INTEGRATED DESIGN UNDER UNCERTAINTY FOR PHARMACEUTICAL PROCESSES

by

David B. Johnson

October 2002

A thesis submitted for the degree of Doctor of Philosophy of the University of London

Department of Chemical Engineering
University College London
London WC1E 7JE, United Kingdom
ABSTRACT

In pharmaceutical process development there is frequently a large element of process uncertainty since knowledge of the mechanisms of production and separation is often limited. The overall objective of this thesis is the development of a general methodology which combines process modelling with uncertainty techniques to support the process development of complete integrated sequences. In a structured approach the uncertainty can be managed and improved process performance may be obtained.

The major concept of this work is the integration of stochastic methods into a general framework for batch and continuous process models, consisting of two main parts. The first combines systematic modelling procedures with Hammersley sampling based Uncertainty Analysis and a range of sample-based Sensitivity Analysis techniques, used to quantify predicted performance uncertainty and identify key uncertainty contributions. In the second, a stochastic optimisation approach is employed to solve different problems under uncertainty. The methodology was implemented on two case studies.

The first study investigated a batch reactor process. Some undesirable performance characteristics were observed when the published nominal optimal isothermal operating policy was implemented in the uncertain system. It was found that a robust operating policy significantly improved the total process time characteristic but not the impurity content and an alternative non-isothermal policy strategy would be a better option. The second study investigated a complete process sequence. As models developed with incoming data, uncertainty in the reaction and crystallisation parameters were critical to the endpoint quality criteria. Expected performance was improved by considering the propagation of uncertainty in the complete process. Four different flowsheets were compared, considering profitability and control tolerance criteria under uncertainty.

The case study results indicate the importance in considering uncertainty systematically and quantitatively when conventional modelling techniques are employed. The methodology showed the opportunity to improve process performance potential and provide more realistic information to support pharmaceutical process development.
ACKNOWLEDGEMENTS

I would like to thank my supervisor, Professor David Bogle, for general guidance throughout the project, my family and friends for their support, the project supervisors at Pharmacia Corp. and all the Pharmacia personnel who expended their time and made significant contributions to my knowledge concerning some of the challenges in pharmaceutical process development. I would also like to thank Pharmacia Corp. for the provision of case study data.

I would also like to acknowledge the financial support from the Engineering and Physical Sciences Research Council and the Centre for Process Systems Engineering.
CONTENTS

Abstract .............................................................................................................................................. 2
Acknowledgements ............................................................................................................................ 3
List of Figures .................................................................................................................................... 10
List of Tables ...................................................................................................................................... 14
List of Problem Formulations ........................................................................................................... 17
1. Introduction and motivation ......................................................................................................... 18
2. A brief view of process development in the pharmaceutical industry ....................................... 20
   2.1 Introduction ................................................................................................................................ 20
   2.2 Some unique challenges ................................................................................................................. 20
   2.3 Modelling applications ................................................................................................................ 23
      2.3.1 Some current advances in process modelling for batch processes ....................................... 25
   2.4 A future vision ............................................................................................................................... 27
   2.5 Conclusions ................................................................................................................................ 29
3. Some aspects concerning uncertainty in models ......................................................................... 31
   3.1 Introduction ................................................................................................................................ 31
   3.2 Defining the problem under uncertainty ........................................................................................ 31
      3.2.1 Relation between data and model parameter uncertainty ...................................................... 33
   3.3 Uncertainty Analysis methods ....................................................................................................... 34
      3.3.1 Screening ............................................................................................................................... 34
      3.3.2 Quantification of uncertainty ................................................................................................. 35
      3.3.3 Construction of uncertainty space ......................................................................................... 37
         3.3.3.1 Limits for parameter uncertainties ................................................................................ 38
         3.3.3.2 Estimation of the parameter covariance matrix ............................................................ 40
         3.3.3.3 Computation of uncertainty space ................................................................................ 41
   3.3.4 Approximation methods for performance under uncertainty ................................................. 41
      3.3.4.1 Differential analysis ...................................................................................................... 42
      3.3.4.2 Response surface methodology ...................................................................................... 43
3.3.4.3 Sampling based techniques ................................................................. 44
3.3.4.4 Numerical integration techniques .................................................. 46
3.3.4.5 FAST method .................................................................................. 47
3.4 Sensitivity Analysis methods ................................................................. 48
  3.4.1 Some sensitivity measure definitions .............................................. 49
    3.4.1.1 Informal sensitivity coefficient ............................................... 49
    3.4.1.2 Differential analysis coefficients ......................................... 50
    3.4.1.3 Response surface coefficient ............................................... 50
    3.4.1.4 Correlation coefficient ......................................................... 50
    3.4.1.5 Standardised regression coefficient .................................. 51
    3.4.1.6 Partial correlation coefficient .......................................... 52
    3.4.1.7 Coefficient of determination ............................................. 52
    3.4.1.8 Stepwise regression analysis ............................................. 53
  3.4.2 Rank transformation ....................................................................... 53
  3.4.3 Measures of importance ................................................................. 53
    3.4.3.1 Correlation ratio .................................................................. 54
    3.4.3.2 Sobol' index ......................................................................... 54
    3.4.3.3 FAST index ......................................................................... 55
3.5 Some past applications ....................................................................... 55
3.6 Discussion ......................................................................................... 57
3.7 Conclusions ...................................................................................... 61
4. Some aspects of optimisation under uncertainty .................................... 63
  4.1 Introduction ...................................................................................... 63
  4.2 Response to uncertainty .................................................................... 64
  4.3 Optimal design and/or operation under uncertainty methods .......... 65
    4.3.1 Flexible plant design and operation ....................................... 66
    4.3.2 Bounded range uncertainty feasible region approaches .......... 67
    4.3.3 Probabilistic uncertainty feasible region methods ................ 74
    4.3.4 Robust partial feasibility approaches .................................... 77
Integrated design under uncertainty for pharmaceutical processes

4.3.5 Operational windows ................................................................. 82
4.4 Discussion .................................................................................. 83
4.5 Conclusions ............................................................................... 84

5. An approach for integrated design under uncertainty ......................... 85
5.1 Introduction ............................................................................... 85
5.2 Conceptual problem for uncertainty in models .............................. 85
5.3 Integrated design objectives ....................................................... 86
  5.3.1 Assumptions ........................................................................ 87
5.4 Combined Risk Analysis and model development approach .......... 88
5.5 Optimisation under uncertainty .................................................. 99
  5.5.1 Optimal input uncertainty reduction ...................................... 101
  5.5.2 Process flowsheet optimisation ............................................ 102
  5.5.3 The value of perfect information .......................................... 104
  5.5.4 Decision variable tolerance ................................................. 107
5.6 Summary ................................................................................... 109

6. Case Study I: A multiphase reactor process ................................. 110
6.1 Introduction ............................................................................... 110
6.2 Process description .................................................................... 110
6.3 Nominal optimisation .............................................................. 111
6.4 Consideration of uncertainty in the stochastic model .................... 112
  6.4.1 Stochastic process optimisation problem ............................. 113
  6.4.2 Optimisation under uncertainty results ......................... 115
6.5 An alternative operating policy ............................................... 118
6.6 Conclusions ............................................................................. 120

7. Case Study II: Risk analysis approach to a process sequence under uncertainty .... 122
7.1 Introduction ............................................................................... 122
7.2 Process Description .................................................................. 123
7.3 Modelling effort ...................................................................... 127
7.4 Risk Analysis .......................................................................... 129
Integrated design under uncertainty for pharmaceutical processes

7.4.1 Stochastic problem for the first generation model set ...........................................................130
7.4.2 First generation model sequence results ................................................................................ 131
7.5 Effect of new data ..........................................................................................................................136
7.6 Optimal uncertainty reduction ........................................................................................................141
7.7 Conclusions ....................................................................................................................................145

8. Case Study II: Base case process flowsheet optimisation under uncertainty .........................147
8.1 Introduction ...................................................................................................................................147
8.2 Base Case problem formulations .................................................................................................147
  8.2.1 Nominal flowsheet optimisation problem ..............................................................................147
  8.2.2 Stochastic flowsheet optimisation problem ...........................................................................149
  8.2.3 Stochastic operating policy tolerance optimisation problem .................................................151
8.3 Base Case flowsheet optimisation results ....................................................................................152
  8.3.1 Robustness Analysis of the input uncertainties .....................................................................155
  8.3.2 The importance of process input specification ......................................................................157
8.4 Base case operating policy tolerance optimisation results ...........................................................158
8.5 Conclusions ....................................................................................................................................159

9. Case Study II: Assessment of alternative process flowsheet options .........................................160
9.1 Introduction ...................................................................................................................................160
9.2 Process flowsheet alternatives .......................................................................................................160
  9.2.1 Alternative 1: Post-reactor back extraction ...........................................................................160
  9.2.2 Alternative 2: solL mother liquor recycle ..............................................................................161
  9.2.3 Alternative 3: No post-filtration crystal wash ........................................................................162
9.3 Comparison of optimal process flowsheets under uncertainty .......................................................162
  9.3.1 Comparison of uncertainty in output criteria predictions ......................................................164
  9.3.2 Comparison of uncertainty contributions ..............................................................................164
9.4 Comparison between flowsheet operating policy tolerances ......................................................166
9.5 Conclusions ....................................................................................................................................168

10. Conclusions and future work ......................................................................................................169
10.1 Conclusions ..................................................................................................................................169
E.1 Optimisation results for uncertainty reduction
E.2 The importance of the state of knowledge of the uncertainties results
E.3 Alternate flowsheet optimisation under uncertainty results
E.5 Alternate flowsheet operating policy tolerance optimisation results
E.6 Computation statistics
LIST OF FIGURES

Figure 2.1. Process development and key roles in the pharmaceutical industry. ........................................... 21
Figure 2.2. A systematic approach to model building (Hangos and Cameron, 2001). ........................................ 29
Figure 3.1. Comparison between orders of error for HSS and MCS sampling. .............................................. 46
Figure 5.1. Management of uncertainty in a model-based approach to integrated design under uncertainty for pharmaceutical processes. ................................................................................................................................. 87
Figure 5.2. Schematic for the systematic model development incorporating the Risk Analysis approach under uncertainty ........................................................................................................................................... 89
Figure 5.3. Schematic for the employed sampling procedure. .......................................................... 92
Figure 5.4. Algorithm for the generation of Hammersley points (Diwekar and Kalagnanam, 1997) .. 95
Figure 5.5. Schematic for the use of input contribution measures in the Sensitivity Analysis. ............... 97
Figure 5.6. Stochastic optimisation algorithm .................................................................................. 99
Figure 5.7. VPI for potentially certain properties ...................................................................................... 105
Figure 5.8. Schematic for the computation of VPI for potential certainty ............................................. 106
Figure 6.1. Reaction process, Case Study I .......................................................................................... 111
Figure 6.2. Cumulative frequency plots for the validated isothermal optimisation results obtained with and without uncertainty consideration, Case Study I .......................................................... 116
Figure 6.3. Sensitivity towards the product yield mean-variance weight, Case Study I ....................... 117
Figure 6.4. Cumulative frequency plots for the validated non-isothermal optimisation results, Case Study I .................................................................................................................................................... 120
Figure 7.1. Process flow diagram for Case Study II .................................................................................. 123
Figure 7.2. Stage 1 Bisphasic chemical reaction, Case Study II ............................................................. 125
Figure 7.3. Cumulative frequency plots for the first generation model set predictions under uncertainty, Case Study II .................................................................................................................................................... 132
Figure 7.4. Uncertain parameter contributor measures for endpoint criteria, Case Study II ............... 134
Figure 7.5. Effect of knowledge incorporation to total yield and conversion predictions, Case Study II .................................................................................................................................................... 137
Figure 7.6. Effect of knowledge incorporation to key impurity composition predictions, Case Study II .................................................................................................................................................... 138
Figure 7.7. Effect of incorporation of knowledge to the total yield uncertainty, Case Study II ................. 139
Figure 7.8. Effect of knowledge incorporation to the endpoint key impurity uncertainty, Case Study II .......... 140
Figure 7.9. Objective function value, $\Sigma \delta_p$, measuring total required levels of input parameter reduction against desired levels of output criteria uncertainty reduction, Case Study II ......................... 143
Figure 7.10. Optimal degree of key input parameter uncertainty reduction (from original uncertain values) required to meet desired levels of output criteria uncertainty reduction, Case Study II .......................................................... 144
Figure 7.11. Change in absolute SRC sensitivities between output criteria and key uncertain parameters with optimal uncertainty reductions, Case Study II .......................................................... 145
Figure 8.1. Cumulative frequency plots for validated results for optimisation with and without uncertainty consideration, Case Study II .......................................................... 154
Figure 8.2. Deviation in optimal results with input uncertainty variations, from the Base Case, Case Study II .................................................................................................................... 156
Figure 8.3. Value of feed purity information to potential profitability, Case Study II ......................... 157
Figure 9.1. Process flowsheet showing the additional back extraction in Alternative 1, Case Study II .................................................................................................................... 161
Figure 9.2. Process flowsheet showing the solL mother liquor recycle in Alternative 2, Case Study II .................................................................................................................... 162
Figure 9.3. Process flowsheet showing the final stages in Alternative 3, Case Study II ......................... 162
Figure 9.4. Deviations in optimal results under uncertainty for alternative flowsheet options, from the Base Case, Case Study II .................................................................................................................... 163
Figure 9.5. Sub-sequence contributions to predicted uncertainty, for flowsheet alternatives, Case Study II .................................................................................................................... 165
Figure 9.6. Tolerance limits for operating policy variables for alternative flowsheets, Case Study II .................................................................................................................... 167
Figure A1. Approximation of cooling capacity constraint, Case Study I ............................................. 196
Figure A2. Evolution of distribution parameter estimates with sample observations for the impurity yield, Case Study I .................................................................................................................... 196
Figure A3. Normalised scatter plots between Yield D and sub-reaction activation energy ($E_{a2}$) sample results, Case Study I .................................................................................................................... 197
Figure A4. Binary variable approximation hyperbolic smoothing curves, Case Study I ......................... 198
Figure B1. Stage I reaction, Case Study II ................................................................. 199

Figure B2. Bench scale drug profile predictions for the first generation Stage I model, Case Study II ........................................................................................................... 202

Figure B3. Reagent Addition, Case Study II ................................................................ 203

Figure B4. Layer separation, Case Study II ............................................................... 204

Figure B5. Batch distillation, Case Study II .............................................................. 205

Figure B6. Crystalliser, Case Study II ......................................................................... 208

Figure B7. Filtration unit, Case Study II ..................................................................... 212

Figure B8. Washing unit, Case Study II ..................................................................... 213

Figure B9. Drying unit, Case Study II ........................................................................ 214

Figure B10. Effect of additional uncertainty in crystallisation 'solute loss' parameter for the key impurity due to violation of desired solL solvent volume to product mass operating range, Case Study II .............................................................................................................................. 218

Figure B11. Evolution of distribution parameter estimates with sample observations for the total yield of the first generation of models (knowledge level 0), Case Study II .............................................................................................................................................. 220

Figure B12. Scatter plots showing the effect of induced rank correlation ($\rho = -0.98$) in the Stage I rate constant parameters for the first generation of models (knowledge level 0), Case Study II .............................................................................................................................................. 220

Figure B13. Scatter plots of the key uncertain parameters with endpoint criteria for the first generation of models (knowledge level 0), Case Study II .............................................................................................................................................. 221

Figure C1. Pilot scale (1000 US gallon, 90 rpm) drug profile, Case Study II ............... 224

Figure D1. Cumulative frequency plots of total yield prediction under uncertainty, Case Study II .............................................................................................................................................. 231

Figure D2. Cumulative frequency plots of key impurity content prediction under uncertainty, Case Study II .............................................................................................................................................. 232

Figure D3. Scatter plots of the key parameters with the endpoint key impurity content, Case Study II .............................................................................................................................................. 232

Figure D4. Cumulative frequency plots of inter-stage key impurity composition under uncertainty, Case Study II .............................................................................................................................................. 233

Figure E1. Cumulative frequency plots for validated optimisation results under input uncertainty variations, Case Study II .............................................................................................................................................. 236

Figure E2. Cumulative frequency plots for potential profitability, Case Study II .......... 237
Figure E3. Cumulative frequency plots for validated optimisation results under input uncertainty variations, Case Study II. .......................................................... 238
LIST OF TABLES

Table 3.1. Summary of propagated uncertainty quantification methods for Uncertainty Analysis. .................. 42
Table 3.2. Summary of Sensitivity Analysis methods and measures. ............................................................. 49
Table 4.1. Summary of work concerning chemical process design and operation defining a feasible region under bounded parameter (deterministic) uncertainty. ...................................................... 68
Table 4.2. Summary of work concerning chemical process design and operation defining the feasible region under probabilistic parameter (stochastic) uncertainty. .............................................. 74
Table 4.3. Summary of work concerning robust (partial feasibility) chemical process design and operation. ........................................................................................................................................ 78
Table 6.1. Comparison between the optimal literature results (determined through repeated simulation) and optimal results obtained with SQP optimisation, Case Study I. ................................. 112
Table 6.2. Key SRRC contributors to predicted uncertainty under the nominal optimum operating policy, Case Study I. .............................................................................................................. 113
Table 6.3. Validated process optimisation results under uncertainty, Case Study I. ....................................... 115
Table 6.4. Nominal optimum operating policy under non-isothermal temperature control, Case Study I. .......................................................................................................................................................... 119
Table 6.5. Uncertainty Analysis results comparing and robust isothermal and non-isothermal temperature control, Case Study I. ........................................................................................................ 119
Table 7.1. Summary of key process operating conditions for Case Study II .................................................. 124
Table 7.2. Possible characterising phenomena for the process operations of Case Study II. .......................... 128
Table 7.3. PPR plant run conditions, Case Study II .......................................................................................... 129
Table 7.4. PPR plant run measurements, Case Study II .................................................................................... 129
Table 7.5. Summary of first generation model characteristics and stochastic problem inputs, Case Study II .................................................................................................................................................... 130
Table 7.6. Summary of first generation model results under uncertainty, Case Study II ................................. 131
Table 7.7. First generation model sub-sequence contributions to predicted uncertainty, Case Study II ................................................................. 133
Table 7.8. Effect of more inter-stage measurements to sub-sequence contributions, Case Study II ......... 135
Table 7.9. List of key parameters regarding first generation model uncertainty in ascending order of priority, Case Study II .............................................................................................................. 136
Table 7.10. Levels of incorporated knowledge with incoming data, Case Study II ........................................137

Table 7.11. List of key parameters and knowledge priorities in the final generation of models (knowledge level 6), in ascending order of priority, Case Study II ................................................................. 141

Table 8.1. Validated Base Case process flowsheet optimisation results under uncertainty, Case Study II ................................................................................................................................................. 153

Table 8.2. Optimal decisions for the Base Case flowsheet, Case Study II ................................................. 153

Table 8.3. SRC ranking of the key input uncertainty contributors to the Base Case flowsheet criteria, Case Study II ................................................................................................................ 155

Table 8.4. Sub-sequence contributions to the uncertainty in the Base Case flowsheet criteria, Case Study II ............................................................................................................................. 155

Table 8.5. Results for tolerance optimisation of the Base Case flowsheet operating policy, Case Study II ........................................................................................................................................... 158

Table 8.6. Tolerance limits for Base Case flowsheet optimal operating policy, Case Study II ................. 158

Table 9.1. Key alternative flowsheet SRC contributors to predicted uncertainty, Case Study II ...... 165

Table 9.2. Alternative flowsheet operating policy tolerance optimisation results, Case Study II ..... 166

Table A1. Uncertainty characterisation in the parameters of Case Study I ............................................. 197

Table A2. Computational statistics for optimisation problems, Case Study I ....................................... 198

Table B1. Parameters for the bench scale Stage I model, Case Study II ................................................. 202

Table B2. Uncertainty characterisation in the parameters of the first generation models, Case Study II ............................................................................................................................................. 216

Table C1. Revised uncertainty characterisation in the parameters of the first generation Stage I model, fitted to new data, Case Study II .......................................................................................................... 222

Table C2. Parameters for the second generation Stage I model, Case Study II ...................................... 224

Table C3. Uncertainty characterisation in the second generation Stage I model parameters, Case Study II ............................................................................................................................................. 224

Table C4. Parameters for the third generation Stage I model, Case Study II ........................................ 226

Table C5. Uncertainty in the parameters of the third generation Stage I model, Case Study II .......... 226

Table C6. Parameters for the second generation layer separation model, Case Study II ...................... 228

Table C7. Uncertainty characterisation in the second generation layer separation models, Case Study II ............................................................................................................................................. 228

Table C8. Uncertainty characterisation in the second generation reagent addition models,
Case Study II........................................................................................................................................ 229

Table D1. Main SRC Sensitivity Analysis results for levels of knowledge incorporation, Case Study II........................................................................................................................................ 234

Table E1. Optimisation results for optimal uncertainty reductions for knowledge level 6 Base Case flowsheet stochastic model, Case Study II........................................................................................................................................ 235

Table E2. Validated alternate process flowsheet optimisation under uncertainty results, Case Study II........................................................................................................................................ 237

Table E3. Optimal decisions for alternate process flowsheet optimisation under uncertainty, Case Study II........................................................................................................................................ 237

Table E4. Key SRC contributors to predicted uncertainty for flowsheet alternatives, Case Study II........................................................................................................................................ 239

Table E5. Optimal decisions for the Base Case operating policy tolerance, Case Study II. ..........239

Table E6. Optimal decisions and tolerance limits for Alternative 1 operating policy tolerance, Case Study II........................................................................................................................................ 240

Table E7. Optimal decisions and tolerance limits for Alternative 2 operating policy tolerance, Case Study II........................................................................................................................................ 240

Table E8. Optimal decisions and tolerance limits for Alternative 3 operating policy tolerance, Case Study II........................................................................................................................................ 240

Table E9. Computational statistics for optimisation problems, Case Study II. .........................241
LIST OF PROBLEM FORMULATIONS

Problem P1. General formulation for input uncertainty space reduction optimisation. ....................... 102
Problem P2. General formulation for stochastic process flowsheet optimisation. ............................... 104
Problem P3. General formulation for operating policy error tolerance optimisation. ......................... 109
Problem P4. Nominal isothermal process optimisation formulation for Case Study I. ....................... 112
Problem P5. Stochastic isothermal process optimisation formulation for Case Study I....................... 114
Problem P6. Nominal non-isothermal process optimisation formulation for Case Study I.................... 118
Problem P7. Stochastic non-isothermal process optimisation formulation for Case Study I. ............... 118
Problem P8. Input uncertainty reduction optimisation formulation for the Base Case, Case Study II. ............................................................................................................................................... 142
Problem P9. Nominal process flowsheet optimisation formulation for the Base Case, Case Study II. ............................................................................................................................................... 149
Problem P10. Stochastic process flowsheet optimisation formulation for the Base Case, Case Study II... ............................................................................................................................................... 151
Problem P11. Operating policy error tolerance optimisation formulation for the Base Case, Case Study II. ............................................................................................................................................... 152
CHAPTER 1

INTRODUCTION AND MOTIVATION

This thesis concerns the management of uncertainty in a model-based approach for the design of pharmaceutical processes. Process development may be aided by the application of a process systems approach based on reliable (accurate) process models. In this work the uncertainty contained in the available process models is considered with respect to the development of more reliable models and the optimisation of complete process sequences.

Process development in the pharmaceutical industry is beset by challenges not encountered in the traditional chemical processing industries. These pressures mean that process design decisions often need to be made despite a scarcity of available process knowledge and with little understanding of the physico-chemical phenomena characterising the process. This can result in the development of low quality processes. A more structured approach to process development, particularly at the earlier stages, has the potential to save development resources, improve the quality of the final process and help ensure added value in the products. In addition, the efficiency of the process development may be enhanced. Therefore, it would be useful to structure the available process information and knowledge within a decision support tool, and where possible transform the knowledge into mechanistic understanding.

A framework based on process models is one way to provide such a structure. In this way, the available information and process knowledge can be systematically documented and retained. This in turn may help improve the understanding of the process and provide opportunities to exploit the potential of a process with the use of computational optimisation techniques. To better exploit the full potential that can be achieved with the use of model-based approaches, they should be incorporated sooner rather than later into process development procedures. Simple models may provide useful information at earlier stages and as data is obtained greater complexity may be incorporated to improve their reliability. However, the usefulness of a model-based approach in supporting and influencing decisions is limited by the confidence which can be placed in the results obtained. Most optimisation techniques do not explicitly account for the presence of model uncertainty and while methods incorporating uncertainty aspects are available, few appear to have been applied to non-linear batch systems let alone integrated sequences of batch process operations typical in the pharmaceutical industry. Since the lack of process knowledge often results in large amounts of uncertainty being contained in the models developed to describe the process, it is important the framework has the capacity to quantify the uncertainty from a wide range of sources in a meaningful manner.

The objective of this work is to exploit the use of a model-based approach in order to provide a general framework within which it is able to quantify the combined influence of a range of uncertainties within an integrated process sequence model, identify the key sources of uncertainty, track the effect on the levels of uncertainty contained when new information is incorporated, determine the best process flowsheet
operating conditions under the uncertainty and which can be used to identify the best flowsheet between a set of topological alternatives according to the desired process criteria. It is believed that such a framework would provide a new and valuable contribution within the scope of process design for integrated process sequences in the pharmaceutical industry.

In Chapter 2, a brief summary of the unique issues associated with process development in the pharmaceutical industry is presented. The current role, limitations and attitudes towards process modelling in the process development cycle is discussed and a future vision for this role is portrayed. Some of the mathematical aspects of uncertainty regarding its characterisation, quantification and sensitivity analysis methods associated with a model-based environment are discussed in Chapter 3. Chapter 4 reviews the main approaches which have been proposed for chemical process design optimisation under uncertainty, the term design may assume both the operating conditions and equipment design parameters. Chapter 5 presents the proposed methodology. The mathematical techniques used to generate the characterisation of the uncertain process system within a systematic model building procedure are described in detail. General formulations describing different optimisation under uncertainty problems are stated. In Chapter 6 an industrial reaction process taken from the literature is investigated. This case study shows the importance and benefits of incorporating uncertainty when using deterministic process models to support the assessment of process performance and in determining optimal operating policies. Chapters 7, 8 and 9 illustrate and verify the features of the methodology with regard to a comprehensive case study investigating a complete process sequence for a pharmaceutical product. This case study is derived from an industrial collaboration with a pharmaceutical company. Chapter 10 presents the conclusions of this thesis and summarises some important avenues of future work within the scope of this project.
CHAPTER 2

A BRIEF VIEW OF PROCESS DEVELOPMENT IN THE PHARMACEUTICAL INDUSTRY

2.1 Introduction

In this chapter a brief background to process development in the pharmaceutical industry is presented. This includes some of the unique aspects which characterise the current working cultures and attitudes. The current advance and utilisation of process modelling to support development is discussed together with a projected vision for process development in the future.

2.2 Some unique challenges

Research based companies within the pharmaceutical industry deliver new therapeutic products to health care providers. To achieve this, vast quantities of time, effort and capital are invested into research and development. The high risks involved are offset by a wide portfolio of projects and the potential pay-back for introducing a novel product to the market first.

With the increasing emphasis for new products, the level of commercial competition is high. A significant proportion of the competitive success of a research based company may stem from the efficient manner in which it can manipulate the available data and execute process development. Basu (1998) defined process development in the pharmaceutical industry as "the process of converting a chemical synthesis into an optimum, economic, robust and reproducible process for the manufacture of a chemical of desired quality at the ultimate desired scale". The roles of process development and scale-up are inter-dependent, where the same author also defined scale-up as "designing for safe and cost-effective operation of a process giving predictable results, by using knowledge and data available at a particular point in time within equipment which may be of similar scale as the manufacturing plant". Ideally a high quality process would be developed with only a minimum number of scale-ups but this is limited by the science and current technology of scale-up (Basu et al., 1999). In addition, regulatory authorities stipulate that once an application for a process for a particular product has been made, opportunities to change the process become severely limited. This means that it is essential to arrive at a high quality process before the application is made or risk the disadvantages of a poor quality process in manufacture and even the need to reapply in case of problems requiring major changes.

Stephanopoulos et al. (1999) interpreted batch process development as a series of phases before technology transfer to the manufacturing plant during which it is speculated that the 'added value' of the process or 'lost opportunity' decreases by several orders of magnitude. These phases involve a largely iterative process between the chemists, the development engineers, the pilot plant engineers and ultimately...
the production engineers. A basic view of process development based on that given by Basu (1998) is shown in Figure 2.1. Here the key roles of the different personnel are loosely defined, though some of these may overlap with the interactions shown.

Good communication between the chemists, the development engineers and the pilot plant engineers is crucial so that problems during scale-up are resolved at earlier stages in the development. This view is endorsed by Mukesh (1999) and Carpenter (2001) who both stated that better processes will only result from a very close interaction between the synthetic chemists and the chemical engineers. It has also been argued that a little more time spent in the early in the development cycle would result in shorter and more productive laboratory and plant runs (Repic, 1998). In this way, the culture of ‘learning by doing’, though still essential, may be minimised so that the number of scale-up failures and the runs required to obtained reproducible scale-up data are reduced. In addition, there would be less reliance on the expertise of individual chemists in determining how much the laboratory process characterisation passed to the pilot plant engineers takes into account the limiting effects of scale-up in the pilot plant equipment.
Process development in the pharmaceutical industry faces a number of unique challenges due to the nature of the products and the market pressures. Some of the key challenges and implications are summarised below:

- limitations on development time and resource due to economic pressure to be first to the market,

  ⇒ requirement for simultaneous process development, scale-up and chemical production for testing (often at short notice) with incomplete process knowledge,

  ⇒ rapid transfer of laboratory processes into the pilot plant so that scale-up problems are identified and resolved quickly (Shah et al., 1999),

  ⇒ early 'freezing' of process options with full scale operations typically based on experimental laboratory processes developed into existing manufacturing equipment,

  ⇒ desire to achieve as much process validation as possible in the pilot plant rather than in the manufacturing plant,

  ⇒ high reliance on experimentation and empiricism to obtain process feasibility with little time for model development and very little or none for process optimisation,

  ⇒ only just enough process characterisation is passed to the pilot plant engineers to operate the process safely and achieve the desired product quality,

  ⇒ when engineers require more data from the chemists or desire further plant measurements to be collected the added value must be apparent,

- shortening product life cycles in a continually evolving market,

  ⇒ a wide portfolio of products and reliance on existing multipurpose equipment in both the pilot and manufacturing plants,

- need for high quality chemical product for efficacy, safety, toxicological tests and clinical trials

  ⇒ additional pressure on the pilot plant and requirement of good documentation,

  ⇒ batch to batch conformity (scale-up reproducibility) is essential,

- high risk of project failure due to adverse tests or trials,

  ⇒ less capital expenditure allocated to pilot plant stages of process development than resolving problems in the manufacturing plant (when project success is more certain),

- complex organic syntheses sometimes with very high cost raw materials,

  ⇒ high number of chemists to chemical engineers with the requirement for effective communication,

  ⇒ high complexity of decision making,

  ⇒ increasing importance of process efficiency in the competitive market such that the need for cost reduction may drive development (Basu, 1998),
restrictions on implementation of process changes after certain regulatory approvals,

⇒ efficient process development and good planning are required to anticipate and validate
performance enhancing process changes earlier,

• adherence to regulatory authority guidelines,

⇒ written evidence for a good understanding of the process (development report),

⇒ iterative process between laboratory and plant to verify reproducibility of scale-up and ensure
success of validation runs,

⇒ documentation of all process changes made during development (process change control system),

• adherence to safety and environmental constraints,

⇒ limitations on the implementation of alternative chemistry, process routes and technology.

Clearly the challenges and implications contribute immensely to the way in which process development is
conducted in the pharmaceutical industry. In this work the benefits that may be derived from the use of a
more system wide process modelling approach is of interest. The next section reviews some of the current
literature reporting the current opinions and contributions of modelling in this field.

2.3 Modelling applications

Applications of process modelling techniques in the batch chemical processing industries have been
previously identified (Wright, 1984, Barrera and Evans, 1984, Rippin, 1993, Terwiesch et al., 1994). A
broad range of potential benefits which may be obtained through the implementation of model-based
systems approaches to support process development in the batch processing and pharmaceutical industries
have been acknowledged from sources within the chemical industry, the engineering software industry and
1999, Stephanopoulos et al., 1999):

• production of higher quality processes,

• increased efficiency of process development by reducing the number of experiments and scale-ups,

• improved scientific understanding of the process in terms of the physico-chemical phenomena and the
interactions with external actions and identification of significant areas where process understanding is
lacking.

• identification of important interactions and tracking of species concentrations within an integrated
process sequence (otherwise expensive and difficult to measure on-line),

• systematic documentation and structure of process knowledge in the provision for a corporate
knowledge warehouse.
• identification of the information necessary for each stakeholder to complete their jobs more effectively,

• provision of more reliable process characterisation for pilot plant and production engineers,

• risk free exploration of a large number of process alternatives and operating and limiting conditions within the complete process sequence,

• identification of critical parameters not apparent from the laboratory or pilot scale work,

• aid trouble shooting in existing operations and study safety problems,

• support control system design.

It has been proposed that these potential benefits may be directed towards supporting decisions in:

• planning of experiments and data collection,

• process feasibility and safety investigations, robust scale-up and technology transfer,

• sensitivity analysis and determining the limits of the process to support process validation,

• optimisation of process operation considering and trading between economic, environmental and/or risk factors.

It is hoped that the introduction of systematic model-based techniques will increase the probability of obtaining good quality processes before application to the regulatory authorities beyond which opportunities to instil process changes are limited. Consequentially, this should also reduce the need to introduce radical process changes due to unpredicted poor performance and problems at later stages, resulting in considerable loss of time and resource during reapplication to the authorities. Some industrial attitudes and potential limitations regarding the use of process modelling are also acknowledged:

• models are only developed in the event of development problems (trouble shooting) or they are considered to add value to the process,

• most knowledge is empirical or based on experience where there is little understanding of the underlying mechanisms and for which modelling from first principles is not generally applicable (e.g. solids processing -crystallisation, filtration, drying),

• implementation must not have an adverse impact on the efficiency of current work methods and there must be some positive quantifiable return for any additional effort,

• the limitation of modelling is that it will only solve a fraction of the problems typically observed due to the unique nature of problems and because some processes are not typical engineering applications and are not conducive to current modelling technologies,

• short term objectives of reaching deadlines may conflict with and override long term objectives for improved process development.
Integrated design under uncertainty for pharmaceutical processes

- models need to be relevant and accurate over realistic ranges of values for the parameters used for scale-up which is more dependent on the experience and expertise of the individual than in the ability of the tools used to solve the mathematical problems,
- models may be unreliable and the results inaccurate due to the number of unmeasurable external influences and the limitations of models applied for scale-up must be realised by the scientists,
- the inability to accurately model certain process aspects, such as physical properties and in particular processes involving solids are major problems,
- on the whole the level of process modelling is low with the exception that the chemistry and reaction stoichiometry is generally well defined though rigorous kinetic reaction models are usually unjustified (for short-lived small-volume products) with operation often based on conversion,

and views towards process optimisation:

- it is largely left to manufacturing due to time constraints, the difficulty in getting a feasible process to begin with and the risk associated with each project,
- it is only considered when there is a basis for a workable and scaleable process and is rarely afforded,
- model-based numerical optimisation solutions may not be suitably sensitive to real life safety issues.

