process is repeated for a total of r repetitions and for each sample the variables selected from each PC and the eigenvalue of the corresponding PC are recorded. In this way, the number of times a variable is selected and the mean eigenvalue of each PC is determined. This will be explored later in this section when we look at the results of Algorithm 1. In

Algorithm 1 $max(PC_j)(min(PC_j))$ are the PC with the largest(smallest) coefficient. We note here that in both Algorithm 1 and Algorithm 3 we use samples of size k where k is less than the total dataset. This is to replicate the acquisition of data. The data from Rosemere is received in groups of 10-30 patients. Therefore, we want to derive the sampling distribution of the ‘population’ from a smaller sample to ensure the variables are being selected sensibly and robustly.

Algorithm 1: AutoPCA variable selection.

Input : Data=Variables to run PCA on, r =number of samples, k = observed sample size, α =significance level for the variation in PC_j , γ = the total required % variation to be explained by the PCs

Output: Variables selected, mean eigenvalues

```

for  $i=1:r$  do
    Select sample of size  $k$ 
    Variance-Explained=eigenvalue(PC $j$ )
    while Variance-Explained <  $\gamma$  do
         $j=1$ 
        if  $Var(PC_j) > \alpha$  then
            if there are significant +ve and -ve variables then
                Choose  $max(PC_j)$  and  $min(PC_j)$ 
            else
                Choose  $max|PC_j|$ 
            end
             $j=j+1$ 
            Variance-Explained=Variance-Explained + eigenvalue(PC $j$ )
        else
             $j=j+1$ 
            Variance-Explained=Variance-Explained + eigenvalue(PC $j$ )
        end
    end
end

```

We now discuss how to determine the parameters α, β and γ . Here we refer to them as significance levels.

Choosing α

To choose a suitable value for α , the significance level for variation in PC j the expected variation in the data is required. This can be found by taking repeated samples of size k to determine the sampling distribution of each PCs' standard deviation. In this way, we can determine the most likely distribution of the PCs variance. In the following let s_j^2 denote the sample variance and s_j the sample standard deviation of the j^{th} PC. We are looking to find the distribution of the sample standard deviation for PC $_j$ here we denote this as $\Pi_j(s)$. This is done via Algorithm 2. Random samples of size k are taken and PCA is performed. The standard deviation of each PC is then calculated and stored. This is repeated r times. This is similar to bootstrapping methodology however here instead of using 66 samples with replacement we here use samples of size k where k is the observed sample size. The number of repetitions r should be chosen to be suitably large such that r samples are unlikely to have many repeated samples.

Algorithm 2: Deriving the sampling distribution of each PCs' standard deviation

Input : Data=Variables to run the PCA on, r =number of samples,
 k =observed sample size

Output: $\Pi_j(s)$ =Distribution of PC $_j$'s standard deviation

for $i=1:r$ **do**

 Randomly select sample of size k from Data

 Run PCA on sample

 Calculate the standard deviation s_j of each PC and store

end

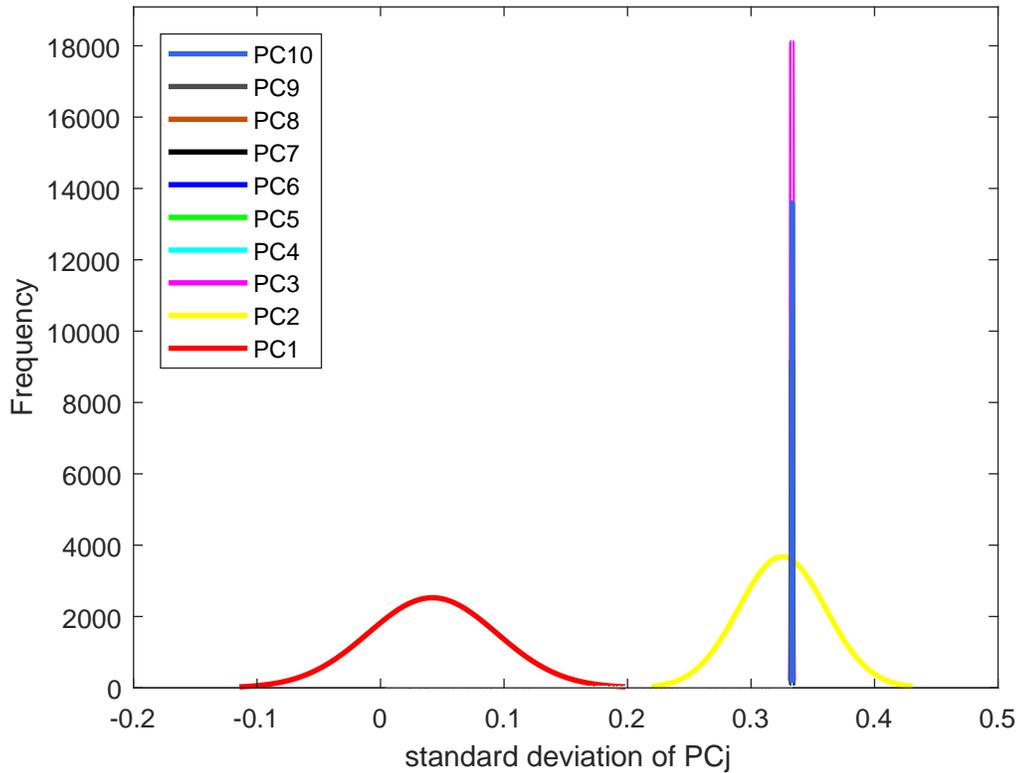


Figure 5.11: Distribution of the inputs PCs' standard deviation, Dataset 2a.

We apply Algorithm 2 to the inputs of Dataset 2a with $k = 35$ and $r = 1,000,000$. We note that there are over 2.5×10^{24} possible unique samples of size 35 from a dataset of 87. Figure 5.11 shows the inputs PCs' standard deviation distributions. Examination of Figure 5.11 suggests that the first PC (red) has a distribution with lower mean than the other PCs. We can see the distribution is further to the left than the other PCs' distribution. We then apply Algorithm 2 to the outputs of Dataset 2a, the results are shown in Figure 5.12. Here the first PC's standard deviation distribution (red) is less different than the other PCs standard deviation distribution compared to Dataset 2a's inputs. Although we can see this difference by examining the plots, we require a variable selection technique that does not require user input or subjectivity. This will allow the process to be implemented easily and automatically for different subsets of variables from different clinical protocols. Therefore, once

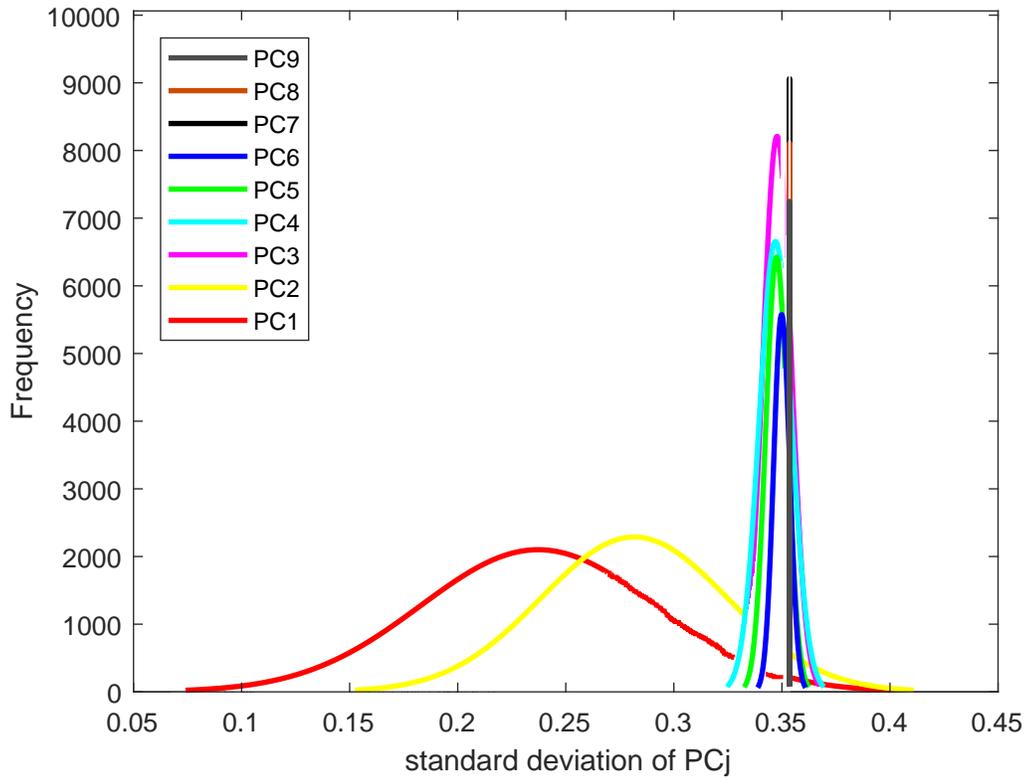


Figure 5.12: Distribution of the outputs PCs' standard deviation, Dataset 2a.

$\Pi_j(s)$, the distribution of the sample standard deviation for PC j ,+ has been found for each PC we must compare the first PC to the remaining PCs to look for a significant difference. To do this we must first define what we consider as a significant difference. We use the derived distributions for the PCs' standard deviation from Algorithm 2 to do this

As r , the number of replications, increases the distribution of the standard deviation becomes normally distributed. For the normal distribution, we expect 99.7% of the values to be within three standard deviations of the mean. Therefore, if there are a significant number of observations for PC 1 inside this region for each PC $_j$ $j \neq 1$ we conclude that they are not significantly different, otherwise we conclude they are. This process is explained in Algorithm 3. Here we assume that the distribution of PC j 's standard deviation has mean μ_j and standard deviation σ_j and roughly follows

$\Pi_j \sim \mathcal{N}(\mu_j, \sigma_j^2)$. Therefore, for each PC j we calculate $\nu_j = \mu_j - 3\sigma_j$ and then count the number of times in our r samples $\sigma_1 > \nu_j$. If this happens in less than 0.05 of the cases then we conclude PC 1 has a standard deviation significantly different from PC j . If PC 1 has a standard deviation that is statistically different from all the remaining PCs then we conclude that we should exclude PC 1 from our variable selection. However, because we are sampling even if the overall dataset has a significantly different s_1 sometimes a sample may not have a significantly different value of s_1 . Hence, we now define α to be used in Algorithm 1 based on the values of ν_j ,

$$\alpha = \min \nu_j \quad j = 2, \dots, N. \quad (5.1)$$

Therefore, during the autoPCA process if a sample has a PC 1 with $\sigma_1 > \alpha$ for that sample a variable will be chosen from the first PC. This will be seen later when the results of PCA are presented for the inputs in Table 5.10 where a variable is selected from PC 1 only 3.8% of the time.

Algorithm 3: Determining the value of α

Input : $\Pi_j(s)$ $j = 1, \dots, N$: the distribution of each PCs standard deviation

Output: Whether or not to include PC 1

for $j=2:N$ **do**

 Calculate: $\nu_j = \mu_j - 3\sigma_j$

κ =number of times $\sigma_1 > \alpha_j$

if $\frac{\kappa}{R} < 0.05$ **then**

 | PC 1's standard deviation is significantly lower than PC j 's

end

end

if PC 1's standard deviation is significantly lower for all PC j , $j = 2, \dots, N$

then

 Calculate α from (5.1)

 Conclude PC 1 has a standard deviation significantly lower than all other

 PCs and so should be excluded from variable selection when $\sigma_1 < \alpha$.

end

Running Algorithm 3 on the inputs of Dataset 2a results in the conclusion that PC 1 has a standard deviation significantly lower than all other PCs and so the first PC is excluded from the results. However, for the outputs we conclude that the standard deviation of the first PC's eigenvalue is not significantly lower than the other PCs' eigenvalues standard deviation and so variable selection proceeds always using the first PC.

Choosing β

Next, we determine a suitable value for β that is used to determine if there are significant +ve and -ve coefficients. From Table 5.3 it can be seen that some PCs

split into two groups of variables. Some will have a positive coefficient and others a negative. For example, PC 2 distinguishes between the bladder variables and the higher rectal coefficients. We need to ensure that we select one from each group if the largest of each are significant. If we use similar significance levels to Jolliffe's (2002) simple +, - notation then we conclude a significant coefficient is one that is half the size of the maximum absolute coefficient, so $\beta = 0.5$. As a result, when the autoPCA results are presented in Tables 5.10 and 5.11 the total percentage of variables chosen from an individual PC j can total over 100%.

Algorithm 4: Determining whether there are significant positive and negative coefficients.

```

Input : PC  $j$ 

Output: Number of variables to keep

for PC  $j$  do
  if  $\max(PC_j) \geq |\min(PC_j)|$  then
    if  $|\min(PC_j)| > \beta \max(PC_j)$  then
      Keep two variables
    else
      Keep one variable
    end
  else
    if  $\max(PC_j) > \beta |\min(PC_j)|$  then
      Keep two variables
    else
      Keep one variable
    end
  end
end

```

Choosing γ

The final significance level to decide for Algorithm 1 is the value for γ , the total percentage variance we wish to be explained. This is very subjective and many papers suggest different cut-off values; a review of them can be found in Peres-Neto et al. (2005). Jolliffe (2002) also offers a review of different methods concluding that most use a cut-off level between 70 and 90% of variation explained. As a result, we explore different values of γ and then select a suitable one that balances the number of variables selected and the total variance explained. This is important as the variables will be used in our DEA analysis in Section 5.3 and in DEA the more variables there are, the less discerning the methodology is.

We now apply the autoPCA variable selection technique as described in Algorithm 1 to Dataset 2a. We use samples of size 35 and repeat the method with $r = 1,000,000$. The results are presented in Tables 5.10 and 5.11.

The values in Tables 5.10 and 5.11 are the percentage of the time the variable is selected for each PC. However, this does not then take into account the different weighting PCs have, this could lead to less significant variables being favoured. To try to account for this the final column of both Tables is a weighted total, which has been calculated in the following way.

Let w_{jq} be the number of times variable q was chosen for PC j in the r repetitions. Note that $0 \leq \sum_{q=1}^Q w_{jq} \leq 2r$ where Q is the total number of variables because up to two variables can be chosen from each PC j . Then to calculate the weighted total for each variable the following formula is used,

$$\frac{\sum_{j=1}^N \lambda_j w_{jq}}{\sum_{j=1}^N \sum_{q=1}^Q \lambda_j w_{jq}},$$

where λ_j is PC j 's average eigenvalue and $\sum_{j=1}^N \sum_{q=1}^Q \lambda_j w_{jq}$ is the weighted total number of variables chosen throughout the r repetitions.

From Table 5.10 for all three values of γ , PC 1 and PC 2 select V_{60}^{Rec} and V_{70}^{Bla} .

Variable	$\gamma = 70$			$\gamma = 80$			$\gamma = 90$				
	PC 1	PC 2	Weighted	PC 1	PC 2	PC 3	Weighted	PC 1	PC 2	PC 3	Weighted
V_{30}^{Rec}	0	0.455	0.241	0	0.472	0.693	0.349	0	0.456	94.936	11.254
V_{40}^{Rec}	0	0.599	0.317	0	0.605	0.007	0.320	0	0.595	2.207	0.512
V_{50}^{Rec}	0.015	12.187	6.463	0.015	12.240	0	6.478	0.016	12.229	0.106	5.264
V_{60}^{Rec}	2.501	54.447	31.032	2.519	54.438	0.001	30.984	2.489	54.468	0.139	25.141
V_{65}^{Rec}	0.180	11.349	6.166	0.187	11.306	0	6.137	0.187	11.334	0.181	5.011
V_{70}^{Rec}	0.035	12.794	6.802	0.034	12.828	0.124	6.824	0.037	12.827	65.195	13.117
V_{50}^{Bla}	0.043	4.444	2.390	0.048	4.450	0	2.393	0.046	4.527	0	1.973
V_{60}^{Bla}	0.088	5.872	3.185	0.094	5.877	0	3.188	0.091	5.899	0	2.593
V_{65}^{Bla}	0.317	11.211	6.214	0.316	11.228	0	6.211	0.323	11.209	0	5.034
V_{70}^{Bla}	0.639	69.199	37.190	0.641	69.190	0	37.115	0.640	69.181	0	30.102

Table 5.10: Input variables selected by autoPCA for $\gamma \in \{70, 80, 90\}$, $r = 1,000,000$, $k = 35$. Table shows the percentage of times the corresponding PC was selected. Numbers in bold represent the largest contributions to the PC. Dataset 2a.

Variable	$\gamma = 70$				$\gamma = 80$				$\gamma = 90$					
	PC 1	PC 2	PC3	Weighted	PC 1	PC 2	PC 3	PC4	Weighted	PC 1	PC 2	PC 3	PC4	Weighted
D ₉₉ ^{PTV3}	1.275	0.420	4.831	1.631	1.287	0.434	25.269	0.004	4.096	1.291	0.422	25.270	18.503	5.112
D ₉₈ ^{PTV3}	5.712	0.200	0.115	3.386	5.681	0.205	1.312	0.000	3.043	5.695	0.204	1.303	2.980	2.960
D ₅₀ ^{PTV3}	0.003	5.877	6.907	3.206	0.003	5.861	84.123	0.106	12.899	0.002	5.871	84.901	73.069	17.356
D ₉₉ ^{PTV1-2}	0.025	50.806	0.001	18.540	0.024	50.826	0.008	0	15.884	0.025	50.822	0.009	0.342	14.230
D ₉₈ ^{PTV1-2}	0.007	40.098	0.008	14.626	0.008	40.068	0.041	0	12.521	0.009	40.058	0.049	0.159	11.207
D ₅₀ ^{PTV1-2}	0.025	8.435	0.110	3.107	0.028	8.436	0.715	0.002	2.742	0.026	8.454	0.726	3.705	2.748
D ₉₉ ^{PTV2-3}	4.780	2.042	1.098	3.671	4.794	2.017	8.627	0.005	4.133	4.739	2.008	8.638	26.404	5.740
D ₉₈ ^{PTV2-3}	63.850	1.826	0.892	37.639	63.917	1.838	3.098	0.002	32.556	63.913	1.825	3.101	7.888	29.729
D ₅₀ ^{PTV2-3}	24.369	0.371	0.004	14.195	24.307	0.375	0.012	0	12.126	24.347	0.369	0.010	0.734	10.918

Table 5.11: Output variables selected by autoPCA for $\gamma \in \{70, 80, 90\}$, $r = 1,000,000$, $k = 35$, Dataset 2a. Table shows the percentage of times the corresponding PC was selected. Numbers in bold represent the largest contributions to the PC.