Some interesting and valid issues covering a range of viewpoints have been raised here. While the vision is worthwhile, the implementation of process modelling techniques appears to be limited by the current technologies available for modelling and/or current working practices (and current skills base). Some of the current advances in this field are discussed in the next section. The accuracy of model predictions, in particular for physical properties, may be viewed as questionable but the industrial perception of the current status of modelling technologies may be lower than it should be.

2.3.1 Some current advances in process modelling for batch processes

Some new ideas and approaches proposed in the literature concerning model-based support for systematic process development are discussed. This is followed by some published examples where the use of process modelling does appear to have made a difference.

Barrera and Evans (1984) introduced a comprehensive approach to the optimal design and operation of multi-product batch process plants using existing equipment. They consider a hierarchical solution approach in recognition of the computational intensity of the optimisation problem for the complete batch process design. The upper level selects the equipment sizes and storage policy from a given inventory, where combinatorial optimisation strategies are proposed. The middle level distributes the horizon time between the products and the lower level selects the operating times and conditions for the individual product processes posed as an NLP problem. An example for the lower level sub-problem demonstrated
the advantages of considering the integrated process sequence through the trade-offs determined with the shifting of process load between operations and with the productivity and operating costs.

The concepts of a comprehensive modelling approach for entire pharmaceutical process sequences are explored by Basu et al. (1999). They concluded that changes to current process development practices are essential and advocate that integration with process modelling is one new approach to bring this change about. Allgor et al. (1996) demonstrated the application of systems process modelling for integrated batch process development and optimisation. In their methodology rigorous mechanistic process models are constructed, validated and used in conjunction with laboratory experimentation and pilot plant data. A first step determines a feasible base case process design (seen as the most challenging aspect) based on these models given a plant superstructure derived from a state task network (STN) representation. This describes the tasks involved in transforming raw materials from one thermodynamic state into the desired products and wastes in other states. Secondly, systematic optimisation of the base case follows according to some economic criteria. Some of the problems recognised by Samsatli and Shah (1996) in their work regarding multi-stage biochemical processes are also applicable to synthetic pharmaceutical processes. They used integrated design procedures which aim to capture the key interacting effects between unit operation models within the optimisation.

The use of process modelling and simulation as a tool to support pharmaceutical process validation is promoted by Wright and Bramfitt (1997). They endorsed the use of dynamic mass and energy balances to provide a detailed description of the intrinsic chemistry, kinetics and thermodynamics with a complete operational specification in the equipment model. Their article refers to illustrating examples for two industrial reaction processes (Wright, 1984 and Bollyn et al., 1995), discussed in the next section.

Shah et al. (1999) questioned the current working practices and also advocate the advanced role of process systems engineering in process design and development, within the specifications of the Britest project (Batch Route Innovative Technology, Evaluation and Selection Techniques). The aim is to improve process efficiency by defining the process in terms of its physicochemical properties and understanding how external conditions affect these in order to determine the best conditions outside the limitations of existing equipment or unit operations. Here model building and model validation purposes drive experimentation and the models are used to determine the process design.

Stephanopoulos et al. (1999) discussed some of the methodologies proposed to support the different phases of batch process development and the need for further progress. The emphasis is on the conceptual design of batch processes using an operating plan approach but also encompasses production route planning and materials selection among others, within the confines of the ‘BatchDesign-Kit’ software.

A new batch process modelling approach is introduced by Martinez (2000) where reactor models are developed driving and making effective use of incoming experimental data. With an uncertain model a sequential experiment design strategy is proposed to locate the optimum operating conditions in a ‘learning by doing’ manner. The modelling and process optimisation are integrated within an experimental data collection loop.
There is a limited number of published examples (possibly due to industrial confidentiality) where process modelling has proved beneficial in real process development. It should be noted that virtually all of these concern reaction processes.

Wright (1984) provide an example where process modelling was applied in conjunction with an industrial pilot process for an equilibrium controlled aqueous phase reaction process (the Williamson Synthesis) operated in semi-batch mode. The aim was to develop a combined kinetic, reactor and controller model (based on very limited kinetic data) in order to determine an operating policy which maintained the primary reaction while minimising the secondary reaction. A reasonable agreement between the model and the limited experimental data was achieved and simulation experiments were able to indicate the possibility for improved yield and reduced reagent loads. However, this article exemplified one of the inherent risks in expending modelling effort during the process development stage in the pharmaceutical industry, in that the process was abandoned by the company.

The multi-step batch reaction and distillation sequence example given by Allgor et al. (1996) provides a useful view of the type of information available in the development of an industrial process for a speciality chemical. It is shown that the use of an integrated process modelling and development methodology allowed the modification of a laboratory operating policy to a feasible one in the plant scale equipment. This required the use of physical property estimation methods and rigorous process models which captured the limitations imposed by the equipment. They justify the need for a systems approach through the determination of optimal inter-unit trade-offs between the key reaction process and downstream separations.

Concerning the development of reaction kinetics, there seems to be some significant advances. Once the reaction kinetics have been ascertained they can be incorporated into a process model and be used to accelerate scale-up and determine the best operating policy. Bollyn et al. (1995) provide one industrial example where reaction calorimetry data was used to determine between four proposed reaction mechanisms and kinetic schemes for a multi-step catalysed oxidation reaction. They developed a process model and validated the best operating policies determined from model simulations with pilot plant data, identified parameter sensitivities which were not obvious from the laboratory experiment and modified and validated the model with respect to the manufacturing equipment. It was claimed that this combined process modelling approach allowed fast scale-up for a safe and economic process with only a few pilot plant runs. Sano et al. (1998) provide a similar account in the development of a reaction process model for a pharmaceutical intermediate. They used this model to determine the optimum operating policy accounting for the limited cooling capacity of the plant vessel.

2.4 A future vision

It is widely agreed that a new approach to process development is required in order to reduce development times and resources, improve the quality of processes and would ultimately determine the success of competing companies in the future (Basu et al., 1999, Shah et al., 1999, Stephanopoulos et al., 1999).
Invariably such an approach would focus on a long term vision for cost-effective process development involving improved interactions between chemists and chemical engineers and/or the use of systematic methods to improve the fundamental understanding of the process for more efficient scale-up and process optimisation.

The perception of a model-based approach in one such vision from within the pharmaceutical industry is summarised. Basu et al. (1999) proposed the use of a process vision in which all the requirements of a good quality process are met and one that is applicable to the efforts of all the stakeholders. They envisage that the integration of computational techniques (process synthesis and process modelling) with the necessary experimental methods would improve the both the efficiency of the process development and the quality of the processes in a more systematic approach. The development of models describing the entire process would be conducted after the embodiment of the process vision and identification of the optimum process from the synthesis. The process modelling application steps involve:

- determining the experimental data required to substantiate the model,
- running the experiments to collect the data, fit the model parameters to the data,
- improving the model,
- validating the model by laboratory and kilogram scale experiments and pilot plant runs,
- coupling the model with pilot and manufacturing plant equipment specifications and operating policies,
- validation and sensitivity analysis of the equipment model to support technology transfer into the manufacturing plant.

These cycles are compatible with the systematic model building procedure presented by Hangos and Cameron (2001) shown in Figure 2.2, where the data required to improve and validate the model can be collected from the necessary scale-up runs and additional experiments. In addition, the data collected from failed runs should be used to rigorously validate the model. The structural characteristics, accuracy and reliability of the models depends on the quality of the data and assumptions upon which they are based.

The development of a tool based on a reliable model of the entire system based on the physico-chemical phenomena and equipment specifications aims to support the answers to some of the process conundrums posed by the complex interactions between upstream and downstream operations, make scale-up more robust and assist in technology transfer and process validation in both the pilot and manufacturing plants.

This is a new approach which would require a significant shift in current attitudes and efforts towards process development. Yet the potential benefits are recognised to be very worthwhile. This view (and that of others) is essentially a long term one, waiting for the development of the tools suitable to the needs of the stakeholders and one which may require the reallocation or introduction of a skills base in model development techniques if such expertise is not already present in the current workforce.
2.5 Conclusions

In this chapter the background and main challenges and implications associated with process development in the pharmaceutical industry have been discussed. It is clear that there are many unique pressures which shape the current culture of process development. A brief review of some of the current applications of modelling approaches observed in development was made. From the literature it appears that actual implementation of process modelling approaches are mainly limited to reaction and distillation processes. This may be due to the poor perception that industry has of the current status of modelling and the quality of predictions, for example in models predicting physical properties and solid phase processes.

The advance of reaction calorimetry techniques now permits the construction of detailed mechanistically derived dynamic equations to describe reaction kinetics. With this understanding it should be possible to use process models to accelerate the scale-up of reactions, reduce scale-up failures and optimise the process with reduced number of pilot plant runs.

An ambitious future vision for the integration of model-based techniques with experimental effort is cited in which modelling of the entire process is envisaged. This vision is dependent on the available modelling technologies, software tools and current attitudes. The aim of the work presented in this thesis is to substantiate the use of a model-based approach which may be appropriate within such a vision with the
explicit consideration and management of model uncertainty, so far not encountered for complete pharmaceutical process sequences.

In the next chapters the expression and analysis of model uncertainty are investigated and a review of the main techniques available for optimisation under uncertainty follow.
CHAPTER 3

SOME ASPECTS CONCERNING UNCERTAINTY IN MODELS

3.1 Introduction

In the previous chapter some of the current issues in process development in the pharmaceutical industries were discussed. Views towards the potential benefits that process modelling could bring to this field indicate their worth but there remain some limiting factors to practical implementation. However, some industrial examples of combined modelling and process development success have been highlighted, though these appear only to concern reaction processes and not complete process sequences. Some recent approaches promoting the use of model-based support in batch process development were identified and a future vision for pharmaceutical process development was made in which process modelling provides a major contribution.

In any decision making process where the data or information available upon which decisions are based may be uncertain and where the entailing course of action involves a risk of significant consequence then quantitative risk assessment studies provide a valuable support tool (Haimes et al., 1994). In this chapter uncertainty in process model systems is considered. With the consensus that systems process modelling can significantly benefit the process development and since the models represent the current knowledge of the process in mathematical form, it becomes clear that assessment of the effects that this uncertainty in the process knowledge may have on the performance should not be overlooked.

The problem under uncertainty is defined in Section 3.2. Possible sources and characteristics of uncertainty are discussed in Section 3.3. Quantitative methods to deal with uncertainty in a stochastic process model system framework are examined. These techniques form the basis for Uncertainty and Sensitivity Analyses in a Risk Analysis type approach (Section 3.3 and Section 3.4). The entailing information provides a platform for the management of uncertainty (or risk management). Some past applications of these methods are discussed, in particular those relating to chemical engineering applications (Section 3.5). The capacity of different techniques in satisfying the identified important properties are compared and the approach and techniques deemed most applicable to the types of problems concerning this thesis are discussed in Section 3.6.

3.2 Defining the problem under uncertainty

The framework for the stochastic system considered in this work consists of a process which in this work is represented by a deterministic model of the chemical process with uncertain parameters. Three assumptions define the Uncertainty Analysis problem considered in this work:

- the deterministic model is not structurally incorrect,
• process knowledge uncertainty represented by uncertainty in the parameters of the deterministic model is more important than natural system variability,
• the uncertainties are random as opposed to systematic.

Models are only imperfect mathematical representations of observed reality and model parameter uncertainties arise from a lack of firm theoretical basis, simplifications and assumptions and/or poor quality modelling data. Uncertainty associated with errors in the model structures are not explicitly considered in this work. To a limited extent uncertainties in the model parameters may be assumed to account for structural errors (Pinto, 1998). In this work the analyses assume parametric errors.

In the context of chemical engineering process systems, Pistikopoulos (1995) classified parametric uncertainties in process engineering into four types:
• model-inherent, including kinetic constants, physical properties, transfer coefficients,
• process-inherent, including flowrates, temperatures, stream qualities,
• external, including feed stream availability, product demand, prices, environmental conditions,
• discrete, equipment availability and failures.

This thesis is concerned with parameter uncertainties associated with the first two types where the operation of the process to obtain a desired product is of interest. Within the confines of the process model, the model inherent uncertainties range from parameters describing physical properties, intrinsic phenomena, extrinsic phenomena and to those factors assumed to describe the possible effects of phenomena which are not explicitly described in the models due to a complete lack of data or understanding (i.e. in a black box approach). These provide the uncertain inputs to the stochastic system considered in this work. The latter two types are associated with the planning and scheduling of multi-product or multi-purpose plants and equipment reliability aspects.

A distinction should be made between uncertainty and variability. The former is due to a lack of fundamental knowledge in the process phenomena or property and can be reduced by increasing the knowledge. This is commonly termed as subjective uncertainty and is a property that is determined by the understanding of the analysts and the level of modelling permitted. The latter is an inherent property of the system and cannot usually be reduced. This is commonly termed as stochastic uncertainty. It is not generally possible to totally separate them. With the emphasis on the management of incomplete process knowledge contained in process models, model parameter uncertainty is the relevant property to which efforts can be made to reduce the uncertainty in the system (by improving the knowledge). In a rigorous stochastic description of the process system both parameter uncertainties and system variabilities (if present and can be realistically characterised) should be considered simultaneously to determine the combined influences and the relative worth of management actions aimed towards reducing the parameter uncertainties.
Computational and time resource may exert limits on the accuracy which can be obtained with some Uncertainty and Sensitivity Analysis methods. In this work the emphasis is not on extremely large computationally intensive models (as may be observed in nuclear, geophysical, environmental or business model applications which may take hours to solve a single realisation) but rather on integrated sequences of deterministic models in a process systems engineering framework for which a single simulation is expected to take the order of seconds to solve. The individual models need to capture the key process behaviour without extreme complexity. This permits a little more relaxation and flexibility in the analysis techniques available, compared to the experimental designs formulated for physical experiments and intense computational models.

3.2.1 Relation between data and model parameter uncertainty

The quality and quantity of the data available to build process models obviously has a large influence on the uncertainty associated with the parameters of the resulting models. Experimental and plant data may contain systematic errors, erroneous data and outliers. While this work does not specifically consider systematic errors, the effect of limited or erroneous data or data outliers contained in the experimental or plant measurements is discussed.

The presence of erroneous data and outliers are associated with the model parameter estimation problem within the model building process (see Figure 2.2, Section 2.4). Their presence in model parameter regressions are passed as uncertainty estimates in the optimal parameter values. They can provide a significant source in the uncertainty estimate, quantified according to the methods employed to estimate parameter confidence intervals and regions (discussed in Section 3.3.3.1). Weighted regression methods can allow for identifiable data outliers. If only limited data is available and the regression is based on only a few data points then it becomes harder to confidently identify and accordingly weight outliers and so a greater degree of uncertainty is likely to be apparent in the resulting parameter estimates. Alternatively, if only a single data point or observation is available from which a model parameter value is assumed then it is clear that error in this data will have a considerable effect on the assumed value. In this case the associated parameter uncertainty is subjective to the analyst and the possible error should be accounted for accordingly.

Since conventional estimation methods assume stable probability distributions (often normal or uniform) for parameter uncertainties then the problem of extreme data points is negated concerning the implementation of Uncertainty Analysis techniques (which place observations based on the characterisation of input probability distributions). Only in the case where input uncertainty distributions are directly developed from raw measurement data, for example some process inherent influences such as feed quality, will the problem of data outliers and associated unstable distributions directly affect the implementation of certain Uncertainty Analysis techniques. Some of the techniques discussed in the following section account for considering extreme parts of input distributions.
3.3 Uncertainty Analysis methods

Conventional error analysis methods can be applied to evaluate the propagation of error and uncertainty in measurement data and parameters to the dependent variables in simple relationships. Specific formulas are available and these may be compounded for more complicated situations. However, with increasing complexity of relationships and time dependent models, large numbers of uncertain parameters and an even larger number of dependent variables the practical application of these formulas are limited. To derive more information about the propagation effects of the uncertainties to the performance variables more flexible and comprehensive methods in the guise of Uncertainty Analyses are discussed.

Uncertainty Analysis aims to provide a quantification of the uncertainty contained in a stochastic system in terms of the output or performance response distribution to the distributions of the inputs (subjective and/or stochastic). Haimes et al. (1994) advocate the use of probability distributions to express a lack of knowledge in Risk Analysis and to provide information to the decision maker. With this information it can be decided whether the observed uncertainty is large enough to affect any subsequent decision and whether this uncertainty is too large to make any meaningful decision in the first instance.

Risk Analysis approaches requires typically involves a number of steps:

- screening,
- definition of measures for the quantification of uncertainty,
- definition of uncertainty space based on the available data,
- approximation of the uncertainty quantification measures within the uncertainty space,
- Sensitivity Analysis measuring contributions of input uncertainty factors.

These steps and associated methods concerning Uncertainty Analysis are discussed in the sub-sections of Section 3.3, with the exception of Sensitivity Analysis which is discussed in Section 3.4.

3.3.1 Screening

Screening is a method which precedes the Uncertainty and Sensitivity Analyses. If there are a large number of potentially uncertain factors in the system (more than a desired number of simulation runs) then it becomes desirable to determine those which may have a significant effect on the system response. A variety of methods are available to accomplish the screening.

Perturbation Analysis is a simple technique in which deviations in the factors are systematically introduced to the deterministic model in turn. This can be used when the number of factors is not too great and the simulation model is not too big. Importance sampling can be used if it is desired to include of high consequence but low probability input scenarios in the screening. The entire input uncertainty space is split into non-overlapping regions (strata) and one observation is sampled from each and is weighted by the probability of each stratum. With a large number of factors group screening techniques are often used.
These are based on the principle of aggregation such that individual factors are combined into groups and are then assessed as individuals. Low-order polynomials are used to approximate the simulation model in a black box manner. This simplifies the system but does not exploit the structure of the actual model as is done in the perturbation and importance sampling methods. Kleijnen (1997) discussed sequential bifurcation (SB) as a novel and efficient group screening technique which proceeds sequentially, splitting the aggregated factors until the most important factors are identified and their effects estimated. The factors identified in the screening procedure may then be incorporated in more rigorous Uncertainty and Sensitivity Analysis methods where a larger sample can be used effectively and more insight obtained.

In this work the Perturbation Analysis approach is considered sufficient since the problem considered is unlikely to contain hundreds of potential factors.

3.3.2 Quantification of uncertainty

In the Risk Analysis approach, to quantify the uncertainty in the performance criteria of the stochastic process system, the uncertain inputs need to be quantified and characterised. The uncertainty in these inputs are modelled by treating them as random variables. In some cases the uncertain inputs are directly based on physical measurement data from which a sample distribution is developed to approximate the population distribution. Here Bayesian techniques can be used to combine prior subjective information on the probability distribution function with new measured data (Draper, 1995). In this work uncertain inputs encompass mainly model parameter uncertainties (as defined in Section 3.2) and are not the measurement data sample distributions. Depending on the information available, the nature of the parameter and the experience of the model developer, normal or uniform distributions are often assumed for model parameter uncertainties. Different distributions may be apparent for other stochastic input properties which are directly based on physical measurements. In order to obtain relevant results to support decisions under uncertainty, a reasonable characterisation which represents the state of knowledge of the input uncertainties is an essential element of any Risk Analysis approach. This topic is discussed in detail by Haimes et al. (1994) in the context of practical risk assessment for decision makers.

Particular measures in the distributed output performance criteria predicted by the stochastic system constitute the risk associated with a process sequence. The general function for a stochastic output criterion measure may be stated in terms of an expected value, E, in some function, f, for a deterministic output performance criterion, D. This is expressed analytically in the probability integral given in Equation 3.1. This is integrated with respect to the cumulative probability distribution of the output criterion, CDF(D), and where θ is the uncertain parameter. The integral in Equation 3.2 is equivalent since the output probability distribution function, PDF(D), is equal to the derivative of the CDF(D) with respect to the output,

\[ E\left[f(D(\theta))\right] = \int_0^1 f(D(\theta))dCDF(D) \] (3.1)
Integrated design under uncertainty for pharmaceutical processes

\[ E\{f(\Phi)\} = \int_{\Theta} f(\Phi) \text{PDF}(\Phi) d\Phi \] (3.2)

Alternatively, since the exact forms of the output CDF(\Phi) and PDF(\Phi) are both unknown, E can be expressed as the multi-dimensional integral over the joint probability distribution function of the stochastic inputs, PDF(\Phi),

\[ E\{f(\Phi(\theta))\} = \int_{\Theta} f(\Phi(\theta)) \text{PDF}(\theta) d\theta \] (3.3)

where \( \Theta \) is the total uncertainty space and PDF(\theta) is estimated from experimental data or assumed from conventional distributions. The two most common measures characterising a distribution are the mean (location) and variance (spread) parameters. The probability integral for the mean or expected value of a performance criterion is given by,

\[ E\{\Phi(\theta)\} = \int_{\Theta} \Phi(\theta) \text{PDF}(\theta) d\theta \] (3.4)

The variance is a common measure for the variability or spread of a distribution about its mean. The probability integral for variance is given by,

\[ \text{Var}\{\Phi(\theta)\} = \int_{\Theta} (\Phi(\theta) - E\{\Phi(\theta)\})^2 \text{PDF}(\theta) d\theta \] (3.5)

Variance can be strongly influenced by the presence of outliers. If outliers are assumed not to be so important then the width of a desired confidence interval or between lower and upper fractiles (e.g. 5% and 95%) can be used as a measure to quantify the uncertainty in only the bulk of the distribution.

The square root of the variance is the standard deviation. This is another measure of variability. It may be interpreted as a measure according to Chebyshev’s rule for any random variable and any distribution: at least \( 1 - 1/k^2 \) of the distribution or observations will fall within \( k \) standard deviations (\( \sigma \)) of the mean (i.e. \( \mu_\Phi - k\sigma_\Phi, \mu_\Phi + k\sigma_\Phi \)) for any number of \( k \) greater than 1.

Taguchi’s quality loss function is an established measure for quality in production engineering. It measures the cost of a loss in quality proportionally to the squared deviation in quality from a desired target relative to the square of the quality deviation at which point the cost is incurred (Taguchi et al., 1988),

\[ L(\Phi) = a(\Phi - \Phi^*)^2 \] (3.6)

where \( a \) is the Taguchi loss proportionality constant,
and $\Delta$ is the permitted tolerance in the performance from the desired threshold value, $\Phi^0$, before a cost, $C_{Tag}$, is incurred for the loss in quality.

A variety of other stochastic criteria may also have application in Uncertainty Analysis. They can include deviation functions from a desired value other than quadratic. These can be linear or one-sided when only uncertainty above a particular threshold is important. Samsatli et al. (1998) recognised a use for one-sided robustness metrics in chemical processes for instances when only the violation either above or below a particular chemical process plant performance threshold is important and any deviations on the opposite side are not important. They introduced a general deviation function from which both one-sided and two-sided performance criteria under uncertainty can be derived. A probability of violation, $Pr_{viol}$, of a (minimum) constraint threshold or desired target, $\Phi^0$, was stated,

$$\begin{align*}
Pr_{viol}\left\{\left(\Phi^0 - \Phi(\theta)\right) > 0\right\} &= \int_{\theta \in \Theta} \beta(\Phi(\theta))PDF(\theta)d\theta \\
\text{where } \beta &= \begin{cases} 
1 & \text{if } \Phi^0 - \Phi(\theta) > 0 \\
0 & \text{if } \Phi^0 - \Phi(\theta) \leq 0
\end{cases}
\end{align*}$$

and a new metric was introduced to measure the average linear extent to which a (minimum) threshold is violated, $E_{viol}$,

$$E_{viol}\left\{\left(\Phi^0 - \Phi(\theta)\right) > 0\right\} = \int_{\theta \in \Theta} \beta(\Phi^0 - \Phi(\theta))PDF(\theta)d\theta$$

One-sided effects can also incorporated into Taguchi’s quality loss function as shown by Bernardo et al. (1999a) where the proportionality constant ($a$) in Equation 3.6 may be preceded by a binary variable, $\beta$.

### 3.3.3 Construction of uncertainty space

Characterisation of the stochastic inputs to the uncertain model system form the uncertainty space from which the system output response is generated. This comprises of probability distribution functions contained within confidence intervals (truncation limits) or joint confidence regions.

The methods discussed in the next sub-section assume the availability of data with which to estimate a characterisation for a particular parameter or set of parameters. In the absence of sufficient data with which to use mathematical methods to estimate parameter uncertainty distributions, then Haimes et al.
(1994) recommend using expert judgement and Bounding Analysis (consideration of sensible values for the limits and the importance of the analytical form of the distribution). This assuming method is employed in this thesis where either normal or uniform distributions are assumed. The effect of different input distributions may be investigated in a Robustness Analysis (Kleijnen, 1997).

3.3.3.1 Limits for parameter uncertainties

If parameters are estimated or assumed independently of each other, the joint sampling space may be described as a hyper-rectangle where each dimension represents one uncertain input bounded by its respective upper and lower confidence limits.

The sampling space for independent uniformly distributed inputs are typically characterised by upper and lower confidence intervals around the nominal value. If no data is available for model parameter estimation, confidence limits around the assumed nominal values are assumed as some percentage of the nominal. For uncertainty in independent parameters of assumed nominal values, desired to be characterised by normal distributions, the standard deviation is assumed at some percentage of the nominal value. Confidence limits around the nominal value can be assumed at some number of standard deviations (typically two or three deviations for approximately 95 or 99.9% probability of containment according to Chebyshev’s rule).

It is a common assumption that model parameter uncertainties arising from parameter estimation procedures based on measurement data, are normally distributed. Least squares regression is a commonly used parameter estimation method for which confidence intervals can be simply stated. For a model linear in its parameters exact confidence intervals around the least squares optimum values of individual parameters, \( \hat{\theta}^* \), may be defined as,

\[
\left| \theta_p - \hat{\theta}_p^* \right| \leq s \sqrt{\left( \frac{1}{N-P} \right) \text{J}^{-1}} \frac{t}{2}
\]

where subscript \( p \) is the index of the input uncertainty (\( \theta \)), \( s \) is the square root of the estimated residual variance computed from the residual sum of squares (RSS) between the regression model predictions, \( \hat{\Phi} \), at the optimum parameter estimates and the measurement data, \( \Phi \),

\[
s = \sqrt{\frac{\text{RSS}(\theta^*)}{N-P}}
\]

\[
\text{RSS}(\theta^*) = \sum_{n=1}^{N} \left( \Phi_n - \hat{\Phi}_n(\theta^*) \right)^2
\]

and \( J \) is the Jacobian matrix of the model predictions with respect to its \( P \) parameters and the values of the confidence limits are defined where the value of \( t \) is taken from the Student’s t-test distribution with \( N-P \)
degrees of freedom (number of regression data points, N, and number of parameters, P, in the regression model), assuming a desired level of confidence, 1-\(\alpha\). For a model non-linear in its parameters, individual confidence intervals can be approximated assuming a linearisation of the model about its optimal estimated parameter values,

\[
\left| \theta_p - \theta_p^* \right| \leq \left( \hat{V}_{pp} \right)^{1/2} t_{N-P,1-\alpha/2}^{N-P,1-\alpha/2}
\]

(3.14)

where \(\hat{V}_{pp}\) is the ppth element of the covariance matrix, \(\hat{V}\), and is the variance estimate of the ppth model parameter (input uncertainty).

In a multi-parameter model where the parameters are estimated simultaneously, a joint confidence region provides a more appropriate measure of the (normally distributed) uncertainty space than would a hyper-rectangle comprising the individual confidence intervals. For a linear model, a hyper-ellipsoidal joint confidence region is defined,

\[
\left( \theta - \theta^* \right)^T J^T J \left( \theta - \theta^* \right) \leq s^2 F_{N,N-P,1-\alpha}
\]

(3.15)

where \(\theta\) is a vector of the model parameters and the value of \(F\) is taken from the F distribution. This is the distribution of a random variable, \(F\), defined as the ratio of two independent chi-squared random variables divided by their respective degrees of freedom. It is commonly used in standard tests of hypotheses in regression. For a non-linear model a hyper-ellipsoidal confidence region is approximated by,

\[
\left( \theta - \theta^* \right)^T \hat{V}^{-1} \left( \theta - \theta^* \right) \leq P F_{N,N-P,1-\alpha}
\]

(3.16)

assuming linearisation of the model about the optimal parameter values. Parameter correlations are contained in the approximation to the covariance matrix.

The likelihood method is another approximate method for confidence intervals and regions. Here the intervals and regions are constructed from contours of constant likelihood which may be expected to provide coverage of the actual confidence intervals or regions more accurately than the linearisation based methods. The problem with this method is the increased computational time in determining a contour of constant likelihood in the model responses and the difficulty in characterising the contour once it has been obtained (Donaldson and Schnabel, 1987).

An exact approach for joint confidence regions is the lack-of-fit method. Donaldson and Schnabel (1987) state that the same disadvantages faced by the likelihood method are present in this method with a further increase in computation requirement in producing a contour based on both the model response and the Jacobian matrix.
Donaldson and Schnabel (1987) conclude from their general study on regression parameter confidence regions that the linearisation methods provide the most concise representation of information required to construct confidence intervals and regions, though not the most accurate. In their work on statistical measures of (joint) parameter estimates from models fitted to respiratory impedance data, Lutchen and Jackson (1986) claim that as a first level of statistical information, the linear approximation methods for confidence intervals and regions can be assumed to be sufficient in capturing the main aspects of the parameter error. This assumption has also been made in chemical engineering applications where linearisation methods have used for confidence intervals and regions. For example Ma et al. (1999) used hyper-ellipsoidal confidence regions for their study of worst case batch process performance as did Rooney and Biegler (1999) in their work on optimal design under uncertainty. In this thesis it is assumed that linearisation methods for confidence intervals and regions are sufficient for parameter uncertainty estimates from least squares parameter estimations.

### 3.3.3.2 Estimation of the parameter covariance matrix

Linearisation methods for the estimation of confidence intervals and regions require the estimation of the parameter covariance matrix (as shown in Section 3.3.3.1). There are three common methods used to approximate the parameter covariance matrix. Donaldson and Schnabel (1987) state that the most common and easily computed estimate for the covariance matrix is,

$$\hat{\sigma} = \sigma^2 \left( J(\theta^*)^T J(\theta^*) \right)^{-1}$$

(3.17)

where $J(\theta^*)$ is the Jacobian matrix of the model predictions at the optimal parameter estimates ($\theta^*$), estimated numerically using the first order Taylor's approximation and $\sigma$ is defined in Equation 3.12. The other linearisation methods require more information,

$$\hat{\sigma} = \sigma^2 H(\theta^*)^{-1}$$

(3.18)

$$\hat{\sigma} = \sigma^2 H(\theta^*)^{-1} \left( J(\theta^*)^T J(\theta^*) \right) H(\theta^*)^{-1}$$

(3.19)

where $H$ is the Hessian matrix of the residual sum of squares at $\theta^*$.

Donaldson and Schnabel (1987) conclude from their study that the linearisation method for confidence regions (see Section 3.3.3.1) should be constructed from the covariance matrix approximated by Equation 3.17 since it is the simplest, most numerically stable and at least as accurate as the other two Hessian based variants (Equation 3.18 and 3.19). The approximation of the covariance matrix by Equation 3.17 for multi-parameter models is assumed in this thesis.
3.3.3.3 Computation of uncertainty space

The uncertain inputs to a system may be characterised by probability distribution functions, as mentioned in Section 3.3. Consideration of a general range of inputs associated with individual process operation models integrated into a single process sequence model can result in a combination of independent distribution functions and a number of multi-variate distributions which may contain correlation structures. Depending on the exact technique used to approximate the system, these distribution functions are either considered independently and the correlation structures implemented after this consideration, or all the distribution functions are combined in a single multivariate function and the corresponding non-zero and zero (independent parameters) correlations combined into a single structure. These issues are further discussed in the following sections.

3.3.4 Approximation methods for performance under uncertainty

Given some definition of the space in the uncertain system inputs, in terms of limits and distribution character between the limits, it is necessary to quantify the propagated output performance criteria of the uncertain system. Due to the common use of highly non-linear input probability distributions (i.e. normal), complex (deterministic) model equations and threshold constraints it is often very difficult or impossible to determine the system output probability distribution function or non-linear probability integral measures (see Section 3.3.2) analytically. Instead approximation methods are used.

For computationally demanding deterministic models within the probability integral (see Equation 3.3), approximation of the actual model can be used. Tatang et al. (1997) proposed a probabilistic collocation method which uses this approach for Uncertainty Analysis of complex geophysical models. The response surface of the actual model is approximated using orthogonal polynomial functions of specified order of the input uncertainties. A sampling technique is then used to approximate the output distribution through simulation of the polynomials. The response surface needs to be well approximated by a low order polynomial expansion and may produce errors if the actual model contains discontinuities. In addition, this method relies on the uncertain inputs being independent.

Alternatively, $f(\Phi(0))$ can be computed directly in scenario based approaches to the multi-dimensional integral approximation. These approximation approaches typically locate observations within the input uncertainty distribution space at which the actual deterministic model is solved. They may be classified into sampling techniques based on either pseudo-random or low discrepancy number generators or on numerical integration approximation formulas.

Different approximation methods for output performance and the main advantages and disadvantages with respect to the issues stated above are discussed next. The different Uncertainty Analysis approaches are summarised in Table 3.1.
Table 3.1. Summary of propagated uncertainty quantification methods for Uncertainty Analysis.

<table>
<thead>
<tr>
<th>Method</th>
<th>Key features</th>
</tr>
</thead>
<tbody>
<tr>
<td>Response surface methodology</td>
<td>Approximate model of the response surface based on a simulated experimental design to place parameter observations. E and Var approximated from the response surface model.</td>
</tr>
<tr>
<td>Monte Carlo sampling (MCS)</td>
<td>Based on pseudo-random number sequence to estimate E and Var. Requires many observations to converge but this number is independent of the dimension.</td>
</tr>
<tr>
<td>Latin hyper-cube sampling (LHS)</td>
<td>Variance reducing stratification approach ensuring full coverage of input distributions to give better performance than MCS. May have poor uniformity properties with increasing dimensions.</td>
</tr>
<tr>
<td>Hammersley sequence sampling (HSS)</td>
<td>Based on low discrepancy (high uniformity) quasi-random number sequence. More efficient than MCS but is dependent on the dimension.</td>
</tr>
<tr>
<td>Equal probability sampling (EPS)</td>
<td>Stratified approach based on sampling from level sets of equal probability in parameter space. More accurate confidence regions for highly correlated parameters.</td>
</tr>
<tr>
<td>Quadrature numerical integration</td>
<td>Numerical approximation technique based on solving Legendre polynomials. High accuracy but number of collocation points is exponentially dependent on the dimension.</td>
</tr>
<tr>
<td>Cubature numerical integration</td>
<td>As for quadrature with reduced accuracy but significantly more efficient requiring fewer collocation points for certain situations (e.g. normal probability distributions, &lt; 10 dimensions).</td>
</tr>
<tr>
<td>Fourier amplitude sensitivity test (FAST)</td>
<td>Transforms multi-dimensional uncertainty space into single dimensional Fourier s-space. Fourier transform pattern search covers s-space for the approximation of E and Var. Increased efficiency and coverage due to integration over one dimension in s-space but not applicable for correlated parameters.</td>
</tr>
</tbody>
</table>

### 3.3.4.1 Differential analysis

Differential analysis is a method used to quantify the local uncertainty about a nominal set of parameter values based on a Taylor series expansion. The effects of perturbations are often approximated with first order terms.

$$
\Phi(\theta) = \Phi(\theta^*) + \sum_{p=1}^{P} \frac{\partial \Phi(\theta^*)}{\partial \theta_p} \Delta \theta_p
$$

(3.20)

where \( \Phi \) is the output performance, \( \theta^* \) is the vector of nominal values of the uncertainties, \( \Delta \theta \) is the vector of perturbations in \( \theta \) and \( p \) is the index of the uncertain factors. Higher order expansions are possible if the necessary partial derivatives can be reasonably obtained. Variance propagation techniques can be used to...
estimate the expected performance value, $E(\Phi)$, and the variance, $\text{Var}(\Phi)$. For first order Taylor series expansion terms, these estimators are given by,

$$E(\Phi) = \Phi(\theta^*) + \sum_{p=1}^{P} \left[ \frac{\partial f(\theta^*)}{\partial \theta_p} \right] E(\Delta \theta_p)$$

(3.21)

$$\text{Var}(\Phi) = \sum_{p=1}^{P} \left[ \frac{\partial f(\theta^*)}{\partial \theta_p} \right]^2 \text{Var}(\theta_p) + 2 \sum_{p=1}^{P} \sum_{r=p+1}^{P} \left[ \frac{\partial f(\theta^*)}{\partial \theta_p} \right] \left[ \frac{\partial f(\theta^*)}{\partial \theta_r} \right] \text{Cov}(\theta_p, \theta_r)$$

(3.22)

where Cov is the covariance.

### 3.3.4.2 Response surface methodology

An approximation to the response surface of the model in the space of the uncertainties can be developed for use in Uncertainty and Sensitivity Analyses. An experimental design is used to obtain input parameter scenarios at which points the original model is simulated. Kleijnen (1997) discussed the use of factorial and fractional factorial designs of simulation experiments to determine which factor combinations are required to maximise the accuracy of the factor effects when the simulation time for large deterministic models is limiting. A surface response is constructed from these simulation results, usually based on least squares techniques. A response surface based on first order terms is given by,

$$\Phi = b_o + \sum_{p=1}^{P} b_p \theta_p$$

(3.23)

More complex response surfaces involving higher orders, cross-products or based on polynomial fits can be used. Equations estimating the expected value and variance are based on the response surface model. Based on the first order response surface the expected value and variance can be estimated,

$$E(\Phi) = b_o + \sum_{p=1}^{P} b_p E(\theta_p)$$

(3.24)

$$\text{Var}(\Phi) = \sum_{p=1}^{P} b_p^2 \text{Var}(\theta_p) + 2 \sum_{p=1}^{P} b_p b_r \text{Cov}(\theta_p, \theta_r)$$

(3.25)

Alternatively, a sampling technique can be used to place scenarios in the response surface (Equation 3.23) and $E$ and $\text{Var}$ can be obtained from the sample estimations. This provides an estimate for the distribution in $\Phi$ at little computational cost but requires a good approximation of the true response surface.
3.3.4.3 Sampling based techniques

Sampling based techniques allow the performance response to be estimated over the entire space of the uncertainties. The selection of successive inputs to obtain the information on the stochastic output is very important when it is desired to keep the number of observations low for computational efficiency reasons. Whatever the sampling technique, the deterministic model is simulated at the observed values of the stochastic inputs located by the sampling strategy. This placement is based on the distribution characteristics of the input uncertainties. Sample based measures are used to approximate the integral quantities defined in Section 3.3.2. Common statistical measures may be applied to quantify different aspects of the generated sample output distribution, such as the sample expected value and variance (sample estimates for the probability integrals given in Equations 3.4 and 3.5, respectively).

\[
E\{\Phi(\theta)\} = \frac{1}{M} \sum_{m=1}^{M} \Phi(\theta_m)
\]  

(3.26)

\[
\text{Var}\{\Phi(\theta)\} = \frac{1}{M-1} \sum_{m=1}^{M} \left(\Phi(\theta_m) - E\{\Phi(\theta_m)\}\right)^2
\]

(3.27)

The sampling based techniques are flexible to the estimation the probability of performance threshold violation and expected extent of (minimum) threshold violation (see Equations 3.8 and 3.10, respectively),

\[
\text{Pr}_{\text{viol}}\left(\left(\Phi^{th} - \Phi(\theta)\right) > 0\right) = \frac{1}{M} \sum_{m=1}^{M} \beta_m(\Phi_m(\theta))
\]  

(3.28)

\[
E_{\text{viol}}\left(\left(\Phi^{th} - \Phi(\theta)\right) > 0\right) = \frac{1}{M} \sum_{m=1}^{M} \beta_m(\Phi_m(\theta)) \left(\Phi^{th} - \Phi_m(\theta)\right)
\]

(3.29)

where the binary variable, \(\beta\), is defined in Equation 3.9.

Monte Carlo, Latin Hyper-cube, Hammersley and Equal Probability Sampling strategies are discussed. These methods differ in the properties of the number sequences generated. There are basically two important properties: randomness and uniformity. Diwekar and Kalagnanam (1997) state that since the error of approximating an integral with a finite sample depends on the equidistant properties and the fact that there is usually no physical significance between successive sample observations in real applications, the uniformity of a sample is more critical than the randomness in the approximation of a uniform distribution. Discrepancy or dispersion is a measure of sequence density and quantifies the deviation of a number sequence from a uniform distribution (Lambert, 1988). Quasi-random sequences provide low discrepancy (good uniformity) and provide more accurate integration approximations than pseudo-random sequences for a specified number of observations.

The most well known and commonly applied sampling technique is based on Monte Carlo pseudo-random number generation methods (MCS). A pseudo-random number sequence is generated to approximate
Integrated design under uncertainty for pharmaceutical processes

uniform coverage over a P-dimensional unit hyper-cube. These observations are inverted over the cumulative distribution functions of the input uncertainties. Typically a large number of observations are required to converge to a reasonable accuracy and on average the number of observations, \( M \), required to maintain a probabilistic integral approximation error within \( \varepsilon \) is proportional to \( \varepsilon^{-2} \) (Wozniakowski, 1991). This is independent of the number dimensions, \( P \).

Latin hyper-cube sampling, LHS (McKay et al., 1979) is a more thorough stratified approach than importance sampling (previously discussed as a screening method in Section 3.3.1). In LHS full coverage of the range of each input uncertainty is ensured. The probability distribution of each stochastic input is split into equal intervals of probability from which one input value is chosen at random and is weighted by the distribution. These weights replace the \( M \) and \( M-1 \) reciprocal weights in the sample mean and variance estimations (Equations 3.26 to 3.29). Values from one distribution are randomly paired with values from the others and once selected they are not replaced. In this way \( M-P \) tuplets of input observations are made. During this selection it is assumed that each input is independent and any correlation structures are implemented afterwards. Median LHS is a similar approach except the input values are not selected randomly from each interval but are the median values. There is no theoretical error bound for stratified methods but studies have shown LHS to be significantly more efficient than MCS (McKay et al., 1979, Diwekar and Kalagnanam, 1997).

The Hammersley sequence sampling (HSS) scheme, introduced by Diwekar and Kalagnanam (1996 and 1997) is based on a variant of the low discrepancy Hammersley sequence. HSS is a quasi-random Monte Carlo scheme which constructs a quasi-random sequence which performs better than pseudo-random Monte Carlo. They present an algorithm which generates the Hammersley points which place \( M \) points of coverage inside a P-dimensional unit hyper-cube in a low discrepancy design. This design has better uniformity properties in P-dimensions than MCS, LHS or median LHS. A disadvantage is that the convergence accuracy is also dependent on \( P \) but in general it is more efficient with the minimal number of samples (\( M \)) required to guarantee an average case error (\( \epsilon \)) is of the order \( \epsilon^{-1/2} \log(\epsilon^{-1}) \)\(^{P-1/2} \) (Wozniakowski, 1991). This relationship for HSS is compared to the MCS for values of \( P \) (in HSS) at 2, 5 and 20, shown in Figure 3.1.

Iman and Conover (1982) introduced a rank correlation technique to induce desired correlation structures between independently generated inputs from sampling based techniques. The advantages of this technique is that it is independent of the types of input distributions, simple to implement and the values of the original (uncorrelated) sample observations are retained so that the structure of the sample is unaffected. This is important for stratified sampling procedures such as LHS. The technique is based on the reordering of independently generated input samples according to the ranking order of a transformed matrix of arbitrary scores which is assumed to have a rank correlation matrix close to the user supplied target correlation matrix. Iman and Conover (1982) applied this to LHS using an arbitrary matrix of van der Waerden scores. Diwekar and Kalagnanam (1997) implemented this technique in the HSS with an arbitrary matrix of Hammersley points.
Vasquez et al. (1999) introduced a new Equal Probability Sampling (EPS) technique for non-linear models. EPS combines a random sampling technique with the generation of level sets of probability (closed hyper-surfaces) in the stratified parameter space, such that each point on a set is equally likely. More accurate confidence regions can be produced than those based only on the covariance matrix, when the parameters are highly correlated. The probability distribution of the parameter estimation regression objective function, the residual sum of squares (RSS), is stratified into intervals of equal probability which are inversed to form the level sets in the parameter space. The EPS method randomly places sample points in each level set to approximate the RSS probability distribution.

3.3.4.4 Numerical integration techniques

Numerical integration techniques can also be used when the distribution function is known or can be approximated. These are not based on the random or low discrepancy generation of input scenarios of equal weighting. The scenarios are simultaneously located and weighted according to the satisfaction of some particular condition. Gaussian quadrature and cubature methods are discussed.

Gaussian quadrature is a common numerical approximation method for multi-dimensional integrals. It has been applied in optimal chemical process design under uncertainty approaches by Straub and Grossman (1990, 1992, 1993), Pistikopoulos and Ierapetritou (1995) and Terwiesch et al. (1998). Quadrature points are located in the [-1, 1] interval at the roots of the relevant order Legendre polynomial (based on the desired number of points). These points are then transformed into the actual space given the bounds. Each point is assigned a weight according to its location and Legendre polynomial solution. The deterministic model is simulated at each weighted quadrature point in the input space and an output is determined to estimate the desired probability integral, E, for some function (possibly a deviation), f, of the output performance criterion, Φ. The product Gauss integration formula with the transformation into the uncertainty space is given by.
integrated design under uncertainty for pharmaceutical processes

\[ E\left[ f(\Phi(\theta_{p,m})) \right] = \prod_{p=1}^{P} \frac{\theta_{UB}^{p} - \theta_{LB}^{p}}{2} \sum_{m=1}^{M} w_{m} \Phi(\theta_{m}) PDF(\theta_{m}) \]

(3.30)

where \( M \) is the total number of quadrature points, \( w \) is weight, \( \nu \) is the location in uncertainty space, \( P \) is the total number of input uncertainties of index \( p \), PDF is a joint probability distribution function characterising the input uncertainties, \( \theta \), and \( \theta_{UB} \) and \( \theta_{LB} \) are the upper and lower bounds. The quadrature approximation can be highly accurate and efficient for low dimensional problems.

Bernardo et al. (1999b) present specialised integration formulas for product Gauss and cubature which can be used to increase the efficiency and accuracy of numerical integration over the multi-dimensional probability distribution if the input uncertainties are normally distributed. They are constructed to integrate over the entire input uncertainty space irrespective of design feasibility constraints. Here cubature is defined as a numerical integration technique which generalises the principles of one-dimensional quadratures to multi-dimensional integration. Cubature formulas can be more efficient than product Gauss but have a reduced accuracy. Specialised formulas are specifically constructed to integrate over multi-dimensional space and are not the products of one-dimensional quadrature and require a lower number of points than the product Gauss formula.

The definition of the multi-variate normal probability distribution function incorporates the correlation structure via the covariance matrix,

\[ N(\mu, \hat{V}) = \frac{1}{(2\pi)^{\frac{1}{2}} (\det V)^{\frac{1}{2}}} \exp \left[ -\frac{1}{2} (\theta - \mu)^{T} \hat{V} (\theta - \mu) \right] \]

(3.31)

where \( \det \hat{V} \) is the determinant of the covariance matrix, \( \Theta \) is the matrix of uncertain input observations (the collocation points in the uncertainty space) and \( \mu \) is the vector of nominal values.

### 3.3.4.5 FAST method

The Fourier amplitude sensitivity test (FAST) procedure introduced by Cukier et al. (1973) was developed to approximate specific criteria such as the performance mean and variance criteria and global sensitivity coefficients. The Uncertainty Analysis aspect is discussed here (mean and variance) and the FAST sensitivity indices are discussed later in Section 3.4.3.3. The FAST method is based on the transformation of the \( P \)-dimensional integral over the joint probability distribution characterising the input uncertainty space into a one-dimensional integral over Fourier s-space. This is achieved by using an appropriate Fourier transformation function, \( G \), for each input factor, \( \theta_{p} \).

\[ \theta_{p} = G_{p}(\sin \omega_{p}s) \]

(3.32)
where $\omega$ is the angular frequency, and $s \in (-\pi, \pi)$ is a scalar. This allows the estimation of the output mean and variance from integrals over a single dimension in $s$.

\[
E\{\Phi\} = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(s)ds
\]  

\[
Var\{\Phi\} = \frac{1}{2\pi} \int_{-\pi}^{\pi} f^2(s)ds - \left[E\{\Phi\}\right]^2
\]  

\[
= 2\sum_{j=1}^{n} \left(A_j^2 + B_j^2\right)
\]

where $f(s)$ is a function of the $G$, and $A$ and $B$ are the Fourier coefficients which are integrals in one-dimensional $s$-space. Application of the FAST method for the Uncertainty Analysis requires the simulation of the deterministic model over a sample of observations in $s$-space for the numerical integration of the Fourier coefficients. The range in $s$ is discretised into equally spaced points for a set of observations which are transformed to the actual input factor values via each Fourier transform $G$ (Equation 3.32). This is a pattern search method as opposed to a random Monte Carlo type placement. The deterministic model is simulated at each input factor co-ordinate. The choice of the Fourier transformation function determines the search curve which aims to sample the input factor space according to the probability distribution of the input. The magnitude of the frequency determines the uniformity and density of the coverage obtained in the input uncertainty space.

A range of techniques for the approximate quantification of performance uncertainty under input parameter uncertainty have been addressed in this section. Related to these Uncertainty Analysis techniques is the application of a variety of methods and measures by which contributions towards the propagated output uncertainty may be identified and quantified, in Sensitivity Analysis. Therefore before the selection as to the most appropriate Uncertainty Analysis approach to the problem considered in this thesis is made, the Sensitivity Analysis techniques are reviewed. The applicability, advantages and disadvantages of the combined analyses are identified in the discussion (Section 3.6).

3.4 Sensitivity Analysis methods

Sensitivity Analysis aims to identify the major contributions to the observed uncertainty in the output predictions as an alternative to the subjective examination of multiple sets of scatter plots. With regard to the contributions of the uncertain inputs to the stochastic system, Sensitivity Analysis provides quantitative measures of the strength of certain relationships between the uncertain parameters to the predicted output variables. There are many different methods for global Sensitivity Analysis (see Saltelli et al., 2000a and 2000b) and these measures are mainly classes of the variance-based methods. Table 3.2 provides a summary of the sensitivity measures. These are defined and discussed in the following subsections.
Table 3.2. Summary of Sensitivity Analysis methods and measures.

<table>
<thead>
<tr>
<th>Method</th>
<th>Key features</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ceteris paribus (one at a time)</td>
<td>Informal approach based on introducing independent deviations in turn to measure (first order) effect on output.</td>
</tr>
<tr>
<td>Differential analysis</td>
<td>Partial derivatives provide a normalised local sensitivity coefficient and an estimate of the fractional contribution of a parameter to the output variance. Often difficult to obtain the partial derivatives.</td>
</tr>
<tr>
<td>Response surface methodology</td>
<td>Normalised coefficients of the response surface model provide first order factor sensitivities. Accuracy depends on how well the response surface predicts reality.</td>
</tr>
<tr>
<td>Correlation coefficient (CC)</td>
<td>Sample based measure for linear relationships. Simple to compute but susceptible to spurious correlations.</td>
</tr>
<tr>
<td>Standardised regression coefficient (SRC)</td>
<td>Sample based linear measure of the ‘standardised global influence’ of an input. It is not susceptible to spurious correlations.</td>
</tr>
<tr>
<td>Partial correlation coefficient (PCC)</td>
<td>Sample based linear measure of input factor importance excluding the effects of other factors.</td>
</tr>
<tr>
<td>Rank transformation</td>
<td>Transformation for sample based measures which negates the strong influence of outliers and resulting first order measures are applicable for non-linear but monotonic relationships.</td>
</tr>
<tr>
<td>Correlation ratio (CR)</td>
<td>Variance based importance measure for global sensitivity analysis. Measures contribution in non-linear and non-monotonic relationships.</td>
</tr>
<tr>
<td>Sobol’ index (So)</td>
<td>Variance based importance measure for global sensitivity analysis. Amenable to computation of total effect indices.</td>
</tr>
<tr>
<td>Fourier amplitude sensitivity test (FAST)</td>
<td>Transforms multi-dimensional uncertainty space into single dimensional Fourier s-space. Fourier transform pattern search covers s-space for the approximation of $E$ and $Var$. Increased efficiency due to integration over one dimension but not applicable for correlated parameters.</td>
</tr>
</tbody>
</table>

3.4.1 Some sensitivity measure definitions

3.4.1.1 Informal sensitivity coefficient

Measurement of the effect on the deterministic model performance criteria under individual perturbations in each uncertain parameter from its nominal value while the others parameters are fixed at their nominal or mean values is also known as the ceteris paribus or one at a time approach. As previously stated this approach has also found application as screening process for the Uncertainty Analysis (see Section 3.3.1), and these effects may be also considered as first order sensitivity measures. This sensitivity or perturbation coefficient, $PC$, may be estimated as a relative change in performance, $\Phi$, for a fixed deviation in a particular uncertain input, $\Delta \theta_p$. 
3.4.1.2 Differential analysis coefficients

The terms computed in the differential analysis approach for Uncertainty Analysis (Section 3.3.4.1) can be used to compute Sensitivity Analysis measures. The partial derivatives from the first order truncation of the Taylor’s series, $\delta f(\theta^*)/\delta \theta_p$ (see Equation 3.20), may be used to form a normalised sensitivity or differential coefficient, $DC$,

$$DC_p = \frac{\delta f(\theta^*)}{\delta \theta_p} \frac{\theta^*_p}{\Phi(\theta^*)}$$ (3.37)

The $DC$ measures the effect on the performance due to a perturbation in an input parameter from its nominal value. Additionally, for a first order Taylor series approximation the fractional contribution to the output variance, $FCV$, of an input uncertainty, $\theta_p$, may be estimated,

$$FCV_p = \left[ \left( \frac{\delta f(\theta^*)}{\delta \theta_p} \right)^2 \frac{\text{Var}(\theta_p)}{\text{Var}(\Phi)} \right]$$ (3.38)

3.4.1.3 Response surface coefficient

Response surface parameter sensitivities, $RSC$, may be obtained from normalisation of the coefficients of the response surface (Equation 3.23),

$$RSC_p = \frac{b_p E(\theta_p)}{E(\Phi)}$$ (3.39)

which are analogous to the differential coefficients given in Equation 3.37. The $RSC$ measures the importance of each $\theta$ with respect to equal sized perturbations from the nominal.

3.4.1.4 Correlation coefficient

Correlation coefficients (CC) provide linear measures of the input contributions. The square of the CC represents the fraction of the variability in the output explained by the total variability in an input. The sensitivity interpretation is lost if the inputs are correlated and accordingly they are susceptible to spurious correlations which may be present in a limited sample over multiple dimensions. The Pearson product
moment sample correlation coefficient (CC) is estimated from the covariance between the output sample and the input, divided by the corresponding standard deviations,

$$CC_p = \frac{SS_{\Phi \theta_p}}{\sqrt{SS_{\Phi}SS_{\theta_p}}}$$  \hspace{1cm} (3.40)

where subscript $p$ is the index of the uncertain inputs, $SS_{\Phi \theta}$ is the sum of products of the distances of the performance, $\Phi$, and input uncertainty, $\theta$, values from their means, $SS_{\Phi}$ and $SS_{\theta}$ are the sum of squares of the distances from the means of $\Phi$ and $\theta$, respectively.

### 3.4.1.5 Standardised regression coefficient

Standardised regression coefficients (SRC) can be estimated in the presence of correlated inputs and may be compared to the CCs to avoid misleading interpretations due to spurious correlations. In contrast to CCs, the square of the SRC represents the fraction of the output variability which is explained by the variability in the input not due to any of the other inputs. SRCs are first order sensitivity measures and may be interpreted as linear measures of the 'standardised global influence' of each input on the output, Hofer (1999). The benefit of standardising the data is the provision of a measure which is independent of the subjective input probability distributions and the values can be directly compared with each other. SRCs are the coefficients of the multi-linear regression problem minimising the sum of squared errors between the standardised output data and the regression model output. They are derived from the regression coefficients, $b$, determined from the following multi-linear regression,

$$\hat{\Phi}_m = b_0 + \sum_{p=1}^{P} b_p \theta_{p,m} + \epsilon_m$$  \hspace{1cm} (3.41)

where $\hat{\Phi}$ is the regression model output prediction, $\epsilon$ is the residual error (from the actual outputs, $\Phi$) due to the linear regression model approximation and subscripts $m$ and $p$ are the indices for the observation number and the uncertain factor. Given $b$, then the regression model can be expressed in standardised form,

$$\Phi_{\text{std},m} = \sum_{p=1}^{P} SRC_p \theta_{p,\text{std},m}$$  \hspace{1cm} (3.42)

where

$$SRC_p = \frac{b_p \sigma_{\theta_p}}{\sigma_{\Phi}}$$  \hspace{1cm} (3.43)

$$\theta_{p,\text{std},m} = \frac{\theta_{p,m} - \bar{\theta}_p}{\sigma_{\theta_p}} \quad , \quad \Phi_{\text{std},m} = \frac{\Phi_m - \bar{\Phi}}{\sigma_{\Phi}}$$  \hspace{1cm} (3.44)
\( \bar{\theta} \) and \( \Phi \) are the sample means of \( \theta \) and \( \Phi \), \( s \) is the sample standard deviation and the subscript std represents a standardised value.

### 3.4.1.6 Partial correlation coefficient

The partial correlation coefficient (PCC) measures the strength of the correlation between the output, \( \Phi \), and a given input, \( \theta_p \), after adjustment for any effect due to correlation between \( \theta_p \) and \( \theta_r \), \( r \neq p \). The largest PCC is associated with the input whose contribution to the multi-linear regression model is least adequately accounted for by the remaining inputs when it is excluded. The PCC can be determined from its relationship with the SRC,

\[
(PCC_p)^2 = (SRC_p)^2 \frac{1 - R_{\theta_p}^2}{1 - R_{\Phi}^2}
\]

where \( R_{\theta_p}^2 \) is the coefficient of determination of the multi-linear regression of \( \theta_p \) on \( \Phi \) and the \( \theta_r, r = 1, 2, \ldots P \) with \( r \neq p \), and \( R_{\Phi}^2 \) is the coefficient of determination of the multi-linear regression of \( \Phi \) on the \( \theta_r, r = 1, 2, \ldots P \).

### 3.4.1.7 Coefficient of determination

The coefficient of determination, \( R^2 \), is not a sensitivity measure but represents the fraction of variability in the sample output which can be explained in a linear function. It is a measure (between 0-1) of how much confidence should be taken from sensitivity measures which assume near linear relationships (CC, SRC, PCC). \( R^2 \) is the square of the multiple correlation coefficient between the multi-linear regression model output and the observed sample output,

\[
R^2 = 1 - \frac{RSS}{TSS}
\]

\[
TSS = \sum_{m=1}^{M} (\Phi_m - \bar{\Phi})^2 = \sum_{m=1}^{M} \Phi_m^2 - M(\bar{\Phi})^2
\]

\[
RSS = \sum_{m=1}^{M} (\Phi_m - \hat{\Phi}_m)^2 = \sum_{m=1}^{M} \Phi_m^2 - b^T \theta^T \Phi
\]

where \( RSS \) is the residual sum of squares between the vector of deterministic model output observations, \( \Phi \), and the vector of regression model outputs, \( \hat{\Phi} \). \( TSS \) is the total sum of squares of the distances between \( \Phi \) and the mean value, \( \bar{\Phi} \). Subscript \( m \) denotes a sample observation in \( \Phi \) of total number \( M \), \( b \)
is vector of the least squares point estimates for the multi-linear regression coefficients (which may be computed analytically from Equation 3.49), $\theta$ is the matrix of uncertain input observations.

3.4.1.8 Stepwise regression analysis

Stepwise regression allows the building of a regression model which does not include all the possible input uncertainties so as to avoid over-fitting problems (where the regression model attempts to fit the predictions of observations rather than the trends). By stepwise addition of the most important inputs as identified in a prior correlation analysis, the regression model is sequentially increased. Each additional input has the largest correlation with the uncertainty in the dependent variable (output) that is not included in the current regression model parameters. The change in the coefficient of determination, $\Delta R^2$, with each parameter addition is an immediate measure of the total fraction of output uncertainty which is accounted for by the added parameter uncertainty. To terminate the addition of inputs, a F-test or t-test of significance is used to determine the probability that the additional regression coefficient has an absolute value larger than a value which would be obtained if there was no relationship at all between the additional parameter and the output. The SRC values for the final regression model can then be compared for contribution ranking.

3.4.2 Rank transformation

Rank transformation of sample observations aims to remove the effect that outliers and/or strongly skewed distributions may have on contributor measures, and in particular those based on regression models (Iman and Conover, 1979). The transformation replaces data values with their ordinal numbers and the ordinal (rank) data is used in the subsequent calculations and regressions. This allows the estimates for some of the sampling based sensitivity measures measuring the strength of linear relationships (e.g. CC, SRC, PCC) to be more dependent on the bulk of the sample values and less dependent on the presence of a small number of outliers. The rank equivalent linear measures (CRR, SRRC, PRCC), are based on the strength of monotonic relationships rather than linear relationships and works with non-linear relationships between the input and output variables if they are monotonic. Non-monotonic non-linear relationships require more sophisticated techniques.

3.4.3 Measures of importance

Since linear relationships may not satisfactorily explain the output variability an alternative measure can be estimated. Measures of importance are global variance-based sensitivity measures which are model independent, being able to measure non-linear and non-monotonic relationships. They can be used to provide measures of global importance which are not based on regression or correlation analysis and are able to avoid the 'curse of dimensionality' for the estimation of total effect indices (exponential increase in number of interaction terms with number of factors). They can be used in conjunction with sampling
based coverage methods. Three measures are discussed: the approximate correlation ratio (CR), the Sobol’ index (So) and the FAST index.

3.4.3.1 Correlation ratio

The correlation ratio, CR, is based on the notion that if an input uncertainty is fixed at its nominal value then it is an important influence if the predicted output variance is reduced by a large amount relative to the output variance when the input is not fixed. The square of the CR is the fraction of the output variability explained only by the variability in a particular input. Analysis of variance components methods to estimate the CR tend to be computationally expensive (Saltelli et al., 2000), requiring many model evaluations. Hofer (1999) approximated the first order CR for a particular input by splitting the input sample set into a number of disjoint intervals which each contain an equal number of observations. In this way the conditional means of the outputs at given values of the inputs can be approximated,

\[
CR^2_p = \frac{\text{Var}_{\theta_p} \left\{ E\Phi|\theta\right\}}{\text{Var}\left\{\Phi\right\}} = 1 - \frac{E_{\theta_p} \left\{ \text{Var}\Phi|\theta\right\}}{\text{Var}\left\{\Phi\right\}} \tag{3.50}
\]

where \( \Phi \) is the vector of deterministic model performance outputs, \( \theta_p \) is the vector of observations in the \( p \)th uncertain input, \( \text{Var}_{\theta_p} \) and \( E_{\theta_p} \) and are the variance and expectation conditioned on \( \theta_p \). A total effect CR index can be estimated by leaving all the \( \theta_p \) undetermined and conditioning on all the \( \theta \), where \( r \neq p \).

3.4.3.2 Sobol’ index

The approach of Sobol’ (1993) is based on a decomposition of the model output function into orthogonal summands of increasing dimensionality. Each index, So, is a ratio of the partial variance (computed from integration over the required decomposition terms), \( D_{p_{1},...}p_{S} \), to the total variance, \( D \),

\[
So_{p_{1},...}p_{S} = \frac{D_{p_{1},...}p_{S}}{D} \quad \text{for} \quad 1 \leq p_{1} < ... < p_{S} \leq P \tag{3.51}
\]

for \( S \) sensitivity measures and \( P \) uncertain inputs. \( D \) and \( D_{p} \) can be approximated from simulation sample observations using an appropriate sampling strategy for the inputs. All the So indices sum to 1 and a first order sensitivity index measures the fractional contribution of each input to the variance of the output (i.e. the fractional contribution of \( \theta_p \) to the variance in the output, \( f(\theta) \)). A second order index measures the interaction effect between the inputs (i.e. the part of the variation in \( f(\theta) \) due to \( \theta_p \) and \( \theta_r \) that cannot be explained by the sum of the individual effects of \( \theta_p \) and \( \theta_r \)).
A total sensitivity index (TSI) which measures the total contribution of a single input factor to the output variance can be estimated given the partial variance due to all the factors not involving the input. The TSI is more reliable than the first order So index in the assessment of the overall effect of each single factor.

### 3.4.3.3 FAST index

As discussed in Section 3.3.4.5, the Fourier amplitude sensitivity test (FAST) procedure of Cukier *et al.* (1973) can be used for Uncertainty and Sensitivity Analyses. FAST Sensitivity Analysis is another variance based approach and the global FAST indices estimate the contribution of individual factors to the output variance. These work for both monotonic and non-monotonic input-output relationships. The Fourier series representation of the original model can be decomposed to obtain these fractional contributions. The partial variance contribution of input factor \( p \), \( D_{wp} \), requires the simulation of the deterministic model at each discrete value in s-space and then computation of the Fourier coefficients, \( A \) and \( B \). The first order FAST indices, \( S_{\text{FAST}} \), is the partial variance relative to the total variance,

\[
D_{wp} = 2 \sum_{j=1}^{m} \left( A_{jwp}^2 + B_{jwp}^2 \right)
\]

\[
S_{\text{FAST}}^p = \frac{D_{wp}}{D_{\text{FAST}}}
\]

providing the angular frequencies, \( \omega_p \), are integers and the total variance estimate, \( D_{\text{FAST}} \), is given in Equation 3.35. The total FAST indices (Saltelli *et al.*, 1999) include the additional higher order effect of an input due to any interactions with the other inputs.

### 3.5 Some past applications

The use of Risk Analysis methods to deal with uncertainty in model systems in a quantitative manner is prevalent in a wide range of applications where models play an important role in structuring the available data into relevant information for the decision maker.

A large amount of the literature concerning Uncertainty and Sensitivity Analysis associates methodologies orientated towards the use of large scale continuous or discrete time event models which contain hundreds of possible uncertain factors and for which a single simulation can take a significant amount of computer time (refer to reviews by Helton, 1993, Kleijnen, 1997, Hofer, 1999, Saltelli *et al.*, 2000a). Methods have been applied to large scale models arising from a variety of applications including the fields of nuclear physics (Helton, 1993), geophysics (Tatang *et al.*, 1997), environmental (Crosetto *et al.*, 2000) and business and logistics (McKay *et al.*, 1999).

In chemical engineering, Risk Analysis approaches involving uncertainty and sensitivity methods of the type discussed have gained some attention for the simulation assessment of continuous processes and in
particular those unit operations employing thermodynamic models and chemical reaction kinetic models subject to physical property and kinetic parameter uncertainties.

Diwekar and Rubin (1991) investigated a complex gasification flowsheet using a sample based (Monte Carlo and Latin hyper-cube) stochastic modelling capability in a continuous chemical process flowsheet simulator (ASPN). They showed that the consideration of internal and external uncertainties in chemical plants was useful for a range of applications. These included performance and economic assessment, risk analysis, feasibility studies and for the comparison between alternative technologies. Frey and Rubin (1992) consolidated these views towards process development involving new technologies, emphasising the importance of uncertainty and stochastic modelling (using MCS) in providing information for research and development decisions.

Several articles have investigated uncertainties in rigorous physical properties in chemical engineering applications. Nelson et al. (1983) tested the sensitivities of distillation column designs to variations in the predicted phase equilibrium behaviour through one at a time deviations in the average relative volatilities. Kubic and Stein (1986) used a fuzzy set approach for uncertainty in thermodynamic models, which was associated more with structural errors. Macchietto et al. (1986) measured first order sensitivities of a VLE flash, a superfractionation column and a multi-stage integrated flowsheet to uncertainties in the relevant thermodynamic parameters and function models. They based this on Taylor series expansion and gradient chain-ruling. A LHS based approach is implemented by Whiting et al. (1993) and Vasquez and Whiting (1998) for the analysis of the combined effect of thermodynamic model parameter uncertainties (correlated binary interaction parameters, accentric factors, critical temperatures and pressures) on the key duties of a distillation column and in the predicted phase equilibria behaviour in liquid-liquid extractions. They used the first order SRC, PCC and the rank equivalent measures to investigate the meaning of the results. In a further study concerning thermodynamic uncertainties, Vasquez and Whiting (1999) incorporate both systematic and random experimental data errors in a MCS based approach. Here, systematic errors are induced into a data set through the random generation of pseudo-experimental data points. The original data set are shifted to within the bias limits (defined from comparison with literature, standard values and from the experimental conditions) according to a representative uniform probability distribution. Random errors are induced from the parameter regressions based on sets of randomly observed experimental data points whose probability distributions are based on instrument statistics and any other information. The authors found that systematic errors can have a significant role in Uncertainty Analysis of thermodynamic applications and the information can support decision making in experimental data measurement and design. Xin and Whiting (2000) and Vasquez and Whiting (2000) applied a new equal probability sampling (EPS) technique (see Section 3.3.4.3 for details) to analyse uncertainty in thermodynamic models. The former work investigated cryogenic air separation and methanol dehydration process flowsheets and the design of a runaway reaction pressure-relief system under thermodynamic BIP uncertainties. SRC, PCC and rank equivalents were used to identify sensitivity contributions. The latter work investigated thermodynamic models for liquid-liquid systems. Maranas (1997) introduced a method for the quantitative assessment of physical property prediction uncertainty in optimal molecular design.
problems under group contribution parameter uncertainties. Chance constrained programming was used to
transform the stochastic optimisation problem to an equivalent deterministic one, bypassing the need for
integration under the input PDF. This is discussed in more detail in Chapter 4. It was noted how important
property prediction uncertainties were in influencing the optimal designs. Tayal and Diwekar (2001)
considered the same problem using Hammersley sequence sampling.

To examine the dynamic propagation of uncertainty, McRae et al. (1982) implemented the FAST
approach to study the uncertainty in the kinetic parameters and operating conditions of a non-isothermal
autocatalytic reaction. The method is used to attribute partial variances of each factor to the total variance
in the product concentration as a function of reaction time. Torvi and Hertzberg (1997) introduced an
orthogonal polynomial expansion technique for the approximation of evolving probability distribution
functions in uncertain dynamic processes. They applied this to a non-isothermal batch reactor with
uncertain kinetic parameters to observe the propagation of uncertainty (distribution moments and fractiles)
in the product concentration with time. This method is limited to small numbers of uncertain factors due to
the use of quadrature for the integration under the multi-variate distribution function.