Variable	PC 1	PC 2	PC3	Weighted	Variable	PC 1	PC 2	PC 3	Weighted
D_{99}^{PTV3}	6.400	0.321	6.669	4.424	D_{98}^{PTV3}	1.412	0.121	5.034	1.644
D_{50}^{PTV3}	0.352	19.681	43.381	14.209	D_{50}^{PTV3}	0.141	21.128	29.511	12.395
D_{99}^{PTV1-2}	0.224	60.682	0.410	20.430	D_{98}^{PTV1-2}	0.066	58.320	0.461	19.962
D_{50}^{PTV1-2}	1.555	34.699	0.058	12.358	D_{50}^{PTV1-2}	0.470	38.523	0.038	13.360
D_{99}^{PTV2-3}	49.840	3.417	0.346	25.880	D_{98}^{PTV2-3}	61.016	3.032	0.031	32.480
D_{50}^{PTV2-3}	43.323	3.731	0.002	22.698	D_{50}^{PTV2-3}	37.426	2.560	0.001	20.158

Table 5.12: Output variables selected from autoPCA for $\gamma = 70$, $r = 1,000,000$, $k = 35$, Dataset 2a. Table shows the percentage of times the corresponding PC was selected. Numbers in bold represent the largest contributions to the PC.

These variables also have the highest weighted total and are consistent with the notion of selecting one bladder and one rectal constraint.

From Table 5.11 if we wish to select a single variable for each value of γ D_{98}^{PTV2-3} would be selected. If a second variable were required, we would choose D_{99}^{PTV1-2} for $\gamma = [70, 80]$ or D_{50}^{PTV3} for $\gamma = 90$. However, if we look at the results in Table 5.12 where we only consider D_{99} or D_{98} variables we see that $D_{99}^{PTV2-3}/D_{98}^{PTV2-3}$ are again selected as the most significant and the corresponding PTV 1-2 is the next most significant.

As mentioned at the beginning of Section 5.2, Rosemere have now changed their treatment process such that BM machines are no longer used. Therefore, we repeat the autoPCA variable selection with dataset 2b. This results in V_{60}^{Rec} and V_{70}^{Bla} being chosen as the input variables and D_{98}^{PTV2-3} being selected as the output variable.

5.2.5 Other variable selection techniques

With only 66 plans the autoPCA variable selection method could be considered subjective despite the repeated sampling. Therefore, we wish to apply further variable selection techniques and compare them to results from the autoPCA. If the variables selected are similar, we can be more confident in our autoPCA method.

We first apply the method introduced in Section 3.5, where we remove the least significant variable from the final PC. PCA is then rerun without this variable until all the variables have been removed. The order to remove the variables in are given in Table 5.13. The last two variables to be removed and hence, the two that arguably explain the most variation in the data are V_{60}^{Rec} and V_{70}^{Bla} . This agrees with the results from the autoPCA method. The last variable to be removed from the outputs is D_{50}^{PTV3} . D_{50}^{PTV3} was the second most significant output variable for autoPCA on Dataset 2a with $\gamma = 90$.

In Section 3.5 we introduced the partial covariance technique for variable selection. Here we apply it to Dataset 2b. From the Rosemere data 10 inputs and 9 outputs were identified from the clinical protocol. We examine how the variables influence the percentage variance explained by considering all of their possible combinations. For DEA it is preferred to have as few variables as possible. We seek the smallest subset of variables that still describes over 76% of the variation in the data (see Section 5.2.5).

For each variable subset size we select the best subset, i.e. the variables which explain most of the variation. The results for the input and output data can be seen in Tables 5.14 and 5.16.

Removal order	Input variable	Removal order	Output variable
1	V_{65}^{Bla}	1	D_{98}^{PTV3}
2	V_{60}^{Bla}	2	D_{99}^{PTV1-2}
3	V_{65}^{Rec}	3	D_{98}^{PTV2-3}
4	V_{60}^{Rec}	4	D_{50}^{PTV2-3}
5	V_{50}^{Bla}	5	D_{50}^{PTV1-2}
6	V_{40}^{Rec}	6	D_{99}^{PTV3}
7	V_{50}^{Rec}	7	D_{99}^{PTV2-3}
8	V_{30}^{Rec}	8	D_{98}^{PTV1-2}
9	V_{70}^{Bla}	9	D_{50}^{PTV3}
10	V_{70}^{Rec}		

Table 5.13: Order to remove variables, Dataset 2b.

Number of Removed Variables	Removed Variables	Percentage Explained
1	9	99.96
2	8,9	99.88
3	5,8,9	99.75
4	5,7,8,10	99.19
5	4,6,7,8,10	98.52
6	2,4,6,7,8,10	97.12
7	1,3,4,6,7,8,10	92.07
8	1,2,3,5,6,7,9,10	80.05
9	1,2,3,4,5,6,7,9,10	41.12
10	1,2,3,4,5,6,7,8,9,10	0

Table 5.14: Table to show the amount of variation explained for the best subset of input variables using the partial covariance method on Dataset 2b.

Key: 1- V_{30}^{Rec} , 2- V_{40}^{Rec} , 3- V_{50}^{Rec} , 4- V_{60}^{Rec} , 5- V_{65}^{Rec} , 6- V_{70}^{Rec} , 7- V_{50}^{Bla} , 8- V_{60}^{Bla} , 9- V_{65}^{Bla} , 10- V_{70}^{Bla}

From Table 5.14 to explain at least 76% of the variation in the input data two variables are required. Table 5.15 lists all combinations of two variables that explain over 76% of the variation in the input data. V_{60}^{Rec} and V_{60}^{Bla} explain the largest percentage variance when only two variables are used. Each pair in Table 5.15 consists of a higher rectal constraint and a bladder constraint which is consistent with our earlier findings. Here, V_{60}^{Rec} and V_{70}^{Bla} , the variables selected from our autoPCA explain 78.7% of the variation. The single input that explains the most variation is V_{60}^{Bla} .

To explain at least 76% of the output variation we require three variables (Table 5.16). There are 22 different combinations of these. The most significant is

Variables		Percentage explained
V_{60}^{Rec}	V_{50}^{Bla}	79.45
V_{60}^{Rec}	V_{60}^{Bla}	80.05
V_{60}^{Rec}	V_{65}^{Bla}	79.99
V_{60}^{Rec}	V_{70}^{Bla}	78.70
V_{65}^{Rec}	V_{50}^{Bla}	77.78
V_{65}^{Rec}	V_{60}^{Bla}	78.37
V_{65}^{Rec}	V_{65}^{Bla}	78.28
V_{65}^{Rec}	V_{70}^{Bla}	76.99
V_{70}^{Rec}	V_{60}^{Bla}	76.01

Table 5.15: All combinations of two variables that account for over 76% of the variation of the inputs for Dataset 2b.

D_{99}^{PTV3} , D_{98}^{PTV1-2} and D_{98}^{PTV2-3} . Alternatively, if we wish to choose the best two output variables that explain 70.01% of the variation D_{99}^{PTV1-2} and D_{98}^{PTV2-3} should be chosen. The single output variables that explain the most variation are D_{98}^{PTV3} and D_{98}^{PTV2-3} explaining 42.05% and 41.96% respectively.

5.2.6 Conclusion

To conclude Section 5.2 we consider the results from the variable selection processes used throughout. A summary of these can be found in Table 5.17.

All three methods in Table 5.17 choose V_{60}^{Rec} as an input variable. Removing the least significant variables and autoPCA selects V_{70}^{Bla} , whereas the partial covariance method selects V_{60}^{Bla} . These two variables are highly correlated with Pearson's correlation coefficient 0.98. From Table 5.15 we see that over 76% of the variation is

Number of Removed Variables	Removed Variables	Percentage Explained
1	2	99.83
2	2,4	99.64
3	2,4,8	99.41
4	1,5,7,9	96.76
5	1,4,6,7,9	91.85
6	2,3,4,6,7,9	81.71
7	1,2,3,5,6,7,9	70.01
8	1,3,4,5,6,7,8,9	42.05
9	1,2,3,4,5,6,7,8,9	0

Table 5.16: Table to show the amount of variation explained for the best subset of output variables using the partial covariance method for Dataset 2b.

Key: 1- D_{99}^{PTV3} , 2- D_{98}^{PTV3} , 3- D_{50}^{PTV3} , 4- D_{99}^{PTV1-2} , 5- D_{98}^{PTV1-2} , 6- D_{50}^{PTV1-2} , 7- D_{99}^{PTV2-3} , 8- D_{98}^{PTV2-3} , 9- D_{50}^{PTV2-3}

explained by using V_{70}^{Bla} and V_{60}^{Rec} and so choose V_{70}^{Bla} and V_{60}^{Rec} as our input variables. This is consistent with our preliminary data analysis in Section 5.2.2, where we find the data splits into two significant groups, higher rectal constraints and bladder constraints.

We discussed the choice of variables with specialists at Rosemere and they believe they are suitable for our analysis. At Rosemere they refer to Michalski et al. (2010) as a general guideline when considering rectal toxicity. Michalski et al. (2010) conclude that “most dose-volume parameters significantly associated with late rectal toxicity consider doses greater than or equal to 60Gy”. This justifies the use of V_{60}^{Rec} . Michalski et al. (2010) also conclude that “ V_x has not been found to be significantly associated

Method	Variables Selected		
	Inputs		Outputs
AutoPCA	V_{70}^{Bla}	V_{60}^{Rec}	D_{98}^{PTV2-3}
Remove least significant	V_{70}^{Bla}	V_{60}^{Rec}	D_{50}^{PTV3}
Partial Covariance	V_{60}^{Bla}	V_{60}^{Rec}	D_{98}^{PTV3}

Table 5.17: Comparison of variables selected for Auto PCA, removing the least significant PC and partial covariance, Dataset 2b.

with differences in rectal toxicity for doses less than or equal to 45 Gy”. This confirms the idea that V_{30} and V_{40} behave differently to the other rectal constraints. They are therefore not as important in plan assessment. Hence, our decision to use a single higher rectal constraint as an input variable can be clinically justified. Clinicians also agreed that V_{70}^{Bla} was a sensible choice as it is the main bladder constraint they check during planning.

The output variables’ results were not as conclusive as the input variables’. This is partly due to the strong correlation between the D_{98} and D_{99} variables. Looking at Table 5.17, from autoPCA we select D_{98}^{PTV2-3} , from removing the least significant variables method we select D_{50}^{PTV3} and from the partial covariance method we select D_{98}^{PTV3} . However, in the partial covariance D_{98}^{PTV3} explains 42.05% and D_{98}^{PTV2-3} explains the next highest percentage with 41.96%. When removing the least significant PC method is applied D_{98}^{PTV2-3} is the third variable removed. At this point for PC 7 D_{98}^{PTV2-3} and D_{99}^{PTV2-3} have eigenvalues 0.73 and 0.67 respectively. PC 7 itself explains only 0.151% of the total variation. Therefore, if we remove D_{99}^{PTV2-3} instead of D_{98}^{PTV2-3} at this stage and proceed with the process, we get the order in Table 5.18. Now D_{98}^{PTV2-3} is the last remaining variable.

Removal order	Output variable
1	D_{98}^{PTV3}
2	D_{99}^{PTV1-2}
3	D_{99}^{PTV2-3}
4	D_{50}^{PTV2-3}
5	D_{50}^{PTV1-2}
6	D_{99}^{PTV3}
7	D_{50}^{PTV3}
8	D_{98}^{PTV1-2}
9	D_{98}^{PTV2-3}

Table 5.18: Order to remove output variables, Dataset 2b. When D_{99}^{PTV2-3} is removed third instead of D_{98}^{PTV2-3} . Number one is the first to be removed.

Over the years Rosemere have improved their treatment planning such that it is now unlikely they will not meet the inner PTV, PTV 3, dose constraints. This can be seen in Figure 5.2 where the range of the PTV 3 variables is smaller than most other variables. It stands to reason that because Rosemere prioritise obtaining the desired dose to PTV 3 and succeed, that a different output is more representative of the variation in the data. Hence using a PTV 2-3 metric is clinically justified. It is unsurprising that we have not selected a D_{50} metric, given that we know D_{98} and D_{99} metrics are prioritised by clinicians. Hence we conclude D_{98}^{PTV2-3} is a suitable output variable to use.

Therefore, the final variables we will use for the DEA models are shown in Table 5.19.

Final variables chosen		
Inputs		Output
V_{70}^{Bla}	V_{60}^{Rec}	D_{98}^{PTV2-3}

Table 5.19: Variables selected for our DEA models.

5.3 Nominal DEA

We now wish to apply DEA to the variables selected in Table 5.19. As Lin et al. (2013) suggest, we use the envelopment form of an input oriented variable returns to scale model to evaluate treatment plans. As a result of the PCA analyses in Section 5.2 we deal with the two input measures V_{70}^{Bla} and V_{60}^{Rec} and the one output measure D_{98}^{PTV2-3} .

As introduced in Section 3.1, (3.8), if we collect all input data for the 66 plans in a 2×66 matrix X , the output data in a 1×66 matrix Y , then the DEA model to assess treatment plan i is the linear optimisation problem

$$\min_{\theta_i, \lambda} \theta_i \quad (5.2a)$$

$$\text{s.t. } Y\lambda - y_i \geq 0 \quad (5.2b)$$

$$X\lambda - \theta_i x_i \leq 0 \quad (5.2c)$$

$$e^T \lambda = 1 \quad (5.2d)$$

$$\lambda, \theta \geq 0. \quad (5.2e)$$

In this model, e is a vector of ones of length 66, x_i and y_i are the i th columns of matrices X and Y respectively. θ_i is a scalar decision variable which represents the efficiency score of treatment plan i and λ is a decision variable of dimension 66.

The linear optimisation problem (5.2) attempts to identify a data point $(x, y) = (X, Y)^T \lambda$ in the set spanned by the OAR and PTV data given by the 66 existing plans

and defined by the constraints (5.2b) to (5.2e) that has output D_{98}^{PTV2-3} higher than or equal to treatment plan i , and at the same time inputs V_{70}^{Bla} and V_{60}^{Rec} lower than or equal to those of plan i . Among all such data points, the objective function (5.2a) together with constraint (5.2c) makes sure that the point with smallest input data is selected. If the optimal value θ_i^* is less than one, plan (DMU) i is called inefficient and the optimal solution of the DEA model provides evidence that it should be possible to scale down the input vector x_i to $\theta_i^* x_i$ while maintaining the same output level. In other words, the data of the 66 plans considered in the study suggest that treatment plan i could be improved by reducing the percentage volume of the bladder and/or rectum receiving 70/60 Gy of radiation dose without also reducing the dose delivered to 98% of PTV 2-3. Moreover, the nonzero entries in an optimal vector λ^* indicate from the data of which plans this suggestion is derived. We note that for all inefficient plans, the corresponding i th entry in λ^* must be zero. Hence, plans with nonzero entries in λ_i are termed peers in DEA. If, on the other hand, $\theta_i^* = 1$ then the data set does not contain evidence that treatment plan i could be improved and plan i is considered efficient.

We note here the use of ratio data in our DEA model arising from the scaling of D_{98}^{PTV2-3} as detailed in Section 5.2.3. There is much discussion in the DEA literature about the validity of DEA models when ratio variables are used. This arises from occurrences of the convexity constraints being broken. Hollingsworth and Smith (2003) note that care must be taken when ratio data are used to ensure that comparison between DMUs is still valid. Emrouznejad and Amin (2009) suggest two methods for adjusting the DEA model to allow for ratio variables. Solution one involves treating the numerator of the variable as an input and the denominator as an output. However this is not suitable for evaluating treatment plans as it would result in the 37 session plans being favoured over the 20 session plans as the 37 session plans have a higher D_{98}^{PTV2-3} value. Additionally, the prescribed dose will be treated as an input variable,

however, we are not trying to reduce the prescribed dose, this is fixed and we cannot change it. Therefore this is not a meaningful input variable as defined in Definition 3.1.3. The second solution Emrouznejad and Amin (2009) suggest is using the ratio variable as a convex combination of the numerator to the denominator. This results in replacing constraint (5.2b) with

$$\sum_{i=1}^I \lambda_i d_i - y_0 \sum_{i=1}^I \lambda_i p_i,$$

where y_0 is the scaled output value for the plan 0 being evaluated, d_i is the original unscaled dose for plan i and p_i is the corresponding prescribed dose. However this again puts undue emphasis on the value of the prescribed dose, favouring a smaller prescribed dose, hence favouring the 20 session plans. Therefore we believe here that neither of these solutions are suitable for our radiotherapy application. However, the variable y_i , the proportion of the prescribed dose achieved for plan i is not a standard DEA ratio variable. Unlike other ratio variables discussed in the DEA literature we only have a single ratio variable not multiple ones, consequently, interaction of ratio variables and their meaning is not a concern. In addition, we have a fixed known denominator that can be one of two values. This is because, the denominator values occur from the clinical application of two clinically comparable treatments. Therefore, we argue that the D_{98}^{PTV2-3} is not a ratio variable in the standard DEA sense and hence, we conclude that using the output variable y in constraint (5.2b) as a standard output variable is sensible here. This ensures that the constraints we include in our model and the conclusions we draw from them are clinically as well as mathematically relevant.

In order to evaluate the quality of all treatment plans, the linear optimisation problem (5.2) needs to be solved once for each plan. Table 5.20 shows the efficiency score θ_i^* and an optimal solution vector λ^* for each of the 66 treatment plans. After running DEA, we find that the five plans 36, 43, 45, 60 and 74 are deemed efficient.