3.6 Discussion

The problem considered in this thesis concerns the analysis of integrated sequences of batch and
continuous deterministic models which maybe subject to model-inherent and process-inherent parameter
uncertainties (defined in Section 3.2). This can result in a reasonably large deterministic process model
but not excessively so (since distributed systems are not considered) and not one which requires a
individual simulation time greater than the order of seconds or tens of seconds. This allows some more
flexibility in the methods available than may be apparent for larger scale models. However, the number of
uncertain inputs may also be quite large, in the order of tens but below probably below fifty. This means
that although more scenario evaluations may be permitted, the dependence of the number of evaluations
required for convergence accuracy should not be strongly dependent on the number of uncertain
parameters for application in this work. Regarding the sensitivity techniques, flexibility of application to a
range of different contribution measures would be useful, since the physico-chemical system maybe
strongly non-linear in its parameters.

A number of properties are important in the consideration between the various Uncertainty and Sensitivity
Analysis methods and measures that have been defined in the previous sections. For the Uncertainty
Analysis these include:

- conceptual simplicity, flexibility and ease of implementation for different analysis situations,
- proper coverage of the full range of the candidate uncertainties,
- dependence of the estimated distribution accuracy to the number of observations (deterministic model
evaluations),
Integrated design under uncertainty for pharmaceutical processes

- dependence of the number of required deterministic model evaluations to the number of input factors,

- capacity to capture the significance of outliers and erroneous data.

The differential analysis approach has two major disadvantages. The first is that it provides only local information on the effect of uncertainty from the nominal point of the analysis. If the relationship between the inputs and outputs are highly non-linear and contains discontinuities then differential analysis will not be useful. No information concerning the distributions of the inputs are conveyed in the results. Also, the calculation of partial derivatives required for the Taylor series can be very difficult, although approximations can be used. These derivatives can be used directly as local measures of sensitivity.

The response surface approach may be straightforward once the surface approximation has been constructed but a number of disadvantages are apparent. These are the difficulty in developing an appropriate experimental design if many input uncertainties are present, the difficulty in detecting non-linearities and discontinuities between the inputs and outputs and the difficulty in including input correlations. Any results from the ensuing analyses are only as good as the approximation of the response surface, for which an adequate construction may be difficult. The response surface approach can also give direct information concerning the sensitivity towards factors.

If there is more flexibility with the number of simulations permitted for the analyses then direct sampling based techniques exhibit a number of desirable advantages for Uncertainty and Sensitivity Analysis. Not only are they conceptually simple but are also flexible to manipulation for specific situations such as the estimation of different statistical measures (not just the mean and variance but also one-sided deviation functions), they exploit the full range of the input uncertainties and the output performance distributions can be estimated without the use of intermediate models (allowing the direct implementation of sequences of models, as observed by Helton, 1993). In addition, the scenario simulation results are readily applicable to the direct estimation of sampling based sensitivity measures such as the CC, SRC, PCC, \( \Delta R^2 \) and their rank equivalents.

An important advantage of sampling techniques based on pseudo-random number sequences is that they do not always require more sample observations as the problem dimension (number of uncertainties) increases. In compromise, a large number (greater than 100) of observations may be needed in the first instance to achieve a reasonable degree of accuracy in the approximation of the integral. The pseudo-random number based approach of MCS is simple to implement, samples from the full range in each input, has a retaining nature (additional observations can be added one at a time to the current sample with disturbing the sample) and the number of observations to reach a reasonable accuracy is more or less independent of the number of uncertainties. The major disadvantages are the large number of observations required to obtain this accuracy and the lack of uniformity which may be apparent in the sample. The stratified approach of LHS is recommended as a variance reduction technique (McKay et al. 1979) and compared to conventional MCS, less variability between samples is expected and a better approximation of the output distribution is obtained for the same number of samples. Diwekar and Kalagnanam (1997) state that the main drawback with LHS is that the stratification (uniformity) is one-dimensional and does
not necessarily provide good uniformity properties on a P-dimensional unit hyper-cube, because the extension to further dimensions is on a random basis. In addition, it is not a retaining method. Similar to MCS, the retaining nature of the Hammersley points in the HSS permits the incorporation of additional input scenarios without disturbance of the original observations. Kalagnanam and Diwekar (1997) show that the convergence rate of samples propagated through different functions are claimed to be between 3 and 100 times faster for HSS than the MCS, LHS and median LHS techniques, for the range of linear and non-linear functions and correlations structures they imposed. A disadvantage may be that the imposition of a correlation structure on the sample appears to change the uniformity properties of the low discrepancy design. The advantages of the EPS method is that it can provide more accurate Uncertainty Analysis predictions when parameters are highly correlated, due to better estimation of the confidence region. However, it is only applicable in the case when the model parameters have been obtained from the regression of experimental data and the residual sum of squares is available. Since the probability distribution of the regression objective function is required in EPS, it does not appear applicable to the case when subjective characterisations of probability distributions in model parameter values are assumed when there is no available data for parameter regression.

Tørvi and Hertzberg (1998) showed that for a given number of observations the convergence accuracy of a stochastic simulation of a batch distillation column, the gaussian quadrature method was better than the median LHS and the Halton sequence methods with MCS being the least accurate. However, the major disadvantage of the quadrature technique is the logarithmic dependency of the sample number with problem dimension (number of input uncertainties). This makes it unsuitable for stochastic systems with a large number (greater than five or six) of uncertain or variable inputs. Bernardo et al. (1999) conclude that when the number of stochastic inputs is less than ten and are normally distributed, specialised cubature formulas can be obtained which are more efficient that HSS and product Gauss formulas. However, since the choice of the most suitable integration formula needs to be selected according to the particular uncertainty problem this is not conducive to a general approach. In addition, application of any of the discussed sensitivity measures would require the re-simulation of new scenarios.

Helton (1993) claim that although the FAST approach allows the full range of each input to be sampled and the original model is simulated without modification (as for the sampling based methods), the main disadvantages are that many observations along the space-filling curve may be required, the underlying mathematical complexity (such as the use of specialised formulas to estimate statistical measures involving Fourier coefficients) renders it inflexible to different analysis situations and it is not possible to specify correlations between the inputs.

A number of sensitivity measures have been identified in Section 3.4. Important considerations regarding the Sensitivity Analysis methods include:

- efficient estimation of sensitivity measures (dependence of the number of required deterministic model evaluations required and dependence on the number of input factors),
• dependence of the sensitivity method accuracy on the additive, linear or monotonic behaviour between the uncertain inputs and outputs of the model,

• capacity to capture the significance of outliers and erroneous data,

• capacity to measure sensitivities of main factor effect and total factor effect.

The limitations of the ceretis paribus (one at a time), differential analysis and response surface approach sensitivity measures have already been identified due to the inherent problems discussed regarding the Uncertainty Analyses.

Use of a sampling based technique allows a range of linear-based sensitivity measures (CC, SRC, PCC, $\Delta R^2$ and their rank equivalents) to be estimated directly from the results of the Uncertainty Analysis without the need for any further simulations. These measures provide first order information on the combined influence of the uncertainties over the entire space. It is important to compare ranking priority of the contributors predicted by the CCs with those predicted by the SRCs. While the SRCs provide more precise information regarding the individual contributions of the uncertainty sources and are not susceptible to spurious correlations, the multi-linear regression from which they are derived, may be susceptible to over-fitting if a large number of inputs are considered. In this case the SRC contributions may be misleading, but careful stepwise regression can be used to avoid over-fitting problems. PCCs can give misleading of input contribution ranking importance since they measure the strength of linear relationships between inputs and the output after corrections have been made for the effects of other inputs (Saltelli et al., 2000). A factor exhibiting a large PCC does not necessarily make a large contribution to the output uncertainty. Hofer (1999) explains the reason is because the PCC is a quotient of parts of the variability in the output explained by the multi-linear regression model, as opposed to fractions of the variability. This means that the relation to the total variability in the output is lost and may make the PCC a less suitable measure for sensitivity ranking. The main limitation of these indices (CC, SRC, PCC and their rank equivalents) is the reliance on the assumption of near linear or near monotonic relationships between the inputs and outputs.

The variance based indices of CR, Sobol' and FAST methods are not limited to near linear or near monotonic relationships. They are global measures of the main effect contribution of each input factor (the fractional contribution of the input factor to the output variance). The first order Sobol' and FAST indices are equivalent to the CR. The total effect indices of Sobol' and FAST provide a more accurate measure of individual factor contribution since they account for effects due to interactions of the considered input with the other inputs (i.e. higher than first order). A disadvantage of the Sobol' sensitivity indices is that a separate integral requiring a new sample set needs to be estimated for any measured effect (first order or higher), which is not required in the FAST method. The Sobol' method requires $n_i(2P+1)$ deterministic model simulations to calculate the first and total sensitivity indices, where $n_i$ is the sample size required to solve each individual variance integral and $P$ is the total number of factors. Unless $P$ is low and the simulation time for each deterministic model is relatively low, computational resource and time could
limit the use of the approach. FAST is a more efficient method for sensitivity than Sobol' indices since it permits the estimation of the first order and higher order sensitivity indices using the same sample set of deterministic model simulations. In contrast to the Sobol' approach, the number of simulations required to calculate the first and total sensitivity indices is less at $P(n_1+1)$.

Of the range of Uncertainty and Sensitivity Analysis approaches discussed, the sampling based methods are preferred in this thesis. In particular Monte Carlo and Latin hyper-cube strategies have been implemented in past chemical engineering applications (see Section 3.5). This is largely due to the flexibility they entail in adaptation to different analysis situations and the more informative analysis they permit in terms of the examination of the complete characterisation of the input uncertainty space. The computational effort required to achieve a sufficient accuracy depends directly on the number of deterministic model simulations and so an efficient sampling method is required. Of the sampling based methods, the Hammersley Sequence Sampling (HSS) that Diwekar and Kalagnanam (1997) implemented for the off-line optimal quality control of chemical process under parameter uncertainty appears to be the most efficient (compared to MCS and LHS) over the range of (linear and non-linear) functions verified in their work. In addition, the authors have already applied HSS successfully to chemical engineering applications (batch distillation, CSTR). Other advantages are the retaining nature of HSS, it remains efficient for a large number of input uncertainty factors and can be used in conjunction with the correlation inducement technique of Iman and Conovor (1982). The disadvantage that as a sampling based approach a reasonable number of observations may still need to be required to attain a reasonable accuracy may be acceptable due to the relatively large number of factors (greater than 5) likely to be employed in the complete process sequence. This makes the use of the numerical integration methods discussed impractical. Regarding sampling based Sensitivity Analysis measures, for essentially linear contributor input-output relationships, the correlation coefficient (CC) and standard regression coefficient (SRC) are proposed. For non-linear but monotonic relationships and to negate the effect of erroneous data and outliers in the input samples, the rank equivalents (RCC and SRCC) are suggested. For strongly non-linear or non-monotonic relationships then the approximate correlation ratio (CR) is suggested since this can also be estimated directly from the sample simulation results of the Uncertainty Analysis. While a limitation is that the sensitivity methods used to compute these measures do not encompass the computation of total effect indices, the essential criterion that the key contributors are identified remains satisfied.

3.7 Conclusions

In this chapter the problem of uncertainty considered in this thesis, is defined as concerning subjective model parameter uncertainties as opposed to stochastic variability or model structure errors. This provides emphasis towards the analysis and improvement of the available models. A variety of different approaches for risk assessment are defined and discussed in this chapter. These encompass both Uncertainty Analysis and Sensitivity Analysis techniques with the aim of quantifying the output performance uncertainty
Integrated design under uncertainty for pharmaceutical processes

propagated from the inputs and identifying the main contributors to this uncertainty. These tools provide more informed support to decisions than would single solution points and provide a basis for the management of uncertainty. Some of these methods have found application in the chemical process literature and in particular regarding the consideration of uncertainty in thermodynamic models for physical property prediction and in reaction kinetics. There appears to be no literature at present concerning the large amounts of uncertainty which may be associated with integrated process sequences such as those found in the pharmaceutical industry.

A sampling based approach using the Hammersley sequence sampling strategy is selected for Uncertainty Analysis in this thesis. This is because of the flexibility in adaptation to different analysis situations (estimation of different measures associated with uncertainty) that sampling methods permit and in particular the improved efficiency that Hammersley sequence sampling (HSS) appears to show over Monte Carlo (MCS) and Latin hyper-cube (LHS) sampling methods. The sampling based approach allows a range of sampling based sensitivity measures to be estimated according to different input-output uncertainty propagation effects, without the need for any additional deterministic model evaluations other than those determined in the Uncertainty Analysis.

In the next chapter optimisation methods under uncertainty are discussed. These provide a form of management response to the process uncertainty. The aim is to provide a more desirable response in the process system performance under uncertainty through the determination of new operating policies or process designs.
CHAPTER 4

SOME ASPECTS OF OPTIMISATION UNDER UNCERTAINTY

4.1 Introduction

In the previous chapter a range of mathematical methods were discussed concerning the role of parameter uncertainties in a stochastic modelling environment. These were associated with Uncertainty and Sensitivity Analyses in relation to internal and external parameter uncertainties in deterministic models. These tools allow the quantification of uncertainty propagated in models and the identification of the main contributors to the observed uncertainty in the performance variables. It was concluded that an efficient sampling based approach would be desirable due to its flexibility for adaptation to different analysis situations.

The analysis approaches discussed in Chapter 3 form the basis of a model based risk assessment to model uncertainty. The next logical step is risk management where actions are considered based on uncertainty information obtained from prior analyses. In this chapter some different approaches to optimisation under uncertainty are reviewed. These may be viewed as a form of management under uncertainty which either consider decisions at the equipment design stage or at the operating stage or both simultaneously. For clarification, this thesis is more concerned with the design of the process in terms of its operating policy and is less concerned with the design of the actual equipment used in the process. This is because in the pharmaceutical industry and particularly in process development it is common to use existing multi-purpose equipment. However, a large proportion of the chemical engineering literature referring to optimisation under uncertainty is concerned with simultaneous equipment design and control variable optimisation. This distinction between design and control variables becomes important with the assumption that the latter can be adjusted for any possible realisations of the uncertainty. The assumption of perfect control adjustment is implicit in optimistic 'wait and see' approaches. The extent of the validity of this assumption depends on the ability to detect the feedback information (from on-line process measurements) and the quality of that information made available for closed loop control adjustment of certain process variables. When design and control variables are considered equivalent only a single operating policy is obtained in a conservative 'here and now' strategy. Since methods proposed for the simultaneous optimisation of equipment design and process operation under uncertainty are characteristic of the evolution of a significant proportion of the problem formulations for optimisation under uncertainty, these are included in the discussion in this chapter.

Similar to the Uncertainty Analysis methods many optimisation under uncertainty approaches involve the evaluation of a deterministic model of the process at particular scenarios (input sets of uncertain parameter values determined from specific observations in the uncertainty space). To reiterate, the deterministic process considered in this thesis consists of a sequence of integrated process models of which the general characteristics are: dynamic or steady-state, lumped parameter and in general non-linear and semi-...
empirical. The prime concern of this thesis are the uncertainties contained in the model inherent parameters since improvement of the fundamental process knowledge associated with the mechanisms for production and separation is a key objective. However, extension to the consideration of uncertainty in process inherent such as feed stream properties and operating conditions is also important since in some cases these factors may also be influential with regard to the overall process system.

4.2 Response to uncertainty

Traditionally, in response to uncertainty, the chemical processing industries have incorporated design margins and/or used one at a time Sensitivity Analyses over specified operating ranges. The latter is implemented in the pharmaceutical industries where it is common to use existing equipment designs and specification of the operating variables is of key importance.

Empirical over-design factors are typically applied to a nominal optimal design obtained from the assumption of a completely deterministic system with nominal parameter values and a single operating condition with the hope to improve operability characteristics. This use of design margins does avoid the complexity of explicitly considering variations, however inadequacies of this convention are:

- lack of a firm rational basis,
- usually not possible to either guarantee optimality or feasible operation for conditions other than the nominal ones,
- achieves a conservatively designed system, but without any way to quantify the degree of conservatism associated with the design.

With the level of computing technology now available, variations and uncertainty in process design can be incorporated in more systematic manners than the use of crude empirical over-design factors. This is apparent in the large amount of chemical engineering literature available concerning uncertainty. Some of reviews on the optimisation of process design and operation under uncertainty are introduced next.

Grossmann et al. (1983) reviewed optimisation strategies for flexible design for steady-state processes under parameter uncertainty. Here, flexibility is defined as the ability of a design to maintain feasible regions of operation under different variations of its parameters. These variations may be due to either parameter uncertainty or multi-product operations. The problem formulations considered in this review focus on minimum cost design with a fixed degree of flexibility and those with an optimal degree of flexibility, quantified via a flexibility index. Rippin (1993) briefly discussed the importance of flexibility and uncertainty in a review of batch process systems engineering. Terwiesch et al. (1994) presented the results of an industrial survey concerning batch unit modelling and operation under uncertainty. They recognised the significant uncertainty which may be apparent in batch reactor process models and reviewed the state of industrial batch reactor practice with respect to design objectives, modelling, instrumentation and operation. Model uncertainty is discussed with respect to off-line (open-loop) control...
profile optimisation under uncertainty (probabilistic and set membership approaches) and in particular regarding optimisation of on-line (closed-loop) feedback control. The former is of interest in this thesis. Pistikopoulos (1995) discussed combined design and operation of this class of problem and provides a classification of uncertainty (model-inherent, process-inherent, external and discrete) upon which a general mathematical formulation for optimisation under uncertainty is given. This formulation is used to address related problems including the value of perfect information (Raiffa, 1968), flexibility, controllability and reliability. From a non-chemical engineering specific perspective, Wets (1996) identified challenges in stochastic programming associated with the design of models for making optimal decisions under uncertainty. This is a general overview but many of the issues are relevant to chemical engineering problems (e.g. maintaining computational tractability while giving probabilistic descriptions of uncertainty, modelling chance constraints, value of information).

The next section summarises some of the key aspects of the approaches discussed in the reviews listed above for management of uncertainty in design and/or operation of chemical processes.

4.3 Optimal design and/or operation under uncertainty methods

Grossmann et al. (1983) state the general form of the problem of design under uncertainty for minimum cost, C, as given by,

\[
\min_{d,z} C(d, z, x, \theta)
\]

\[
s.t. \quad h(d, z, x, \theta) = 0
\]

\[
g(d, z, x, \theta) \leq 0
\]

(4.1)

where d, z, x and \( \theta \) are vectors of design, control, state and uncertain parameters, and h and g are vectors of the equalities and inequalities. The problem may be stated as selecting d to minimise C while ensuring an optimum level of feasible operation with the manipulation of z. If only optimal process operation is considered and d is already fixed then the problem complexity is reduced since only decisions in z are made. Many of the methods which have been proposed to solve the type of problem shown in Equation 4.1 are associated with determining the limits of the feasible region in some manner. Other methods permit some infeasibility in the optimum process without explicitly defining the feasible region.

Under the parameter uncertainties the evaluation of the objective requires the multiple integration under the multi-variate probability distribution, as shown for the stochastic quantities defined in Section 3.3.2. Different methods have been used to approximate this multiple integral and solve the optimisation under uncertainty problem. The problem can be solved directly through the use of sampling based approaches such as Monte Carlo, Latin hyper-cube or Hammersley sampling (as discussed in Section 3.3.4.3) which explicitly discretise the uncertainty space. Alternatively, scenario based optimisation approaches can be
Integrated design under uncertainty for pharmaceutical processes

used based on implicit discretisation using numerical integration techniques (see Section 3.3.4.4). Sometimes the scenarios are expressed with pre-assigned probabilities of occurrence.

The choice of parameter uncertainty characterisation is a key assumption which strongly influences the way in which many methods for design under uncertainty have developed,

- parameter uncertainty within bounded ranges of values where the problem may be transformed to a deterministic one,

- knowledge or prediction of the uncertain parameter probability distributions and the expected value objective is optimised in a stochastic optimisation.

The discussion in the following sub-sections covers methods which include optimal design and/or operation under uncertainty which are either associated with:

(i) scenario-based methods which determine a feasible region under parameter uncertainty characterised only by bounded ranges,

(ii) stochastic methods which determine the feasible region under probabilistic characterisations of parameter uncertainty,

(iii) stochastic methods which permit partial feasibility without explicitly determining the limits of the feasible region (robust approaches).

Another sub-section discusses the concept of operational windows for integrated batch processes where ranges of control variables are determined for feasible operation.

4.3.1 Flexible plant design and operation

Flexibility analysis is a tool which has been commonly applied to design under uncertainty problems in chemical engineering in the past. Here, design specifically means the selection of the equipment design variables. Grossmann et al. (1983) discussed two types of generic problem:

- design for a fixed degree of flexibility, where the plant is designed for optimal economics while maintaining operational feasibility over a pre-specified range of parameter uncertainties,

- design for an optimal degree of flexibility, where a design is optimised for both economics and flexibility and the degree of flexibility is quantified using a flexibility index.

The former class of problems is also associated with the deterministic multi-period optimisation problem in which a plant is optimally designed to operate under a fixed set of sequential operating conditions for multi-product or multi-purpose type plants (e.g. oil refining, pharmaceuticals). The parameter uncertainty problem can be reduced to the general form of the deterministic multi-period problem with discretisation of the parameter uncertainty space to a set of alternate scenarios for which the control variables are independently adjustable.
The difference between design and control variables is implied from the common assumption that flexibility permits the (perfect) adjustment of control variables according to particular realisations of the uncertainty whereas the design variables remain fixed. This assumption corresponds to the ‘wait and see’ operating strategy. Conversely, the ‘here and now’ operating strategy treats the operating variables as equivalent to design variables which are not subject to adjustment at particular realisations. This assumption has important implications on the solution the design and operation optimisation problem. A two-stage programming formulation has been considered very effective in characterising optimum chemical plant design under uncertainty assuming perfect control manipulation. For a cost objective, $C$, the general two-stage problem is given as:

Design Stage,

$$\min_d E_{\theta \in R(d)} \{ C(d, \theta) \}$$  \hspace{1cm} (4.2)$$

Operating Stage,

$$C(d, \theta) = \min_z C(d, z, \theta)$$  \hspace{1cm} (4.3)$$

s.t. $f(d, z, \theta) \leq 0$

and the feasible region $R$ associated with design $d$ is given by,

$$R(d) = \{ \theta | \forall \theta \in R \exists z : f(d, z, \theta) \leq 0 \}$$  \hspace{1cm} (4.4)$$

The design is selected in the design stage and the operating stage aims to ensure design feasibility through control variable manipulation. This formulation poses a difficult problem and a variety of techniques have been proposed to solve it which are included in the discussion in the following sections.

4.3.2 **Bounded range uncertainty feasible region approaches**

The methods discussed in this section are associated with guaranteeing the existence of a hyper-cube (or rectangle) description of a feasible region inside bounded ranges of uncertain parameters. A summary of optimal design and/or operation methods based on this approach to uncertainty is given in Table 4.1.

Two-stage design and operation approaches to optimisation for the consideration of the dependence of the control on the design have been investigated extensively for the case of (deterministic) parameter uncertainties characterised by bounded ranges. Nishida *et al.* (1974) used a minimax strategy for process synthesis in which the best value of the objective function is determined at the worst parameter value in the bounded range. The continuous design parameters are chosen once the flowsheet structure parameters have been modified. The problem with the minimax approach is that the design is only optimal for the worst case set of parameter values which may not be very representative and may not be feasible at other parameter values.
# Integrated design under uncertainty for pharmaceutical processes

Table 4.1. Summary of work concerning chemical process design and operation defining a feasible region under bounded parameter (deterministic) uncertainty.

<table>
<thead>
<tr>
<th>Author</th>
<th>Key features</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grossmann and Sargent (1978)</td>
<td>Two-stage multi-period design stage and operating stage assuming vertex critical parameter points based on local gradients.</td>
</tr>
<tr>
<td>Halemane and Grossmann (1983)</td>
<td>Design for fixed degree of flexibility using two-stage multi-period design stage and operating stage feasibility test assuming vertex solutions for critical parameters.</td>
</tr>
<tr>
<td>Swaney and Grossmann (1985a and 1985b)</td>
<td>Deterministic flexibility index definition for design for optimal degree of flexibility, improved implicit vertex enumeration search.</td>
</tr>
<tr>
<td>Chacon-Mondragon and Himmelblau (1988)</td>
<td>New definition of flexibility index in the space of the control variables, exact solution for linear systems or approximate solutions for non-linear systems.</td>
</tr>
<tr>
<td>Pistikopoulos and Grossmann (1988a)</td>
<td>Optimal retrofit design for increased flexibility of linear systems with an analytical expression for the flexibility index.</td>
</tr>
<tr>
<td>Pistikopoulos and Grossmann (1989a, 1989b)</td>
<td>Optimal retrofit design for fixed degree and optimal degree of flexibility of non-linear systems.</td>
</tr>
<tr>
<td>Bansal <em>et al.</em> (2000a)</td>
<td>Integration of flexibility and controllability for synthesis of process structures and control systems.</td>
</tr>
<tr>
<td>Ma <em>et al.</em> (1999)</td>
<td>Optimal design for fixed degree of flexibility, using critical parameter scenarios based on joint confidence region vertices.</td>
</tr>
<tr>
<td>Rooney and Biegler (1999 and 2001)</td>
<td>Solution of deterministic flexibility analyses of linear systems with explicit dependence of the flexibility index on the design variables.</td>
</tr>
<tr>
<td>Bansal <em>et al.</em> (2000b)</td>
<td>Simplified reformulations of feasibility test and flexibility index problems for design optimisation.</td>
</tr>
</tbody>
</table>
Grossmann and Sargent (1978) introduced a two-stage formulation which aims to guarantee design feasibility (equivalent to an infinite penalty for infeasibility). The design and operating variables are selected by a multi-period optimisation of the expected performance comprising of a discrete set of parameter scenarios. This is subject to maximising each individual inequality constraint with respect to the uncertain parameters. This does not guarantee feasibility for all parameter values since the selection of the parameter scenarios in the constraint maximisations are based on local gradient signs from which the extreme (bounding) values are assumed. A stochastic programming approach of Malik and Hughes (1979) used a two stage approach for flexible design and operation based on Monte Carlo sampling. Halemane et al. (1983) claim that the disadvantages with this method is there is no guarantee that the design will be feasible under the parameter uncertainties and the MCS requires considerable computational effort.

A two-stage formulation which does guarantee design feasibility over bounded ranges of parameter uncertainties, from which many later approaches are based, was proposed by Halemane and Grossmann (1983). This formulation aims to solve the design under uncertainty problem for a fixed degree of flexibility, as discussed by Grossmann et al. (1983). A multi-period (or multi-stage) optimisation is solved in the design stage,

\[ \max_{d, z_m} \text{Pft} = f_0(d) + \sum_{m=1}^{M} w_f f_m(d, z_m, \theta_m) \]

\[ \text{s.t. } f(d, z_m, \theta_m) \leq 0 \quad \text{for } \forall m = 1, \ldots, M \]

\[ \theta \in T = \{ \theta | \theta^{LB} \leq \theta \leq \theta^{UB} \} \]

where \( \text{Pft} \) is a multi-period profit function, \( f_0 \) is a function for the fixed costs, index \( m \) represents a period or scenario and \( w_f \) is a weight factor corresponding to the discrete probability of each period. The original equalities, \( h \), and inequalities, \( g \), of Equation (4.1) are reformulated into a new vector of inequalities, \( f \), which express the implicit elimination of the state variables \( x \) from the problem. A key assumption is that the vector of uncertain parameters \( \theta \) are of the deterministic type, characterised by bounded values, \( \theta^{LB} \) and \( \theta^{UB} \), in a region \( T \), which contains all possible values of the parameters. Grossman and Sargent (1978), Varvarezos et al. (1992), Paules and Floudas (1992) and Subrahmanyam et al. (1994) have presented methods to solve multi-period design optimisation problems in applications for design feasibility under uncertainty, multi-period multi-product batch plants, distillation sequence synthesis and scheduling under uncertainty.

The optimal design, \( d^* \), obtained from the design stage, which is guaranteed to be feasible only at the specified periods, is subjected to a feasibility test which aims to select for every realisation of \( \theta \in T \) a vector of operating variables which is both optimal and feasible. The assumption is that with knowledge of the exact realisations of the parameters perfect control can be achieved. Halemane and Grossmann (1983) showed that the feasibility test is equivalent to the sub-problem,
Integrated design under uncertainty for pharmaceutical processes

\begin{equation}
\chi(d^*) = \max_{\theta \in \Theta} \min_{z} \max_{j \in J} f(d^*, z, \theta) \leq 0 \tag{4.6}
\end{equation}

where \( J \) is the number of reformulated inequality constraints. This constraint states that if for a given design, \( d^* \), a control, \( z \), can be selected to satisfy the critical parameter scenario, \( \theta^* \), which gives the maximum violation of the worst constraint (maximum valued constraint for given \( d, z, \theta \)), then \( \chi(d^*) \leq 0 \) for all \( \theta \in \Theta \) and the feasibility test is passed for design \( d^* \). This sub-problem allows the possibility of circumventing an infinite number of inequality constraints.

The problem is to select a finite number (\( M \)) of discrete periods for the multi-period design problem such that the assurance of operational feasibility at these points also ensures feasibility at all the other possible points which are not considered. Determination of the critical points (parameter scenarios which violate the constraints the most) using the feasibility test aims to find these periods. A limitation is that in the case where the uncertain parameters represent parameterisations of physico-chemical phenomena models the assumption that exact knowledge of the uncertain parameters at a particular time can be determined is unlikely to be valid even with on-line measurement of the dependent variables.

A range of methods have been proposed for the feasibility test problem. Halemane and Grossmann (1983), Swaney and Grossmann (1985a and 1985b) and Ostrovsky et al. (1994) suggested vertex search methods (explicit and implicit branch and bound enumeration techniques) assuming deterministic uncertain parameters for which the critical points are located at vertices of the hyper-rectangle. This may not be true for non-convex constraints. Swaney and Grossmann (1985a and 1985b) introduced a flexibility index in response to the design for optimal flexibility problem under deterministic uncertainty. They quantify this index, \( F \), as the maximum fractional deviation, \( \delta \), from the nominal parameter value, \( \theta^N \), in any of the uncertain parameter dimensions within which a hyper-rectangle feasible region can be inscribed,

\[ F = \max \delta \]

\[ \max_{\theta \in \Theta} \min_{z} \max_{j \in J} f(d^*, z, \theta) \leq 0 \]

\[ T(\delta) = \{ \theta | (\theta^N - \delta^- \Delta \theta^-) \leq \theta \leq (\theta^N + \delta^+ \Delta \theta^+) \} \]

\[ \emptyset \]

where

\[ \delta^- = \frac{\theta_p - \theta_p^N}{\Delta \theta_p^+}, \quad \delta^+ = \frac{\theta_p^N - \theta_p}{\Delta \theta_p^-} \]

\[ \text{for } \forall p = 1, \ldots, P \tag{4.8} \]

and \( \Delta \theta \) is the deviation from the upper or lower uncertainty bound to the nominal value. Similarly, the solution was based on vertex points and does not guarantee correct solutions in the presence of non-convex constraints which result in non-vertex critical points.
Integrated design under uncertainty for pharmaceutical processes

In an attempt to overcome this problem and avoid the enumeration of all vertices Grossmann and Floudas (1987) suggested an active constraint set strategy (ACS) based on the a priori identification of potential active constraints (from the stationary conditions) which limit flexibility. This does not rely on the assumption that critical points rely on vertices. The feasible region is projected into the space of the feasibility functions (functions of $d$ and $\theta$) of which one function is associated with each potential active set. They take advantage of the fact that the feasibility function is a piecewise continuous function and general mixed integer formulations are provided for both the design for fixed degree of flexibility and flexibility index problems. For the case of linear constraints the problem either reduces to a mixed integer linear programming problem (MILP) which may be solved by branch and bound enumeration approaches or to a series of linear programming problems (LP) using the ACS. For the case of non-linear constraints the ACS decomposes the mixed integer non-linear programming problem (MINLP) into non-linear programming (NLP) sub-problems but the possibility of a large number of active sets requires the solution of many NLPs. Global solutions can be guaranteed for non-linear constraints quasi-concave in the uncertain parameters. To improve the computational efficiency and extend the globality of solutions for different problem classes, Ostrovsky et al. (1994, 1997, 1999 and 2002) proposed different bounding methods for the solution the deterministic flexibility analysis chemical process design problems. Raspanti et al. (2000) present reformulations which simplify the feasibility test and flexibility index problem formulations of Halemane and Grossmann (1983), Swaney and Grossmann (1985a and 1985b) and Grossman and Floudas (1987). This is achieved by aggregating inequality constraints or using smoothing functions which simplify the optimality conditions.

Cabano (1987) identified that plant retrofit is inherently associated with uncertainty and flexibility. Pistikopoulo and Grossmann (1988a) formally considered the retrofit design of a linear plant for a specified degree of flexibility using ACS. They defined the retrofit problem as determining the parameter and/or structural changes which are required in an existing process so as to increase its flexibility with the least investment cost. They exploited the linear properties to provide an analytical expression for the deterministic flexibility index which allows the identification of all the critical parameter points and the inclusion of an explicit inequality constraint on the design flexibility. While this allows a more compact and efficient representation a disadvantage is that this approach still requires the a priori identification of all the active sets which despite systematic enumeration procedures can still involve the solution of a large number of MILPs. To overcome this problem Varvarezos et al. (1995) used sensitivity information (of the feasibility function) to identify all the critical points that limit flexibility (i.e. the non-redundant active sets) associated with a linear design. They used this information to evaluate the flexibility index for the optimal design of linear systems under fixed degree of flexibility without the enumeration of all the possible active sets. Pistikopoulos and Grossmann (1989a) extended their previous work to the retrofit design of non-linear systems and an efficient strategy is proposed for special classes (linear in the design variables and bi-linear in the uncertain parameters and control variables, i.e. where the state variables have been eliminated to give a reformulated set of linear inequalities). Pistikopoulos and Grossmann (1989b) extended their work on optimal retrofit under fixed degree of flexibility to optimum flexibility in non-
linear systems. Their approximation of the two-stage design and operating problem (Equation 4.1) involved solving the economic cost/profit optimisation and the design feasibility separately. An iterative parametric analysis is performed under two sub-problems. The first is a trade-off between the retrofit cost and the deterministic flexibility index and the second computes the expected revenue over the hyper-cube definition of the feasible region for each of the retrofit redesigns. The optimal degree of flexibility is not necessarily the largest value of flexibility but that which optimises the difference between the expected revenue and the retrofit cost. However, the expected revenue is likely to be under estimated since the hyper-cube is an over conservative definition.

Chacon-Mondragon and Himmelblau (1988) provided an alternative definition of the flexibility index. In this work the maximum hyper-rectangle is inscribed inside the feasible region defined in the space of the control variables and not the uncertain parameters. They claimed it is easier to compute than the index of Swaney and Grossmann (1985a and 1985b) but the exact measure can only be computed for systems linear in the control variables otherwise a linearisation approximation is required.