Plan	Efficiency score θ_i	Optimal values λ_j^i for efficient plans $j =$				
		36	43	45	60	74
30	0.517	0.65	0	0.35	0	0
31	0.766	0.63	0	0	0.37	0
32	0.720	0.22	0	0	0.78	0
33	0.530	1	0	0	0	0
34	0.781	0	0.33	0.33	0	0.34
35	0.440	0.50	0	0	0.50	0
36	1	1	0	0	0	0
37	0.487	1	0	0	0	0
38	0.888	0.78	0	0	0.22	0
41	0.386	0.94	0	0.06	0	0
42	0.655	0	0.53	0.15	0	0.32
43	1	0	1	0	0	0
44	0.414	0	0	0	0	1
45	1	0	0	1	0	0
46	0.916	0.40	0	0.60	0	0
47	0.604	0	0	0	0	1
48	0.754	0	0	0	0	1
49	0.580	0.59	0	0.41	0	0
50	0.414	0.72	0	0.26	0	0.02
51	0.671	0.70	0	0.10	0	0.20
52	0.418	0.22	0	0	0.78	0
53	0.656	0.63	0	0.37	0	0
54	0.538	0.85	0	0.15	0	0

Plan	Efficiency score θ_i	Optimal values λ_j^i for efficient plans $j =$				
		36	43	45	60	74
55	0.472	0.50	0	0	0.50	0
56	0.584	0.63	0	0	0.37	0
57	0.424	0.22	0	0	0.78	0
58	0.831	0.17	0	0.83	0	0
59	0.680	0.67	0	0	0.33	0
60	1	0	0	0	1	0
61	0.863	0.19	0	0.44	0	0.37
62	0.713	0.37	0	0	0.42	0.21
63	0.302	1	0	0	0	0
64	0.689	0.77	0	0	0.23	0
65	0.374	0.67	0	0	0.33	0
66	0.348	0.52	0	0	0.04	0.44
67	0.843	0.69	0	0.29	0	0.02
68	0.790	0.22	0	0	0.78	0
70	0.519	1	0	0	0	0
71	0.496	1	0	0	0	0
72	0.863	0.04	0	0.27	0	0.69
73	0.584	0.42	0	0.24	0	0.34
74	1	0	0	0	0	1
75	0.536	0.77	0	0	0.23	0
76	0.578	0.91	0	0	0.09	0
77	0.800	0.54	0	0.13	0	0.33
78	0.508	0.91	0	0	0.09	0

Plan	Efficiency score θ_i	Optimal values λ_j^i for efficient plans $j =$				
		36	43	45	60	74
80	0.581	0.75	0	0.25	0	0
81	0.491	0.86	0	0.14	0	0
82	0.678	0.62	0	0.36	0	0.01
83	0.388	0.22	0	0	0.78	0
84	0.613	1	0	0	0	0
85	0.493	0.95	0	0.05	0	0
86	0.446	0.92	0	0.06	0	0.03
87	0.607	0.45	0	0.55	0	0
88	0.935	0.10	0	0.90	0	0
89	0.407	1	0	0	0	0
90	0.511	0.56	0	0.44	0	0
91	0.486	0.87	0	0.13	0	0
92	0.919	0	0	1	0	0
93	0.749	0.36	0	0	0.64	0
94	0.710	0	0	0	0	1
95	0.764	0.50	0	0	0.50	0
96	0.687	0.34	0	0.63	0	0.02
97	0.403	0.97	0	0.03	0	0
98	0.903	0.44	0	0.56	0	0
99	0.423	0.77	0	0	0.23	0

Table 5.20: Nominal DEA results, Dataset 2b.

Observe that for an efficient plan i , λ^* is a unit vector with the one in position i . For each inefficient plan we obtain a list of its peer efficient plans. These peers define the target input values $\theta_i^* x_i$ for each inefficient plan i . If the plan can be modified in a way that these target values are achieved without deteriorating the dose to the tumour, i.e. the D_{98}^{PTV2-3} value, the plan would become efficient. Hence, peers and target values provide suggestions for improving the plan. For example from Table 5.20 we can see that for plan 30 the data suggest that (via a combination of the data of plans 36 and 45) there might potentially exist a better plan. This plan would achieve a D_{98}^{PTV2-3} value not worse than that of plan 30 and it would achieve this despite a 42 % reduction in its V_{70}^{Bla} and V_{60}^{Rec} values, i.e. $0.65x_{36} + 0.35x_{45} \leq 0.52x_{30}$.

Peers also allow us to check the reliability of the efficient plans. If an efficient plan is only a peer to itself this means it is only efficient because it has the lowest value for one of the inputs. From Table 5.21 we can see that every efficient plan is a peer to at least one inefficient plan so we know that they are all strongly efficient.

We can also see that plan 36 is the plan most often featured as a peer of an inefficient plan: it is a peer to 55 inefficient plans, therefore it is not a peer to only 6 of the 61 inefficient plans. We note that plan 36 is the plan with the lowest V_{60}^{Rec} . The number of times a plan features as a peer gives an indication of which plans are particularly good ones. This alludes to the idea of gold standard plans defined relative to the specific hospital. This information is summarised in Table 5.21.

5.3.1 Bladder volume as an environmental variable

In Section 3.1.5 we introduced the use of environmental variables in DEA. We now consider the use of them for the radiotherapy application. In Table 5.2 we identified six different possible volume variables. In Figure 5.4 we noted that the largest variability in the volume of a structure occurred in the bladder. We suggested that this may be a suitable variable to consider as an environmental variable in our DEA models.

Efficient plan i	Number of plans i is a peer to
36	55
43	3
45	30
60	22
74	19

Table 5.21: Nominal DEA results peer frequency.

For the environmental variable selection we consulted with clinicians at Rosemere to advise us on suitable variables. We relied more on their advice than analytic methods for choosing our variables. In addition, the volumes are somewhat accounted for by the nature of the V_x and D_x constraints. These are metrics based on the % of the volume of an OAR receiving x Gy or the dose received by $x\%$ of the volume of a PTV respectively. After consulting with clinicians at Rosemere we concluded that in addition to the variables selected in Table 5.19, the total bladder volume should be considered, (for more details see Section 2.1.2). This is a measure that has an influence on how well a treatment can meet the constraints set out in the clinical protocol: a larger bladder volume means that a smaller percentage of the bladder volume is in close proximity to the prostate. Hence, a larger bladder volume makes it easier to achieve low values of V_{70}^{Bla} for the same D_{98}^{PTV2-3} value or vice versa, a higher D_{98}^{PTV2-3} value for the same V_{70}^{Bla} value. Therefore, treatment plans incorporating a small bladder volume should not be compared to ones with a large bladder volume, since they are probably outperformed by the latter. In DEA terminology, bladder volume is an environmental variable, see Section 3.1.5. Due to the effect of the environmental variable on treatment plan quality, it can be incorporated in the DEA model in a similar way as an input variable. We add the environmental data in a

1×66 matrix Z to the DEA model (5.2) using the following constraint

$$Z\lambda - z_i \leq 0. \quad (5.3)$$

In this way, (5.2) together with (5.3) attempts to identify a data point $(x, y, z) = (X, Y, Z)^T \lambda$ in the set spanned by the OAR, PTV, and volume data given by the 66 existing plans. Constraint (5.3) ensures that this data point would also represent a plan with bladder volume at most as big as that of plan i . Notice that θ_i is omitted in constraint (5.3) so that it plays no direct role in the determination of the optimal value of θ_i .

We apply the DEA model (5.2) together with the bladder volume constraint (5.3) to Dataset 2b. The results are in Table 5.22.

In Table 5.22 we note that there are now ten efficient plans as opposed to five in Table 5.20. We see that when Vol^{Bla} is included as an environmental variable plans 37 and 48 are no longer efficient and have peers 36 and 38 and 74 and 94 respectively. Similarly, plans 31, 45, 92 and 98 are now deemed efficient. We note here that plan 38 has the smallest Vol^{Bla} and plan 72 has the largest. This means plan 38 is efficient because it cannot be compared to any other plan in Dataset 2b, although it was efficient when Vol^{Bla} was not a variable so we do not need to be concerned by this. Similarly, plan 72 can be compared to all the plans in Dataset 2b and we see here it is not deemed efficient.

Plan	Efficiency score θ_i	Optimal values λ_j^i for efficient plans $j =$									
		31	36	38	43	45	60	74	92	94	98
30	0.517	0	0.65	0	0	0.35	0	0	0	0	0
31	1	1	0	0	0	0	0	0	0	0	0
32	0.890	0	0	0.33	0	0	0	0.26	0	0.41	0
33	0.542	0	0.86	0.14	0	0	0	0	0	0	0
34	0.942	0	0	0	0.46	0.06	0	0.48	0	0	0
35	0.550	0.74	0	0	0	0	0	0.24	0	0	0.02
36	1	0	1	0	0	0	0	0	0	0	0
37	0.552	0	0.18	0.82	0	0	0	0	0	0	0
38	1	0	0	1	0	0	0	0	0	0	0
41	0.386	0	0.94	0	0	0.06	0	0	0	0	0
42	0.655	0	0	0	0.53	0.15	0	0.32	0	0	0
43	1	0	0	0	1	0	0	0	0	0	0
44	0.676	0	0	0	0	0	0	0.14	0	0.86	0
45	1	0	0	0	0	1	0	0	0	0	0
46	0.916	0	0.40	0	0	0.60	0	0	0	0	0
47	0.926	0	0	0	0	0	0	0.34	0	0.66	0
48	0.891	0	0	0	0	0	0	0.56	0	0.44	0
49	0.580	0	0.49	0	0	0.39	0	0.12	0	0	0
50	0.414	0	0.72	0	0	0.26	0	0.02	0	0	0
51	0.671	0	0.70	0	0	0.10	0	0.20	0	0	0
52	0.519	0	0	0.33	0	0	0	0.25	0	0.42	0
53	0.656	0	0.63	0	0	0.37	0	0	0	0	0
54	0.539	0	0.16	0	0	0.06	0	0.77	0	0	0
55	0.541	0.35	0.12	0	0	0	0	0.53	0	0	0
56	0.757	0.88	0	0	0	0	0	0.11	0	0	0.01
57	0.595	0.23	0	0	0	0	0	0.31	0	0.22	0.24

Plan	Efficiency score θ_i	Optimal values λ_j^i for efficient plans $j =$									
		31	36	38	43	45	60	74	92	94	98
58	0.831	0	0.17	0	0	0.83	0	0	0	0	0
59	0.703	0	0.46	0.27	0	0	0.27	0	0	0	0
60	1	0	0	0	0	0	1	0	0	0	0
61	0.876	0	0	0	0	0.35	0	0.52	0	0	0.13
62	0.713	0	0.37	0	0	0	0.42	0.21	0	0	0
63	0.400	0.83	0.08	0.09	0	0	0	0	0	0	0
64	0.689	0	0.77	0	0	0	0.23	0	0	0	0
65	0.408	0.13	0.35	0	0	0	0	0.52	0	0	0
66	0.360	0.04	0.37	0	0	0	0	0.59	0	0	0
67	0.843	0	0.69	0	0	0.29	0	0.02	0	0	0
68	0.790	0	0.22	0	0	0	0.78	0	0	0	0
70	0.519	0	1	0	0	0	0	0	0	0	0
71	0.496	0	1	0	0	0	0	0	0	0	0
72	0.863	0	0.04	0	0	0.27	0	0.69	0	0	0
73	0.775	0.09	0	0	0	0	0	0.06	0	0.14	0.71
74	1	0	0	0	0	0	0	1	0	0	0
75	0.683	0.50	0	0	0	0	0	0.46	0	0	0.04
76	0.716	0.65	0.30	0	0	0	0	0.04	0	0	0
77	0.800	0	0.52	0	0	0.13	0	0.35	0	0	0
78	0.638	0.68	0.23	0.09	0	0	0	0	0	0	0
80	0.597	0	0	0	0	0.03	0	0.73	0	0	0.24
81	0.491	0	0.86	0	0	0.14	0	0	0	0	0
82	0.717	0	0	0	0	0.01	0	0.47	0	0	0.52
83	0.392	0	0.16	0.08	0	0	0.76	0	0	0	0
84	0.859	0.58	0	0	0	0	0	0.31	0	0	0.11
85	0.493	0	0.95	0	0	0.05	0	0	0	0	0

Plan	Efficiency score θ_i	Optimal values λ_j^i for efficient plans $j =$									
		31	36	38	43	45	60	74	92	94	98
86	0.446	0	0.92	0	0	0.06	0	0.03	0	0	0
87	0.607	0	0.34	0	0	0.53	0	0.13	0	0	0
88	0.990	0	0	0	0	0.41	0	0	0.37	0	0.22
89	0.507	0.55	0.23	0	0	0	0	0.22	0	0	0
90	0.519	0	0	0	0	0.30	0	0.58	0	0	0.12
91	0.486	0	0.87	0	0	0.13	0	0	0	0	0
92	1	0	0	0	0	0	0	0	1	0	0
93	0.749	0	0.36	0	0	0	0.64	0	0	0	0
94	1	0	0	0	0	0	0	0	0	1	0
95	0.764	0	0.50	0	0	0	0.50	0	0	0	0
96	0.731	0	0	0	0	0.29	0	0.11	0	0	0.60
97	0.403	0	0.92	0	0	0.03	0	0.06	0	0	0
98	1	0	0	0	0	0	0	0	0	0	1
99	0.942	0	0	0.72	0	0	0	0	0	0.28	0

Table 5.22: Nominal DEA results including Vol^{Bla} , Dataset 2b.

Efficient plan i	Number of plans i is a peer to
31	14
36	37
38	10
43	3
45	27
60	8
74	35
92	2
94	9
98	13

Table 5.23: Nominal DEA results peer frequency including Vol^{Bla} .

In this chapter we have extended the work of Lin et al. (2013) to apply DEA to evaluate the quality of radiotherapy treatment plans for prostate cancer. From the clinical protocol used at Rosemere, we have extracted 25 variables that are relevant to the decision of whether or not a plan can be approved for treatment or should be improved via replanning. In Section 5.2, initial data analyses highlighted structural differences related to two different treatment regimes used at Rosemere. Once these structural differences were accounted for, we developed a method to select suitable variables to be used in the DEA model. In Section 5.3 we used our DEA model to evaluate treatment plan quality using the variables identified by our autoPCA method. This was done both with and without considering the bladder volume as an environmental variable. Next we consider the effect of uncertainty on these results.

Chapter 6

Uncertainty in radiotherapy

It is well known in radiation oncology that the outcomes of radiotherapy differ from the plans, i.e. the doses delivered to structures are usually (slightly) different from those calculated during treatment planning. This uncertainty in predicting radiation dose delivered to PTVs and OARs has many sources and is researched well in the clinical literature, see Section 2.3. This discussion may cast doubt on the DEA based evaluation of treatment plans in Section 5.3 and reported in Tables 5.20 and 5.22. Here we assumed the planning data to be exact and classified treatment plans as efficient or inefficient based on these data. However, it is likely that the values for the measures listed in Table 5.1 are imprecise. Hence, it is also possible that an inefficient plan does actually perform well in practice. Thus, we need to take uncertainty into account when evaluating the quality of treatment plans. In the rest of this section we explore this opportunity to leverage uncertainty in order to identify treatment plans that are only considered inefficient due to the precise computation of plan data but that would perform well when considering the uncertainty.

We therefore assume that the measures listed in Table 5.19 are in fact realisations from a range, called an uncertainty set. We consider how uncertainties in treatment planning affect treatment quality as measured by the efficiency score from DEA.

We must first determine a suitable amount of uncertainty, i.e. how uncertain the data are. In radiotherapy treatment planning the standard assumption is that uncertainty is proportional to the dose. The international commission on radiation units and measurements conclude that the available evidence for certain types of tumour suggests an accuracy of $\pm 5\%$ is required (Andreo et al., 2004). Combining the standard uncertainty value for dose determination and the uncertainty associated with Pinnacle for multileaf collimators, Henríquez and Castrillón (2008) suggest an uncertainty of 3.6% is used. Therefore, we wish to consider uncertainty in the range of $\pm 5\%$ with a larger focus on the range $\pm 3.6\%$.

To model the uncertainty in the data we must consider the possible changes to each variable. The amount of uncertainty we consider here is denoted by u . The two inputs selected in Table 5.19 are V_x metrics. They are the percentage of the OAR receiving x Gy or more. Therefore, to consider a change in the data of $\pm 3.6\%$ we require $u = 3.6$, where this is an absolute value. The output from Table 5.19 is D_{98}^{PTV2-3} , this is the dose received by 98% of the PTV 2-3. To account for 20 or 37 fractions we divided the D_{98}^{PTV2-3} by the prescribed dose, here 54.60/67.34 Gy for 20/37 fractions. Therefore, our output, D_{98}^{PTV2-3} is now a measure of the proportion of the prescribed dose achieved by 98% of the volume and we require $u = 0.036$ to model uncertainty of 3.6% . However, DEA is not affected by scaling and hence we can multiply the D_{98}^{PTV2-3} by 100 and in this way using a value of $u = 3.6$ will be equivalent to a 3.6% change for both the inputs and outputs.

Therefore, by using an absolute value of $u = 3.6$ in our simulations and robust DEA model we are consistent with the notion of an uncertainty interval in Definition 4.0.3.

Note that we do not consider uncertainty for total bladder volume. Instead, we consider that the volume of the bladder of a patient has the same effect on treatment plan quality for each realisation of the data.

In this chapter, we consider two methods for handling uncertainty. In Section 4.3.1 we concluded that when the value of uncertainty is not fixed the uDEA model is a non-convex problem. Therefore, alternative techniques such as simulation are required. In Section 6.1 we simulate changes in the data and calculate the efficiency score for each set to find the maximum possible efficiency score a DMU can obtain as done in Stubington et al. (2019). In Section 4.2 we concluded that for fixed values of uncertainty the overall best improvement in efficiency score is obtained by solving the robust DEA model (4.19). In Section 6.2 we apply this model for fixed values of u . In Section 6.3 we then compare the two methods to determine whether or not the results are the same and apply uDEA to the small subset of the data to see how this compares. To conclude, in Section 6.4, we identify a set of plans for Preston to replan and see if they have improved.

6.1 Simulation

First, we use simulation to explore the effect of uncertain data on the DEA results. To do this we assume that each OAR and PTV variable v can take any value in the interval $[v - \epsilon, v + \epsilon]$, where ϵ is uniformly distributed in $[0, u]$. Then we can sample from these distributions and simulate uncertainty in the treatment plans. To begin with we consider $u \in \{0.3, 0.6, \dots, 3.6, 5\}$. For each plan i and each value of u , we simulate the DEA data 10,000 times, i.e. we randomly generate the entries of X and Y uniformly within an interval of the plan value $+/- u$ and compute an efficiency score every time. We record the largest efficiency score among the 10,000 repetitions, this is because we are interested in the capability of the DMU. The results are summarised in Table 6.1.