Soroush and Kravaris (1993) and Dimitriadis and Pistikopoulos (1995) have considered flexibility in the optimal design and operation of dynamic batch systems. The former work defined flexibility with the assumption that the controller is always able to force the design independent operating conditions to follow their nominal optimal profiles. This may be conservative as it does not account for any operating conditions which may be more appropriate under uncertainty than the nominal set. Dimitriadis and Pistikopoulos (1995) extended the feasibility test and flexibility analyses problems of Halemane and Grossmann (1983) and Swaney and Grossmann (1985a and 1985b) to application with dynamic systems under time varying deterministic uncertainties. The dynamic feasibility test is stated as the problem of being able to establish if there is at least one control variable profile which satisfies the feasibility constraints over the entire time horizon for every possible dynamic profile of the uncertainties. The dynamic flexibility index is the largest scaled deviation of the uncertain parameter profile that the design can tolerate while remaining feasible over the entire time horizon. They utilise dynamic optimisation techniques to solve the proposed two-stage formulations for these problems. One limitation is the large size of the optimisation problems for even small systems. The index problem also assumes that the direction in parameter space for the location of the critical point is known or is at one of the vertices in the dynamic hyper-rectangle.

The parametric programming approach provides a systematic method for the analysis of parameter uncertainty on the optimal solution in linear programming problems. The parametric solution is a function of the uncertain parameters and provides a map of the optimal decisions over the uncertainty space. Bansal et al. (2000b) solved both deterministic and stochastic flexibility analyses for linear systems using parametric programming. The use of parametric programming provides the explicit dependence of the deterministic feasibility test measure and the flexibility index on the design variables so reducing the size of the test and index problems (similar to the approach of Varvarezos et al. (1995), for a fixed degree of flexibility in linear systems). In the case of stochastic parameters the approach reduces the size of the analysis problems and allows the explicit expression of the system cost as a function of a target flexibility.
index and also the dependence of the flexibility on the design variables. Parametric programming has found more extensive application to MILP and MINLP process synthesis problems under uncertainty (Pertsinidis et al., 1998, Acevedo and Pistikopoulos, 1996 and 1997, Dua et al., 1999, Hene et al., 2002).

One of the underlying assumptions in the two-stage design and operation methods for optimal design for fixed degree of flexibility is that the parameter uncertainties are characterised by bounded ranges or individual confidence intervals. This may lead to overly conservative estimates of the influence the uncertainties have on the design since the resulting hyper-rectangle may be a poor estimate of the actual parameter uncertainty space. Rooney and Biegler (1999) recognised this problem with particular concern regarding correlated model parameters. To overcome this they incorporated elliptical joint confidence region information into the two stage multi-period design formulation. They selected the two extremums along the longest axis of the joint confidence region of the parameters as the initial periods for the multi-period design optimisation. The feasibility test is then solved (using the active constraint strategy) to find additional critical scenarios to add to the design stage. The design optimisations contained far fewer periods than otherwise obtained in the hyper-rectangle approach, and converged more rapidly. Limitations are the requirement of the parameter covariance matrix usually from the regression of experimental data, the reliability on the non-global feasibility test and the computational expense in determining the confidence region extremums for large problems. To determine critical scenarios from confidence regions not well approximated by hyper-ellipsoids (due to highly non-linear models), Rooney and Biegler (2001) modified their prior approach by using a likelihood ratio test to derive non-linear parameter confidence regions.

Model parameter uncertainty has also received attention in the design and control of chemical plants and process operation. Bahri et al. (1997) discussed a systematic approach to consider controllability and flexibility of a plant. Flexibility is introduced into the plant using the concept of back-off as a measure of flexibility (ensuring feasible operation by moving the optimum point to a point inside the feasible region where the worst disturbance set will not cause constraint violation). In a two-stage dynamic optimisation approach, the outer loop computes the best flowsheet design and/or control structures with the combination of disturbances which gives the most constraint violation (determined from the inner loop). Mohideen et al. (1996) introduced a framework for optimal integrated process design and control system design under parametric uncertainty and process disturbances. This aimed to provide an optimum design and control scheme at minimum annualised cost over the entire time horizon under the specified uncertainty. A multi-period design sub-problem is solved for the process and control structure and design to which critical scenarios are added iteratively following the solution of a dynamic feasibility test. The resulting MINLP problems were very large for even relatively small scale systems and were limited to the consideration of multi-loop proportional integral (PI) controllers and continuous decisions. Bansal et al. (2000a) provided a more efficient approach for the solution of a similar type of problem but for discrete and continuous decisions. It does not guarantee global solutions. Iterative addition of critical scenarios in a two-stage multi-period problem is also used by Kuhlmann et al. (1998) for the optimal control of fed-batch fermenters. They determine the critical scenario inside bounded ranges of parameter uncertainty.
from a constraint maximising sub-problem and amend the current set of scenarios in the optimal control problem used to determine the robust control profile providing feasible operation under the uncertainty.

Ma et al. (1999) provided an alternative approach to quantify the effect of uncertainty in model parameter and control implementation on the performance of optimal open-loop control policies for batch chemical processes. Instead of attempting to estimate the integral quantities denoted in Equations 3.1 to 3.10, they used an analytical method to measure the sensitivity of both the optimal control policy and the performance objective deviations from their nominal values at an estimate of the worst case uncertain parameter scenario.

4.3.3 Probabilistic uncertainty feasible region methods

The methods discussed in this section are associated with the attempt to determine the feasible region more accurately than the hyper-rectangle definition. These methods are typically associated with probabilistic definitions of the uncertainty and are summarised in Table 4.2. Stochastic or probabilistic uncertainty characterisation incorporates a greater level of knowledge regarding the uncertainty (through

<table>
<thead>
<tr>
<th>Author</th>
<th>Key features</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kubic and Stein (1988)</td>
<td>Design reliability (probability of feasible design operation) assuming no control variables.</td>
</tr>
<tr>
<td>Pistikopoulos and Grossmann (1988b)</td>
<td>Retrofit design for optimal degree of flexibility in linear systems.</td>
</tr>
<tr>
<td>Straub and Grossmann (1990)</td>
<td>Stochastic flexibility index for flexible and reliable linear systems subject to combined discrete and continuous parameter uncertainties.</td>
</tr>
<tr>
<td>Terwiesch et al. (1998)</td>
<td>Optimisation of probabilistic success measures for robust dynamic semi-batch process operating policies.</td>
</tr>
<tr>
<td>Bansal et al. (2000b)</td>
<td>Solution of stochastic flexibility analyses of linear systems with explicit dependence of the flexibility index on the design variables.</td>
</tr>
</tbody>
</table>
probability distribution functions) than is posed in bounded range deterministic uncertainty approaches.

As mentioned in Section 4.3.1, stochastic parameters are usually associated with the design objective for optimal degree of flexibility, where economic trade-offs with flexibility (quantified through a stochastic flexibility index) are explored. Flexibility is associated with the size of the feasible region and accurate determination of the feasible region is a key element of many approaches.

Stochastic flexibility analyses are concerned with the probability of feasible operation of a process subject to uncertainties which are described by probability distributions and for which perfect control is assumed. Kubic and Stein (1988) used a stochastic flexibility measure based on the integration of the probability distribution function of the uncertainty over the feasible region. They called this design reliability, i.e. the likelihood that a design will operate, and was based on random and fuzzy uncertainties. The concept of fuzzy set theory was used for non-random parameter uncertainties for which distinct bounds are not known. Membership functions define the fuzzy set used to describe the possible realisations of the uncertain parameters. However, the explicit consideration of degrees of freedom (decisions) was not possible in their work. Pistikopoulos and Grossmann extended their work on optimal retrofit under fixed degree of flexibility (bounded ranges of uncertainty) for linear systems (1988a) to optimum flexibility in linear systems under stochastic uncertainty (1988b). A two-phase retrofit cost-flexibility-expected revenue trade-off approach as used for non-linear systems (1989b) and previously described in Section 4.3.2, is implemented with the difference being that the stochastic flexibility index (1988b) provides a more accurate estimation of the feasible region than the hyper-cube index definition (1989b).

Pistikopoulos and Mazzuchi (1990) introduced a systematic approach for the evaluation of a flexibility index for process systems under normally distributed uncertainties. The advantages are that the transformation of the feasible region into the space of only the uncertain parameters explicitly accounts for the perfect adjustment of the degrees of freedom and that large numbers of correlated uncertain parameters can be handled. However, their approach is restricted to linear systems and normal parameter distributions. Straub and Grossmann (1990) introduced a stochastic flexibility index (probability of feasible operation) for linear systems subject to combined discrete and continuous parameter uncertainties. The work combines flexibility with reliability into an expected stochastic flexibility which is the sum of products of the probability for each discrete state and the associated stochastic flexibility under the continuous uncertainties (which may be correlated and characterised by a range of distributions). The application to linear models allows an effective analytical scheme for the quadrature integration of the distribution function. It is limited to a modest number of constraints and it is not easily extended to design optimisation. Straub and Grossmann (1992) similarly combine flexibility and reliability for the evaluation and optimisation of expected flexibility in multi-product batch plants under normally distributed uncertainty in demand and discrete uncertainty in equipment availability. They show how parametric optimisation can be used to trade-off between optimal flexibility or expected profit against cost. The assumption of normal distributions and the fact that the time horizon constraint is linear in demand uncertainty allows simplified problem formulations which would otherwise not be possible.
The extension of stochastic flexibility analysis to non-linear models was made by Straub and Grossmann (1993). They presented an approach to evaluate the stochastic flexibility of a non-linear feasible region under perfect control which can be extended to design optimisation. The stochastic flexibility index, $SF$, for a non-linear feasible region in a total of $P$ uncertain parameter dimensions is defined as,

$$SF = \int \int \int \ldots \int PDF(\theta) d\theta_1 d\theta_2 d\theta_3 \ldots d\theta_P$$

(4.9)

where $PDF$ is the truncated joint probability distribution function of the uncertain parameters (which may be correlated), $\theta$, and $\theta^{UB}$ and $\theta^{LB}$ are the upper and lower bounds of the feasible region. As proposed by Straub and Grossmann (1990) gaussian quadrature is used to accurately approximate the multi-dimensional integral by determining the bounds of the feasible region in each dimension in turn and then locating the collocation points inside. The determination of these bounds assumes perfect control and a convex feasible region. A non-convex region will be incorrectly defined unless the region is one-dimensional convex in each uncertain parameter. As discussed in Section 3.3.4.4, any approach which uses gaussian quadrature to approximate integrals is limited by the number of uncertain parameters which can be reasonably handled given the exponential increase in the size of the problem with the number of dimensions. The extension to design optimisation involves iterating between a design (master problem) and an operating stage (sub-problem) by applying Benders decomposition. The authors showed that the $SF$ metric is superior to Taguchi’s quadratic quality loss metric (defined in Section 3.3.2) since the latter designs may produce designs which ignore the effect of hard inequality constraints, whereas the $SF$ produces feasible designs but which may exhibit larger quadratic loss.

Pistikopoulos and Ierapetritou (1995) guarantee design feasibility in their simultaneous design feasibility and economic optimisation approach for convex non-linear systems under stochastic parameter uncertainty. This does not require a priori discretisation of the uncertainty and implements the sequential quadrature technique of Straub and Grossmann (1993) to determine and integrate over the feasible region. The two-stage Benders decomposition algorithm used in this work has also been used for the stochastic optimisation of process planning problems (Ierapetritou and Pistikopoulos, 1994). The approach is restricted by the number of uncertain parameters permitted due to the exponential increase in the number of operational feasibility sub-problems. In addition, it may not find the global optimum for non-convex problems. Ahmed et al. (2000) avoid the problem of entrapment in local solutions and the solution of the feasibility test sub-problems in a reformulated algorithm for optimal planning under uncertainty.

Bansal et al. (1998) applied stochastic flexibility analysis to linear dynamic systems under normally distributed uncertain parameters. They presented a procedure for the evaluation of stochastic flexibility over time and a framework for design optimisation which are both analytical. A single stage design problem is posed to determine the optimal design that meets a desired stochastic flexibility target over the entire time horizon. While Bansal et al. (1998) determined the dynamic feasible region throughout the process time for linear systems Terwiesch et al. (1998) determined the feasible region for non-linear batch
processes in terms of the desired endpoint performance criteria for an optimal robust operating policy. They recognised the importance of utilising robust objectives and their approach permitted a variety of risk objectives and inequality constraints to be expressed (maximum probability of feasible product, threshold level of risk, expected value, variance). To evaluate the probabilistic measure of success they applied the gaussian quadrature integration approximation of Straub and Grossmann (1993) to determine the size of the feasible region for the dynamic system under constant (correlated) stochastic parameter uncertainties without assuming perfect control. Results from tightly controlled experiments were shown to qualitatively confirm the improved success rate with an operating policy determined by the probabilistic method over the nominal and heuristical operating policies. The limitation is associated with the size of the problem depending on the number of uncertain parameters due to reliance on gaussian quadrature. Only simple operating policies can be optimised with their algorithm but this is stated to be no great drawback given the industrial techniques in batch process operation.

In the consideration of both deterministic and stochastic parameter uncertainties the vector of uncertain parameters in the general problem (Equation 4.1) is partitioned into two subsets of deterministic and stochastic parameter uncertainties. This results in a combined multi-period and stochastic programming problem to which Pistikopoulos (1995) and Ierapetritou et al. (1996) provided a general formulation. The design selected in the design stage is passed to the operating stage which aims to identify an optimal vector of control variables for all realisations of the stochastic uncertainties and for every period in the deterministic uncertainties. This can result in a particularly large optimisation problem.

4.3.4 Robust partial feasibility approaches

In this section methods are discussed for which the solutions do not guarantee feasibility over a defined feasible region but rather permit partial feasibility by evaluating the entire uncertainty space. Robust approaches are often based on the concept of quality management and robust design and operation, introduced by Taguchi (1986). In this theory every departure of the measurable performance quality from the nominal value causes some economic loss according to a quadratic penalty function (see Equation 3.6, Section 3.3.2).

A stochastic constraint is defined as a constraint on a stochastic variable. The constraint may be treated as either hard (either it cannot be violated or is violated but with severe consequences) or soft (for which the decision maker permits some violation at the expense of a finite penalty). Robust approaches allow for partial feasibility of processes in certain inequality constraints under uncertainty by applying soft stochastic inequality constraints to allow for a limited degree of failure in certain variables. Such approaches differ from flexibility analyses which treat all constraints as hard and attempt to guarantee feasibility within the uncertainty space assuming perfect control in the face of the uncertainties. In this respect robust approaches tend to lead to less conservative solutions than flexibility analysis approaches but at the expense of some level of failure. On the other hand, perfect 'wait and see' control is not always assumed. A summary of literature concerning robust approaches is given in Table 4.3.
Table 4.3. Summary of work concerning robust (partial feasibility) chemical process design and operation.

<table>
<thead>
<tr>
<th>Author</th>
<th>Key features</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maranas (1997)</td>
<td></td>
</tr>
<tr>
<td>Petkov and Maranas (1997, 1998)</td>
<td>Unconstrained process optimisation where the probabilistic constraint violations are expressed as penalty functions in the objective.</td>
</tr>
<tr>
<td>Weisman and Holzman (1972)</td>
<td>Two-stage plant design with finite penalty functions for infeasibility in the operating stage.</td>
</tr>
<tr>
<td>Pai and Hughes (1987)</td>
<td></td>
</tr>
<tr>
<td>Ruppen et al. (1995)</td>
<td>Optimal control profile formulation over the entire space of stochastic parametric uncertainty, with stochastic performance constraints.</td>
</tr>
<tr>
<td>Kalagnanam and Diwekar (1997)</td>
<td></td>
</tr>
<tr>
<td>Bhatia and Biegler (1997)</td>
<td>Feedback correction policy for dynamic processes in simultaneous batch design and scheduling under process model uncertainty.</td>
</tr>
<tr>
<td>Uesbeck et al. (1998)</td>
<td></td>
</tr>
<tr>
<td>Bernardo and Saraiva (1998)</td>
<td>Robust design and operating policy optimisation with uncertainty in model parameters and operating policy implementation.</td>
</tr>
<tr>
<td>Darlington et al. (1999)</td>
<td>Mean-variance objective optimisation problems for non-linear systems with soft stochastic constraints.</td>
</tr>
<tr>
<td>Michel et al. (1999)</td>
<td>A robust design approach which analyses the propagation of Taguchi's quality loss function through the process unit sequence.</td>
</tr>
<tr>
<td>Bernardo et al. (1999a and 2001)</td>
<td>Incorporation of Taguchi's penalty functions into a single level design and operating policy stochastic optimisation.</td>
</tr>
<tr>
<td>Bernardo et al. (2000)</td>
<td>Optimal reduction of model parameter uncertainties for economic objective, in a two-stage design formulation.</td>
</tr>
</tbody>
</table>

Early optimisation under uncertainty approaches treated design and operating variables equivalently. The stochastic approach of Weisman and Holzman (1972) optimised an unconstrained problem where the expected value of the cost objective function incorporated penalty functions for the probability of violation for individual constraints. A drawback is that this formulation does not rigorously ensure a lower limit on the probability of constraint failure. The stochastic chance-constrained programming approach of Charnes and Cooper (1959, 1962, 1963) does guarantee a lower limit on the probability of failure in
inequality constraints in an alternative approach to the use of penalty functions. A chance constraint is of the following general type,

$$\Pr[g(x, \theta) \leq 0] \geq \alpha$$  \hspace{1cm} (4.10)

where the probability, \(\Pr\), that the uncertain constraint, \(g\), which is a function of the uncertain parameters, \(\theta\), and the variables, \(x\) (which may include the decision variables), is satisfied by some probability, \(\alpha\). Chance-constrained optimisation is one method of stochastic programming which reconciles the optimisation over uncertain constraints. If the probability distribution function of \(\theta\) is known and is stable then the probabilistic constraint given in Equation 4.10 can be reformulated as a deterministic equivalent (i.e. one which is not a function of \(\theta\)). The form of the original constraint \((g)\) defines the difficulty in reformulating a deterministic equivalent. Chance constraints linear in the uncertain parameters can be readily reformulated. If the uncertain parameters follow stable two parameter probability distributions (such as the normal distribution), then \(g\) is rearranged into the standard normal form by subtracting the mean and dividing by the square root of the variance of \(g(x, \theta)\). Since \(\theta\) follows a stable distribution then the chance-constraint is replaced by applying the standardised normal cumulative distribution function \((\text{SNCDF})\) to the standardised \(g\).

$$\text{SNCDF}\left(\frac{0 - \mu[g(x, \theta)]}{\sqrt{\text{Var}[g(x, \theta)]}}\right) \geq \alpha$$  \hspace{1cm} (4.11)

Inverting the SNCDF and rearranging leads to the deterministic equivalent,

$$\mu[g(x, \theta)] + \text{SNCDF}^{-1}\alpha\left(\sqrt{\text{Var}[g(x, \theta)]}\right) \leq 0$$  \hspace{1cm} (4.12)

where the mean and variance terms of \(g\) are represented using the mean \((\mu_\theta)\), variance \((E[\theta-\mu_\theta]^2)\) and covariance \((E[\theta_1-\mu_\theta][\theta_2-\mu_\theta])\) terms for \(\theta\) combined with explicit evaluations of \(g\) as a function of \(x\) only (i.e. \(g(x)\)).

The advantage of chance-constrained programming is that this results in a convex programming problem and avoids the multi-variate integration which defines the problem size in other stochastic programming approaches. Chance-constrained programming is limited by the requirement that the uncertain parameters are linear in the uncertain constraints and that their probability distribution functions are stable. However, it does allow the incorporation of a high number of correlated parameters. In chemical engineering applications this approach has been implemented by Maranas (1997) for the optimal molecular design and Petkov and Maranas (1997 and 1998) for the planning and scheduling of multi-product batch plants under demand uncertainty.
The two-stage design and optimisation approach of Pai and Hughes (1987) under deterministic parameter uncertainty acknowledged that it is not always desirable to always guarantee feasibility. Permanent feasibility is not required but rather a finite cost penalty is introduced into the operating stage objective function. This penalises the realisations of uncertainties in designs for which no feasible solution of the operating stage is apparent.

Ruppen et al. (1995) discuss the computational aspects of dynamic optimisation of ‘here and now’ batch reactor operating policies under parameter uncertainty. They formulate the problem by discretising the entire uncertainty space to create a deterministic multi-scenario model equivalent of the stochastic problem. An objective function of an expectation operator of a deterministic performance quality and stochastic inequality constraints are applied in the multi-scenario optimisation.

Instead of creating a deterministic equivalent to the stochastic model, Diwekar and Kalagnanam (1996, 1997) and Kalagnanam and Diwekar (1997) used a stochastic optimisation algorithm to determine robust ‘here and now’ design and operating policies under uncertainty in model parameters and the operating variables. As discussed in Section 3.3.4.3 they introduced a Hammersley number sequence based sampling strategy (HSS) to efficiently sample the entire uncertainty space. Direct solution of the stochastic model places the sampling sub-routine inside the optimisation loop and scenarios are placed until some convergence criterion is met. Bernardo and Saraiva (1998) also used stochastic optimisation based on HSS for the simultaneous robust design and operating policy under uncertainty in model parameters and operating policy implementation. The operating policy decisions are ‘here and now’ in the parameter uncertainty space (single stage optimisation problem) but are themselves subject to error. These errors are expressed as operating regions characterised by the relative error (standard deviation divided by the mean). The economic objective penalises narrow operating regions as control costs and explores interactions between quality robustness (Taguchi loss function) design, operating and control costs.

Uesbeck et al. (1998) introduced a robust multi-scenario dynamic optimisation formulation (for fermentation processes) which provides an operating policy which attempts to maintain good performance in the nominal (no uncertainty) case while constraining product quality to a high extent of satisfaction under the uncertainties. They discretised the parameter space a priori (to the implementation of the optimisation algorithm) to provide a deterministic equivalent to the stochastic model. A one-sided maximum expected violation soft constraint is imposed by introducing constraint violation variables which are properly expressed using a switching technique. This comprised of one additional degree of freedom (a deviation variable) and two additional inequality constraints (one determining the deviation variable and the other forcing the deviation variable to be either greater or smaller than zero) per scenario. The drawback is that the size of the model is increased as a result. Samsatli et al. (1998) solved a similar problem in a more comprehensive formulation. They provided a general function for robustness metrics applicable to any one-sided or conventional two-sided stochastic constraints. Instead of using additional constraints to handle one-sided deviation quantities (Bhatia and Biegler, 1997, Uesbeck et al. 1998) they used a hyperbolic smoothing function to approximate binary variables. Gaussian quadrature is used to evaluate the expectation form of the integral under the joint parameter PDF and the operating policy can
be either ‘here and now’ (scenario independent) or ‘wait and see’ at each collocation point. As mentioned in Section 3.3, the number of points that quadrature schemes require, grows exponentially with the number of uncertain parameters. They state that despite the non-convexity of the smoothing functions most industrially relevant dynamic models are highly non-convex such that entrapment in local optimum solutions would be a problem anyway.

A multi-scenario robust optimisation approach incorporating Taguchi quadratic loss penalty function and robustness metrics (mean, variance, quantiles) was introduced by Bernardo et al. (1999a and 2001). They transformed the two-stage design and operating policy optimisation into a single level optimisation using cubature integration to place parameter scenarios and allowing the operating policy (but not the design variables) to vary with each scenario. They used the same approach as Uesbeck et al. (1998) to account for asymmetric loss functions. The efficiency of the cubature integration under the entire uncertainty space may be limiting for problems with many uncertainties and requires they all be normally distributed.

As an alternative to multi-scenario methods to approximate the uncertainty space, Darlington et al. (1999) used Taylor series expansion to approximate mean-variance objectives in their optimisation approach for robust process operation. Penalty functions are used to allow partial feasibility in non-linear stochastic constraints. Michel et al. (1999) optimised the design and operation of process flowsheets using factorial designs in the uncertain parameters and the decisions. They decomposed the flowsheet into sub-systems. The system is optimised minimising the deviations in each sub-system output quality variable (interconnecting stream variable) from its nominal value via a Taguchi quadratic loss function. A measure of the relative contribution of the quality loss penalty associated with each sub-system to the final product penalty is provided.

Bernardo et al. (2000) differentiated between inherent system variabilities and model parameter uncertainties. They incorporated information R&D costs for the optimal reduction of key model parameter uncertainties into a two-stage process robust design optimisation problem under uncertainty. The uncertainty space is redefined with each optimisation iteration of new decisions concerning the magnitude of the standard deviations in the reducible and significant parameter uncertainties. A strategy allows for the fact that the true mean of the reduced uncertainties may not be known. In addition to the pros and cons of using the cubature integration technique, it could be difficult to provide a realistic R&D cost to levels of uncertainty reduction. Pinto (1998 and 2001) also considered the cost of parameter uncertainties assuming linear Taylor series expansion of uncertainty around the nominal parameters and the nominal optimal decision variables. An economic value was assigned to parameter uncertainties and used for process optimisation. It represents the total average cost which may be expected for not knowing actual optimum operating conditions due to the uncertainties. Sensitivity Analysis is used to identify which parameter uncertainties strongly influence the objective. This information is used to support decisions in sequential experimental design, providing the objective can be well defined. The economic cost function is proposed as the experimental design criterion used to determine experimental design procedures based on an expectation of how experimentally regressed parameter variances change when additional data is added to the prior set (Bard, 1974).
Some approaches take into account the fact that process knowledge becomes available during operation through measurements and can be used to reduce the effect of the uncertainty in the remaining duration of the process. Solutions based on operating policies implemented with feedback control determined under model uncertainty provide a more realistic compromise between the optimistic ‘wait and see’ solutions and the conservative ‘here and now’ solutions, discussed in the introduction to this chapter (Section 4.1). Terwiesch et al. (1994) discussed on-line feedback control optimisation methods under uncertainty based on knowledge of current states. They state that the lack of on-line measurements limits practical implementation in industrial practice. Dynamic optimisation using feedback correction policies under model parameter uncertainty was implemented by Bhatia and Biegler (1997) for simultaneous operating policy and planning of batch plants. They assumed a small number of a priori known instances of parameter uncertainties in a multi-period formulation, with the acknowledgement of the computational problem with increasing scenarios. Pinto (2001) incorporated the evolutionary operation procedures (EVOP) to optimise a continuous plant during operation. This rectifies the over-estimation of the cost of parameter uncertainty by evaluating the partial recovery of the losses which were introduced during the design phase due to uncertainties.

While most of the discussed approaches only considered static uncertainties, Abel and Marquardt (2000) incorporate time varying uncertainties into the robust optimisation of operating policies for dynamic hybrid (discrete-continuous) systems. Optimisation approaches are formulated for uncertainty in the model parameters but also in the process model structure (e.g. implementation of safety systems) resulting in either single-level or bi-level scenario-integrated optimisation problems. These provide solutions comprising of different control profiles for the various switching times (at which point the model structure and uncertainties may change). Limitations are the complexity of the problem and the associated application of current numerical solution techniques.

4.3.5 Operational windows

Although the concept of operational windows is not an approach which necessarily directly implements uncertainty knowledge, they are particularly useful for batch processes since they recognise that in real life these are never run only at one set of operating conditions and some flexibility should be allowed. The idea is to combine the knowledge contained in process models with operational windows in the control variables so that any co-ordinate of control values inside the window results in feasible operation in meeting the constraints and any desired product qualities. This is similar to the notion of feasibility under uncertainty except here the feasible region is expressed in the space of the control parameters and uncertainties are not considered.

Woodley and Titchener-Hooker (1996) used successive model simulations to construct (non-linear) feasible operational windows in multi-step biochemical engineering processes. In addition, they built trade-off windows for varying constraints on the process criteria. They used these to indicate the optimal balance between operational flexibility and process quality while accounting for interactions between
process stages. Zhou et al. (1997) attempted to formulate operational windows through the correct selection of control variables and understanding the dependence of the process constraints on them. They visualised complex bioprocess problems by plotting windows at different design stages, from conceptual design (based on heuristics) to detailed process design (where specific correlations are known), and used them as a tool for Sensitivity Analysis to assess interactions and changes in operating variables. Samsatli et al. (1999 and 2001) adapted the multi-period method of Halemane and Grossmann (1983) to maximise the volume of a hyper-rectangle characterisation of the feasible control space. Critical scenarios in the control variable space are iteratively added to the problem but one difference is that this volume is not centred on a nominal point since this may already be on the edge of the feasible region. They applied this to sequenced operations so that trade-offs between control variables in different stages can be made. As previously discussed concerning robust approaches, Bernardo and Saraiva (1998) solved process design and control optimisation problems combining parameter uncertainty with windows of control variable tolerance to account for errors in implementation. They optimised the size of the control tolerance windows by attaching a cost for narrow windows.

4.4 Discussion

The robust approaches which explore the entire space of the uncertainties are considered the most applicable to the type of problem posed in this thesis. The robust approaches discussed in Section 4.3.4 have been successfully applied to one or two step batch process examples though not to complete integrated batch/continuous process sequences. While the guarantee of a defined feasible region is an important issue in some conceptual problems it is also important to contemplate the effect that regions of non-feasibility may have on the process as considered in the robust approaches. The assumption of fixed plant equipment design bypasses the need for the design and operating two-stage optimisation strategies employed in many of the feasible region approaches. An approach which can provide optimal decisions under all the uncertainty while capturing the dynamic/continuous and non-linear effects modelled in the system of integrated processes appears to be most satisfied by the robust approaches discussed.

It is desirable that a general optimisation framework has the capacity to provide optimal decisions for a complete pharmaceutical process accounting for the following properties:

- capacity to implement a range of different uncertainty distribution models,
- capacity to implement a range of different deterministic process models in an integrated sequence,
- capacity to implement a range of different robustness metrics/stochastic inequality constraints,
- capacity to implement a number of uncertain input factors (reaching the order of tens),
- capacity to implement different conceptual optimisation opportunities (decision variable application to operating policies, operating windows/tolerance, parameter uncertainty reductions).
The first four issues have already been portrayed in the context of Uncertainty Analysis (see Section 3.6) and remain relevant in the extension to optimisation under uncertainty problems. The last issue is important for a generic approach. A robust stochastic optimisation approach based on the direct solution of the general optimisation problem under uncertainty using an efficient sampling based approach is one method which can allow the provision for the issues posed above which limits the eventuality of excessive problem sizes and computational requirements. The capacity to directly implement the knowledge structures provided in the Uncertainty Analysis allows the decisions to be determined by the optimisation algorithm without the enforced loss of any of the information contained. It is proposed that such an approach within a general optimisation under uncertainty framework will produce solutions which are meaningful in the context of pharmaceutical systems.

4.5 Conclusions

In this chapter, the use of optimisation under uncertainty methods to determine better solutions in the desired decision space of deterministic process models subject to parameter and external process uncertainty, is postulated as a management response tool to the specified uncertainty. A range of methods under parameter uncertainty are discussed. These included formulations constructed to determine equipment design assuming perfect control, not assuming perfect control and those only concerning operating policies. The discussion was split between methods which determine hyper-rectangle feasible regions in bounded ranges of uncertain parameter space, those which determine feasible regions in uncertainty space characterised by stochastic probability distributions, robust approaches which consider the entire uncertainty space and assess partial feasibility, and windows of operation which consider feasible regions in the control variable space (without specifically accounting for model uncertainty).

A robust approach was selected as being the most suitable for the type of problem considered with the assumption that the equipment design is fixed given the allocation of pre-existing plant equipment and emphasis is placed on determining robust operating policies. A sampling based approach (using an efficient sampling technique) for the direct solution of stochastic optimisation problems is considered for the following key reasons: it is able to capture the non-linear and dynamic behaviours of integrated dynamic/continuous process sequences, permits the incorporation of a reasonably large number of uncertain factors (more than ten) into the optimisation without exponentially increasing the problem size and provides a framework which is flexible to the implementation of different robustness metrics and conceptual optimisation objectives.

The next chapter defines both the problem considered in this thesis for integrated pharmaceutical process design under uncertainty and the methodology proposed to solve it. The associated mathematical techniques implemented in the methodology are described in detail.
CHAPTER 5

AN APPROACH FOR INTEGRATED DESIGN UNDER UNCERTAINTY

5.1 Introduction

Various methods to deal with uncertainty in models and in the optimisation of uncertain process models have been discussed in Chapter 3 and Chapter 4, respectively. The range of optimisation methods under uncertainty, discussed in Chapter 4, originated from deterministic and stochastic programming problem viewpoints. As would be expected, the interpretation of the problem under uncertainty is important in defining the context to which the results may be useful in supporting process design and development decisions.

In Section 5.2 the conceptual problem is posed and objectives of this work are clarified in Section 5.3. In Section 5.4 an approach is presented where Risk Analysis methods associated with a stochastic representation of the uncertain process system are integrated within systematic model development procedures. This approach aims to quantify the process response to uncertainty in the current understanding of the system resulting from incomplete process knowledge characterised by model parameter uncertainties. In doing so, this aims to focus effort towards those parts of the system which have the greatest effect on the response. In Section 5.5 four stochastic optimisation problems are formulated to address the problems of determining (1) the required stochastic input parameter uncertainty reductions to meet desired levels of uncertainty in the response, (2) optimal operating policy decisions for the uncertain system, (3) the value of perfect information regarding potential certainty in measurable properties and (4) permitted tolerances to errors or uncertainty in the implementation of operating policies.

5.2 Conceptual problem for uncertainty in models

The objective of this work is to introduce a systematic model-based approach to support the process development of pharmaceutical processes, with the specific aim of accounting for the large amounts of uncertainty in the process knowledge. Given a conceptual process design, a lack of understanding in the underlying physico-chemical mechanisms is transcribed to uncertainty in the imperfect models used to describe the operations of the integrated process sequence. As mentioned in Section 3.2, this work does not aim to indicate what a better model structure should be but rather where parameter uncertainties in a current set of models appear to be relatively important. Pinto (1998) suggests that the presence of significant structural inadequacies might be inferred from these uncertainties. An important distinction becomes apparent when independent data is available for model validation purposes. Both the location of the distribution of a predicted output performance variable relative to the data point(s) and its spread are important aspects. The quantification of the spread and sensitivity to the spread are the main
considerations of the model-based Risk Analyses proposed in this thesis. However, the model structure may be more relevant than parameter uncertainty to the relative location of the distribution to the data.

5.3 Integrated design objectives

This work aims to introduce some form of management of the uncertainty associated with the model representations of the current process knowledge. In the face of large amounts of uncertainty predicted in the important process output criteria, three management responses are considered:

(i) reduction of the uncertainty by improving current models/parameter estimations associated with the key contributing uncertainty factors identified,

(ii) manipulation of the available process decisions (operating policy) to improve process robustness to model uncertainty,

(iii) consideration of process alternatives.

Response (i) concerns the gathering of additional information for systematic model development for more reliable models. Response (ii) and (iii) concern the optimisation and comparison of uncertain but complete processes sequences. In order to formulate a framework to allow the implementation of the above management responses, methods are provided which couple the use of a model-based approach under uncertainty with the vision of complete integrated sequences of pharmaceutical processes. The following issues are considered regarding the explicit consideration of uncertainty in systematic model development:

- quantification of the effect of model uncertainty on the important output predictions,
- identification of the key process sub-sequences and key contributors to the observed effects under model uncertainty,
- identification of required reductions in key input uncertainties to meet desired limits on output prediction uncertainty,
- the capacity to incorporate new information.

The issues addressed for the optimisation of integrated process flowsheets under uncertainty are:

- quantification of the effect of model uncertainty on the objective function and constraints,
- optimisation of uncertain process performance criteria with manipulation of operating policy variables,
- the value of potential measurements on external process inherent uncertainty sources,
- the maximum process operating policy variable error tolerance permitted to meet desired limits on the uncertainty in output performance predictions.