Plan	Uncertainty													
	0	0.3	0.6	0.9	1.2	1.5	1.8	2.1	2.4	2.7	3	3.3	3.6	5
58	0.83	0.90	1	1	1	1	1	1	1	1	1	1	1	1
59	0.68	0.82	1	1	1	1	1	1	1	1	1	1	1	1
60	1	1	1	1	1	1	1	1	1	1	1	1	1	1
61	0.86	1	1	1	1	1	1	1	1	1	1	1	1	1
62	0.71	1	1	1	1	1	1	1	1	1	1	1	1	1
63	0.30	0.34	0.40	0.48	0.57	0.66	0.72	1	1	1	1	1	1	1
64	0.69	0.80	1	1	1	1	1	1	1	1	1	1	1	1
65	0.37	0.47	0.60	1	1	1	1	1	1	1	1	1	1	1
66	0.35	0.52	1	1	1	1	1	1	1	1	1	1	1	1
67	0.84	0.91	1	1	1	1	1	1	1	1	1	1	1	1
68	0.79	1	1	1	1	1	1	1	1	1	1	1	1	1
70	0.52	0.59	0.70	0.80	0.91	0.98	1	1	1	1	1	1	1	1
71	0.50	0.54	0.62	0.72	0.83	0.94	0.98	1	1	1	1	1	1	1
72	0.86	1	1	1	1	1	1	1	1	1	1	1	1	1
73	0.58	0.81	1	1	1	1	1	1	1	1	1	1	1	1
74	1	1	1	1	1	1	1	1	1	1	1	1	1	1
75	0.54	0.64	0.87	1	1	1	1	1	1	1	1	1	1	1
76	0.58	0.66	0.79	0.90	1	1	1	1	1	1	1	1	1	1
77	0.80	1	1	1	1	1	1	1	1	1	1	1	1	1
78	0.51	0.58	0.71	0.85	0.98	1	1	1	1	1	1	1	1	1
80	0.58	0.62	0.82	1	1	1	1	1	1	1	1	1	1	1
81	0.49	0.52	0.56	0.65	0.86	1	1	1	1	1	1	1	1	1
82	0.68	0.74	1	1	1	1	1	1	1	1	1	1	1	1
83	0.39	1	1	1	1	1	1	1	1	1	1	1	1	1
84	0.61	0.70	0.85	0.98	1	1	1	1	1	1	1	1	1	1
85	0.49	0.53	0.56	0.65	0.72	0.78	0.85	0.90	0.93	0.96	0.99	1	1	1

Plan	Uncertainty													
	0	0.3	0.6	0.9	1.2	1.5	1.8	2.1	2.4	2.7	3	3.3	3.6	5
86	0.45	0.47	0.67	0.81	1	1	1	1	1	1	1	1	1	1
87	0.61	0.68	1	1	1	1	1	1	1	1	1	1	1	1
88	0.94	1	1	1	1	1	1	1	1	1	1	1	1	1
89	0.41	0.43	0.45	0.49	0.54	0.57	0.61	0.64	0.67	0.71	0.75	0.78	0.86	1
90	0.51	0.54	0.73	0.87	1	1	1	1	1	1	1	1	1	1
91	0.49	0.51	0.54	0.57	0.62	0.68	0.73	0.79	0.85	0.90	0.95	1	1	1
92	0.92	1	1	1	1	1	1	1	1	1	1	1	1	1
93	0.75	1	1	1	1	1	1	1	1	1	1	1	1	1
94	0.71	1	1	1	1	1	1	1	1	1	1	1	1	1
95	0.76	1	1	1	1	1	1	1	1	1	1	1	1	1
96	0.69	1	1	1	1	1	1	1	1	1	1	1	1	1
97	0.40	0.43	0.45	0.49	0.55	0.60	0.67	0.72	0.77	0.81	0.85	0.89	0.92	1
98	0.90	1	1	1	1	1	1	1	1	1	1	1	1	1
99	0.42	0.50	0.64	0.89	1	1	1	1	1	1	1	1	1	1

Table 6.1: Maximum efficiency score after 10,000 repetitions, $u \in \{0.3, 0.6, \dots, 3.6, 5\}$.

If an inefficient plan from Table 5.20 can be seen as efficient for some u greater than zero in Table 6.1 this suggests that the plan was previously deemed inefficient due to the inherent uncertainty in the data, rather than being a bad treatment plan. In Table 6.1 we see that with only a very small amount of uncertainty, $u = 0.3$, 28 plans are efficient and when we have $u = 3.6$ this increases to 64 plans being efficient, i.e. only two are classified as inefficient. When we have $u = 5$ all the plans are efficient, which is in-line with clinical plans needing to be within 5% uncertainty. In Table 6.1 the five plans that require the most uncertainty are plans 37, 85, 89, 91 and 97 with $u = 3, 3.3, 5, 3.3$ and $u = 5$ respectively.

Due to the structure of the data set, it is by no means guaranteed that an inefficient plan with an originally high efficiency score becomes efficient when considering a small amount

of uncertainty. For example, the efficiency score of plan 91, was $\theta_0^* = 0.49$ but requires a value of $u = 3.3$ to be deemed efficient. In contrast, plan 57 has a lower efficiency score of $\theta_0^* = 0.42$ to begin with but with $u = 0.3$ becomes efficient. On the other hand, a plan that has a low efficiency score for certain data and does not reach an efficiency score of one, even with a large amount of uncertainty, is a good candidate for further improvement as its perceived inefficiency is not due to uncertainty. Hence, improved plan quality by replanning this treatment is likely to be beneficial. We note that the efficiency scores in Table 6.1 are consistent with the theoretical finding of Ehrgott et al. (2018) that efficiency scores must increase with increasing uncertainty.

The efficiency scores for the different values of u in Table 6.1 are for 10,000 repetitions of the DEA model. In Table 6.2 we compare the number of efficient DMUs at each value of u with the number of repetitions, $r \in \{1, 10, 100, 1000, 10,000, \}$ in the simulation. We see that for small values of r , $r \in \{1, 10\}$ the number of efficient plans does not necessarily increase as the uncertainty increases. For example, with $r = 10$ and $u = 2.1$ there are 26 efficient plans but when $u = 3.6$ there are only 24 efficient plans. Note here that the reduction in the number of efficient plans is not because a plan stops being efficient when more uncertainty is introduced but because we have a small number of repetitions. This means the realisation of the data for an earlier plan resulted in a plan being deemed efficient but a different realisation was selected when there was more uncertainty that rendered the plan inefficient. We see that as the number of repetitions increases this does not occur. This demonstrates the need for a suitably large value of r . As expected there is a large difference between the number of plans that are efficient when $r = 1$ and $r = 10,000$. For example, for $u = 3.6$ we go from 9 to 64 efficient plans.

There are an infinite number of scenarios that can occur when we simulate the uncertainty in this manner. Therefore, we would need an infinite number of repetitions to guarantee that we have considered all scenarios. However, by increasing the value of r until the results begin to converge we can be more confident of our results.

Repetitions	Uncertainty												
	0.3	0.6	0.9	1.2	1.5	1.8	2.1	2.4	2.7	3	3.3	3.6	5
1	6	6	6	7	8	8	8	8	9	9	10	9	10
10	12	16	19	21	22	24	26	25	24	24	24	24	29
100	18	27	33	35	36	38	40	44	46	48	48	49	55
1,000	23	32	37	42	48	51	53	55	56	58	60	60	64
10,000	28	41	47	52	54	55	58	59	61	62	63	64	66

Table 6.2: Number of efficient DMUs for $u \in \{0.3, 0.6, \dots, 3.6, 5\}$ for increasing number of replications.

6.2 Fixed uncertainty: Robust DEA

In Section 4.2 we concluded that for fixed values of uncertainty the overall best improvement in efficiency score is obtained by solving model (4.19). We wish to determine whether the plans in Dataset 2b are efficient for the same values of u used in Section 6.1. Solving model (4.19) gives the results in Table 6.3. With $u = 0.6$ there are only seven plans that are inefficient, all of which are efficient at $u = 0.9$. Therefore, the results for $u > 0.9$ have been omitted from Table 6.3. The plans that appear to be very inefficient in the nominal DEA may require less uncertainty to be deemed efficient, similar to the simulation results. For example, plan 66 has $\theta_0^* = 0.35$ when $u = 0$ but is efficient when $u = 0.3$, whereas, plan 58 has $\theta_0^* = 0.83$ at $u = 0$ but is not efficient until $u = 0.6$. This is because different facets of the efficient frontier have different rates at which their inputs change with their outputs. This can be easily seen in the case of a single input and output in Theorem 4.3.5. Here as the rate at which the outputs change compared to the inputs on the efficient frontier decreases, the required uncertainty for an inefficient DMU to become efficient when compared to that facet decreases. This can also be caused by plans being projected to sections of the efficient frontier that are weakly efficient. We note that the amount of uncertainty needed is not monotonic with the original efficiency score. To find the value of u to two decimal places required for each plan to become efficient we calculate the efficiency scores from the robust

DEA model (4.19) for $u \in \{0, 0.01, \dots, 0.9\}$ to determine for what value of u each plan first becomes efficient. The results are in Table 6.4. There are six plans that require $u > 0.65$ to become efficient, they are plans 30, 37, 85, 89, 91 and 97. These six plans include the five least efficient plans from the simulation results in Table 6.1. Unlike the simulation, for each value of u these results are very fast to obtain as each requires a single LP to be solved for each DMU as opposed to r LPs for each DMU in the simulation.

6.3 Comparison

From our simulation results with $r = 10,000$ and $u = 0.9$, only 47 plans are efficient. However, after applying robust DEA to Dataset 2b we find that all the plans are efficient when $u \leq 0.9$. We wish to determine whether this is because we have simply not run the simulation for large enough r , or if the results do actually differ. We now repeat the simulation for $r \in \{100,000, 200,000, 400,000\}$ to see if the number of efficient DMUs increases. After 100,000 there are 51 plans that are efficient, 200,000 52 plans, 400,000 53 plans. Therefore, with a large enough number of repetitions we believe the results of the simulation and robust DEA are consistent. The large number of replications required for the plans to become efficient with $u = 0.9$ show that these are the plans that may not become efficient with small amounts of uncertainty. Therefore, they may be more likely to benefit from replanning.

We now test the hypothesis that with enough repetitions, the robust DEA results agree with the simulation by using a small subset of the data with a single input and output so that we can also calculate the exact amount of uncertainty required using methodology from Section 4.3.

Plan	Uncertainty			
	0	0.3	0.6	0.9
30	0.52	0.56	0.60	1
31	0.77	1	1	1
32	0.72	1	1	1
33	0.53	0.58	0.82	1
34	0.78	1	1	1
35	0.44	1	1	1
36	1	1	1	1
37	0.49	0.52	0.65	1
38	0.89	1	1	1
41	0.39	0.42	1	1
42	0.66	1	1	1
43	1	1	1	1
44	0.41	1	1	1
45	1	1	1	1
46	0.92	1	1	1
47	0.60	1	1	1
48	0.75	1	1	1
49	0.58	0.65	1	1
50	0.41	0.65	1	1
51	0.67	1	1	1
52	0.42	1	1	1
53	0.66	1	1	1

Plan	Uncertainty			
	0	0.3	0.6	0.9
54	0.54	0.58	1	1
55	0.47	1	1	1
56	0.58	1	1	1
57	0.42	1	1	1
58	0.83	0.91	1	1
59	0.68	1	1	1
60	1	1	1	1
61	0.86	1	1	1
62	0.71	1	1	1
63	0.30	0.36	1	1
64	0.69	0.99	1	1
65	0.37	0.66	1	1
66	0.35	1	1	1
67	0.84	1	1	1
68	0.79	1	1	1
70	0.52	0.61	1	1
71	0.50	0.56	1	1
72	0.86	1	1	1
73	0.58	1	1	1
74	1	1	1	1
75	0.54	0.77	1	1
76	0.58	0.69	1	1

Plan	Uncertainty			
	0	0.3	0.6	0.9
77	0.80	1	1	1
78	0.51	0.60	1	1
80	0.58	0.71	1	1
81	0.49	0.53	1	1
82	0.68	1	1	1
83	0.39	1	1	1
84	0.61	0.72	1	1
85	0.49	0.53	0.86	1
86	0.45	0.49	1	1
87	0.61	1	1	1
88	0.94	1	1	1
89	0.41	0.43	0.47	1
90	0.51	0.58	1	1
91	0.49	0.52	0.57	1
92	0.92	1	1	1
93	0.75	1	1	1
94	0.71	1	1	1
95	0.76	1	1	1
96	0.69	1	1	1
97	0.40	0.43	0.47	1
98	0.90	1	1	1
99	0.42	0.61	1	1

Table 6.3: Efficiency scores from robust DEA model.

Plan	First efficient	Plan	First efficient	Plan	First efficient
30	0.69	54	0.58	77	0.21
31	0.25	55	0.28	78	0.40
32	0.13	56	0.29	80	0.35
33	0.64	57	0.19	81	0.51
34	0.03	58	0.44	82	0.27
35	0.28	59	0.28	83	0.19
36	0	60	0	84	0.43
37	0.68	61	0.09	85	0.63
38	0.20	62	0.17	86	0.42
41	0.47	63	0.45	87	0.25
42	0.04	64	0.31	88	0.22
43	0	65	0.34	89	0.90
44	0.10	66	0.28	90	0.35
45	0	67	0.26	91	0.68
46	0.17	68	0.12	92	0.16
47	0.08	70	0.44	93	0.17
48	0.04	71	0.54	94	0.06
49	0.35	72	0.03	95	0.21
50	0.34	73	0.21	96	0.16
51	0.28	74	0	97	0.79
52	0.19	75	0.35	98	0.20
53	0.30	76	0.38	99	0.37

Table 6.4: Uncertainty required for each plan to become efficient from robust DEA results.

Plan	Input	Output	Efficiency score
41	14.3286	101.4194	0.2461
42	13.4308	102.2660	0.4144
43	12.3384	102.3403	1
44	17.9491	102.1520	0.1964
45	3.5256	102.2463	1

Table 6.5: Data for D_{98}^{PTV2-3} and V_{70}^{Bla} , plans 41-45, Example 11.

Example 11: Plans 41-45 Here we use data for plans 41-45 with input V_{70}^{Bla} and output D_{98}^{PTV2-3} . The data for these are in Table 6.5 along with the nominal DEA efficiency scores. Here plans 43 and 45 are efficient.

First we consider $u \in \{0.3, 0.6, \dots, 3.6, 5\}$, and simulate the results for $r \in \{10, 100, 1000\}$. The results are shown in Table 6.6. From Table 6.6 plan 42 and 44 are efficient when $u = 0.1$ and plan 41 by $u = 0.7$.

For robust DEA if we have a fixed value of uncertainty we can calculate whether or not a DMU is efficient. When $u \in \{0.1, 0.2, \dots, 0.7\}$ we get the results in Table 6.7. In Table 6.7 we find that plan 41 is efficient when $u = 0.5$. Therefore, we then try robust DEA for $u \in \{0.4, 0.41, \dots, 0.5\}$ and find plan 41 is first efficient when $u = 0.47$. Further robust DEA application gives $u = 0.461$. However, from the simulation results in Table 6.6 plan 41 is not efficient until $u = 0.7$. Therefore, we now repeat the simulations for a larger number of repetitions with $u = 0.461$; the results are in Table 6.8. As the number of repetitions increases the maximum efficiency score for plan 41 with $u = 0.461$ increases.

Plan	r=10					r=100				r=1000			
	Uncertainty					Uncertainty				Uncertainty			
	0.1	0.3	0.5	0.7	0.9	0.1	0.3	0.5	0.7	0.1	0.3	0.5	0.7
41	0.25	0.26	0.32	0.72	1	0.25	0.27	0.38	1	0.25	0.27	0.53	1
42	1	1	1	1	1	1	1	1	1	1	1	1	1
43	1	1	1	1	1	1	1	1	1	1	1	1	1
44	0.40	1	1	1	1	0.48	1	1	1	1	1	1	1
45	1	1	1	1	1	1	1	1	1	1	1	1	1

Table 6.6: Maximum efficiency score for $u \in \{0.1, 0.3, \dots, 0.9\}$ and $r \in \{10, 100, 1,000\}$, Example 11.

Plan	Uncertainty						
	0.1	0.2	0.3	0.4	0.5	0.6	0.7
41	0.25	0.26	0.27	0.28	1	1	1
42	1	1	1	1	1	1	1
43	1	1	1	1	1	1	1
44	1	1	1	1	1	1	1
45	1	1	1	1	1	1	1

Table 6.7: Efficiency score robust $u \in \{0.1, 0.2, \dots, 0.7\}$, Example 11.

r	Plan 41 efficiency
1	0.2221
10	0.2791
100	0.3028
1000	0.3105
10,000	0.4350
100,000	0.5682
1,000,000	0.6567
10000000	0.8711

Table 6.8: Simulation results, efficiency score for plan 41 with $u = 0.461$ for increasing numbers of repetitions, Example 11.

Finally, we wish to compare the robust DEA and simulation results to the exact value of u from uDEA which we can calculate because we have a single input and output. From Theorem 4.3.6 plan 41 should be compared to the efficient frontier facet $y = y_{43}$. Then from (4.26) the value of u can be calculated to be 0.46045. Therefore, the first time we have enough uncertainty for plan 41 to be considered efficient to three decimal places is $u = 0.461$ which agrees with our robust DEA results. Table 6.9 summarises the results of the three methods.

Although a large number of repetitions of the simulation does give similar results to the robust DEA and uDEA models, it does not fully capture the effect of uncertainty. The robust DEA model agrees with the uDEA model and can be used to find the amount of uncertainty required to the desired precision. However, unlike the uDEA model it does not give us the exact amount of uncertainty. Although this is not a concern for the current application, as a suitable precision can be determined, there may be other applications that require exact solutions. This would require further research into solving the uDEA model for different uncertainty sets and for methods to solve it exactly as $M + N$ and I increase.

Plan	Simulation	robust DEA	uDEA
41	0.871	0.461	0.46045
42	1	0.038	0.03715
43	1	0	0
44	1	0.095	0.0941
45	1	0	0

Table 6.9: Maximum efficiency score from simulation, robust DEA and uDEA, Example 11.

6.4 Replanning

We now test our method clinically. To do this we select a small subset of plans for Rosemere to replan. This subset contains plans we believe can be improved and ones we believe are already efficient. Planners at Rosemere are not aware that some plans are efficient to avoid bias and check our measure of efficiency is clinically important. We choose a wide selection of plans including plans 36 and 43 which are efficient for the nominal DEA, plan 89 which requires the largest amount of uncertainty to become efficient and plan 63 which had the smallest efficiency score for the nominal DEA (without bladder volume included). All the plans in the dataset have been used for patient treatment and therefore meet mandatory tolerances. This means we do not expect to see large improvements as these plans are already clinically acceptable according to current practice.