A schematic linking the different parts of the methodology used to address these issues is shown in Figure 5.1. The individual aspects are examined in more detail in the following sections of this chapter.
5.3.1 Assumptions

The assumptions associated with the problem to which the objectives of this methodology are directed are reiterated:

- **status of conceptual pharmaceutical process**
  - conceptual process (including feed materials) is already specified,
  - equipment (of fixed design) is already allocated,

- **deterministic process model**
  - integrated sequence of individual process models (representing the unit operations specified in the conceptual process),
  - linear and/or non-linear models,
  - time-dependent and/or steady-state models,
  - lumped parameter and not distributed systems,

- **stochastic system**
  - uncertainty in the parameters of the deterministic process model system and no specific consideration of natural variability or structural uncertainty,
- model inherent parameter uncertainty (including kinetic constants, transfer coefficients, thermodynamic and physical property parameters),
- process inherent parameter uncertainty (measurable properties of external feed streams),
- parameter uncertainties are static in time unless associated with an external action,
- model parameter uncertainties may be characterised by uniform, normal or other types of stable probability distributions (stability is assumed since parameter distributions are either regressed or assumed due to a lack of data and are not directly estimated from distributions of raw data points),
- uncertain parameters may be correlated,

- optimisation under uncertainty
  - decisions (degrees of freedom) are determined simultaneously for the integrated flowsheet,
  - equipment design parameters are not degrees of freedom but take fixed values according to the available equipment allocated in the conceptual design,
  - decisions are determined without recourse knowledge of the uncertain properties except that their values are assumed to reside somewhere in the defined uncertainty space (i.e. scenario independent),
  - controller operation is perfect giving instantaneous application of the desired operating conditions (i.e. controller design is not considered).

5.4 Combined Risk Analysis and model development approach

The technical Risk Analysis approach comprises of methods to quantify the uncertainty contained in a stochastic representation of the process model system (Uncertainty Analysis) and to identify and rank the most important contributors to the uncertainty in the system response (Sensitivity Analysis). In this methodology elements of the Risk Analysis approach (see Chapter 3) are combined with systematic model development procedures (see Figure 2.2, Section 2.4), so that the most important ranked parameters to the uncertain (but structured) system can be identified. The suggestion is that this information can be used to drive the general direction for data collection to improve the models and reduce the uncertainty. As more data becomes available the methodology allows the tracking of the effect of increased knowledge in individual process models and the effect on the complete uncertain system, in an iterative manner.

A schematic of the combined approach is shown in Figure 5.2. The steps within the dashed line box in Figure 5.2 comprise the Uncertainty Analysis. The stochastic characterisation of the system is defined by the assumptions stated previously in Section 5.3.1. The steps of the proposed model development scheme are discussed next. The mathematical methods are stated for those steps specifically associated with the Risk Analysis approach. Less emphasis is placed on the steps associated with conventional model development procedures. The reader is referred to Hangos and Cameron (2001) for further details of steps specifically concerning procedures in the systematic approach to model development.
Figure 5.2. Schematic for the systematic model development incorporating the Risk Analysis approach under uncertainty.
The first six steps are general to all conventional process model development problems. These six steps are applied in turn for the development of process models of individual operations comprising a sequence. In Step 1 the modelling problem is defined regarding the desired complexity, range, accuracy, time characteristics and spatial distribution. In Step 2 the proposed controlling mechanisms and characteristics of the process believed to be important to the problem definition are selected. The type, amount and quality of the data available to support the problem definition and desired model characteristics is evaluated in Step 3. The problem definition may be revised accordingly. The equations for the assumed physico-chemical phenomena, physical property models with the equipment model and any significant external actions (operating policy) are assimilated in Step 4 and a deterministic model of the process is constructed in Step 5. The model may be mechanistic (white box), semi-empirical (grey box) or empirical (black box). Ideally the model-based process development approach should be aiming towards the derivation of mechanistic models for which a good underlying understanding of the process is displayed (but may require an unrealistic amount of effort). The viability of this aim depends on the complexity of the process and underlying mechanisms (of particular concern in pharmaceutical processes) as well as the data which can be reasonably obtained.

If no quantitative data is immediately available concerning a particular process operation then the methodology assumes the formulation of a simple mass balance model in Step 5, as a minimum pre-requisite level of model complexity. Recognition of this inadequacy is an acknowledgement that there is a need to obtain data to improve the process understanding required to develop a more accurate model structure if justified. Unit operation models and material specific physical property, thermodynamic, kinetic and/or transfer models may be available from engineering literature, for which parameter values need to be fitted. Physical property data banks may contain estimates of the required parameter values for the more common components. However, in the case of pharmaceutical processes for new synthetic chemicals, data concerning the active pharmaceutical chemicals will not be available from existing data banks and it is necessary either to conduct further experiments or assume the properties of substitute materials, for those process models requiring property knowledge.

With the available data or observations and the assumed model structure, values for the parameters of the individual unit operation models are estimated (Step 6a) or assumed (Step 6b). Non-linear least squares is a common method for parameter estimation when experimental or process data is available for parameter fitting. Different forms of least squares regression can be utilised according to different situations. Weighted or robust regressions are useful in the presence of outlying data points. In addition, examination of the profile predictions compared to the trends in the data points can indicate the presence of errors in the assumed model structure. If quantitative data is not available then engineering judgement must be used which may be based on qualitative observations to define simple mass balances. The resulting model of each process operation are integrated to form a deterministic model of the complete sequence.

In Step 7 a Perturbation Analysis (one at a time approach, see Section 3.3.1) is used as an initial screening procedure to determine which of the parameter uncertainties in the complete sequence may be potentially significant in the integrated system. Positive and negative deviations from the nominal parameter
estimates, based on the judgement of the developer (for example these could be the approximate precision ranges for different types of model parameters as suggested by Hangos and Cameron, 2001), in all the parameters of the process sequence model are simulated in turn. The parameters whose deviations return significant responses in the important process output criteria are selected to characterise the stochastic system to the Uncertainty Analysis. A significant response is considered to be a deviation in any of the important performance criteria greater than 1% from its nominal value.

In Step 8 the approach used for the quantitative estimation of the uncertainties in these parameters is determined by the data available for parameter estimation. The development of parameter uncertainty estimations can be based on three different information sources:

- analysis of the performance of the model building based on experimental measurement data (Step 8a), allows the estimation of uncertainty in the optimum parameter estimates using confidence intervals or regions from least squares regression (or other parameter fitting methods), often assuming normal probability distributions in the parameter distributions,

- expert technical judgement is needed when quantitative data is not available for systematic model building and models are assumed whose parameter values are instead based on observations and/or assumptions and associated confidence intervals and probability distributions need to be assumed (Step 8b), for which conservative estimates may be elicited given the lack of data,

- published information either quantitatively specifying parameter uncertainties or from which judgements can be inferred (Step 8a or 8b).

Step 9 completely defines the stochastic system considered in this problem and the sampling procedure is invoked in Step 10. The sampling procedure is discussed in more detail with regard to the schematic given in Figure 5.3, where the step numbers relate to those in Figure 5.2.

If model parameters are estimated simultaneously in multi-parameter models (Step 6a), correlations between the parameters may be apparent. The stochastic system is approximated using an efficient sampling scheme over the defined uncertainty space which may include correlated uncertain parameter space. Approximate correlation structures can be obtained from the estimation of the parameter covariance matrix. In this methodology, estimation of the covariance matrix in Step 10b, $\hat{\Sigma}$, is based on the first order Taylor’s linearisation, as discussed in Section 3.3.3.2,

$$\hat{\Sigma} = s^2 \left( J(\theta^*)^T J(\theta^*) \right)^{-1}$$

where $J$ is the Jacobian matrix of the output $\Phi(\theta)$ at the optimal parameter values, $\theta^*$, (i.e. the $N \times P$ matrix with the $(n, p)$th element estimated by $\partial f(x_n, \theta) / \partial \theta_p$ at $\theta^*$, for $N$ data points and $P$ parameters) and $s^2$ is the estimated residual variance, previously defined in Equation 3.12 (Section 3.3.3.1). $J$ is estimated by introducing deviations into each optimal parameter value in turn and re-evaluating the change in the
Integrated design under uncertainty for pharmaceutical processes

predicted dependent variable at each data point (Step 10a). Given the covariance matrix it is straightforward to determine the correlation matrix, \( \hat{C} \) (Step 10c),

\[
\hat{y} = \begin{bmatrix}
\sigma_1^2 & \rho_{12}\sigma_1\sigma_2 & \cdots & \rho_{1p}\sigma_1\sigma_p \\
\rho_{21}\sigma_2\sigma_1 & \sigma_2^2 & \cdots & \rho_{2p}\sigma_2\sigma_p \\
\vdots & \vdots & \ddots & \vdots \\
\rho_{p1}\sigma_p\sigma_1 & \rho_{p2}\sigma_p\sigma_2 & \cdots & \sigma_p^2
\end{bmatrix}
\Rightarrow \hat{C} = \begin{bmatrix}
1.0 & \rho_{12} & \cdots & \rho_{1p} \\
\rho_{21} & 1.0 & \cdots & \rho_{2p} \\
\vdots & \vdots & \ddots & \vdots \\
\rho_{p1} & \rho_{p2} & \cdots & 1.0
\end{bmatrix}
\] (5.2)

where \( \rho \) is the correlation coefficient and \( \sigma \) is the parameter standard deviation (determined from the square root of the parameter variances, \( \sigma^2 \), in the leading diagonal of the covariance matrix). In Step 10e a unit hyper-rectangle of dimension \( P \) is sampled using the quasi-Monte Carlo Hammersley sequence.

---

**Figure 5.3.** Schematic for the employed sampling procedure.

---

92
sampling scheme (HSS) of Diwekar and Kalagnanam (1997). This scheme is selected since the significant improvement in efficiency which can be obtained has been verified for stochastic chemical processes in previous work (Diwekar and Kalagnanam, 1997, Kalagnanam and Diwekar, 1997). In addition, the quasi-Monte Carlo nature of the Hammersley points means that the number of observations required to get a reasonable convergence of the output performance variance (e.g. 1% of the true value) was shown to be less strongly related to the number of input uncertainties than a stratified technique such as LHS. Their results showed the decrease in the required observations for a given number of uncertainty factors was of the order of 2-200 between HSS and LHS for the linear and non-linear functions employed. The HSS and the sample correlation inducement techniques are described in the following paragraphs.

Diwekar and Kalagnanam (1997) define the M points of the Hammersley sequence variant, $e_p(m)$ in a P-dimensional hyper-cube as,

$$e_p(m) = 1 - c_p(m)$$  \hspace{1cm} (5.3)

$$c_p(m) = \left( \frac{m}{M} \phi_{R_1}(m), \phi_{R_2}(m), ..., \phi_{R_{P-1}}(m) \right)$$  \hspace{1cm} for $m = 1, 2, ..., M$  \hspace{1cm} (5.4)

where integer $m$ is written in radix-$R$ notation, resulting in the expansion of integer $m$ into (radix) base $R$ scale with integer coefficients, $a_k$,

$$m = a_v a_{v-1} ... a_2 a_1 a_0$$ \hspace{1cm} \{0 \leq a_k < R_p\}  \hspace{1cm} (5.5)

$$= a_0 + a_1 R_p^1 + a_2 R_p^2 + ... + a_v R_p^v$$ \hspace{1cm} \{0 \leq a_k < R_p\}  \hspace{1cm} (5.6)

and $R_1$, $R_2$, ..., $R_{P-1}$ are the first $P-1$ prime numbers (integers) and $v = \log_R m = \lfloor (\ln m) / (\ln R) \rfloor$ which is rounded down to the nearest integer. The inverse radix number, $\varphi$ (a unique fraction between 0 and 1), for the $m$th observation and $p$th factor, is constructed by reversing the order of the digits of $m$ (integers $a_k$) around the decimal point,

$$\varphi_{R_{p-1}}(m) = .a_0 a_1 a_2 ... a_v$$ \hspace{1cm} \{0 \leq a_k < R_p\}  \hspace{1cm} (5.7)

$$= a_0 R_{p-1}^1 + a_1 R_{p-1}^2 + ... + a_v R_{p-1}^{v-1}$$ \hspace{1cm} \{0 \leq a_k < R_p\}  \hspace{1cm} (5.8)

For example, to place the first six points in the second dimension of a unit hyper-cube ($p=2$), then the binary scale is used (first prime number $R_1=2$) and the inverse radical numbers, $\varphi_2(m)$, are obtained from,
and so on for \( M \) points, where the decimal scale value for \( \varphi_2(m) \) is determined from Equation 5.8. It is clear from Equation 5.8 that the addition of points does not disturb the positions of those already placed. To generate the Hammersley points Diwekar and Kalagnanam (1997) presented an algorithm, shown in Figure 5.4, which is implemented in this thesis. The sub-routine (invradix) used in this thesis for the generation of the inverse radix notation for \( \varphi_R(m) \) is based on long division and is valid for any magnitude of \( P \) assuming \( P \) is less than \( M \). Since a large number \( M \) is defined a priori to the sampling routine, such that the convergence criterion on the output is met for an observation number (\( m \)) less than \( M \), then the first dimension (\( p=1 \)) associated with the first column of \( \varphi_R(m) \) in Equation 5.4 (\( m/M \)) will not be covered for the range between \( m/M \) and \( M \). To avoid this the first column of the Hammersley sequence denoted in Equation 5.3 is not used in the sampling routine employed in this methodology.

As mentioned in Section 3.3.4.3, Diwekar and Kalagnanam (1997) used the rank correlation technique of Iman and Conover (1982) to induce any correlation structures between independently generated inputs which may be apparent due to the simultaneous estimation of parameters. This approach is implemented in this methodology as depicted in Step 10f of the sampling scheme (Figure 5.3). The basis of this approach assumes that rank correlation coefficients are a meaningful way to describe dependencies between stochastic inputs. The desired rank correlation matrix, \( C^* \), of a matrix of independently generated sample input (column) vectors, \( X \), is set as equal to \( C \) (the desired correlation matrix of \( X \)). The elements of the columns of \( X \) are rearranged so that the desired rank correlation structure results between the inputs. This rearrangement is achieved by defining a new matrix in Step 10e, \( K \), which has the same dimension as \( X \) but is independent of \( X \). Iman and Conover (1982) used random permutations of the van der Waerden scores for each column in \( K \) which were found to give a correlation matrix close to the identity. Diwekar and Kalagnanam (1997) used the Hammersley points which is likewise implemented in this methodology. These are inverted over the standard normal cumulative distribution (mean and standard deviation equal to one) and the elements in \( K \) are rearranged to obtain the correlation structure defined by \( C \), to give a matrix, \( K^* \). Since \( C \) is positive definite (i.e. \( x^T C x > 0 \) for any non-zero vector \( x \)) Cholesky factorisation is used to obtain a lower triangular matrix, \( L \), such that,

\[
C = LL^T
\]

(5.10)

\( K \) is multiplied by \( L^T \) to give a matrix, \( K^* \).
If the correlation matrix for K, E, is the identity matrix (i.e. zero correlations between the columns of K) then K* has a correlation matrix which is C. Not only is it necessary that the correlation matrix of K is close to the identity matrix but also the correlation and rank correlation matrices of K* should be close to each other. If these two conditions are met then the desired correlation structure can be induced into X by rearranging each column vector of X into the same rank order as the vectors of K* to give the desired input matrix X*. In this way, the sample rank correlation structure of X* will be the same as that for K* and
approximates $C$ to the same extent that the rank correlation matrix for $K^*$ approximates $C$. Iman and Conover (1982) provide a correction for the situation when $E$ is not close to the identity matrix,

$$E = QQ^T$$

$$K^* = K(Q^{-1})^T L^T$$

where $Q$ is the lower triangular matrix of $E$ and the correlation matrix of $K^*$ is $C$. The elements of $X$ are reordered according to the rank of the corrected matrix $K^*$. This correction factor is incorporated in the sampling routine provided in this methodology.

In Step 10h the columns of the reordered input matrix $X^*$ are inverted over their respective cumulative distribution functions according to their distribution parameters. The distribution parameters are determined either from multi-parameter regressions (Step 10d) or from single parameter regressions or assumed values as already defined in the stochastic model (Step 10g). The stochastic system is solved in the Uncertainty Analysis in Step 11 (Figure 5.3), to obtain probability distributions and distribution parameters (expected value, $E$, and variance, $Var$, from the sample estimates given in Equation 3.26 and 3.27, Section 3.3.4.3) for the desired process variables. This is achieved by sequential simulation of the deterministic model in Step 13 at each observation of the uncertain parameters and at the initial conditions and operating conditions fixed in Step 12, given the matrix of stochastic input observations with any induced correlation structures ($X^*$, from Step 10h in Figure 5.3) and the deterministic model of the complete flowsheet defined in Step 5 (Figure 5.2).

In Step 14 a convergence test is employed to terminate the solution of the stochastic model when the sample estimates of the parameters characterising the distributions (mean and variance) of the stochastic model output criteria are deemed to remain sufficiently unchanged with increasing samples. The convergence test used in this methodology is a tolerance limit on the average squared deviation measured in a distribution parameter, $w$, from the estimate at the current iteration observation, $m_a$, over a specified number of preceding iteration estimates, $m_b$. This limit, $\Delta w$, for the $q$th predicted process output quality criterion is shown in Equation 5.14,

$$\Delta w_q = \frac{1}{m_a} \sqrt{\sum_{m=m_b}^{m_a} (w_{q,m} - w_{q,m_b})^2} \leq \varepsilon_{w,q}$$

where $w$ is the mean or variance estimate over $m_a$ observations and $\varepsilon$ is a permitted tolerance. In this work a value of $1\%$ of the current iteration value of $w$ is suggested for $\varepsilon$, and a value of 100 scenarios for $m_a$. The retaining nature of the employed Hammersley sequence variant allows extra observations to be made without disturbing the current sample. This permits sequential placement of observations in Step 10a (Figure 5.3) in order to satisfy the convergence test in Step 14 without resimulation of the entire sample set. Termination of the sampling loop requires the multiple fulfilment of the convergence requirements for the mean and variance estimates for all of the desired process performance criteria.
Referring to Figure 5.2 the stochastic results of the Uncertainty Analysis can be compared to independent data to validate the uncertain model (Step 15). As discussed in the conceptual problem (Section 5.2) both the location and spread of the distributions in relation to the data are important in the validation. Independent data may already be available from other runs or it can be obtained from model validation runs. As stated by Basu et al. (1999) there should be plenty of opportunities to obtain more data for this purpose given the nature of pharmaceutical process development.

In Step 16 Sensitivity Analysis is used to estimate the ranking priority of the key stochastic inputs contributing to the uncertainty in the stochastic process output criteria. In an efficient manner the sensitivity techniques employed in this methodology reuse the sample results of the Uncertainty Analysis to avoid the need for any further simulations of the deterministic model. The appropriate use of specific statistical measures which may be calculated in this analysis is discussed next.

A schematic proposing the recommended measures according to the estimated form of the relationships between the inputs and outputs is shown in Figure 5.5. The mathematical definitions of these measures have been previously defined in Section 3.4.1. The presence of outliers or non-linear but monotonic input-output relationships can have a strong effect on the estimation of linear contributor measures and the use of rank transformed sample data may negate this effect in these estimations. This transformation is made in Step 16b to account for this eventuality. Hofer (1999) states that as a rule, if the coefficient of

Figure 5.5. Schematic for the use of input contribution measures in the Sensitivity Analysis.
determination of the unranked data \( (R^2) \) is greater than that of the ranked data \( (RR^2) \) then the unranked data is used for Sensitivity Analysis. This rule is implemented in Step 16c. In addition, the use of linear measures may be unsuitable in the presence of highly non-linear models. The \( R^2 \) is a measure of the proportion of the total variation in the output data that is explained by the linear regression model. If the \( R^2 \) (or \( RR^2 \)) is greater than 0.5 (Step 16d) then linear contributor measures estimated in Step 16e can be assumed to explain an adequate proportion of the observed variability in the output data. The correlation coefficient \( (CC) \) provides an estimation of the total variability in the input associated in a linear manner with the variability in the output. The standardised regression coefficient \( (SRC) \) measures only the variability in the input that is unexplained by the other inputs. This removes the effect of any spurious correlations to which the CCs are susceptible. If the \( R^2 \) (or \( RR^2 \)) is less than 0.5 then linear contributor measures explain too small a proportion of the observed variability in the output data and non-linear measures should be used in Step 16f. The non-directional correlation ratio \( (CR) \) is used since it is not restricted to the estimation of linear or monotonic contributors. If there is any element of doubt then scatter plots between individual input-output pairs can be viewed, though these may also be susceptible to spurious correlations.

In addition to the direct identification of the key stochastic input-output associations as shown in Figure 5.5, process sub-sequence contributions to uncertainty are estimated within an integrated flowsheet. These contributions provide a measure of the accumulation and propagation of uncertainty observed in particular process properties over certain operational sub-sequences throughout the complete sequence. A process sub-sequence criterion may either be related to the operational effectiveness of a stage/sub-sequence or it may be a particular stream property. In this way the cumulative influence that a number of minor contributors within a particular sub-sequence may exert on important inter-stage criteria is estimated. Key sub-sequences may be identified which would not necessarily correspond with the location of key input contributions identified using the ranking procedures based on the input-output associations.

For this analysis it is assumed that the inter-stage criteria comprise of potential inter-stage plant data measurements. Each sub-sequence contribution is defined as the difference between the relative uncertainty in the two criteria defining a sub-sequence. These contributions are normalised with respect to the relative uncertainty in the final sub-sequence criterion to give a sum of the normalised sub-sequence uncertainty contribution factors (SSC) equal to one over the complete process. Each SSC estimates the fraction of the endpoint uncertainty which has accumulated over a sub-sequence. A normalised sub-sequence uncertainty contribution of an inter-stage sub-sequence criterion, \( Q \), over a sub-sequence, \( ss \), is estimated,

\[
SSC_{ss} = \frac{\left(Q_{ss}^UB - Q_{ss}^LB\right) - \left(Q_{ss-1}^UB - Q_{ss-1}^LB\right)}{\left(Q_{ss}^UB - Q_{ss}^LB\right)} = \frac{\mu_{Q_{ss}}}{\mu_{Q_{ss-1}}} \quad \text{for} \quad ss = 1...SS \quad (5.15)
\]
where the uncertainty in each sub-sequence criterion is quantified by the width between the 5 and 95\% fractiles, $Q_{5}^{LB}$ and $Q_{95}^{UB}$, relative to the mean value, $\mu$. 

In summary, Sensitivity Analysis (Step 16) is used to estimate the key input uncertainties and process sub-sequences which influence the uncertainty observed in the important process performance predictions. The inadequacies of the current deterministic process sequence model and associated parameter uncertainties are observed from the comparison between available validation data and these predictions. The optimal reduction in key parameter uncertainties (Step 17) is discussed in the next section. It is suggested that this combined information could be used to (a) focus relative efforts towards improving a specific process model within the sequence by inferring the key uncertain phenomena associated with the identified process sub-sequence and parameter uncertainties and (b) support decisions concerning what data to collect or what experiments to plan (Step 18 - see the data driver feedback loop in Figure 5.2). The specific details regarding these decisions are not part of the problem considered in this thesis.

5.5 Optimisation under uncertainty

The typical algorithm for the solution of a stochastic optimisation problem is represented by the schematic shown in Figure 5.6. This algorithm provides the basis for the stochastic optimisation problems posed in this thesis. In a convenient manner the HSS sampling procedure implemented in the Uncertainty Analysis (Figure 5.3) is also implemented in the stochastic optimisation algorithm as was proposed by Diwekar and Kalagnanam (1997) and Kalagnanam and Diwekar (1997). The uncertain factors in the stochastic problem are sampled within the optimisation loop and may be subject to a convergence checking sub-routine for

\[\text{Initial decisions} \rightarrow \text{Nonlinear constrained optimisation routine} \rightarrow \text{Optimal solution}\]

\[\downarrow\]

\[\text{New decisions} \leftarrow \text{Stochastic objective function and constraints}\]

\[\downarrow\]

\[\text{Sampling routine} \leftarrow \text{Convergence check on output distribution parameters}\]

\[\downarrow\]

\[\text{Stochastic model} \leftarrow \text{Deterministic model}\]

\[\downarrow\]

\[\text{Fixed parameters Initial conditions}\]

Figure 5.6. Stochastic optimisation algorithm.
the parameters characterising the distributions of the output. If the convergence checking sub-routine is selected the algorithm evaluates the deterministic model sequentially at the current decisions determined by the optimisation routine, until the specified tolerance in the convergence checking sub-routine is satisfied for each function evaluation. In this way the stochastic problem is posed as a non-linear programming (NLP) optimisation which can be solved using a gradient-based optimisation technique. An alternative option is to assume a fixed sample size (the convergence checking option is not selected) which permits the sampling routine to be placed prior to the optimisation loop. In this case the problem can be reformulated as a multi-scenario deterministic optimisation problem in which the scenarios are evaluated simultaneously at each stochastic function evaluation. This may also be solved as a NLP problem. This can save optimisation time by providing better gradient information for the (gradient-based) optimiser and by using less scenarios at the expense of solution accuracy (since convergence to a particular tolerance level is not guaranteed).

In this thesis the former approach (convergence checking invoked) is used for reasons based on the desire for solution accuracy with efficient approximation of the P-dimensional input uncertainty space where the number of factors may be large (e.g. greater than 10). The quality of the gradient information supplied to the optimiser may be poor due to the ‘random’ representation of the stochastic model at each optimisation iteration which may increase the number of optimisation iterations and computational time.

In this work a gradient-based NLP optimisation program (Matlab Optimisation Toolbox fmincon function, The MathWorks, Inc., USA) employing successive quadratic programming (SQP) is used. In SQP, at each major iteration the NLP problem is modelled as a quadratic programming sub-problem (QP) which is solved to obtain a new approximate solution at which a new QP sub-problem is generated.

To reduce the size of the NLP problem the uncertainties which do not appear to provide any significant contribution in the full stochastic process evaluated in the Uncertainty Analysis are not included in the optimisation. The important parameters are determined from the Sensitivity Analysis. In addition, less strict convergence criteria are employed in the stochastic optimisation as compared to the Uncertainty Analysis. This reduces the number of scenarios in the stochastic model, evaluated in each function evaluation of the optimisation algorithm. Grossmann and Sargent (1978) justify a reduced accuracy approach for the uncertain performance estimation in their work for design under uncertainty, since the input parameter distribution functions will rarely be known sufficiently well.

In a general stochastic programming formulation the decision variables may be scenario independent as in the ‘here and now’ context, or scenario dependent as in the ‘wait and see’ context. In this methodology the former is selected. The reason for this is to reduce the over optimistic uncertain performance solutions which would be obtained under the assumption of perfect control with knowledge of uncertainty realisations. This assumption is not made in this methodology since perfect knowledge is assumed to be an unrealistic situation. On the other hand ‘here and now’ solutions can be more conservative. The compromise of incorporating feedback controller models is not employed in this methodology. One reason is that the control of pharmaceutical operations is often manual. Validation of the optimal solution is
performed by simulating the stochastic model under the original number of input uncertainties using stricter convergence limits on the performance distribution parameter estimates.

The four different stochastic optimisation problems associated with the methodology in Figure 5.1 (Section 5.3) are defined in the following sub-sections.

### 5.5.1 Optimal input uncertainty reduction

Desired levels of reduction in the uncertainty of predicted output criteria can be achieved with reductions in the important uncertain parameters of the stochastic problem. A quantitative estimate of the extent of reductions required in the uncertainty sources to meet the desired output uncertainty levels can be provided. This can be posed as a stochastic optimisation problem which is solved using the algorithm shown in Figure 5.6. The solution to this problem constitutes Step 17 in the schematic for combined systematic model development and Risk Analysis (Figure 5.2, Section 5.4). The decision variables are fractions of the original values (before uncertainty reduction) of the parameters which characterise the spread of the input uncertainties. The values of these decisions are passed to the sampling sub-routine (from the optimiser) which locates observations from within the redefined uncertainty space. The resulting observations form the matrix of parameter scenarios. The deterministic model is solved at each scenario given the set of fixed parameters, initial conditions and operating policy within the stochastic model. The objective is to maximise the sum of the decision variable fractions whilst maintaining desired levels of output uncertainty reduction with the incorporation of inequality constraints on the uncertainty in the desired output distributions. The objective function value and constraints are estimated from the simulation of the stochastic model and returned to the optimiser. The general formulation for this optimisation problem is shown in Problem P1 for input uncertainties which are stochastic (assumed to be normal) and/or deterministic (uniform).

Objective function

$$\max_{\delta_t, \delta_d} \sum_{st=1}^{ST} \delta_{st} + \sum_{dt=1}^{DT} \delta_{dt}$$

subject to:

**Deterministic equations (process models)**

$$J_{o,s,m}(\hat{x}_{o,s,m}, x_{o,s,m}, y_{o,s,m}, z_d, \theta_{st,m}, \theta_{dt,m}) = 0 \quad \forall \ s = 1 \ldots S, \ m = 1 \ldots M$$

$$f_{i,m}(\hat{x}_{i,m}, x_{i,m}, y_{i,m}, z_d, \theta_{st,m}, \theta_{dt,m}) = 0 \quad \forall \ s = 1 \ldots S, \ m = 1 \ldots M$$

**Deterministic inequality constraints**

$$g_{s,m}(\hat{x}_{s,m}, x_{s,m}, y_{s,m}, z_d, \theta_{st,m}, \theta_{dt,m}) \leq 0 \quad \forall \ s = 1 \ldots S, \ m = 1 \ldots M$$
stochastic inequality constraints

\[ FW(\Phi_q) \leq \alpha_q FW(\Phi_q) \]  \quad \forall \quad q = 1 \ldots Q

decision variable bounds

\[ 0 < \delta_{st} \leq 1 \]  \quad \forall \quad st = 1 \ldots ST

\[ 0 < \delta_{dt} \leq 1 \]  \quad \forall \quad dt = 1 \ldots DT

uncertainty space characterisation

\[ \Theta_N = \{ \theta_n \mid N(\mu_n, \delta_n, \sigma'_n) \} \]  \quad \forall \quad st = 1 \ldots ST

\[ \Theta_U = \{ \theta_u \mid U(\mu_u - \delta_u \Delta \theta_u, \mu_u + \delta_u \Delta \theta_u) \} \]  \quad \forall \quad dt = 1 \ldots DT

\[ \Delta \theta_{dt} = \theta_{dt}^{UB} - \mu_{dt} = \mu_{dt} - \theta_{dt}^{LB} \]  \quad \forall \quad dt = 1 \ldots DT

where \( I, f, g \) are the vectors of initial conditions, equalities and inequalities. The indices \( o, s, m, d, q, st \) and \( dt \) are associated with the initial conditions, process stages, parameter scenarios, operating policy variables (\( z \)), stochastic output performance criteria (\( \Phi \)) and the stochastic and deterministic uncertainties (\( \Theta \)). \( \dot{x} \) are the time derivatives of the differential variables, \( x \), and \( y \) are the algebraic variables. For a dynamic optimisation problem, the operating policy variables, \( z \), comprise the set of the time invariant decisions, \( u \), time dependent decisions, \( v \), and operation time, \( t_f \). \( FW(\Phi) \) is the width between the 5 and 95% fractiles for the performance criteria, \( \Phi \), and \( \alpha \) is the desired fraction of the original value of the criteria before uncertainty reduction. The prime represents the original value before uncertainty reduction.

The total uncertainty space, \( \Theta \), is defined as the stochastic (assumed normal) and the deterministic (uniform) space, \( \Theta_N \) and \( \Theta_U \), combined. The former are characterised by the mean and standard deviation (\( \mu \) and \( \sigma \)) and the latter by the mean placed equidistant between upper and lower limits (\( \Theta^{UB} \) and \( \Theta^{LB} \)).

It is assumed that the same distribution means (nominal input parameter values) as determined or assumed in the original stochastic problem, are maintained. If the stochastic problem contains decisions in correlated inputs, it is assumed that a change in the spread of one of the correlated parameters gives an equivalent change in the others, while maintaining the same correlation structure. A trade-off curve may be constructed between the extent of optimal input reduction and desired level of performance uncertainty by solving Problem P1 at different values of \( \alpha \).

5.5.2 Process flowsheet optimisation

In this problem formulation, process optimisation of the flowsheet performance under uncertainty is considered. Operating policy optimisation of the stochastic model gives a solution which accounts for the combined influence of all the significant uncertainties included in the problem. It does not rely on the
identification of limiting scenarios with little chance of occurring in the probabilistic sense and which lead to more conservative solutions.

The operating policy decision variables of a given process flowsheet under uncertainty are optimised for performance using the stochastic optimisation algorithm shown in Figure 5.6. The set of decision variables can include time dependent, time independent and duration time operating variables for the process stages, depending on the dynamic nature of the individual process models and the perceived suitability of conceptual operating policies with regard to the equipment (i.e. limited implementation of dynamic control profiles). The general formulation for this stochastic optimisation is shown in Problem P2 for a stochastic objective in a deterministic quality criterion $\Phi$.