When we first requested plans to be replanned, Rosemere started by trying to reduce rectal doses using the usual methods they employ when a plan’s rectal dose is outside mandatory tolerances. This involves two steps: reducing the posterior border of PTV 1 from a 1 cm expansion of CTV 1 to a 0.7 cm expansion, (see Section 2.1.2) and outlining the volume of the rectum that lies outside the PTV1, they call this the ‘rectum remaining’. They can then apply an additional maximum dose constraint to rectum remaining. Clinically these two steps would only be taken if the mandatory rectal constraints were not met and

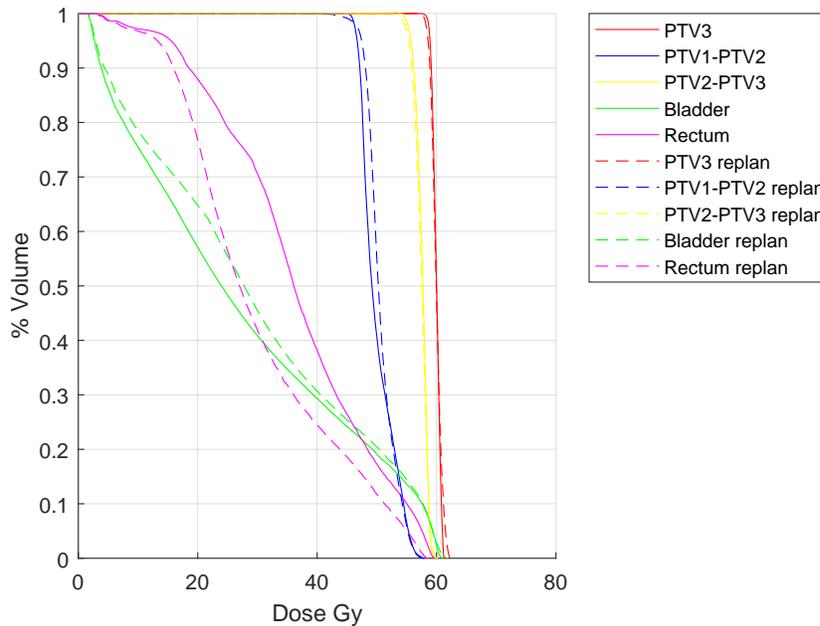


Figure 6.1: Plan 85 DVH comparing original plan and after replanning.

only with consultation with a clinician. Unfortunately, altering the volume of PTV 1 means plans are no longer clinically comparable as the structure definitions are not the same. These plans cannot be used as comparable DMUs. Therefore, Rosemere replanned the plans without modifying the PTV 1, only applying a rectum remaining constraint. They pointed out that this replanning method is more akin to what they might attempt if the rectal doses were within the mandatory tolerances but not achieving the desired tolerances. This highlights the fact that some plans in our dataset had already had their PTV 1 redefined and hence should no longer be included. These plans have not been included in Dataset 2b. After replanning we are hoping that the V_x metrics for the OARs will have decreased and the D_x metrics for the PTVs will still meet the prescribed doses. Table 6.10 shows the difference between the replan and the original plan.

We can also plot the DVH curves of the original and replanned plans to see the changes. An example for plan 85 can be seen in Figure 6.1. The dotted lines show the replanned versions. Here, an improvement in the plan will be shown by dashed lines lower than the solid for the rectum and the bladder (purple and green lines) and little change in the PTV

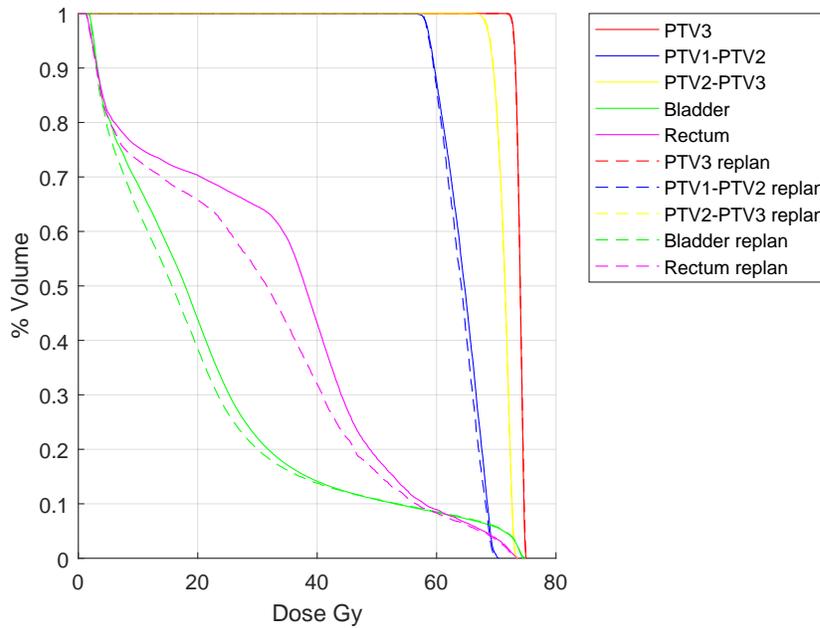


Figure 6.2: Plan 53 DVH comparing original plan and after replanning.

lines. Many of the replans showed a good reduction in the bladder and rectum DVH curves for the lower doses, 20-50 Gy, see Figure 6.2. Planners at Rosemere showed the DVH curves to two clinicians. They said that when comparing two plans they would rarely look at the DVH curves to determine which was a better plan, instead they would rely on the V_x metrics as hard limits. As a result, we then asked what change in these hard constraints would be considered as a significant clinical improvement and hence a better plan for a patient? They concluded that a 2% change in a V_x metric would be regarded as significant and indicate a ‘better’ plan if the PTV target doses were still met. A 2% change in a V_{70} metric would be more significant than a 2% change in other DVH points. This way of reporting significant changes reinforces our decision to use an input-oriented DEA model as they are interested in the change in the inputs, V_x metrics, while the outputs remain fixed. In Table 6.10 we highlight the plans that have improved by over 2%. Here we see that plan 42 has improved by over 2% for V_{60}^{Rec} , V_{65}^{Rec} and V_{70}^{Rec} , however the trade-off is that the D_{98}^{PTV1-2} has reduced by 0.9 Gy, the largest decrease in D_x metric. Further consultations are required to determine which would be preferable here.

Plan	Variables: difference between replan and original								
	D_{98}^{PTV1-2}	D_{98}^{PTV2-3}	D_{98}^{PTV3}	V_{60}^{Rec}	V_{65}^{Rec}	V_{70}^{Rec}	V_{60}^{Bla}	V_{65}^{Bla}	V_{70}^{Bla}
30	-0.10	-0.05	-0.05	0.33	-0.25	-0.51	0.10	0.04	0.02
36	-0.05	0.15	-0.05	-0.28	-0.38	-0.63	0.14	0.13	0.12
37	-0.05	0.20	0.05	-1.17	-1.79	-1.86	-0.18	-0.37	-1.71
41	-0.50	-0.30	-0.10	-0.26	-0.24	-1.00	-0.45	-0.36	-1.36
42	-0.90	-0.70	-0.15	-2.29	-2.80	-6.25	-0.73	-0.69	-0.93
43	-0.20	-0.05	0.00	0.02	0.03	0.28	-0.03	-0.08	-0.04
50	-0.45	-0.40	-0.10	-0.22	0.15	-0.07	-0.55	-0.47	-0.49
51	-0.15	-0.10	0.05	-0.10	-0.13	-0.09	0.01	-0.02	0.03
53	-0.30	-0.10	0.05	0.40	-0.41	-0.41	-0.12	-0.10	-0.22
63	0.35	-0.35	-0.55	0.96	1.04	1.07	-3.15	-0.38	0.14
85	0.35	-0.45	-0.45	-4.64	-5.14	-5.84	0.49	1.04	1.25
89	0.20	-0.15	-0.15	-0.69	-0.40	-0.64	-0.18	-0.09	-0.12
91	-0.05	0.10	0.10	0.05	0.14	-0.27	0.04	0.06	-0.05
97	0.05	0.15	0.35	-0.04	0.10	-0.44	0.39	0.17	0.11

Table 6.10: Difference in variables replan-original for selected inputs (%) and outputs (Gy). Bold numbers represent an improvement of over 2Gy in the V_x metrics.

Plan	Nominal efficiency score		After replanning			
	without bladder	with bladder	without bladder		with bladder	
			original	replan	original	replan
30	0.52	0.52	0.52	0.52	0.52	0.52
36	1.00	1.00	1.00	1.00	1.00	1.00
37	0.49	0.55	0.45	0.50	0.55	0.61
41	0.39	0.39	0.39	0.40	0.39	0.40
42	0.66	0.66	0.66	0.45	0.66	0.45
43	1.00	1.00	1.00	0.82	1.00	0.89
50	0.41	0.41	0.41	0.43	0.41	0.43
51	0.67	0.67	0.67	0.67	0.67	0.67
53	0.66	0.66	0.66	0.67	0.66	0.67
63	0.30	0.40	0.28	0.27	0.40	0.42
85	0.49	0.49	0.49	0.60	0.49	0.60
89	0.41	0.51	0.38	0.39	0.50	0.51
91	0.49	0.49	0.49	0.49	0.49	0.49
97	0.40	0.40	0.40	0.40	0.40	0.40

Key	Improved	Same	Worse
------------	----------	------	-------

Table 6.11: Comparing replans efficiency scores.

Replanning was undertaken by different staff, each spending varying amounts of time trying to improve the plan. This is representative of a normal replanning procedure due to staff shifts, experience level and varying workloads. Planners reported that they found it impossible to improve two plans in particular, plan 36 and 85. From Table 6.11, plan 36 was efficient in the nominal solution, so this is expected. The other plan we gave to Rosemere that was efficient in the nominal DEA was plan 43. In trying to improve plan 43 planners have in fact reduced the efficiency score. From Table 5.21, we see that plan 36 was the efficient plan that was a peer to the most plans, 55 plans, whereas plan 43 was a peer to only 3 plans. This suggests that plan 36 was a particularly good plan, alluding to the idea of a gold standard plan defined for Rosemere. This highlights that care should be taken during replanning as it can sometimes reduce a plan's efficiency. This reinforces the work we have proposed in this Thesis and the use of EvaluatePlan we introduce in Chapter 7 to ensure the final plan selected for a patient is the most favourable.

The second plan Rosemere struggled to improve was plan 85. However, in Table 6.11 we identify plan 85 as having improved its efficiency score. In Table 6.10 we see that the V_{70}^{Rec} has improved by 5.84%, which would be counted as a clinical improvement. Unfortunately, this plan is no longer clinically acceptable as the prescribed dose to the D_{98}^{PTV1-2} and D_{98}^{PTV2-3} are no longer acceptable. This means the plan is no longer comparable as we currently define our DMUs to be clinically acceptable treatment plans. It will be beneficial to adapt our methodology to account for plans that do not meet the prescribed doses and are hence deemed clinically unacceptable. This can be done by adding additional constraints to the DEA model.

In Table 6.11 the nominal efficiency scores in column two (three) and the efficiency scores in column four(six) for the same plans when we consider replans may differ in value. This is because the reference sets are different. Columns two and three only include the original plans and columns four and six have the additional replans included, which can change the PPS. The nominal efficiency scores are included as an indication of which plans were originally efficient, not as a comparison with columns four to seven. However, the efficiency scores in columns four and five and similarly, six and seven can be compared with

one another as they are obtained from the same DEA model and hence the same reference set.

Plan 89 was the plan that required the most uncertainty to be deemed efficient. In Table 6.11 it is one of the plans that the efficiency score improves both with and without bladder volume. Similarly, plan 63 was the plan that had the lowest efficiency score in the nominal DEA problem and improves when we include bladder volume.

We conclude that, as expected, there was very little improvement in the plans after replanning. This reinforces the notion that planners at Rosemere are consistently able to produce good treatment plans for patients, that are both clinically acceptable and require very little uncertainty to be deemed efficient. The results highlight the problem that can occur when planners try and improve an already efficient plan and reinforces the need for a quantitative tool to evaluate the efficiency of a plan.

Chapter 7

Software

Here we provide details of a simplified proof of concept piece of software, here called EvaluatePlan to be used in treatment centres. Developing this is a vital area of further work. Currently EvaluatePlan utilises results from Chapter 4 on fixed uncertainty and does not use uDEA methodology. As we improve the heuristics and the exact methods for solving the uDEA problem, this will allow us to develop EvaluatePlan further.

We have chosen to develop EvaluatePlan using R (R Core Team, 2013) because it has a number of advantages. Firstly, it is open source so purchasing of software licenses by Rosemere is not required. This is an important consideration as we wish the research to be easily adopted and do not wish funding costs to be a barrier. The use of the R shiny package (Chang et al., 2019) allows the development of a simple GUI that can be interfaced with Excel[®]; Rosemere currently output their plan evaluation data to Excel[®]. Furthermore, the GUI can be packaged on a memory stick with a minimal version of R, again making uptake of EvaluatePlan as simple as possible. Finally, R has a large statistical package library allowing further developments to include our autoPCA method. In the remainder of this chapter, we discuss EvaluatePlan, provide users with simple instructions for its use and suggest further adaptations to future versions.

7.1 Interfacing with Rosemere

Currently before a treatment plan is finalised the metrics from a plan are exported to an Excel[®] spreadsheet that highlights the constraints that have met preferred/mandatory or failed mandatory doses with green, orange, red respectively. This allows the clinician in charge of the plan to quickly assess the plan. Therefore, to interface EvaluatePlan easily with Rosemere we wish to use the evaluation spreadsheet as the input data.

From Chapter 4 Theorem 4.1.16, for inefficient DMUs, there exists an optimal solution to the uncertain problem (4.3) such that only DMUs that are efficient in the nominal DEA problem are peers to the inefficient DMUs. This means we only need to store plans that are efficient in the nominal DEA problem. This reduces the size of the dataset required, thus speeding up the computation time and reducing the memory required for the software. The input data required is a dataset of the nominal efficient plans, from Tables 5.20 and 5.22 they are plans 31, 36, 38, 43, 45, 60, 74, 92, 94 and 98. Here we include those that are efficient in the nominal solution both with and without the inclusion of the bladder volume as an environmental variable.

The user input required is to select the evaluation spreadsheet plan they wish to assess and whether they wish to include the bladder volume as an environmental variable or not.

7.2 Software functionality

The aim of EvaluatePlan is to assess the efficiency of the treatment plan. If it is efficient in the nominal DEA, EvaluatePlan will notify the user of this and the analysis is complete. If the plan is inefficient in the nominal DEA problem, EvaluatePlan will produce the following information:

- The peers of the plan
- The minimum amount of uncertainty to two decimal places required for the plan to become efficient if $u < 3.6$.

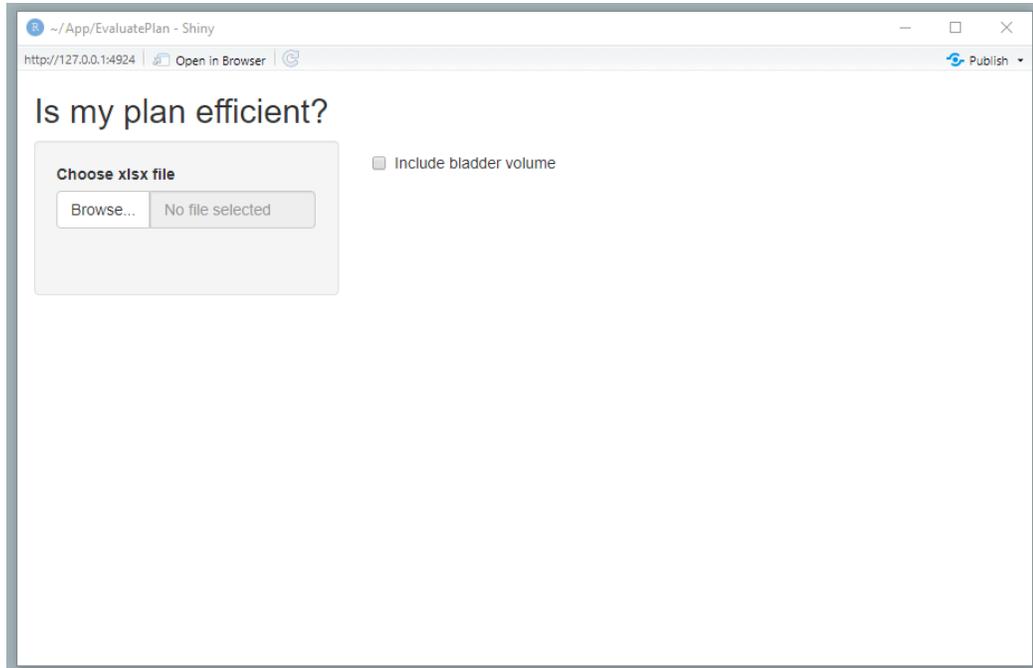
The user can also decide if they wish to include Bladder Volume as an environmental variable or not.

7.3 How to use EvaluatePlan

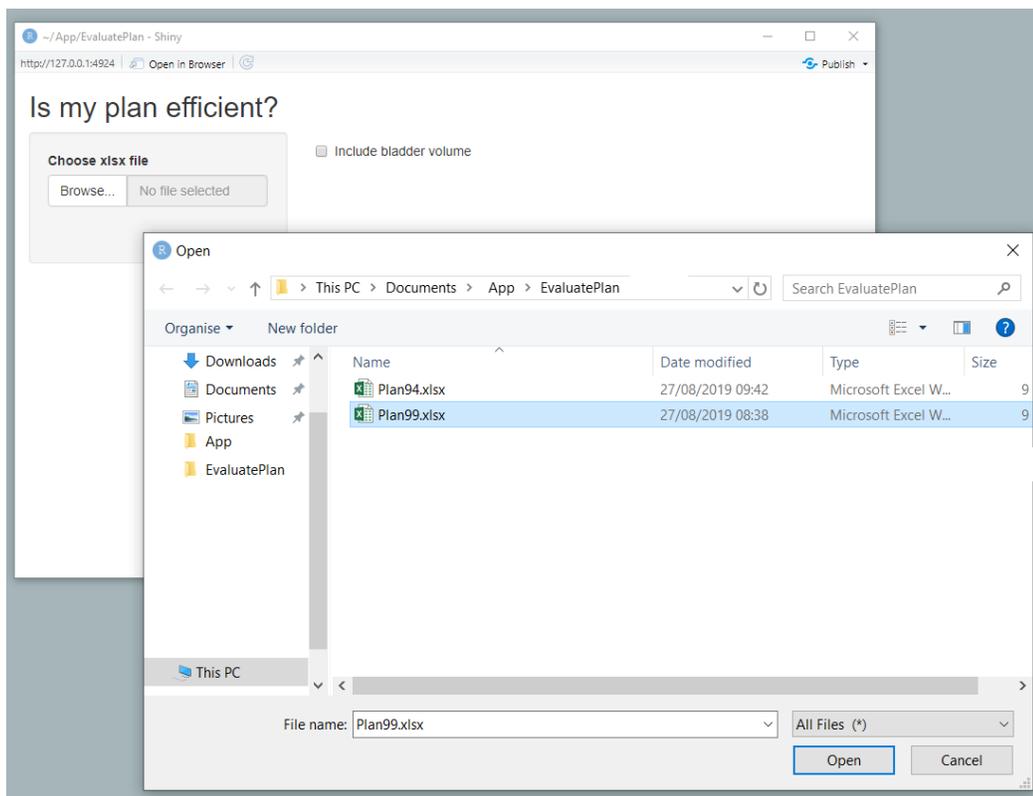
EvaluatePlan can be run from a memory stick loaded with a small version of R (R Core Team, 2013) and the R shiny package (Chang et al., 2019). After clicking launch, EvaluatePlan loads the screen shown in Figure 7.1a. The user can then select an evaluation spreadsheet for the plan they wish to assess, which can be stored anywhere on their computer. By clicking on the browse button, in Figure 7.1b, the user selects plan 99.

Once a plan has been uploaded, EvaluatePlan outputs the results from the nominal DEA (3.8) and the robust DEA model (3.19) with $u \in \{0, 0.01, \dots, 3.6, 5\}$. If the plan is not efficient in model (3.8), the nominal efficiency score, the amount of uncertainty required to become efficient and the peers of the plan are displayed. This can be seen in Figure 7.2a where plan 99 has a nominal efficiency score of 0.423, requires $u = 0.37$ to become efficient and has peers plan 53 and 60. The user can also compare the efficiency of the plan when the bladder volume is included. This is done by selecting the tick box at the top of EvaluatePlan, highlighted in Figure 7.2b. By selecting (or deselecting) this box, EvaluatePlan re-evaluates the plan with the bladder volume included(excluded). In Figure 7.2b, we see that by including the bladder volume, plan 99 requires less uncertainty to become efficient and the plan's peers change. If a plan is efficient in the nominal DEA problem, EvaluatePlan will display "This plan is efficient". This can be seen in Figure 7.3b where plan 94 is efficient when the bladder volume is included. However, it is not efficient when the bladder volume is excluded, Figure 7.3a. If a plan is inefficient when $u = 5$, EvaluatePlan will output this message and output its peers (see Chapter 6 for justification of this).

EvaluatePlan is easy to use, and by solving the robust DEA model gives a value of uncertainty accurate to two decimal places. It allows planners to quickly assess whether a plan is efficient based on plans achieved previously. EvaluatePlan has the option of including the

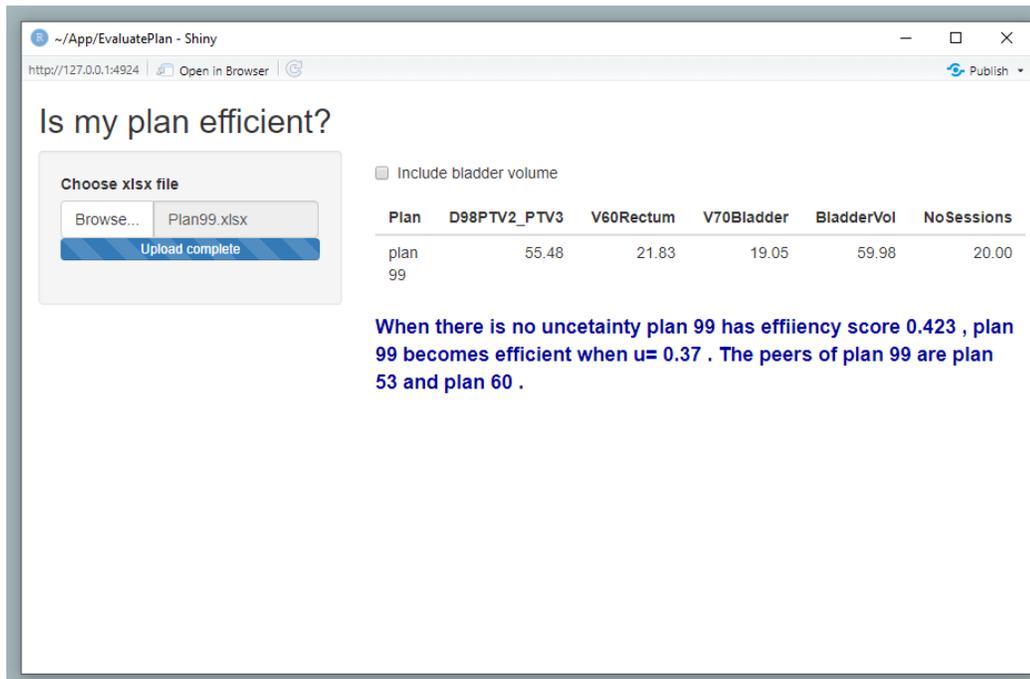


(a) Load EvaluatePlan.

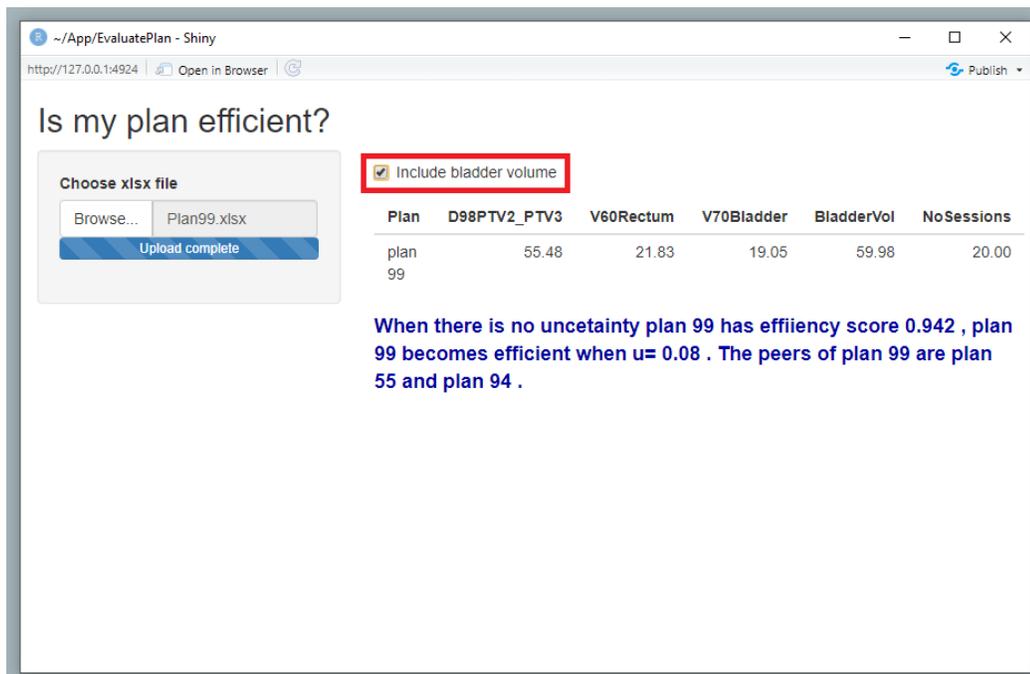


(b) Select the plan to evaluate.

Figure 7.1: Launching and selecting the plan to evaluate in EvaluatePlan.

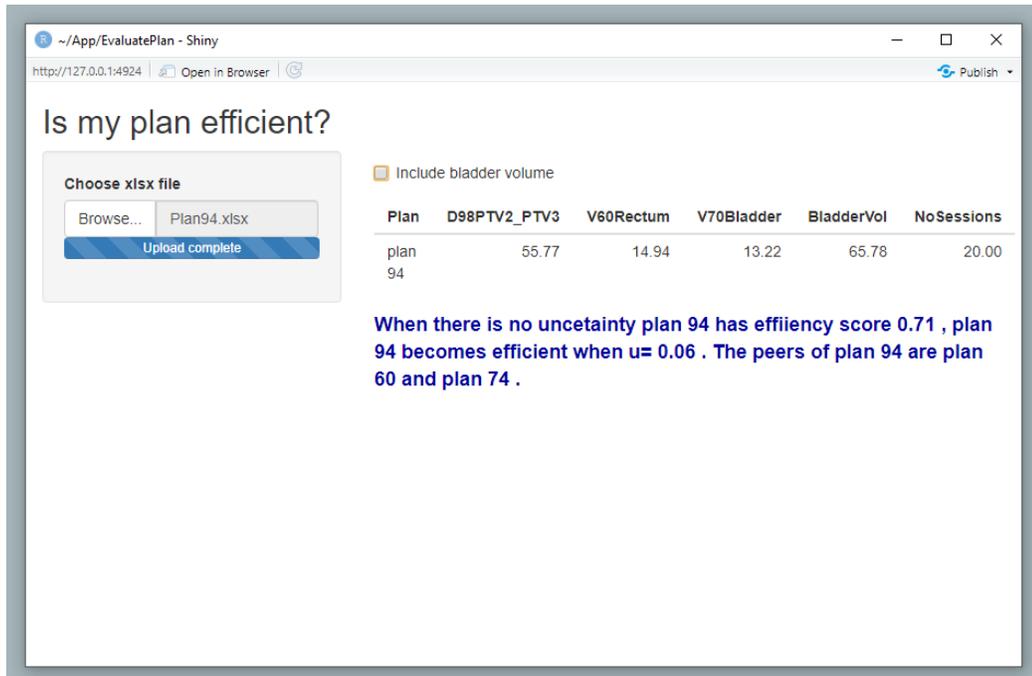


(a) No bladder Volume.

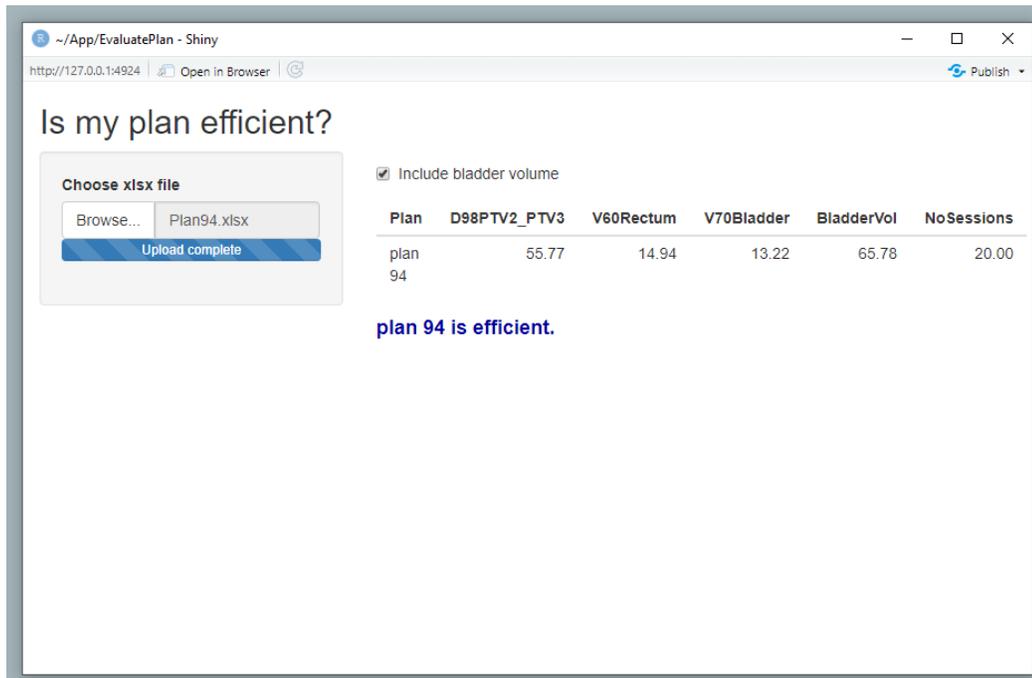


(b) Bladder volume included.

Figure 7.2: Results for plan 99 from EvaluatePlan.



(a) No bladder Volume.



(b) Bladder volume included.

Figure 7.3: Results for plan 94 from EvaluatePlan.

bladder volume as an environmental variable in the analysis. This is because EvaluatePlan has been developed specifically for Rosemere. They are interested in assessing the difference the bladder volume has on plan creation. In future versions different environmental variables can easily be added for different centres.

7.4 Future software improvements

EvaluatePlan, discussed in this chapter, has the potential to be extended in many directions.

A few improvements we hope to make in the future include:

- User defined inputs and outputs to allow users to explore the effect of certain constraints on the efficiency of a plan.
- Added user inputs to include ranges of uncertainty they wish to consider further.
- Exact amounts of uncertainty by incorporating uDEA methodology.
- Batch processing of plans where multiple evaluation forms can be inputted at once.
- Integration with other hospitals evaluation spreadsheet/software.
- Conversion to Python code to allow direct integration into some treatment plan software.

Continued feedback and evaluation from users of EvaluatePlan at Rosemere will also help to develop this further.

Chapter 8

Conclusion and further work

External beam radiotherapy is one of the major forms of cancer treatment. About two thirds of all cancer patients undergo a course of radiotherapy at some stage of their treatment. Treatment planning involves delivering high doses of radiation to the tumour site while sparing surrounding organs. Most cancer centres worldwide do not have access to multi-objective treatment planning systems. Therefore, plans are made on single objective planning systems many of which use a weighted sum method to incorporate different conflicting objectives. To improve a plan, multiple iterations of the planning process are required where practitioners change parameters based on prior knowledge and experience. As a result, although a plan is deemed acceptable due to it meeting certain clinical criteria it is very hard for planners to tell if it is in fact the best plan they can achieve for that patient. We evaluate the quality of treatment plans generated by an existing system and aim to provide recommendations to planners on possible improvements. This is done using DEA. We assess how well the plans perform in terms of converting inputs to outputs that is protecting the OARs while delivering the prescribed dose to the tumour. This study is motivated by data from Rosemere Cancer Centre, Royal Preston Hospital, Lancashire and specialist knowledge from clinicians therein. Variables from Rosemere's clinical protocol are extracted and a variable selection technique, autoPCA, is developed to alleviate the dependence of PCA methods on the sample at hand.

DEA is then applied using the variables selected from autoPCA. On Rosemeres recommendation, DEA is applied without the bladder volume as an environmental variable to contribute to their ongoing investigation into the effect of bladder volume during treatment. To account for the inherent uncertainty in the treatment process simulation, Robust DEA and uDEA techniques are compared. We extend the uDEA methodology for the specific case of box uncertainty after identifying its suitability from the medical literature. This allows us to assess how well the plans perform in converting inputs to outputs, i.e. protecting the OAR while delivering the prescribed dose to the tumour. We determine plans that may benefit from replanning and conclude that the plans for prostate cancer at Rosemere are consistently good compared to one another. This suggests that prostate cancer planning is an area in which Rosemere succeed in tailoring plans. Only plans that have been deemed acceptable for treatment and used on a patient are included in our dataset. Therefore, they are not necessarily the first plans created on the treatment planning software. By using our methodology on the first treatment plan created for a patient we can advise planners of the potential improvement possible and the amount of uncertainty required for a plan to be deemed efficient. This will allow them to streamline their planning process by prioritising plans that are able to improve the most.

Previous research focused on prostate cancer, which many hospitals count as their ‘bread and butter’ due to it being the most common cancer type for men and being relatively easy to treat. They perform many similar treatments each year and they have seen rapid improvement in the past few decades. This work was an extension of the method proposed in Lin et al. (2013). To ensure it is applicable to current treatment planning methods in the UK, and is suitable for different centres, a large body of this work focused on variable selection and considering the uncertainties arising throughout the treatment process. Our findings reinforce the idea that these plans are in fact all good as only minor improvements were possible. By considering uncertainty in the planning variables, we can account for the effects of uncertainties present throughout the treatment process.

We conclude by considering the possible areas of further work that are linked to the research presented throughout this thesis. Many of these areas have been introduced in the

previous chapters, we discuss them further here.

A multicentre study Throughout this thesis we have used data from a single treatment centre, Rosemere. Their clinical protocol is based on the CHHiP clinical trial and guidelines from the Royal College of Radiologists (RCR, 2018), Clinical Oncology Information Network (Mason et al., 1999) and British Association of Urological surgeons (BAUS, 2019). We would like to extend our methodology to other treatment centres. Initially this will involve discussions with specialists to identify differences in treatment techniques and the suitability of our methodology. The correct variables to represent their plans can then be determined via autoPCA and discussions with the relevant clinicians. These may differ between treatment centres as they may have different constraints and prescribed doses in their clinical protocol and different treatment plan creation software.

A subset of northern hospitals are currently in consultation over unifying their clinical protocols. This would mean that the plans from different treatment sites could be compared as the plans would be clinically comparable. This would highlight differences between treatment plan quality at different centres and would help identify planning weaknesses.

Linking treatment plan quality with follow up results for prostate cancer

Previously radiotherapy treatment planning assessment has been based on metrics extracted from the treatment plan and how these relate to clinically prescribed recommended doses. In the future, we wish to consider the post-treatment prognosis of the patients. Two to three months after patients are treated with radiotherapy, they will have a follow up appointment to assess their side effects and treatment success, these appointments continue every three to six months for at least three years. The follow up appointments may also include blood tests to check their prostate specific antigen (PSA) level.

There have been many studies on the effectiveness of radiotherapy as a treatment for prostate cancer using follow up data (see Lane et al. (2010)) but no studies that link the treatment plan assessment and final outcome. Lambin et al. (2013) discuss how using the planned dose-distribution is not enough to base treatment assessment on due to the

deviations from the original plan and the need to create ‘multifactorial decision support systems’ to assess treatment quality.

The original plans from Rosemere are now over three years old. Rosemere are looking into obtaining PSA and side effect data for our original patients. In this way it is hoped that we can link the efficiency scores derived from our DEA methodology with the clinical outcome. We hope that an efficient plan translates to a better clinical outcome for the patient.

Extension to head and neck cancer Head and neck tumours are a broad category of tumours that include several sub-sites. Each sub-site presents its own unique planning challenges. They are particularly difficult to plan due to the large number of OARs that surround the tumours. These include the brain-stem, spinal-canal and the parotid arteries. To begin with we would consider a single subcategory of head and neck tumour, oropharynx, as it is one of the more common types so data are more widely available. The clinical protocol for some head and neck types are fairly similar due to the need to irradiate nodes that could potentially include tumour cells. Involved nodes are hard to identify on medical scans so often similar areas are identified on each patient to ensure all key areas are irradiated. We would like to extend the methods developed in this thesis to compare treatment plans across different head and neck sub-sites. In this way, less common cancer types can borrow information from other similar types. This would involve adding additional constraints to the DEA model in the form of modified environmental variables to account for different head and neck sub-sites.

Proton therapy Throughout this thesis we have used plans from IMRT prostate cancer patients. However, research into the use of proton therapy is at the forefront of cancer research. The methodology developed here would be easily extended to proton therapy. To do this we require comparable plans created from the same clinical protocol. As this is a growing area of research, the methodology and techniques are not as developed as IMRT for prostate cancer and hence the potential benefit of a method to selectively compare plans has

great clinical potential. Mohan and Grosshans (2017) highlight that “treatment planning and plan evaluation of passively-scattered proton therapy and intensity modulated proton therapy require special considerations compared to the processes used for photon treatment planning. The differences in techniques arise from the unique physical properties of protons but are also necessary because of the greater vulnerability of protons to uncertainties, especially from inter- and intra-fractional variations in anatomy. These factors must be considered in designing as well as evaluating treatment plans”.