Objective function

$$\max_{x_d} F\left( E\left[ \Phi(\mathbf{x}_{s,m}, \mathbf{y}_{s,m}, \mathbf{z}_{d,m}, \theta_{st}, \theta_{dt,m}) \right] \right)$$

subject to:

deterministic equations

$$J_{a,s,m}(\mathbf{x}_{a,s,m}, \mathbf{x}_{a,s,m}, \mathbf{y}_{a,s,m}, \mathbf{z}_{d,m}, \theta_{st}, \theta_{dt,m}) = 0 \quad \forall \ s = 1..S, \ m = 1..M$$

$$f_{a,s,m}(\mathbf{x}_{a,s,m}, \mathbf{y}_{a,s,m}, \mathbf{z}_{d,m}, \theta_{st}, \theta_{dt,m}) = 0 \quad \forall \ s = 1..S, \ m = 1..M$$

deterministic inequality constraints

$$g_{a,s,m}(\mathbf{x}_{a,s,m}, \mathbf{y}_{a,s,m}, \mathbf{z}_{d,m}, \theta_{st}, \theta_{dt,m}) \leq 0 \quad \forall \ s = 1..S, \ m = 1..M$$

deterministic quality criteria

$$\Phi_{q,a,s,m} = f_{q,a,s,m}(\mathbf{x}_{q,a,s,m}, \mathbf{y}_{q,a,s,m}, \mathbf{z}_{q,a,s,m}, \theta_{st}, \theta_{dt,m}) \quad \forall \ q = 1..Q, \ s = 1..S, \ m = 1..M$$

quality criteria deviation

$$\Delta \Phi_{q,a,s,m} = \Phi_{q,a,s,m}^{th} - \Phi_{q,a,s,m}$$

$$\forall \ q = 1..Q, \ s = 1..S, \ m = 1..M$$

continuous deviation function

$$f_{dev,q,a,s,m}(\Phi_{q,a,s,m}, \Phi_{q,a,s,m}^{th}) = \beta_{q,a,s,m} d_{1,q,a,s,m}(\Delta \Phi_{q,a,s,m}) + (1 - \beta_{q,a,s,m}) d_{2,q,a,s,m}(\Delta \Phi_{q,a,s,m}) \quad \forall \ q = 1..Q, \ s = 1..S, \ m = 1..M$$

binary variable smoothing approximation

$$\beta_{q,a,s,m} = \frac{1}{2} \left[ \tanh(\xi_{q,a,s,m}) + 1 \right] \quad \forall \ q = 1..Q, \ s = 1..S, \ m = 1..M$$

stochastic inequality constraints
Integrated design under uncertainty for pharmaceutical processes

\[
F_{q,s}\left(\mathbb{E}\left\{ f_{dev,q,s,m}(\phi_{q,s,m},\phi_{q,s,m}^{th})\right\} \right) \leq \gamma_{q,s} \quad \forall \quad q = 1..Q \quad s = 1..S
\]

decision variable bounds

\[
z^{LB}_d \leq z_d \leq z^{UB}_d \quad \forall \quad d = 1..D
\]

uncertainty space characterisation

\[
\Theta_{N} = \left\{ \theta_{st} \right\} | N(\mu_{st},\sigma_{st}) \quad \forall \quad st = 1..ST
\]

\[
\Theta_{U} = \left\{ \theta_{dt} \right\} | U(\theta^{LB}_{dt},\theta^{UB}_{dt}) \quad \forall \quad dt = 1..DT
\]

(Problem P2)

where the general continuous deviation function, \( f_{dev} \), introduced by Samsatli et al. (1998), represent a variety of one or two-sided deviations of a deterministic quality criterion, \( \phi \), from a desired constraint, \( \phi^{th} \). The binary variables, \( \beta \), allow the representation of both one-sided (non-symmetrical) and two-sided deviation functions, \( d_1 \) and \( d_2 \), to be contained in \( f_{dev} \). The use of the hyperbolic smoothing function (Samsatli et al., 1998) to approximate the binary variables in the continuous deviation function, \( f_{dev} \), means that the stochastic problem including non-symmetrical deviation functions does not contain discontinuities due to these deviations. This allows the formulation of a problem which can be solved as a NLP optimisation problem using gradient-based methods. A general stochastic function, \( F \), is used to represent the objective function and any stochastic inequality constraints. These can represent robustness metrics which may be comprised of the expected value (E) of the deterministic deviation functions. These could include probabilistic quantities such as those discussed in Section 3.3.2, including the expected value, or typical two-sided measures of variance and Taguchi’s quality loss, or one-sided measures such as probability of constraint failure and expected extent of constraint violation (Samsatli et al., 1998). They are estimated from the sample estimates given in Equations 3.26 to 3.29 (Section 3.3.4.3). The input parameter uncertainty space remains fixed, defined by the distribution characterising parameters passed to the sampling routine (Step 10d and Step 10g in Figure 5.3, Section 5.4), specified from prior model development and Uncertainty Analysis. The decision variables are operating policy variables which are specified in the ‘here and now’ context (scenario independent). In Problem P2 the algorithm passes the decisions from the optimiser directly to the stochastic model simulator, bypassing the sampling subroutine since the input uncertainty space remain unaffected by the decisions (as compared to Problem P1). The sampling sub-routine is then invoked by the stochastic model simulator.

5.5.3 The value of perfect information

The concept of the value of perfect information (Raiffa, 1968), VPI, can have a useful application to process design optimisation under uncertainty as suggested by Pistikopoulos (1995). The provision of perfect knowledge in replacement of certain sources of input uncertainty can be quantified using the VPI if physical measurement permits. For instance, physical characteristics of important external influences
which may be uncertain can be potentially certain if they could be accurately measured. It is assumed that this does not extend to the parameter uncertainties associated with the deterministic process models for which perfect information of the values is unrealistic.

The complete range of uncertainties considered in this problem are the stochastic and deterministic parameter uncertainties (as in Problem P1 and P2), \( \theta_{st} \) and \( \theta_{dt} \), and in the potential certainties, \( \theta_{pc} \). A general definition of the value of information, \( VI \), is defined in Equation 5.16 as a stochastic performance objective in some criterion, \( F(E(\Phi)) \), determined over some definition of the uncertainty space, \( \Theta \).

\[
VI = F\left( E_\Theta \{ \Phi \} \right) \tag{5.16}
\]

The VPI is defined as the performance objective achieved with the perfect knowledge of the potentially certain properties (\( VI_{\text{wait}} \)) less the performance objective achieved under uncertainty in these properties (\( VI_{\text{here}} \)), as shown in Equation 5.17. This is illustrated in Figure 5.7 where the VPI is the average difference between \( VI_{\text{wait}} \) and \( VI_{\text{here}} \) over the range of potential knowledge in \( \theta_{pc} \) (Equation 5.18). \( VI_{\text{wait}} \) is defined as the value of information objective (\( F_{\text{wait}} \)) achieved with perfect knowledge of \( \theta_{pc} \) but without perfect knowledge of \( \theta_{st} \) or \( \theta_{dt} \). \( VI_{\text{here}} \) is defined as the value of information objective (\( F_{\text{here}} \)) achieved without perfect knowledge of \( \theta_{pc} \) or \( \theta_{st} \) or \( \theta_{dt} \). The VPI can be approximated as the expected value of \( F_{\text{wait}} \) less \( F_{\text{here}} \) over \( \theta_{pc} \) space as shown in Equation 5.19,

\[
VPI = E_{\theta_{pc}} \left\{ VI_{\text{wait}} \right\} - E_{\theta_{pc}} \left\{ VI_{\text{here}} \right\} \tag{5.17}
\]

\[
= E_{\theta_{pc}} \left\{ VI_{\text{wait}} - VI_{\text{here}} \right\} \tag{5.18}
\]

\[
= E_{\theta_{pc}} \left\{ F_{\text{wait},m}(E_{\theta_{st},\theta_{dt}} \{ \Phi \}) - F_{\text{here},m}(E_{\theta_{st},\theta_{dt},\theta_{pc}} \{ \Phi \}) \right\} \tag{5.19}
\]

where index \( m' \) represents the observations in \( \theta_{pc} \) space. The VPI is computed using the approximation given in Equation 5.19. This involves a number of steps shown in Figure 5.8. Independent stochastic optimisation problems are solved for \( VI_{\text{wait}} \) and for \( VI_{\text{here}} \). In the case of \( VI_{\text{wait}} \), the entire uncertainty space

![Figure 5.7. VPI for potentially certain properties.](image)
comprises of uncertainties which are viewed by the operating policy decision variables as scenario independent (i.e. \( \theta_{st} \) and \( \theta_{st} \)) and uncertainties which are viewed as scenario dependent (i.e. \( \theta_{pc} \)). This means that is necessary to differentiate between the space of the uncertain and potentially certain parameters (Step 1a) and perform a number of 'here and now' optimisations under \( \theta_{st} \) and \( \theta_{st} \) space at specific realisations in \( \theta_{pc} \) space. The total number of observations, \( M' \), and their placement in \( \theta_{pc} \) space are determined by the sampling sub-routine in Step 2. \( F_{\text{wait}} \) and \( F_{\text{here}} \) are both computed at these values of \( \theta_{pc} \). A relatively slack tolerance is recommended for Step 2 to limit the number of observations in \( \theta_{pc} \) space to maintain the solution of 'wait and see' problems to a reasonable number. With the solution of \( M' \) 'wait and see' optimisation problems in Step 3a (one at each observation of \( \theta_{pc} \)), \( M' \) optimal values of \( F_{\text{wait}} \) and \( M' \) vectors of optimal decisions are obtained. In the case of \( V_{\text{here}} \), a single 'here and now' optimisation is solved in Step 1b, under the entire uncertainty space comprising \( \theta_{st} \), \( \theta_{st} \) and \( \theta_{pc} \). This yields a single 'here and now' optimal decision policy. This policy is implemented in Step 3b to evaluate \( M' \) values of \( F_{\text{here}} \) at the same realisations in \( \theta_{pc} \) space as were used for the \( F_{\text{wait}} \) solutions. The VPI can then be approximated in Step 4 with the values for \( F_{\text{wait}} \) and \( F_{\text{here}} \) using Equation 5.19.

It is possible that some of the stochastic simulations for \( F_{\text{here}} \) at the realisations of \( \theta_{pc} \) may predict violations in any stochastic inequality constraints employed, since the operating policy decisions are determined from the 'here and now' optimisation without perfect knowledge of \( \theta_{pc} \). It is desirable to consider these violations as some form of penalty in \( F_{\text{here}} \) so that the \( V_{\text{here}} \) reflects the poor behaviour of

---

**Figure 5.8. Schematic for the computation of VPI for potential certainty.**
these stochastic criteria. In this methodology a penalty function, $\Psi_{\text{here}}$, is used to penalise the original value of information objective ($F_{\text{here}}$) by some function of the extent of any stochastic inequality constraint violations, $[(F_{\text{here},q,s} - \gamma_{q,s}) \leq 0]$, as shown in Equation 5.20,

$$VI_{\text{here}} = \Psi_{\text{here}} \left( F_{\text{here}} \cdot \beta_{q,s} \left( F_{\text{here},q,s} - \gamma_{q,s} \right) \right) \tag{5.20}$$

where $\gamma_{q}$ is the desired constraint limit in the $q$th criterion $F_{\text{here},q,s}$ and $\beta$ is a binary variable which is one in the case of a violation and zero otherwise.

### 5.5.4 Decision variable tolerance

It is unrealistic that a single optimal operating policy can always be implemented due to process variability and safety reasons. The maximum amount of tolerance permitted in a 'here and now' optimum operating policy can be ascertained from an optimisation determining the maximum deviation in the limits around the 'here and now' optimal operating policy, subject to stochastic performance constraints. The general formulation for this stochastic optimisation is given in Problem P3, where the tolerance is modelled as uncertainty in the 'here and now' optimum operating policy values which are assumed to be uniformly distributed.

Problem P3 combines the aspect of the determination of uncertainty space of the input uncertainty reduction optimisation in Problem P1, with the stochastic performance constraints of the process flowsheet optimisation in Problem P2. In Problem P3 an extra set of dimensions, $\Theta_{U}$, in the overall uncertainty space, $\Theta$, represents the tolerance in the operating policy variables. The decision variables, $\delta^U$ and $\delta^L$, are the disjoint fractions of the available operating policy range in each direction from the single 'here and now' optimum policy values, $z^*$,

$$\delta^U_d = \frac{z^{UB}_d - z^*_d}{z^{UB}_d - z^*_d}, \quad \delta^L_d = \frac{z^*_d - z^{LB}_d}{z^*_d - z^{LB}_d} \tag{5.21}$$

where the superscripts $UB^*$ and $LB^*$ represent the upper and lower tolerance bounds determined in the optimisation and $UB$ and $LB$ represent the bounds on the operating variables. The tolerance is defined as the fraction of the total available operating policy range which is occupied by the maximum feasible operating policy range. The objective function is the average tolerance over the set of operating policy variables. The stochastic optimisation algorithm (Figure 5.6) is implemented for the solution of Problem P3. The upper and lower bounds on the policy uncertainties are constrained by the desired stochastic performance targets and the deterministic model equations. The stochastic constraint performance targets ($\gamma$) may be specified as some function of the performance achieved in the optimal 'here and now' solution associated with the $z^*$. 
Objective function

\[
\max_{\delta_d^L, \delta_d^U} \frac{1}{D} \sum_{d=1}^{D} \left( \frac{\delta_d^L (z_d^* - z_d^L)}{z_d^L - z_d^L} + \delta_d^U (z_d^U - z_d^*) \right)
\]

subject to:

deterministic equations

\[
J_{o,s,m}(x_{o,s,m}, y_{o,s,m}, z_{d,m}, \theta_{st,m}, \theta_{dt,m}) = 0 \quad \forall \ s = 1..S, \ m = 1..M
\]

\[
f_{s,m}(x_{s,m}, y_{s,m}, z_{d,m}, \theta_{st,m}, \theta_{dt,m}) = 0 \quad \forall \ s = 1..S, \ m = 1..M
\]

deterministic inequality constraints

\[
b_{s,m}(x_{s,m}, y_{s,m}, z_{d,m}, \theta_{st,m}, \theta_{dt,m}) \leq 0 \quad \forall \ s = 1..S, \ m = 1..M
\]

deterministic quality criteria

\[
\phi_{q,s,m} = f_{q,s,m}(x_{s,m}, y_{s,m}, z_{d,m}, \theta_{st,m}, \theta_{dt,m}) \quad \forall \ q = 1..Q, \ s = 1..S, \ m = 1..M
\]

quality criteria deviation

\[
\Delta \phi_{q,s,m} = \phi_{q,s,m}^{th} - \phi_{q,s,m} \quad \forall \ q = 1..Q, \ s = 1..S, \ m = 1..M
\]

continuous deviation function

\[
f_{dev,q,s,m}(\phi_{q,s,m}, \phi_{q,s,m}^{th}) = \beta_{q,s,m} d_{q,s,m}(\Delta \phi_{q,s,m}) + (1 - \beta_{q,s,m}) d_{q,s,m}(\Delta \phi_{q,s,m}) \quad \forall \ q = 1..Q, \ s = 1..S, \ m = 1..M
\]

binary variable smoothing approximation

\[
\beta_{q,s,m} = \frac{1}{2} \left( \tanh(\varepsilon_{q,s} \Delta \phi_{q,s,m}) + 1 \right) \quad \forall \ q = 1..Q, \ s = 1..S, \ m = 1..M
\]

stochastic inequality constraints

\[
F_{q,s}(E[f_{dev,q,s,m}(\phi_{q,s,m}, \phi_{q,s,m}^{th})]) \leq Y_{q,s} \quad \forall \ q = 1..Q, \ s = 1..S
\]

decision variable bounds

\[
0 < \delta_d^L \leq 1 \quad \forall \ d = 1..D
\]

\[
0 < \delta_d^U \leq 1 \quad \forall \ d = 1..D
\]

uncertainty space characterisation
Integrated design under uncertainty for pharmaceutical processes

\[ \Theta_N = \left\{ \theta_t \mid N(\mu_t, \sigma_t) \right\} \quad \forall \ st = 1 \ldots ST \]

\[ \Theta_{U,z_d} = \left\{ \theta_d \mid \left[ (z_d^* - \delta_d^L (z_d^* - z_d^{LB}), z_d^* + \delta_d^U (z_d^* - z_d^{UB})) \right] \right\} \quad \forall \ d = 1 \ldots D \]

\[ \Theta_U = \left\{ \theta_{dt} \mid U(\theta_{dt}^{LB}, \theta_{dt}^{UB}) \right\} \quad \forall \ dt = 1 \ldots DT \quad \text{(Problem P3)} \]

5.6 Summary

A methodology and mathematical techniques are presented for the systematic and quantitative management of uncertainty in process sequences described with deterministic models. It is split into two major areas of study, both involving stochastic descriptions of the system. The first is associated with systematic model development and the second is associated with process optimisation under uncertainty.

A combined Risk Analysis and systematic model development approach is suggested to systematically account for model parameter uncertainties during model building. This aims to provide information to focus and support data collection and model development decisions with respect to the predicted stochastic response to the quantified uncertainty in the process model of the complete sequence. An efficient Hammersley sampling strategy is implemented to provide coverage of the defined uncertainty space and generate the stochastic model. Uncertainty Analysis and Sensitivity Analysis sampling based techniques are proposed to quantify the uncertainty in important process performance criteria and identify the sources of key contributions which could be used to focus the direction of further modelling effort. The iterative nature of the systematic modelling procedure permits the tracking of the effects of the uncertainty contained in the models as more data becomes available and is incorporated in the models.

Stochastic optimisation problem formulations are stated for four different problem cases which can be solved using a conventional stochastic optimisation algorithm. The methodology implements the same sampling strategy (HSS) as employed for the Risk Analysis. The first problem case quantifies the required levels of parameter uncertainty reductions necessary to achieve desired levels of uncertainty in the output criteria. The second problem provides a robust optimal flowsheet operating policy under uncertainty. In addition, the value of perfect information is solved using this second formulation. In this methodology the VPI is concerned with those external process inherent uncertain properties which could feasibly be directly measured a priori. The final problem formulation determines the level of tolerance permitted in operating policy variables whilst maintaining target levels of performance uncertainty.

In the following chapters two case studies are presented to verify the methodology and techniques which have been proposed here. The first case study (Chapter 6) investigates a published industrial multi-phase reactor process. The Uncertainty and Sensitivity Analysis methods are implemented and the benefits of the robust optimisation method (second problem case) are examined. The second case study (Chapter 7, 8, and 9) is more comprehensive, comprising of a complete integrated pharmaceutical process sequence. It aims to provide a rigorous examination of the methods proposed in this chapter.
CHAPTER 6

CASE STUDY I: A MULTIPHASE REACTOR PROCESS

6.1 Introduction

In the previous chapter an approach to integrated design under uncertainty was proposed. Risk Analysis methods for the quantification of uncertainty and identification and ranking of major contributors in modelled process sequences are presented within a systematic model development scheme. The stochastic problem generated is reduced in size for the stochastic optimisation of process performance under uncertainty and tolerance to operating policy variable error. In this chapter, aspects of the Uncertainty Analysis and Sensitivity Analysis with the optimisation for process performance are applied to a reaction process under uncertainty. Case Study I is based on a multiphase reaction process developed with the aid of process modelling and computer simulation (Sano et al., 1998). This chapter aims to verify the aspects of the methodology mentioned above and to express the pitfalls and benefits associated with the consideration of uncertainty in process models.

6.2 Process description

Case Study I is based on an exothermic multiphase reaction process and kinetic model investigated by Sano et al. (1998). The process is for the production of a pharmaceutical intermediate, formed from the amination of a bromopropyl compound. The authors developed a kinetic model based on reaction calorimetry data obtained under laboratory conditions in order to determine the optimum feasible and safe operating policy.

Solid particles of the active pharmaceutical ingredient (API) bromopropyl feed compound (A) reside in an organic solvent (methanol) inside the reaction vessel. A fixed volume of a 50 wt% aqueous dimethylamine reagent (B) is added to the vessel at a constant flowrate under continuous agitation. The solids gradually dissolve and react with the dimethylamine. A diagram of the process is shown in Figure 6.1. The exact physico-chemical phenomena for this process are not known. The reaction consists of a parallel-series reaction in which the dimethylamine reacts with the dissolved API feed to form the desired intermediate (C), Equation 6.1, which in turn reacts with the active feed (A) to form a dimeric byproduct (D) in parallel, Equation 6.2. D is stated to be very difficult to remove in the downstream purification stages.

\[
A + B \rightarrow C \quad \text{Main reaction} \\
A + C \rightarrow D \quad \text{Sub-reaction}
\] (6.1)

Intrinsic first order reaction kinetics are assumed in the deterministic process model proposed by Sano et al. (1998). An initial rate limiting period due to the dissolution of solids B. was observed to be
independent of solvent concentration and agitation speed within the range of conditions approved. A crude approximation of first order kinetics (with Arrhenius constant and activation energy) is assumed in the model for this dissolution controlled period. This period was observed to last until approximately 55% conversion of A for all the conditions considered, at which point the reaction appeared to be limited by the intrinsic reaction kinetics.

The kinetic model is integrated with a standard semi-batch reactor model with constant volume addition (of reagent B). Consideration of the cooling capacity of the reactor resulted in a limiting relationship between the operating policy variables of feed B addition time, \( t_{\text{add}} \), and isothermal temperature, \( T_{\text{iso}} \). For the purposes of this study, this relationship is well approximated with \( T_{\text{iso}} \) as a quadratic function of \( t_{\text{add}} \) (see Figure A1, Appendix A), since data regarding the energy balance is unavailable. The model equations used to describe this process are given in Model A1 (Appendix A).

6.3 Nominal optimisation

Sano et al. (1998), state that one of the objectives for the development of the model was to help determine the best operating conditions for maximum product yield, \( Y_C \). A reaction time, \( t_r \), of less than 8 hours (terminated when the rate of conversion of A falls below 0.1%) and a final yield in the impurity, \( Y_D \), of below 2% are desired be maintained. The optimal results the authors determined through repeated simulation of their model are given in Table 6.1. Optimisation of an equivalent model for maximum product yield subject to the stated constraints (Problem P4), using a non-linear constrained SQP optimisation software, are shown to compare reasonably well (Table 6.1).

Objective function,
Integrated design under uncertainty for pharmaceutical processes

\[
\max_{t_{iso}, t_{add}} Y_C
\]

subject to,

model equations (Model A1, Appendix A),

\[ t_f \leq 8 \]

\[ Y_D \leq 2\% \]

\[ 288 \leq T_{iso} \leq 313 \]

\[ 0.5 \leq t_{add} \leq 3.0 \]

\[ T_{iso} \geq 7.06(t_{add})^2 - 43.50(t_{add}) + 352.67 \]  

(Problem P4)

Table 6.1. Comparison between the optimal literature results (determined through repeated simulation) and optimal results obtained with SQP optimisation, Case Study I.

<table>
<thead>
<tr>
<th></th>
<th>Sano et al. (1998)</th>
<th>SQP optimisation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Criteria</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( Y_C ) (%)</td>
<td>Not given</td>
<td>97.1</td>
</tr>
<tr>
<td>( Y_D ) (%)</td>
<td>~1</td>
<td>1.4</td>
</tr>
<tr>
<td>( t_f ) (hr)</td>
<td>~7</td>
<td>6.7</td>
</tr>
<tr>
<td>Operating policy</td>
<td>( T_{iso} ) (K)</td>
<td>298.0</td>
</tr>
<tr>
<td></td>
<td>( t_{add} ) (hr)</td>
<td>1.7</td>
</tr>
</tbody>
</table>

6.4 Consideration of uncertainty in the stochastic model

Uncertainty in the model parameters could have a large effect on any results predicted by the model. This may be of particular importance regarding the optimal operating policy determined subject to the safety constraint and the desired limits on process performance.

Perturbation Analysis indicates that the uncertain parameters which have a non-negligible influence on yield of C, \( Y_C \), yield of D, \( Y_D \) and the final time, \( t_f \), are the kinetic rate law parameters (\( E_{a1}, A_1, E_{a2}, A_2, E_{a\text{diss}}, A_{\text{diss}} \)), the transition point from dissolution controlled kinetics to intrinsically controlled kinetics, \( X_{\text{diss}} \), and the isothermal temperature, \( T_{iso} \). These parameters are assumed to be uncorrelated and their assumed uncertainties are quantified in Table A1 (Appendix A). A total of 456 scenarios were required to satisfy the convergence criterion of 1% error in the mean and variance parameters over a backlog of 25 observations for both \( Y_D \) and \( t_f \). Graphs of the evolution of the mean and variance estimates of \( Y_D \) are shown in Figure A2 (a) and (b), Appendix A.
The main results from Sensitivity Analysis of the Uncertainty Analysis sample generated under the SQP nominal optimum conditions are shown in Table 6.2. Since the coefficients of determination computed on rank transformed data (RR^2) are greater than those on the unranked data (R^2), then sensitivity measures are based on rank transformed data (according to the Sensitivity Analysis schematic shown in Figure 5.5, Section 5.4). The effect of rank transformation is shown in the normalised scatter plot (see Figure A3 (a), Appendix A) between Yield D and the sub-reaction activation energy (Ea2). This plot shows that the ranking transforms the non-linear but monotonic relationship shown in the unranked data (dots in Figure A3 (a)) to points (circles in Figure A3 (a)) which may be better quantified by linear sensitivity measures. The main Sensitivity Analysis results in Table 6.2 indicate that the activation energy parameters in the intrinsic reaction rate laws (Ea1 and Ea2) are the most strongly related to the observed uncertainty in the output criteria (Table 6.2).

### Table 6.2. Key SRRC contributors to predicted uncertainty under the nominal optimum operating policy, Case Study I.

<table>
<thead>
<tr>
<th></th>
<th>Yield C</th>
<th>Yield D</th>
<th>Final time</th>
</tr>
</thead>
<tbody>
<tr>
<td>R^2</td>
<td>0.65</td>
<td>0.65</td>
<td>0.73</td>
</tr>
<tr>
<td>RR^2</td>
<td>0.85</td>
<td>0.80</td>
<td>0.99</td>
</tr>
<tr>
<td>Ea1</td>
<td>-0.62</td>
<td>0.60</td>
<td>0.99</td>
</tr>
<tr>
<td>A1</td>
<td>0.02</td>
<td>-0.02</td>
<td>-0.03</td>
</tr>
<tr>
<td>Ea2</td>
<td>0.74</td>
<td>-0.76</td>
<td>0.03</td>
</tr>
<tr>
<td>A2</td>
<td>-0.03</td>
<td>0.03</td>
<td>-0.00</td>
</tr>
</tbody>
</table>

#### 6.4.1 Stochastic process optimisation problem

The proposed optimisation problem under uncertainty aims to maximise a mean-variance objective in the product yield (Y_C). The operating policy decisions, t_add and T_iion, are scenario independent, assuming the a priori ‘here and now’ mode of control where knowledge of particular realisations of the uncertainties is not assumed in the optimal solution.

One-sided stochastic constraints are incorporated to maintain the desired limitations on the impurity yield (Y_D less than 2.0%) and the final time (t_f less than 8 hours). Since certain realisations within the uncertainty space may result in values of the endpoint impurity yield and final time exceeding the desired limits, expected violations of these limits of 1.5% and 1 hour, respectively, are permitted. This allows some tolerance to the desired limits (E_viol the summed extent of violation of those observations failing divided by the total number of observations) which permits the determination of optimal solutions which are not overly conservative.
To solve this problem the stochastic optimisation formulation is given in Problem P5. This is derived from the general formulation, Problem P2 (Section 5.5.2). The binary variable approximations, \( \beta \), return zero for scenarios failing the desired limits and one otherwise. The general deviation function, \( d_1 \), is zero for passing scenarios and \( d_2 \), is a linear deviation (\( \phi = \phi^0 \)) for failing scenarios. The stochastic optimisation algorithm shown in Figure 5.6 (Section 5.5) is used to solve the problem, where the Hammersley sequence sampling scheme is used to place observations in the uncertainty space. A reduced convergence criterion of ± 2% deviation in the output distribution parameters is permitted to reduce the number of observations per objective function evaluation.

Objective function

\[
\max \left\{ \kappa \frac{\mu_{\text{robust}}}{\mu_{\text{nominal}}} - (1 - \kappa) \frac{\sigma_{\text{robust}}^2}{\sigma_{\text{nominal}}^2} \right\}
\]

subject to:

process model equations (Model A1, Appendix A) \( \forall \ m = 1 \ldots M \)

binary variable approximations

\[
\beta_{Y_{D,m}} = \frac{1}{2} \left[ \tanh \left( c_{Y_D} \left( 2.0 - Y_{D,m} \right) \right) + 1 \right] \quad \forall \ m = 1 \ldots M
\]

\[
\beta_{t_f,m} = \frac{1}{2} \left[ \tanh \left( c_{t_f} \left( 8.0 - t_{f,m} \right) \right) + 1 \right] \quad \forall \ m = 1 \ldots M
\]

stochastic inequality constraints

\[
E_{\text{viol}} \left\{ Y_D \right\} = \frac{1}{M} \sum_{m=1}^{M} (1 - \beta_{Y_{D,m}}) (Y_{D,m} - 2.0) \leq 1.5
\]

\[
E_{\text{viol}} \left\{ t_f \right\} = \frac{1}{M} \sum_{m=1}^{M} (1 - \beta_{t_f,m}) (t_{f,m} - 8.0) \leq 1.0
\]

decision bounds

\( 0.50 \leq t_f \leq 3.00 \)

\( 288 \leq T_{iso} \leq 313 \)

uncertainty space

\[
E_{a_1, a_2, a_3, a_4, a_{\text{diss}}, a_{\text{diss}}, x_{\text{diss}}, T_{\text{iso}}} = \left\{ \theta_p | N \left( \mu_p, \sigma_p \right) \right\}
\]

(defined in Table A1, Appendix A)

\( \text{where} \quad p = \text{diss} \cdot 1.2 \) (Problem P5)
The values for the constants of the hyperbolic smoothing functions for the impurity yield and final time binary approximations, \( \xi_{yD} \) and \( \xi_{tf} \), were selected to be 160 and 40, respectively. With these values the binary approximation smoothing functions calculate zero and one for criteria values to within approximately 1% of the threshold values (see Figure A4, Appendix A).

### 6.4.2 Optimisation under uncertainty results

The results for the optimisations under uncertainty in the key parameters are given in Table 6.3, where the value for the product yield objective mean-variance weight (\( \kappa \)) is 0.5. These have been validated in an Uncertainty Analysis under the uncertainties identified in the original Perturbation Analysis (see Section 6.4). It is clear that under the nominal optimal operating policy decisions, the expected violation of the final time, \( E_{\text{viol}}(t_f) \), is significantly greater (4.49 hr) then the desired limit (1 hr). In addition, the expected yield of impurity, \( E(\gamma_D) \) at 2.75% with an expected violation, \( E_{\text{viol}}(\gamma_D) \), of 1.39% is not satisfactory.

Comparing the results obtained when uncertainty was considered, shown in Table 6.3, it can be immediately seen that an improvement in the \( Y_C \) mean-variance objective is achieved (largely due to the 47% reduction in the variance) with maintenance of both the stochastic constraints for \( E_{\text{viol}}(t_f) \) and \( E_{\text{viol}}(\gamma_D) \). Huge reductions in the both the expected final time (2.35 hr) and the 5-95% fractile width (5.17 hr) are observed, with no significant loss in the expected product yield or increase in the impurity yield.

Cumulative frequency plots, Figure 6.2, show the great improvement achieved with respect to the final time but only very slight reductions in the spread of the predicted product and impurity yields (5-95% fractile widths of 15.92 and 7.55 respectively). The plots show that the presence of uncertainty in the model parameters result in long tailed cumulative distributions for \( Y_C \), \( Y_D \) and \( t_f \) under the mid-range

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Nominal optimal operation</th>
<th>Uncertain optimal operation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scenarios</td>
<td>456</td>
<td>418</td>
</tr>
<tr>
<td>Mean-variance ( Y_C )</td>
<td>( [E(Y_C) %], \text{Var}(Y_C) ]</td>
<td>( [94.35, 50.4] )</td>
</tr>
<tr>
<td>( E(\gamma_D) %)</td>
<td>2.75</td>
<td>2.77</td>
</tr>
<tr>
<td>( E(t_f) ) (hr)</td>
<td>9.34</td>
<td>2.35</td>
</tr>
<tr>
<td>( FW(Y_C) %)</td>
<td>18.73</td>
<td>15.92</td>
</tr>
<tr>
<td>( FW(\gamma_D) %)</td>
<td>8.83</td>
<td>7.55</td>
</tr>
<tr>
<td>( FW(t_f) ) (hr)</td>
<td>24.97</td>
<td>5.17</td>
</tr>
<tr>
<td>([P_{\text{pass}}(\gamma_D \leq 2.0)], P_{\text{pass}}(t_f \leq 8.0)])</td>
<td>( [0.59, 0.59] )</td>
<td>( [0.53, 0.98] )</td>
</tr>
<tr>
<td>([E_{\text{viol}}(\gamma_D \leq 2.0)], E_{\text{viol}}(t_f \leq 8.0)])</td>
<td>( [1.39, 4.49] )</td>
<td>( [1.25, 0.05] )</td>
</tr>
<tr>
<td>Decisions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( t_{\text{add}} ) (hr)</td>
<td>1.79</td>
<td>1.12</td>
</tr>
<tr>
<td>( T_{\text{inc}} ) (K)</td>
<td>296.8</td>
<td>312.4</td>
</tr>
</tbody>
</table>
values of the nominal optimum operating policy (isothermal temperature, $T_{iso}$ of 296.8 K and feed addition time, $t_{add}$ of 1.79 hr). Hence, the poor performance in the one-sided $E_{viol}$ and $Pr_{pass}$ criteria (Table 6.3). The implementation of a higher $T_{iso}$ (312.4 K) and a shorter $t_{add}$ (1.12 hr) result in a greater rate of formation for both the product and impurity with a corresponding lower $t_f$. Under the model uncertainty a significant leftwards shift in the $t_f$ distribution and reduction in the length of the tail is observed in Figure 6.2 (c). This results in the vast improvement in $E_{viol}(t_f)$ and $Pr_{pass}(t_f)$ shown in Table 6.3. However, this operating policy does not significantly reduce the long tails in the $Y_C$ and $Y_D$ distributions.

Sensitivity Analysis under the robust optimal operating policy indicate the same key contributing

![Cumulative frequency plots for the validated isothermal optimisation results obtained with and without uncertainty consideration, Case Study 1.](image)

Key: • = nominal optimisation, o = optimisation under uncertainty, — = desired upper limit.
uncertain kinetic parameters as was observed under nominal policy (Ea1 and Ea2). The significant changes from the results under the nominal operating policy (given in Table 6.2) are the increasing strength of the linear ranked relations between the uncertainty in Ea2 with that in YC (0.88 from 0.74) and YD (-0.88 from -0.76), and the decreasing strength of the linear ranked relations between the uncertainty in Ea1 with that in YC (-0.42 from -0.62) and YD (0.41 from 0.60). This effect is shown in a normalised scatter plot between YD and Ea2 in Figure A3 (b), Appendix A. The effect of a shorter process completion time induced by a shorter addition feed time and higher isothermal temperature in the robust operating policy results in YD (and its uncertainty) being more strongly related to the uncertainty in Ea2 and less so to that in Ea1. It is postulated that this is probably due to the shorter periods in which the concentrations of A and C (strongly dependent on the uncertainty in Ea1) are at significant levels with the result that the uncertainty in the rate constant k2 propagated from Ea2 has an increasingly stronger influence on rate of formation of D.

Figure 6.3. Sensitivity towards the product yield mean-variance weight, Case Study I.

Key: \( \ast = \text{E}(Y_D) \), \( \circ = \text{Var}(Y_D) \).

The insensitivity of the value used for the mean-variance weight, \( \kappa \), to the robust optimum solution is shown in Figure 6.3. Only at a value of 1.00, for a maximisation of \( \text{E}(Y_C) \) with no variance consideration, is any significant change in the optimum solution observed. \( t_{add} \) increases to 1.50 hr and \( T_{iso} \) drops to 303.3 K resulting in a slight increase in \( \text{E}(Y_C) \) to 94.81% and an increase in \( \text{Var}(Y_C) \) to 30.91. In benefit, a reduced value for \( E_{viol}(Y_D) \) (1.13 from 1.25%) is obtained but at the expense of an increase in \( E_{viol}(t_i) \) (0.52 from 0.05). Clearly for this uncertain process some consideration of the variance in \( Y_C \) in the objective function is an important factor in the resulting decisions and stochastic criteria determined.
6.5 An alternative operating policy

From the optimal results under uncertainty shown in Table 6.3, the key area in which further improvement would be desired appears to be in $Y_D$. A robust optimisation minimising $E_{\text{sol}}(Y_D)$ determined a minimum value of 1.12% under the isothermal 'here and now' mode of temperature control. This indicates that the possibilities of reducing the expected violation in the $Y_D$ constraint under the specified uncertainties are limited by the isothermal operating policy. Sano et al. (1998) state that a higher quality process may be achieved with non-isothermal temperature control. Such a policy is considered in this sub-section, in order to increase the probability of passing the $Y_D$ constraint (only 53% under the robust isothermal optimal decisions) and to reduce the expected violation under model uncertainty. The dynamic optimisation problem without and with the consideration of uncertainty is formulated as an optimal control problem (Problem P6 and Problem P7, respectively) with piecewise constant temperature control, $T_{\text{non-iso},p,e}$, discretised over four intervals in the time horizon, $t_{pe}$.