Develop EvaluatePlan As discussed in Section 7.4, we wish to develop EvaluatePlan further. In particular, to develop EvaluatePlan to be compatible with different hospitals’ treatment planning software. This will involve adding additional user defined inputs and outputs, different ranges of uncertainty, batch processing of plans and ultimately incorporating the variable selection process into the software.

Refine autoPCA Throughout this thesis we have based our DEA analysis on a sample of 66 treatment plans. Although we were given over 100 plans in total, only 66 were truly comparable and hence suitable for our study. To help address our small sample size, the autoPCA method explores repeated sampling and derives the sampling distribution of the PCs. However, we would like to trial our method on a larger sample of data. This would increase our confidence in the conclusions we draw by validating our autoPCA method. If the variables selected represent the true variability of all plans at Rosemere, then the addition of new plans to our dataset should not change the variables identified. If there is a large change in the variables selected this may indicate a change in planning procedure: either a new treatment planning machine, such as the switch from AG to BM, or a change in the clinical protocol or software that creates the treatment plans. It would be beneficial to develop methodology that can account for future changes in the planning treatment process and update the variables accordingly.

uDEA areas of further research Building upon Ehrgott et al. (2018) who considered ellipsoidal uncertainty, we refined the concept of uDEA for the specific case of box uncer-

tainty. Ellipsoidal uncertainty is more general than box uncertainty and is less influenced by the extreme points of the uncertainty set (the corner points). The use of box uncertainty was motivated by the radiotherapy treatment planning application where uncertainty is assumed to lie in a set interval. Box uncertainty also has the advantage of being one of the most widely used uncertainty sets and results in tractable solutions to robust optimisation problems.

We now wish to use historical data from Rosemere to extend this methodology. We hope to identify if there is another uncertainty set that models the specific uncertainty at Rosemere, as opposed to the predicted uncertainty across the International Congress of Radiology network. If we extend the discussed methodology to other cancer sites we may need to model the uncertainty differently. Therefore, extending the uDEA methodology to other uncertainty sets would be beneficial.

We have compared the results of uDEA, robust DEA and simulation and conclude that although simulation results are converging to the results of uDEA, an extremely large number of replications are required. This makes simulation unsuitable for a clinical setting as it is too slow. Therefore, further work into fast, exact uDEA methods or a heuristic that gives results to a suitable degree of accuracy are required.

Appendix A

Plan	37 Session plans			Plan	20 Session plans		
	$gEUD^{Rec}$	D_{95}^{PTV1}	D_{95}^{PTV3}		$gEUD^{Rec}$	D_{95}^{PTV1}	D_{95}^{PTV3}
1	63.38	60.33	72.13	27	50.38	48.63	58.78
2	59.13	59.13	72.53	28	48.16	49.33	58.73
3	62.78	59.08	72.83	29	50.72	48.03	58.83
4	56.93	59.58	72.73	30	50.80	47.88	58.78
5	60.88	60.43	72.73	31	49.41	47.78	58.88
6	62.78	59.08	72.83	32	46.69	48.53	58.93
7	64.42	61.03	72.28	33	48.74	53.63	58.88
8	53.35	64.08	72.63	34	50.22	48.13	59.23
9	60.72	60.03	72.73	35	50.12	49.93	58.93
10	62.70	66.03	72.78	36	50.17	48.98	59.03
11	60.62	60.53	72.73	37	50.26	48.38	58.88
12	61.04	59.98	72.73	38	49.45	48.13	59.03
13	59.55	58.58	72.53	39	48.35	47.83	58.83
14	59.09	59.18	72.58	40	48.49	48.83	59.03
15	60.91	59.68	72.53	41	48.07	48.33	59.03
16	63.45	59.78	72.48	42	49.65	48.23	58.73
17	59.72	66.18	72.63	43	49.57	48.03	59.08
18	61.68	62.08	72.78	44	50.41	50.28	59.08
19	60.69	64.43	72.68	45	50.37	48.18	59.18
20	61.89	59.83	72.88	46	49.70	48.18	59.28
21	59.96	58.78	72.93	47	48.43	48.48	59.18
22	61.50	59.58	72.88	48	48.71	49.33	59.18
23	63.61	59.78	72.83	49	48.18	48.08	59.18
24	59.97	65.73	73.13	50	49.86	47.93	59.08
25	62.21	62.33	73.03	51	49.83	47.38	59.28
26	61.05	64.13	72.73				

Table 1: Dataset 1. published in Stubington et al. (2019)

Bibliography

- N. Adler and B. Golany. PCA-DEA. In *Modeling data irregularities and structural complexities in data envelopment analysis*, pages 139–153. Springer, 2007.
- N. Adler and E. Yazhensky. Improving discrimination in data envelopment analysis: PCA-DEA or variable reduction. *European Journal of Operational Research*, 202(1):273–284, 2010.
- A. I. Ali and L. M. Seiford. Translation invariance in data envelopment analysis. *Operations research letters*, 9(6):403–405, 1990.
- R. Allen and E. Thanassoulis. Improving envelopment in data envelopment analysis. *European Journal of Operational Research*, 154(2):363–379, 2004.
- R. Allozi, X. A. Li, J. White, A. Apte, A. Tai, J. M. Michalski, W. R. Bosch, and I. El Naqa. Tools for consensus analysis of experts contours for radiotherapy structure definitions. *Radiotherapy and Oncology*, 97(3):572–578, 2010.
- P. Andersen and N. C. Petersen. A procedure for ranking efficient units in data envelopment analysis. *Management Science*, 39(10):1261–1264, 1993.
- P. Andreo, J. Cramb, B. Fraass, F. Ionescu-Farca, J. Izewska, V. Levin, B. Mijnheer, J. Rosenwald, P. Scalliet, K. Shortt, et al. Commissioning and quality assurance of computerized planning systems for radiation treatment of cancer. Technical Report 430, International Atomic Energy Agency technical report series, 2004.

- J. A. Antolak, I. I. Rosen, C. H. Childress, G. K. Zagars, and A. Pollack. Prostate target volume variations during a course of radiotherapy. *International Journal of Radiation Oncology* Biology* Physics*, 42(3):661–672, 1998.
- J. Aparicio. A survey on measuring efficiency through the determination of the least distance in data envelopment analysis. *Journal of Centrum Cathedra*, 9(2):143–167, 2016.
- R. D. Banker and R. C. Morey. The use of categorical variables in data envelopment analysis. *Management Science*, 32(12):1613–1627, 1986.
- R. D. Banker, A. Charnes, and W. W. Cooper. Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management Science*, 30(9):1078–1092, 1984.
- BAUS. British association of urological surgeons homepage, 2019. URL <https://www.baus.org.uk/>.
- A. Ben-Tal and A. Nemirovski. Robust convex optimization. *Mathematics of Operations Research*, 23(4):769–805, 1998.
- A. Ben-Tal and A. Nemirovski. Robust solutions of uncertain linear programs. *Operations research letters*, 25(1):1–13, 1999.
- A. Ben-Tal and A. Nemirovski. Robust solutions of linear programming problems contaminated with uncertain data. *Mathematical Programming*, 88(3):411–424, 2000.
- A. Ben-Tal, L. El Ghaoui, and A. Nemirovski. *Robust optimization*, volume 28. Princeton University Press, 2009.
- S. M. Bentzen. Dose-response relationships in radiotherapy. *Basic Clinical Radiobiology*, 4th ed. London: Hodder Arnold, 2009.
- D. Bertsimas and J. Tsitsiklis. Introduction to linear programming. *Athena Scientific*, 1: 997, 1997.

- J. Bohsung, S. Gillis, R. Arrans, A. Bakai, C. De Wagter, T. Knöös, B. J. Mijnheer, M. Paiusco, B. A. Perrin, H. Welleweerd, et al. IMRT treatment planning a comparative inter-system and inter-centre planning exercise of the ESTRO QUASIMODO group. *Radiotherapy and Oncology*, 76(3):354–361, 2005.
- R. Bokrantz and A. Forsgren. An algorithm for approximating convex Pareto surfaces based on dual techniques. *INFORMS Journal on Computing*, 25(2):377–393, 2013.
- V. Boljunčić. Sensitivity analysis of an efficient DMU in DEA model with variable returns to scale. *Journal of Productivity Analysis*, 25(1-2):173–192, 2006.
- J. Booth and S. Zavgorodni. Set-up error & organ motion uncertainty: A review. *Australasian physical & engineering sciences in medicine*, 22(2), 1999.
- J. Booth and S. Zavgorodni. Modelling the variation in rectal dose due to inter-fraction rectal wall deformation in external beam prostate treatments. *Physics in Medicine & Biology*, 50(21):5055, 2005.
- J. T. Booth and S. F. Zavgorodni. The effects of radiotherapy treatment uncertainties on the delivered dose distribution and tumour control probability. *Australasian Physics & Engineering Sciences in Medicine*, 24(2):71, 2001.
- T. Bortfeld. IMRT: A review and preview. *Physics in Medicine & Biology*, 51(13):363–379, 2006.
- M.-L. Bournol and J. H. Dulá. Anchor points in DEA. *European Journal of Operational Research*, 192(2):668–676, 2009.
- S. Breedveld, P. R. Storchi, and B. J. Heijmen. The equivalence of multi-criteria methods for radiotherapy plan optimization. *Physics in Medicine & Biology*, 54(23):7199–7209, 2009.
- W. Briec and H. Leleu. Dual representations of non-parametric technologies and measurement of technical efficiency. *Journal of Productivity Analysis*, 20(1):71–96, 2003.

- P. Brockett, A. Charnes, W. Cooper, Z. M. Huang, and D. Sun. Data transformations in DEA cone ratio envelopment approaches for monitoring bank performances. *European Journal of Operational Research*, 98(2):250–268, 1997.
- E. Budiarto, M. Keijzer, P. Storchi, M. Hoogeman, L. Bondar, T. Mutanga, H. de Boer, and A. Heemink. A population-based model to describe geometrical uncertainties in radiotherapy: applied to prostate cases. *Physics in Medicine & Biology*, 56(4):1045, 2011.
- G. Cabrera, M. Ehrgott, A. Mason, and A. Philpott. Multi-objective optimisation of positively homogeneous functions and an application in radiation therapy. *Operations Research Letters*, 42(4):268–272, 2014.
- Cancer-Research-UK. Cancer incidence for common cancers, 2014. URL <https://www.cancerresearchuk.org/health-professional/cancer-statistics/incidence/common-cancers-compared>.
- W. Chang, J. Cheng, J. Allaire, Y. Xie, and J. McPherson. *shiny: Web Application Framework for R*, 2019. URL <https://CRAN.R-project.org/package=shiny>.
- A. Charnes and L. Neralic. Sensitivity analysis of the simultaneous proportionate change of inputs and outputs in data envelopment analysis. Technical report, Texas University at Austin center for Cybernetic Studies, 1991.
- A. Charnes, W. Cooper, and E. Rhodes. Measuring the efficiency of decision making units. *European Journal of Operational Research*, 1978.
- A. Charnes, W. Cooper, and E. Rhodes. Short communication: Measuring the efficiency of decision making units. *European Journal of Operational Research*, 3(4):339, 1979.
- A. Charnes, W. W. Cooper, and E. Rhodes. Evaluating program and managerial efficiency: an application of data envelopment analysis to program follow through. *Management Science*, 27(6):668–697, 1981.

- A. Charnes, W. W. Cooper, A. Y. Lewin, R. C. Morey, and J. Rousseau. Sensitivity and stability analysis in DEA. *Annals of Operations Research*, 2(1):139–156, 1984.
- A. Charnes, W. Cooper, B. Golany, L. Seiford, and J. Stutz. Foundations of DEA for pareto-koopmans efficient empirical production functions. *Journal of Econometrics*, 30: 91–107, 1985.
- A. Charnes, S. Haag, P. Jaska, and J. Semple. Sensitivity of efficiency classifications in the additive model of data envelopment analysis. *International Journal of Systems Science*, 23(5):789–798, 1992.
- A. Charnes, W. W. Cooper, A. Y. Lewin, and L. M. Seiford. *Data envelopment analysis: Theory, methodology, and applications*. Springer Science & Business Media, 1994.
- A. Charnes, J. J. Rousseau, and J. H. Semple. Sensitivity and stability of efficiency classifications in data envelopment analysis. *Journal of Productivity Analysis*, 7(1):5–18, 1996.
- Y. Chen and J. Du. Super-efficiency in data envelopment analysis. In *data envelopment analysis*, pages 381–414. Springer, 2015.
- R. Chowdhury, I. Wilson, C. Rofe, and G. Lloyd-Jones. *Radiology at a Glance*. John Wiley & Sons, 2010.
- T. Coelli, D. Rao, C. O’Donnell, and G. Battese. *An introduction to efficiency and productivity analysis*. Springer, 2005.
- W. Cooper, K. S. Park, and G. Yu. IDEA and AR-IDEA: Models for dealing with imprecise data in DEA. *Management Science*, 45(4):597–607, 1999.
- W. W. Cooper, R. G. Thompson, and R. M. Thrall. Chapter 1 introduction: Extensions and new developments in DEA. *Annals of Operations Research*, 66(1):1–45, 1996.
- W. W. Cooper, S. Li, L. Seiford, K. Tone, R. M. Thrall, and J. Zhu. Sensitivity and stability analysis in DEA: some recent developments. *Journal of Productivity Analysis*, 15(3):217–246, 2001.

- W. W. Cooper, L. Seiford, and J. Zhu. Data envelopment analysis. In *Handbook on data envelopment analysis*, pages 1–39. Springer, 2004.
- D. Craft and M. Monz. Simultaneous navigation of multiple Pareto surfaces, with an application to multicriteria IMRT planning with multiple beam angle configurations. *Medical Physics*, 37(2):736–741, 2010.
- D. Craft and C. Richter. Deliverable navigation for multicriteria step and shoot IMRT treatment planning. *Physics in Medicine & Biology*, 58(1):87–103, 2013.
- D. L. Craft, T. F. Halabi, H. A. Shih, and T. R. Bortfeld. Approximating convex Pareto surfaces in multiobjective radiotherapy planning. *Medical Physics*, 33(9):3399–3407, 2006.
- S. Daneshvar, G. Izbirak, and A. Javadi. Sensitivity analysis on modified variable returns to scale model in data envelopment analysis using facet analysis. *Computers & Industrial Engineering*, 76:32–39, 2014.
- G. Dantzig. Inductive proof of the simplex method. *IBM Journal of research and development*, 4(5):505–506, 1960.
- G. B. Dantzig and M. N. Thapa. *Linear Programming. 1, Introduction*. Springer-Verlag New York Incorporated, 1997.
- I. J. Das, C.-W. Cheng, K. L. Chopra, R. K. Mitra, S. P. Srivastava, and E. Glatstein. Intensity-modulated radiation therapy dose prescription, recording, and delivery: Patterns of variability among institutions and treatment planning systems. *Journal of the National Cancer Institute*, 100(5):300–307, 2008.
- D. Dearnaley, I. Syndikus, H. Mossop, V. Khoo, A. Birtle, D. Bloomfield, J. Graham, P. Kirkbride, J. Logue, Z. Malik, et al. Conventional versus hypofractionated high-dose intensity-modulated radiotherapy for prostate cancer: 5-year outcomes of the randomised, non-inferiority, phase 3 CHHiP trial. *The Lancet Oncology*, 17(8):1047–1060, 2016.
- J. O. Deasy, A. I. Blanco, and V. H. Clark. CERR: A computational environment for radiotherapy research. *Medical Physics*, 30(5):979–985, 2003.

- D. Despotis and Y. Smirlis. Data envelopment analysis with imprecise data. *European Journal of Operational Research*, 140(1):24–36, 2002.
- W. Dinkelbach. On nonlinear fractional programming. *Management Science*, 13(7):492–498, 1967.
- R. Drzymala, R. Mohan, L. Brewster, J. Chu, M. Goitein, W. Harms, and M. Urie. Dose-volume histograms. *Int. Journal of Radiation Oncology.biology.physics*, 21(1):71–78, 1991.
- R. Dyson, R. Allen, A. Camanho, V. Podinovski, C. Sarrico, and E. Shale. Pitfalls and protocols in DEA. *European Journal of Operational Research*, 132(2):245–259, 2001.
- M. Ehrgott and I. Winz. Interactive decision support in radiation therapy treatment planning. *OR Spectrum*, 30(2):311–329, 2008.
- M. Ehrgott, A. Holder, and J. Reese. Beam selection in radiotherapy design. *Linear Algebra and its Applications*, 428(5-6):1272–1312, 2008.
- M. Ehrgott, Ç. Güler, H. W. Hamacher, and L. Shao. Mathematical optimization in intensity modulated radiation therapy. *Annals of Operations Research*, 175(1):309–365, 2010.
- M. Ehrgott, A. Holder, and O. Nohadani. Uncertain data envelopment analysis. *European Journal of Operational Research*, 268(1):231–242, 2018.
- L. El-Ghaoui and H. Le Bret. Robust solutions to least-square problems to uncertain data matrices. *Sima Journal on Matrix Analysis and Applications*, 18:1035–1064, 1997.
- L. El Ghaoui, F. Oustry, and H. Le Bret. Robust solutions to uncertain semidefinite programs. *SIAM Journal on Optimization*, 9(1):33–52, 1998.
- A. Emrouznejad and G. R. Amin. Dea models for ratio data: Convexity consideration. *Applied Mathematical Modelling*, 33(1):486–498, 2009.
- A. Emrouznejad and G.-l. Yang. A survey and analysis of the first 40 years of scholarly literature in DEA: 1978–2016. *Socio-Economic Planning Sciences*, 61:4–8, 2018.

- M. Falkinger, S. Schell, J. Müller, and J. J. Wilkens. Prioritized optimization in intensity modulated proton therapy. *Zeitschrift für Medizinische Physik*, 22(1):21–28, 2012.
- M. J. Farrell. The measurement of productive efficiency. *Journal of the Royal Statistical Society: Series A (General)*, 120(3):253–281, 1957.
- J. Ferlay, I. Soerjomataram, M. Ervik, R. Dikshit, S. Eser, C. Mathers, M. Rebelo, D. Parkin, D. Forman, and F. Bray. Cancer incidence and mortality worldwide: IARC cancerbase no. 1, 2013.
- L. Feuvret, G. Noël, J.-J. Mazon, and P. Bey. Conformity index: A review. *Int. Journal of Radiation Oncology.biology.physics*, 64(2):333–342, 2006.
- F. Frei and P. Harker. Projections onto efficient frontiers: theoretical and computational extensions to DEA. *Journal of Productivity Analysis*, 11(3):275–300, 1999.
- H. Fukuyama and K. Sekitani. Decomposing the efficient frontier of the DEA production possibility set into a smallest number of convex polyhedrons by mixed integer programming. *European Journal of Operational Research*, 221(1):165–174, 2012.
- M. Ghadimia and M. Ahadzadeh Namin. New approach in sensitivity analysis and identification of the region of efficiency for an efficient DMU. *Mathematical Sciences*, 2009.
- A. Gholam Abri, N. Shoja, and M. Fallah Jelodar. Sensitivity and stability radius in data envelopment analysis. *International Journal of Industrial Mathematics*, 1(3):227–234, 2009.
- J. F. Hair Jr, R. E. Anderson, R. L. Tatham, and C. William. Black (1995), multivariate data analysis with readings. *New Jersey: Prentice Hall*, 1995.
- F. He, X. Xu, R. Chen, and N. Zhang. Sensitivity and stability analysis in DEA with bounded uncertainty. *Optimization Letters*, 10(4):737–752, 2016.

- F. C. Henríquez and S. V. Castrillón. A novel method for the evaluation of uncertainty in dose-volume histogram computation. *International Journal of Radiation Oncology*Biophysics*, 70(4):1263–1271, 2008.
- A. Holder. Radiotherapy treatment design and linear programming. In *Operations research and health care*, pages 741–774. Springer, 2005.
- B. Hollingsworth and P. Smith. Use of ratios in data envelopment analysis. *Applied Economics Letters*, 10(11):733–735, 2003.
- S. Holloway, M. Holloway, and S. Thomas. A method for acquiring random range uncertainty probability distributions in proton therapy. *Physics in Medicine & Biology*, 63(1):01NT02, 2017.
- Z. Huang, S. Li, and J. Rousseau. Determining rates of change in data envelopment analysis. *Journal of the Operational Research Society*, 48(6):591–593, 1997.
- R. Jacobs. Alternative methods to examine hospital efficiency: data envelopment analysis and stochastic frontier analysis. *Health Care Management Science*, 4(2):103–115, 2001.
- G. Jahanshahloo, F. Hosseinzadeh, N. Shoja, M. Sanei, and G. Tohidi. Sensitivity and stability analysis in DEA. *Applied Mathematics and Computation*, 169(2):897–904, 2005a.
- G. Jahanshahloo, F. Lotfi, N. Shoja, G. Tohidi, and S. Razavyan. A one-model approach to classification and sensitivity analysis in DEA. *Applied Mathematics and Computation*, 169(2):887–896, 2005b.
- G. Jahanshahloo, F. Lotfi, H. Rezai, and F. Balf. Finding strong defining hyperplanes of production possibility set. *European Journal of Operational Research*, 177(1):42–54, 2007.
- G. Jahanshahloo, F. H. Lotfi, N. Shoja, A. G. Abri, M. F. Jelodar, and K. J. Firouzabadi. Sensitivity analysis of inefficient units in data envelopment analysis. *Mathematical and Computer Modelling*, 53(5-6):587–596, 2011.

- K. W. Jee, D. L. McShan, and B. A. Fraass. Lexicographic ordering: Intuitive multicriteria optimization for IMRT. *Physics in Medicine & Biology*, 52(7):1845–1861, 2007.
- I. Jolliffe, B. Jones, and B. Morgan. Utilising clusters: A case-study involving the elderly. *Journal of the Royal Statistical Society: Series A (General)*, 145(2):224–236, 1982.
- I. T. Jolliffe. Discarding variables in a principal component analysis. i: Artificial data. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 21(2):160–173, 1972.
- I. T. Jolliffe. Discarding variables in a principal component analysis. ii: Real data. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 22(1):21–31, 1973.
- I. T. Jolliffe. *Principal component analysis*. Springer series in statistics. Springer, second edition. edition, 2002. ISBN 9780387954424.
- H. F. Kaiser. The application of electronic computers to factor analysis. *Educational and psychological measurement*, 20(1):141–151, 1960.
- S. Kantz, M. Söhn, A. Troeller, M. Reiner, H. Weingandt, M. Alber, C. Belka, and U. Ganswindt. Impact of MLC properties and IMRT technique in meningioma and head-and-neck treatments. *Radiation Oncology*, 10(1):184, 2015.
- C. P. Karger. Biological models in treatment planning. In *New Technologies in Radiation Oncology*, pages 221–235. Springer, 2006.
- M. Kessler. Image registration & data fusion in radiation therapy. *The British Journal of Radiology*, 79(special_issue.1):S99–S108, 2006.
- F. Khan. *The Physics of Radiation Therapy 3rd ed*, volume 4. Lippincott Williams & Wilkins, 2003.
- T. I. Kozlovska, P. F. Kolisnik, S. M. Zlepko, N. V. Titova, V. S. Pavlov, W. Wójcik, Z. Omiotek, M. Kozhambardiyeva, and A. Zhanpeisova. Physical-mathematical model of

- optical radiation interaction with biological tissues. In *Photonics Applications in Astronomy, Communications, Industry, and High Energy Physics Experiments 2017*, volume 10445, page 104453G. International Society for Optics and Photonics, 2017.
- G. F. Kuder and M. W. Richardson. The theory of the estimation of test reliability. *Psychometrika*, 2(3):151–160, 1937.
- C. Kurokawa. Accuracy requirements and uncertainties in radiotherapy. *Igaku Butsuri. Supplement*, 38(suppl. 2):7–32, 2018.
- P. Lambin, R. G. Van Stiphout, M. H. Starmans, E. Rios-Velazquez, G. Nalbantov, H. J. Aerts, E. Roelofs, W. Van Elmpt, P. C. Boutros, P. Granone, et al. Predicting outcomes in radiation oncologymultifactorial decision support systems. *Nature reviews Clinical oncology*, 10(1):27, 2013.
- Lancashire-Teaching-Hospitals. The cancer centre, 2017. URL <https://www.lancsteachinghospitals.nhs.uk/cancer/>.
- J. Lane, F. Hamdy, R. Martin, E. Turner, D. Neal, and J. Donovan. Latest results from the uk trials evaluating prostate cancer screening and treatment: the cap and protect studies. *European Journal of Cancer*, 46(17):3095–3101, 2010.
- J. Lian and L. Xing. Incorporating model parameter uncertainty into inverse treatment planning. *Medical Physics*, 31(9):2711–2720, 2004.
- K.-M. Lin and M. Ehrgott. Multiobjective navigation of external radiotherapy plans based on clinical criteria. *Journal of Multi-Criteria Decision Making*, In print., 2017.
- K.-M. Lin, J. Simpson, G. Sasso, A. Raith, and M. Ehrgott. Quality assessment for VMAT prostate radiotherapy planning based on DEA. *Physics in Medicine & Biology*, 58(16):5753, 2013.
- K.-M. Lin, M. Ehrgott, and A. Raith. Integrating column generation in a method to compute a discrete representation of the non-dominated set of multi-objective linear programmes. *4OR*, pages 331–357, 2016.

- F. Liu and C. Lai. Stability of efficiency in data envelopment analysis with local variations. *Journal of Statistics and Management Systems*, 9(2):301–317, 2006.
- L. Marks, E. Yorke, A. Jackson, T. H. L. Constine, A. Eisbruch, S. Bentzen, J. Nam, and J. Deasy. Use of normal tissue complication probability models in the clinic. *International Journal of Radiation Oncology* Biology* Physics*, 76(3):S10–S19, 2010.
- M. Mason, P. Malone, N. Clarke, D. Kirk, L. Moffat, C. Chapple, D. Neal, D. Dearnaley, N. James, C. Tyrell, et al. The royal college of radiologists’ clinical information network and british association of urological surgeons. guidelines on the management of prostate cancer. *Clinical Oncology*, 1999.
- G. P. McCabe. Principal variables. *Technometrics*, 26(2):137–144, 1984.
- S. J. McMahon and K. M. Prise. Mechanistic modelling of radiation responses. *Cancers*, 11(2):205, 2019.
- J. M. Michalski, H. Gay, A. Jackson, S. L. Tucker, and J. O. Deasy. Radiation dose–volume effects in radiation-induced rectal injury. *International Journal of Radiation Oncology Biology Physics*, 76(3):S123–S129, 2010.
- R. Mohan and D. Grosshans. Proton therapy–present and future. *Advanced drug delivery reviews*, 109:26–44, 2017.
- V. Moiseenko, M. Liu, S. Kristensen, G. Gelowitz, and E. Berthelet. Effect of bladder filling on doses to prostate and organs at risk: a treatment planning study. *Journal of Applied Clinical Medical Physics*, 8(1):55–68, 2007.
- M. Monz, T. R. Küfer, K.-H. and Bortfeld, and C. Thieke. Pareto navigation–algorithmic foundation of interactive multi-criteria IMRT planning. *Physics in Medicine & Biology*, 53(4):985, 2008.
- J. Moore, K. Evans, W. Yang, J. Herman, and T. McNutt. Automatic treatment planning implementation using a database of previously treated patients. In *Journal of Physics: Conference Series*. IOP Publishing, 2014.

- K. Moore. Data based metrics for the evaluation of autosegmentation algorithms in clinical radiotherapy workflow. *Medical Physics*, 2009.
- K. L. Moore, R. S. Brame, D. A. Low, and S. Mutic. Experience-based quality control of clinical intensity-modulated radiotherapy planning. *International Journal of Radiation Oncology* Biology* Physics*, 81(2):545–551, 2011.
- K. L. Moore, R. S. Brame, D. A. Low, and S. Mutic. Quantitative metrics for assessing plan quality. In *Seminars in radiation oncology*, volume 22, pages 62–69. Elsevier, 2012.
- D. F. Morrison. *Multivariate statistical methods*. McGraw-Hill series in probability and statistics. McGraw-Hill, 3rd ed. edition, 1990.
- L. Neralić and R. Wendell. Generalized tolerance sensitivity and DEA metric sensitivity. *Croatian Operational Research Review*, 6(1):169–180, 2015.
- L. Neralić and R. E. Wendell. Enlarging the radius of stability and stability regions in data envelopment analysis. *European Journal of Operational Research*, 2019.
- A. Niemierko. Reporting and analyzing dose distributions: a concept of equivalent uniform dose. *Medical Physics*, 24(1):103–110, 1997.
- A. Niemierko. A generalized concept of equivalent uniform dose (EUD). *Medical Phys*, 26(6):1100, 1999.
- O. Olesen and N. Petersen. Identification and use of efficient faces and facets in DEA. *Journal of Productivity Analysis*, 20(3):323–360, 2003.
- O. B. Olesen and N. C. Petersen. Facet analysis in data envelopment analysis. In *Data Envelopment Analysis*, pages 145–190. Springer, 2015.
- I. Patel, A. Glendinning, and M. Kirby. Dosimetric characteristics of the Elekta Beam Modulatortm. *Physics in Medicine & Biology*, 50(23):5479, 2005.
- P. C. R. S. G. PCRSG. Comparing the results of modern treatments, 2013.

- P. R. Peres-Neto, D. A. Jackson, and K. M. Somers. How many principal components? stopping rules for determining the number of non-trivial axes revisited. *Computational Statistics & Data Analysis*, 49(4):974–997, 2005.
- Philips. Pinnacle treatment planning, 2009. URL <https://www.philips.co.uk/healthcare/solutions/radiation-oncology/radiation-treatment-planning>.
- M. Phillips and C. Holdsworth. When is better best? a multiobjective perspective. *Medical physics*, 38(3):1635–1640, 2011.
- R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, 2013.
- R. Ramanathan. *An introduction to data envelopment analysis: a tool for performance measurement*. Sage, 2003.
- RCR. The royal college of radiologists homepage, 2018. URL <https://www.rcr.ac.uk/>.
- J. C. Roeske, J. D. Forman, C. F. Mesina, T. He, C. A. Pelizzari, E. Fontenla, S. Vijayakumar, and G. T. Chen. Evaluation of changes in the size and location of the prostate, seminal vesicles, bladder, and rectum during a course of external beam radiation therapy. *International Journal of Radiation Oncology*Biophysics*, 33(5):1321–1329, 1995.
- H. E. Romeijn, J. F. Dempsey, and J. G. Li. A unifying framework for multi-criteria fluence map optimization models. *Physics in Medicine & Biology*, 49(10):1991–2013, 2004.
- H. Ruotsalainen. *Interactive multiobjective optimization in model-based decision making with applications*. PhD thesis, University of Kuopio, Kuopio, Finland, 2009.
- M. Ruschin, A. Sahgal, S. Iradji, H. Soliman, C. Leavens, and Y. Lee. Investigation of two linear accelerator head designs for treating brain metastases with hypofractionated volumetric-modulated arc radiotherapy. *The British journal of radiology*, 89(1063):20160093, 2016.

- W. Schlegel and Mahr. 3d-conformal radiation therapy: A multimedia introduction to methods and techniques, 2002.
- L. Seiford and J. Zhu. Sensitivity analysis of DEA models for simultaneous changes in all the data. *Journal of the Operational Research Society*, 49(10):1060–1071, 1998a.
- L. Seiford and J. Zhu. Stability regions for maintaining efficiency in data envelopment analysis. *European Journal of Operational Research*, 108(1):127–139, 1998b.
- L. Seiford and J. Zhu. Infeasibility of super-efficiency data envelopment analysis models. *INFOR: Information Systems and Operational Research*, 37(2):174–187, 1999.
- L. Shao and M. Ehrgott. Discrete representation of non-dominated sets in multi-objective linear programming. *European Journal of Operational Research*, 255(3):811–836, 2016.
- E. Shaw, R. Kline, M. Gillin, L. Souhami, A. Hirschfeld, R. Dinapoli, and L. Martin. Radiation therapy oncology group: radiosurgery quality assurance guidelines. *International Journal of Radiation Oncology Biology Physics*, 27(5):1231–1239, 1993.
- D. M. Shepard, M. C. Ferris, G. H. Olivera, and T. R. Mackie. Optimizing the delivery of radiation therapy to cancer patients. *Siam Review*, pages 721–744, 1999.
- R. Siegel, A. Jemal, and S. . H. S. R. Program. Cancer facts and figures 2015. Technical report, American Cancer Society, 2015.
- L. Simar and P. W. Wilson. Sensitivity analysis of efficiency scores: How to bootstrap in nonparametric frontier models. *Management Science*, 44(1):49–61, 1998.
- W. P. Smith, M. Kim, C. Holdsworth, J. Liao, and M. H. Phillips. Personalized treatment planning with a model of radiation therapy outcomes for use in multiobjective optimization of IMRT plans for prostate cancer. *Radiation Oncology*, 11(1):38, 2016.
- A. L. Soyster. Convex programming with set-inclusive constraints and applications to inexact linear programming. *Operations research*, 21(5):1154–1157, 1973.

- E. Stubington, M. Ehrgott, G. Shentall, and O. Nohadani. Evaluating the quality of radiotherapy treatment plans for prostate cancer. In *Multiple Criteria Decision Making and Aiding*, pages 41–66. Springer, 2019.
- T. Sueyoshi and M. Goto. Slack-adjusted DEA for time series analysis: performance measurement of japanese electric power generation industry in 1984–1993. *European Journal of Operational Research*, 133(2):232–259, 2001.
- E. Thanassoulis. *Introduction to the theory and application of DEA*. Springer, 2001.
- E. Thanassoulis and R. Allen. Simulating weights restrictions in data envelopment analysis by means of unobserved DMUs. *Management Science*, 44(4):586–594, 1998.
- C. Thompson, S. Weston, V. Cosgrove, and D. Thwaites. A dosimetric characterization of a novel linear accelerator collimator. *Medical Physics*, 41(3):031713, 2014.
- R. Thompson, F. Singleton Jr, R. Thrall, and B. Smith. Comparative site evaluations for locating a high-energy physics lab in texas. *INFORMS Journal on Applied Analytics*, 16(6):35–49, 1986.
- R. Thompson, P. Dharmapala, and R. Thrall. Sensitivity analysis of efficiency measures with applications to kansas farming and illinois coal mining. In *Data Envelopment Analysis: Theory, Methodology, and Applications*, pages 393–422. Springer, 1994.
- P. Uva, L. Aurisicchio, J. Watters, A. Loboda, A. Kulkarni, J. Castle, F. Palombo, V. Viti, G. Mesiti, V. Zappulli, et al. Comparative expression pathway analysis of human and canine mammary tumors. *BMC genomics*, 10(1):135, 2009.
- D. van der Merwe, J. Van Dyk, B. Healy, E. Zubizarreta, J. Izewska, B. Mijnheer, and A. Meghzifene. Accuracy requirements and uncertainties in radiotherapy: a report of the international atomic energy agency. *Acta oncologica*, 56(1):1–6, 2017.
- M. R. Waddle, R. Landy, K. Ryan, K. S. Tzou, W. C. Stross, T. Kaleem, D. Miller, D. M. Trifiletti, S. Herchko, C. Serago, et al. Bladder filling during radiation therapy for prostate cancer treatment: Assessment via bladder ultrasound scanner., 2018.

- A. Wald. Statistical decision functions which minimize the maximum risk. *Annals of Mathematics*, pages 265–280, 1945.
- Wikipedia and Narayanese. Linac radiotherapy, 2011. URL https://en.wikipedia.org/wiki/File:Linac_radiotherapy.png.
- J. J. Wilkens, J. R. Alaly, K. Zakarian, W. L. Thorstad, and J. O. Deasy. IMRT treatment planning based on prioritizing prescription goals. *Physics in Medicine & Biology*, 52(6):1675–1692, 2007.
- C. Winsten. Discussion on mr. farrells paper. *Journal of the Royal Statistical Society*, 120(3):282–284, 1957.
- B. Wu, F. Ricchetti, G. Sanguineti, M. Kazhdan, P. Simari, M. Chuang, R. Taylor, R. Jacques, and T. McNutt. Patient geometry-driven information retrieval for IMRT treatment plan quality control. *Medical Physics*, 2009.
- Y. Yang, E. Ford, B. Wu, M. Pinkawa, B. van Triest, P. Campbell, D. Song, and T. McNutt. An overlap-volume-histogram based method for rectal dose prediction and automated treatment planning in the external beam prostate radiotherapy following hydrogel injection. *Medical Physics*, 40(1):011709, 2013.
- P. Zamani and M. Borzouei. Finding stability regions for preserving efficiency classification of variable returns to scale technology in data envelopment analysis. *Journal of Industrial Engineering International*, 12(4):499–507, 2016.
- F. Zhang, Y. Wang, W. Xu, H. Jiang, J. Gao, Q. Liu, N. Lu, D. Chen, B. Yao, J. Hou, et al. Comparison of the effects of two types of multileaf collimators on tumor control probability in radiotherapy for breast cancer after conservative surgery based on the EUD model. *Oncology*, 3(2):P77–P81, 2017.
- J. Zhu. Super-efficiency and DEA sensitivity analysis. *European Journal of Operational Research*, 129(2):443–455, 2001.