Problem P4 (nominal) / Problem P5 (robust) with revised piecewise constant decision bounds.

\[
0.2 \leq t_{pe} \leq 3 \quad \forall \quad p,e = 1,4
\]

\[
288 \leq T_{\text{non-iso},p,e} \leq 313 \quad \forall \quad p,e = 1,4
\]

\[
t_f = \sum_{p,e=1}^{4} t_{pe}
\]

(Problem P6 / Problem P7)

For this system the optimal non-isothermal operating policy obtained under uncertainty was the same as that determined in the nominal optimal control problem (see Table 6.4) with the exception that the duration of the final control interval for each observation in the uncertain optimisation varies according to the 0.001 conversion rate termination criterion. The robust solution was verified with a number of different starting points. Subsequent Uncertainty Analysis results under the non-isothermal policy are compared with the robust isothermal case in Table 6.5. The cumulative frequency plots in Figure 6.4, show the improvement in the distributions of $Y_C$ and $Y_D$ and the deterioration in $t_f$ that is achieved with the non-isothermal policy.

Non-isothermal control allows significant improvements in the expected values of $Y_C$ (96.37%) and $Y_D$ (1.81%) coupled with reductions in the uncertainty (quantified by the widths between the 5-95% fractiles) in all three output criteria and a significant increase in $Pr_{\text{pass}}(Y_D)$ to 0.74 and an $E_{\text{sol}}(Y_D)$ of almost half the isothermal value. The only drawback is the increase in $E(t_f)$ to 3.96 hr from 2.35 hr coupled with slight deteriorations in $Pr_{\text{pass}}(t_f)$ and $E_{\text{sol}}(t_f)$. However, the latter is still well within the desired stochastic constraint limit of 1 hr with 96% of the distribution below the 8 hr limit.
Table 6.4. Nominal optimum operating policy under non-isothermal temperature control, Case Study I.

<table>
<thead>
<tr>
<th>Time interval</th>
<th>T_{non-iso}</th>
<th>Duration (hr)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>288.0</td>
<td>0.61</td>
</tr>
<tr>
<td>2</td>
<td>288.0</td>
<td>1.24</td>
</tr>
<tr>
<td>3</td>
<td>288.0</td>
<td>0.50</td>
</tr>
<tr>
<td>4</td>
<td>313.0</td>
<td>1.06</td>
</tr>
<tr>
<td>t_f (hr)</td>
<td></td>
<td>3.41</td>
</tr>
<tr>
<td>t_{add} (hr)</td>
<td></td>
<td>2.50</td>
</tr>
</tbody>
</table>

Table 6.5. Uncertainty Analysis results comparing and robust isothermal and non-isothermal temperature control, Case Study I.

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Robust isothermal</th>
<th>Nominal/robust non-isothermal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scenarios</td>
<td>418</td>
<td>532</td>
</tr>
<tr>
<td>Mean-variance [Y_C]</td>
<td>0.234</td>
<td>0.278</td>
</tr>
<tr>
<td>[E(Y_C) (%), Var(Y_C)]</td>
<td>[94.30, 26.9]</td>
<td>[96.37, 23.48]</td>
</tr>
<tr>
<td>E(Y_D) (%)</td>
<td>2.77</td>
<td>1.81</td>
</tr>
<tr>
<td>E(t_f) (hr)</td>
<td>2.35</td>
<td>3.96</td>
</tr>
<tr>
<td>FW(Y_C) (%)</td>
<td>15.92</td>
<td>11.89</td>
</tr>
<tr>
<td>FW(Y_D) (%)</td>
<td>7.55</td>
<td>5.85</td>
</tr>
<tr>
<td>FW(t_f) (hr)</td>
<td>5.17</td>
<td>4.85</td>
</tr>
<tr>
<td>[Pr_{pass}(Y_D \leq 2.0), Pr_{pass}(t_f \leq 8.0)]</td>
<td>[0.53, 0.98]</td>
<td>[0.74, 0.96]</td>
</tr>
<tr>
<td>[E_{viol}(Y_D \leq 2.0), E_{viol}(t_f \leq 8.0)]</td>
<td>[1.25, 0.05]</td>
<td>[0.68, 0.08]</td>
</tr>
</tbody>
</table>

The optimal non-isothermal policy consists of an initial minimum temperature period (288 K for 2.35 hr) followed by a maximum temperature period (313.0 K of scenario dependent duration according to the fulfilment of the process termination criterion). This means that virtually the entire dissolution rate limiting period (2.38 hr) occurs during the initial low temperature interval (2.35 hr) in the nominal case. Uncertainty Analysis indicates that for 38% of the total observations (532) the transition between the dissolution and intrinsic kinetic control is achieved before the end of the low temperature interval and 78% before the end of the feed addition (2.50 hr). The longer feed addition time and lower rate of product formation result in a reduced concentration of product in the reactor during the initial low temperature period, which conspires to minimise the driving force for the impurity formation. A shorter period at a high temperature achieves the termination criterion with a reduced opportunity for impurity formation. Hence, the desired movement of the distributions for Y_D and Y_C under the model parameter uncertainty, with the compromising movement of the t_f distribution to the right, as shown in Figure 6.4. The
Integrated design under uncertainty for pharmaceutical processes

(a) Yield of product, C.

(b) Yield of impurity, D.

(c) Final time.

Figure 6.4. Cumulative frequency plots for the validated non-isothermal optimisation results, Case Study 1.

Key: ● = non-isothermal policy (nominal and robust), ○ = isothermal policy (robust), — = desired upper limit.

Postulation behind the optimal non-isothermal policy is supported for the considerations both with and without model uncertainty, since Problem P6 and Problem P7 converged to the same policy.

6.6 Conclusions

Responses to the presence of uncertainty in a model-based approach have been investigated for a multiphase reaction process comprising Case Study 1. Uncertainty Analysis was used to quantify the effect of model uncertainty on the output performance predictions. Subsequent Sensitivity Analysis identifies the
uncertainty in the estimated intrinsic reaction kinetic parameters as being critical to the uncertainty in the output criteria. In response to the first management option (see Section 5.3) it may be surmised that focusing of experimental effort towards increasing the confidence in these parameters would provide the greatest reduction in the output uncertainties as predicted by the process model. Secondly, a stochastic process optimisation problem was solved which determined an operating policy providing a significantly improved prediction for the distribution location and spread of the final process time under uncertainty compared to the nominal optimum policy. However, the predicted product and impurity yield distributions remained relatively unaffected by this solution. Thirdly, an alternative non-isothermal operating policy was considered. Uncertainty Analysis showed that the nominal optimal operating policy managed to improve the location and reduce the uncertainty in the distributions of both the product and impurity yield criteria at the expense of an adverse response in the process time distribution. For this case, optimisation under uncertainty did not provide an improved solution and the same solution as for the nominal non-isothermal optimisation was obtained.

The information obtained using the proposed Risk Analysis approach aims to allow development engineers to make more informed decisions as to how to best improve the potential process performance and exploit the process opportunities despite the uncertainty in the models developed to structure the available process knowledge. While it is prudent to consider uncertainty when using models to aid process development, it can be essential if optimisation techniques are also used to determine processes better able to provide desired performances. For Case Study I, this was shown to be true for the isothermal operating policy but not for the non-isothermal case. In addition, since it was possible to explain why the optimal operating policies might be better with regard to the uncertainty in the assumed physico-chemical phenomena and the required manipulation of the output criteria distributions, some confidence in the results under uncertainty may be derived.

In Chapters 7, 8 and 9 the proposed approach for design under uncertainty is extended in a more comprehensive case study concerning an industrial process for a therapeutic product, comprised of a complete sequence of integrated operations.
Chapter 7

CASE STUDY II: RISK ANALYSIS APPROACH TO A PROCESS SEQUENCE UNDER UNCERTAINTY

7.1 Introduction

Application of a method for the robust optimisation of a 'here and now' operating policy under model parameter uncertainty for a published pharmaceutical reaction process is portrayed in the previous chapter. It shows that a more robust process can be obtained with respect to some criteria but not in others. In Chapter 7, 8 and 9, a case study comprising a complete integrated process sequence is presented. In this chapter Uncertainty and Sensitivity Analysis techniques are employed to manage the evolving uncertainty in developing models with incoming process modelling information. Then optimisation for desired levels of prediction uncertainty reduction is used to show the required levels of input parameter uncertainty reduction. With respect to the methodology, the objectives of the case study are to provide:

1. verification of the methodology, comprising of,
2. a quantification of uncertainty (Chapter 7),
3. a priority list for process knowledge (Chapter 7),
4. a quantification of the effect of increased process knowledge (Chapter 7),
5. analysis of optimal trade-off between the reduction in performance criteria uncertainty and that required in the uncertain parameters (Chapter 7),
6. an optimal robust operating policy in the available decision variables (Chapter 8),
7. the value of information for potentially measurable process inherent uncertainties, (Chapter 8),
8. a measure of tolerance to error in the optimal robust operating policy variables (Chapter 8),
9. a basis for the assessment of process flowsheet alternatives under uncertainty (Chapter 9).

The process operations which constitute the integrated sequence are described in Section 7.2. This is followed by a brief evaluation of the modelling effort justifiable in relation to the possible characterising phenomena and the available data, Section 7.3. In Section 7.4, the Risk Analysis problem is defined for a first generation of models based on the initial data available and assumptions. The results of the Uncertainty Analysis and the Sensitivity Analysis are discussed. Iterative results of the effect of new data on the modelling effort in the Risk Analysis approach are summarised in Section 7.5. Minimum reductions in input model parameter uncertainty required to achieve desired reductions in the uncertainty predicted in the important output criteria are determined in Section 7.6.
7.2 Process Description

Case Study II is based on data provided by a pharmaceutical company. It is derived from a process for the production of a chemical drug. The process objective is the production of a crystalline drug product of key component actB and of consistent purity with respect to the by-product impurities, actC and actE, from feed solids comprising the active pharmaceutical ingredient (API), actA, and a stereo-isomer impurity.
Integrated design under uncertainty for pharmaceutical processes

actD. The considered sequence comprising of 15 stages is shown in Figure 7.1. The key process conditions are summarised in Table 7.1, where LOD is the level of dampness in solids (defined in the filtration model -Model B6, Appendix B).

An aqueous-organic liquid phase chemical reaction takes place in Stage 1. The description for this process stage is based on the details obtained from private communication with a pharmaceutical company. The stoichiometric reactions believed to be occurring are shown in Equations 7.1 to 7.4. The reaction objective is to produce the drug product, actB, from the feed. Feed solids of actA, stereo-isomer actD and

Table 7.1. Summary of key process operating conditions for Case Study II

<table>
<thead>
<tr>
<th>Stage</th>
<th>Key process conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Reaction</td>
<td>77 wt% ± 3 wt% actA purity API feed solids and catalyst. 2 mol eq. (ratio to feed actA) reG solids. 10.4 mol eq. (ratio to feed actA) 30% aq. reH solution, controlled addition rate to maintain a constant temperature, ( T_1 ). Maintain a constant temperature, ( T_1 ) ± 1 °C, throughout entire reaction. Termination at ~90-95% conversion of actA (typically 2.5-3 hours).</td>
</tr>
<tr>
<td>2. Dilution</td>
<td>1 volume eq. (ratio to Stage 1 reH) distilled water. 15 min agitation period.</td>
</tr>
<tr>
<td>3. Layer separation</td>
<td>30 min settling period. Drain heavy organic phase to parallel vessel.</td>
</tr>
<tr>
<td>4. reH destruction</td>
<td>0.4 mol eq. (ratio to feed actA) 6% aq. baseJ solution per shot, pending litmus paper test for residual reH presence. 15 min agitation period.</td>
</tr>
<tr>
<td>5. reG destruction</td>
<td>0.7 mol eq. (ratio to feed actA) 50% aq. baseJ solution. Agitate mixture for 120 min at a constant temperature, ( T_5 ).</td>
</tr>
<tr>
<td>6. Layer separation</td>
<td>30 min settling period. Drain heavy organic phase to original vessel.</td>
</tr>
<tr>
<td>7. pH neutralisation</td>
<td>0.7 mol eq. (ratio to actA feed) baseK solids 15 min agitation period.</td>
</tr>
<tr>
<td>8. Layer separation</td>
<td>30 min settling period. Drain heavy organic phase to parallel vessel.</td>
</tr>
<tr>
<td>9. solL distillation</td>
<td>Distil solL until vessel minimum stir volume is reached. 1 bar pressure and zero reflux.</td>
</tr>
<tr>
<td>10. 1st solL distillation</td>
<td>Add a fraction of the total solL volume:product ratio between 14 and 15. Distil a fraction of the solL. 1 bar pressure and zero reflux.</td>
</tr>
<tr>
<td>11. 2nd solL distillation</td>
<td>Add remaining fraction of the total solL volume:product ratio. Distil solL to achieve a final solL volume:product ratio between 7 and 8. 1 bar pressure and zero reflux.</td>
</tr>
<tr>
<td>12. Crystallisation in solL</td>
<td>Cool boiling mixture to 25 °C and hold for 60 min. 1 bar pressure.</td>
</tr>
<tr>
<td>13. Filtration</td>
<td>Vacuum filter the slurry at a constant temperature, ( T_{13} ), until ~W_{13}% LOD is achieved.</td>
</tr>
<tr>
<td>14. Washing</td>
<td>Rinse the damp solids with a 2 volume:product ratio of pure solL at a constant temperature, ( T_{14} ), and refilter to the prior LOD.</td>
</tr>
<tr>
<td>15. Drying</td>
<td>Dry with pure N₂ at a high temperature to a low LOD value ~W_{15}%.</td>
</tr>
</tbody>
</table>
reagent reG dissolve in the organic solvent, solF. Controlled addition of the aqueous reagent, reH, leads to the production of an oxidant, oxG, in a reaction between aqueous reH and dissolved reG, believed to occur at the aqueous-organic phase interface, see Figure 7.2. This oxidant reacts with the dissolved drug components in the organic phase. The key feed API (actA) is oxidised to the desired product (actB). Overoxidation leads to the formation of actC from actB in a consecutive reaction. actC is the critical impurity believed to cause problems in the morphology during the crystallisation of the final product. In parallel, the feed impurity (actD) is oxidised to a secondary impurity (actE).

\[
\begin{align*}
\text{reG} + \text{reH} & \rightarrow \text{oxG} \\
\text{actA} + \text{oxG} & \rightarrow \text{actB} \\
\text{actB} + \text{oxG} & \rightarrow \text{actC} \\
\text{actD} + \text{oxG} & \rightarrow \text{actE}
\end{align*}
\]

(7.1) \hspace{1cm} (7.2) \hspace{1cm} (7.3) \hspace{1cm} (7.4)

The following seven operations, Stages 2 to 8, provide a termination of the reaction and treatment of the residual reagents. The termination is precipitated by rapid dilution of the aqueous solution of reH with addition of distilled water followed by a period of stirring, Stage 2. An aqueous-organic layer separation is conducted after a period of settling, Stage 3. The heavy organic phase is drained to another vessel and the aqueous phase is sent to waste. Residual reH and reG are destroyed using aqueous solutions of baseI and baseJ in Stage 4 and Stage 5, respectively. The organic phase is drained back to the original vessel in another layer separation, Stage 6. An aqueous solution of baseK is charged to effect a pH neutralisation, Stage 7, followed by a final layer separation, Stage 8.

The next three operations effect a solvent exchange from solF for a crystallisation from an organic solvent, solL. solF is distilled from the vessel to a minimum concentrate, Stage 9. A fraction of a

![Figure 7.2. Stage 1 Bisphasic chemical reaction, Case Study II.](image-url)
Integrated design under uncertainty for pharmaceutical processes

A predetermined total volume of solL is added in Stage 10. Some fraction of this is removed in the Stage 10 distillation, after which the second solL fraction is added and distilled (Stage 11). This aims to maintain a desired initial and final solL volume to product mass ratio, without violating the maximum boiling volume limit of the vessel.

The final four operations involve the purification of the product. A crystallisation from solL aims to remove actC and actE to acceptable levels in the product solids (Stage 12). This also acts to remove unreacted actA and actD. The solids are filtered and then washed in pure solL to remove residual solution containing dissolved impurities before drying (Stages 13, 14 and 15).

Some important process issues concerning the product yield and final quality are indicated:

- The controlled addition of reH to the reactor is necessary to prevent a potentially strong exothermic reH reaction and maintain a constant low temperature, T₁.
- Maintenance of a constant low temperature in the reaction helps prevent the possibility of increased impurity formation.
- The molar charges for the chemical reagents are estimated based on the initial moles of actA feed in Stage 1, and not based on measurements of the species to be quenched. The exception is Stage 4, where the complete destruction of residual reH is ensured with the utilisation of a litmus paper test, and additional basel charges.
- There is a potential for drug loss in Stage 5, up to 1-2 wt% of the product yield. The presence of excess baseJ may lead to a product decomposition reaction, due to either an incorrect baseJ molar charge, or a prolonged stirring period.
- The desired values for initial and final solL volume to product mass ratios are obtained from statistical design of experiments. These ranges are believed to be important to the behaviour of the subsequent crystallisation process.
- The addition and distillation split of the total solL make up is imposed due to the maximum volume limit allowed for boiling within the given vessel. The split may be an important factor regarding the quantity of residual solF present in the Stage 12 crystallisation stream and the total process sequence time.
- A higher crystallisation temperature may lead to yield loss due to higher solubility of the product while a lower temperature may mean increased impurity content in the solids.
- Low solubility of product in solL at the controlled wash temperature, T₁₄, means that dissolution of solid product in the solL wash is assumed to be negligible.

These observations provide limitations on the operating policy and the underlying issues should provide incentives for the use of a modelling-based approach. It is important that such observations can be considered in such an approach though there may be little mechanistic understanding behind them.
7.3 Modelling effort

For this case study the modelling effort aims to predict the quantity of product and impurities obtained at the end of the sequence. For optimisation purposes it is important that key control variables are characterised in the models. Lumped parameter models are used which may or may not be dynamic in nature. The identification of the important controlling factors, the key physico-chemical phenomena which characterise the desired state transformations, is an important step in a systematic model building process, after the definition of the modelling objective. A list of the physico-chemical phenomena which might characterise the processes of Case Study II, are shown in Table 7.2.

To evaluate the realistic characterisation of these phenomena within a model-based approach, the following factors need to be considered:

- the benefits which may be obtained from using the models,
- the difficulty in developing the models due to the complexity of the phenomena,
- availability of related literature,
- availability of relevant process data and ease of measurement of relevant process variables,
- opportunities and availability of resource to obtain more data for model validation and refinement.

Given these factors, an effort was made to evaluate the realistic modelling opportunity for these phenomena, Table 7.2. Where no modelling opportunity is indicated for the characteristic phenomena assumed to be present, a very simple mass balance model is implemented instead. Data from a pilot plant run at the R&D facilities of a pharmaceutical company (process run name PPR), is available for comparison. Specific operating conditions for the PPR plant run given in Table 7.3, are in relation to the general process statements in Table 7.1. PPR measurement data is given in Table 7.4. A first generation of process models, based on bench scale data where available and otherwise engineering assumptions, are provided in Appendix B.
Table 7.2. Possible characterising phenomena for the process operations of Case Study II.

<table>
<thead>
<tr>
<th>Stage</th>
<th>Physico-chemical phenomena</th>
<th>Modelling opportunity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Reaction</td>
<td>reG feed solids dissolution</td>
<td>√ Trivial level, limited data</td>
</tr>
<tr>
<td></td>
<td>Reaction stoichiometry</td>
<td>√ Profile data</td>
</tr>
<tr>
<td></td>
<td>Intrinsic reaction kinetics</td>
<td>√ Profile data</td>
</tr>
<tr>
<td></td>
<td>Interfacial mass transfer/mixing</td>
<td>√ Empirical correlation</td>
</tr>
<tr>
<td></td>
<td>Drug mass transfer to aqueous phase</td>
<td>√ Simple tie-line model</td>
</tr>
<tr>
<td></td>
<td>Heat transfer</td>
<td>√ Energy balance possible, no temperature related reaction rate data</td>
</tr>
<tr>
<td>2. Dilution</td>
<td>Reaction stoichiometry</td>
<td>× Unclear reagent roles</td>
</tr>
<tr>
<td></td>
<td>Intrinsic reaction kinetics</td>
<td>? Depends on understanding concerning reagents</td>
</tr>
<tr>
<td></td>
<td>Interfacial mass transfer/mixing</td>
<td>× No available data</td>
</tr>
<tr>
<td></td>
<td>Drug mass transfer to aqueous phase</td>
<td>√ Simple tie-line model</td>
</tr>
<tr>
<td>3. Layer separation</td>
<td>Drop sedimentation rate</td>
<td>× No available data, complex</td>
</tr>
<tr>
<td></td>
<td>Drop coalescence rate</td>
<td>× No available data, complex</td>
</tr>
<tr>
<td>4. reH destruction</td>
<td>reH/baseJ reaction stoichiometry</td>
<td>× No available data</td>
</tr>
<tr>
<td></td>
<td>reH/baseJ intrinsic reaction kinetics</td>
<td>× No available data</td>
</tr>
<tr>
<td></td>
<td>Interfacial mass transfer/mixing</td>
<td>× No available data</td>
</tr>
<tr>
<td></td>
<td>Drug mass transfer to aqueous phase</td>
<td>√ Simple tie-line model</td>
</tr>
<tr>
<td>5. reG destruction</td>
<td>reG/baseJ reaction stoichiometry</td>
<td>× No available data</td>
</tr>
<tr>
<td></td>
<td>reG/baseJ intrinsic reaction kinetics</td>
<td>× No available data</td>
</tr>
<tr>
<td></td>
<td>Drug/baseJ reaction stoichiometry</td>
<td>× No available data</td>
</tr>
<tr>
<td></td>
<td>Drug/baseJ intrinsic reaction kinetics</td>
<td>× No available data</td>
</tr>
<tr>
<td></td>
<td>Interfacial mass transfer/mixing</td>
<td>× No available data</td>
</tr>
<tr>
<td></td>
<td>Drug mass transfer to aqueous phase</td>
<td>√ Simple tie-line model</td>
</tr>
<tr>
<td>6. Layer separation</td>
<td>As for Stage 3</td>
<td></td>
</tr>
<tr>
<td>7. pH neutralisation</td>
<td>Acid-base reaction stoichiometry</td>
<td>× No available data</td>
</tr>
<tr>
<td></td>
<td>Acid-base intrinsic reaction kinetics</td>
<td>× No available data</td>
</tr>
<tr>
<td></td>
<td>Interfacial mass transfer/mixing</td>
<td>× No available data</td>
</tr>
<tr>
<td></td>
<td>Drug mass transfer to aqueous phase</td>
<td>√ Simple tie-line model</td>
</tr>
<tr>
<td>8. Layer separation</td>
<td>As for Stage 3</td>
<td></td>
</tr>
<tr>
<td>9. solF distillation</td>
<td>Vapour-liquid equilibria</td>
<td>√ Energy balance possible</td>
</tr>
<tr>
<td></td>
<td>Heat transfer</td>
<td></td>
</tr>
<tr>
<td>10. 1&lt;sup&gt;st&lt;/sup&gt; solL distillation</td>
<td>As for Stage 9</td>
<td></td>
</tr>
<tr>
<td>11. 2&lt;sup&gt;nd&lt;/sup&gt; solL distillation</td>
<td>As for Stage 9</td>
<td></td>
</tr>
<tr>
<td>12. Crystallisation</td>
<td>Crystal growth kinetics</td>
<td>√ Assume simple seeded</td>
</tr>
<tr>
<td></td>
<td>Nucleation kinetics</td>
<td>× No available data, complex</td>
</tr>
<tr>
<td></td>
<td>Heat transfer</td>
<td>× No available data</td>
</tr>
<tr>
<td></td>
<td>Crystal birth/death phenomena</td>
<td>× No available data, complex</td>
</tr>
<tr>
<td></td>
<td>Impurity effects</td>
<td>√ Trivial model, lack of understanding and data</td>
</tr>
<tr>
<td>13. Filtration</td>
<td>Liquid mass transfer</td>
<td>× No available data</td>
</tr>
<tr>
<td>14. Washing</td>
<td>Liquid mass transfer/displacement</td>
<td>× No available data</td>
</tr>
<tr>
<td>15. Drying</td>
<td>Heat transfer</td>
<td>× No available data, complex</td>
</tr>
<tr>
<td></td>
<td>Liquid mass transfer</td>
<td>× No available data, complex</td>
</tr>
</tbody>
</table>
Table 7.3. PPR plant run conditions, Case Study II.

<table>
<thead>
<tr>
<th>Stage</th>
<th>PPR plant run conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Operation time</td>
</tr>
<tr>
<td></td>
<td>Total drug feed mass</td>
</tr>
<tr>
<td></td>
<td>Drug feed purity</td>
</tr>
<tr>
<td></td>
<td>solF</td>
</tr>
<tr>
<td></td>
<td>30% aq. reH</td>
</tr>
<tr>
<td></td>
<td>reH addition time</td>
</tr>
<tr>
<td></td>
<td>Agitation speed</td>
</tr>
<tr>
<td>2</td>
<td>Distilled water addition</td>
</tr>
<tr>
<td>4</td>
<td>6% aq. baseJ solution</td>
</tr>
<tr>
<td>5</td>
<td>50% aq. baseK solution</td>
</tr>
<tr>
<td>7</td>
<td>35% aq. baseK solution</td>
</tr>
<tr>
<td>9</td>
<td>Minimum stir volume</td>
</tr>
<tr>
<td>10/11</td>
<td>Total sol</td>
</tr>
<tr>
<td></td>
<td>Total sol</td>
</tr>
<tr>
<td>12</td>
<td>Cooling rate (assumed)</td>
</tr>
</tbody>
</table>

Table 7.4. PPR plant run measurements, Case Study II.

<table>
<thead>
<tr>
<th>Stage</th>
<th>PPR plant run measurements</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>X_{actA}</td>
</tr>
<tr>
<td></td>
<td>actB content</td>
</tr>
<tr>
<td></td>
<td>actC content</td>
</tr>
<tr>
<td></td>
<td>actE content</td>
</tr>
<tr>
<td></td>
<td>Post reactor crude product mass</td>
</tr>
<tr>
<td>10/11</td>
<td>Cumulative initial sol</td>
</tr>
<tr>
<td>11</td>
<td>Final sol</td>
</tr>
<tr>
<td></td>
<td>Pre-crystallisation crude product mass</td>
</tr>
<tr>
<td>14</td>
<td>Post filtration LOD</td>
</tr>
<tr>
<td>15</td>
<td>Dry product mass</td>
</tr>
<tr>
<td></td>
<td>Post drying LOD</td>
</tr>
<tr>
<td></td>
<td>actB dry content</td>
</tr>
<tr>
<td></td>
<td>actC dry content</td>
</tr>
<tr>
<td></td>
<td>actE dry content</td>
</tr>
</tbody>
</table>

7.4 Risk Analysis

In the previous section the sequence of operations comprising Case Study II is described. The first generation of process models, given in Appendix B, provide the first outcome of the iterative process of model development for the complete sequence. In this section, the effect of the large amounts of uncertainty concerning the first generation of process models in the integrated sequence is quantified.
using Uncertainty and Sensitivity Analyses. The effect that additional process data may bring to the models and the assumed uncertainty within the sequence is also considered.

In the hypothetical situation that data from complete process sequence runs at subsequent scale-ups becomes available, it is proposed that the methodology is used to analyse the predicted uncertainty in the current level of models with respect to this data, to determine where the models need to be developed further. This corresponds to the validation step in the systematic model development process shown in Figure 5.2 (Section 5.4). Since such data is not available for this case study, the PPR plant data (Table 7.4) is used as the benchmark with which to assess the effect of the uncertainty in the predicted criteria using models and parameter uncertainty estimated or assumed to correspond to the current levels of knowledge available. The corresponding PPR plant conditions given in Table 7.3 are used in this assessment. Estimation of the model parameters and quantification of the parameter uncertainty is based on actual data where available.

7.4.1 Stochastic problem for the first generation model set

The first generation stochastic system to which the Uncertainty and Sensitivity Analyses are applied with respect to the schematic shown in Figure 5.2 (Section 5.4) is defined in this section. The main characteristics of the first generation set of deterministic models and the uncertainty inputs to the problem

<table>
<thead>
<tr>
<th>Stage</th>
<th>Operation</th>
<th>Main deterministic model characteristics</th>
<th>Uncertainty sources</th>
<th>Reference (Appendix B)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Reaction</td>
<td>First order reaction kinetics.</td>
<td>Kinetic rate parameters</td>
<td>Model B1</td>
</tr>
<tr>
<td>2,4,5,7</td>
<td>Reagent addition</td>
<td>Two-phase mass balance with fractional drug loss assumption due to aqueous phase solubility.</td>
<td>Fractional drug loss parameter from organic phase</td>
<td>Model B2</td>
</tr>
<tr>
<td>3,6,8</td>
<td>Layer separation</td>
<td>Two-phase mass balance with fractional organic phase loss due to imperfect phase cut.</td>
<td>Fractional organic phase cut loss parameter</td>
<td>Model B3</td>
</tr>
<tr>
<td>9,10,11</td>
<td>Distillation</td>
<td>Batch distillation assuming ideal VLE with specified reboiler duty for estimation of operation time.</td>
<td>VLE coefficients</td>
<td>Model B4</td>
</tr>
<tr>
<td>12</td>
<td>Crystallisation</td>
<td>Crystal growth kinetics based on solute saturation in a seeded batch cooling regime, with drug impurity solute concentration 'loss'.</td>
<td>Crystal growth rate and impurity 'loss' parameters and saturation data</td>
<td>Model B5</td>
</tr>
<tr>
<td>13</td>
<td>Filtration</td>
<td>Two-phase mass balance attaining a desired moisture hold-up.</td>
<td>Moisture hold-up and filtration rate</td>
<td>Model B6</td>
</tr>
<tr>
<td>14</td>
<td>Washing</td>
<td>Two-phase mass balance with moisture displacement with wash.</td>
<td>Moisture hold-up and displacement</td>
<td>Model B7</td>
</tr>
<tr>
<td>15</td>
<td>Drying</td>
<td>Two-phase mass balance attaining a desired moisture hold-up.</td>
<td>Moisture hold-up and drying rate</td>
<td>Model B8</td>
</tr>
</tbody>
</table>
are summarised in Table 7.5.

The uncertain inputs to the stochastic system comprise of those model parameters which generate a significant response in important output criteria when individually perturbed from their nominal values. Characterisations of the 29 parameter uncertainties assumed in the first generation models are shown in Table B2 (Appendix B). For the consecutive reaction rate constants, $k_1$ and $k_2$, estimated simultaneously using non-linear least squares, the degree of correlation and the parameter standard deviations are estimated using a first term Taylor series expansion for the covariance matrix (see Section 5.4 and Appendix B for details). For uncertain parameters not estimated simultaneously, standard deviations are estimated from the relevant data or assumed (based on the standard deviation being a percentage of the nominal value) or for uniformly distributed uncertainty, range limits are estimated or assumed.

The non-maintenance of initial (14-15) and final (7-8) desired volumes of soll solvent to expected mass of actB prior to crystallisation, identified in Section 7.2 as an important observation with regard to the final crystal impurity content, is given some account in the stochastic problem. Violations of these desired operating ranges are assumed to give a greater uncertainty in parameters characterising downstream criteria believed to be related but for which there is no mechanistic understanding (see Equation B3, Appendix B, for mathematical definition). The effect of this added uncertainty is shown in Figure B10 (Appendix B) where violation of these desired solvent to product ranges leads to an increase in the uncertainty in the impurity concentration 'loss' parameter (from solution) and a corresponding increase in the uncertainty in endpoint key impurity content of the crystals.

### 7.4.2 First generation model sequence results

A total of 431 scenarios were required to satisfy the multiple 1% mean and variance parameter error convergence criteria (Equation 5.14, Section 5.4) for both the total yield ($Y_T$) and endpoint key impurity content ($w_{actC}$). Graphs of the evolution of the mean and variance of the total yield are shown in Figure

<table>
<thead>
<tr>
<th>Stage</th>
<th>Criteria</th>
<th>PPR Plant data</th>
<th>Predicted mean</th>
<th>Predicted fractiles [5%, 95%]</th>
<th>Data proximity to fractile interval, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$X_{actA}$</td>
<td>0.955</td>
<td>0.993</td>
<td>[0.987, 0.997]</td>
<td>307</td>
</tr>
<tr>
<td></td>
<td>actB product composition, wt%</td>
<td>71.5</td>
<td>72.9</td>
<td>[72.1, 73.4]</td>
<td>49</td>
</tr>
<tr>
<td></td>
<td>actC impurity composition, wt%</td>
<td>0.74</td>
<td>1.21</td>
<td>[0.89, 1.48]</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>actE impurity composition, wt%</td>
<td>2.88</td>
<td>8.53</td>
<td>[7.45, 9.51]</td>
<td>221</td>
</tr>
<tr>
<td>11</td>
<td>Final soll:product ratio, dm$^3$ kg$^{-1}$</td>
<td>7.2</td>
<td>6.94</td>
<td>[6.83, 7.04]</td>
<td>76</td>
</tr>
<tr>
<td>15</td>
<td>actB product content, wt%</td>
<td>89.4</td>
<td>90.52</td>
<td>[88.0, 93.5]</td>
<td>inside</td>
</tr>
<tr>
<td></td>
<td>actC impurity content, wt%</td>
<td>0.24</td>
<td>0.33</td>
<td>[0.19, 0.45]</td>
<td>inside</td>
</tr>
<tr>
<td></td>
<td>actE impurity content, wt%</td>
<td>1.4</td>
<td>2.78</td>
<td>[1.72, 3.70]</td>
<td>21</td>
</tr>
<tr>
<td>1-15</td>
<td>Total yield, %</td>
<td>84.2</td>
<td>90.2</td>
<td>[87.2, 92.4]</td>
<td>59</td>
</tr>
</tbody>
</table>

Table 7.6. Summary of first generation model results under uncertainty, Case Study II.
Total yield, % Stage I Conversion

Figure 7.3. Cumulative frequency plots for the first generation model set predictions under uncertainty, Case Study II. Key: * = predicted results, --- = PPR data, --- = 5%, 95% fractiles.

B11 (a) and (b), Appendix B. Scatter plots showing the effect of the induced rank correlation procedure for the Stage I \( k_1 \) and \( k_2 \) uncertain parameters, are shown in Figure B12 (Appendix B). The results of the Uncertainty Analysis are summarised in Table 7.6. 5% and 95% fractiles are used to quantify the predicted uncertainty in the endpoint output and certain inter-stage criteria. If the data does not fall within the 5-95% fractiles of the predictions then the proximity values (the final column in Table 7.6) show how close the data is to the nearest fractile (5 or 95%), relative to the fractile width.

For example, in this case study the cumulative frequency plot in Figure 7.3 (a) shows the predicted distribution in endpoint total yield under uncertainty in the first generation process model sequence relative to the independent PPR plant data (the solid vertical line in the Figure). Since the plant data for this criterion does not fall within the predicted uncertainty as enclosed within the 5% and 95% fractiles (the dashed lines in the Figure), clearly some process models of the first generation model set may not be suitable for prediction of the total yield at the PPR plant scale (as would be expected). Since model parameter error has quantitatively been accounted, an element of structural error may be suspected. The extent of the error in the prediction distribution indicated by the proximity values, are particularly large for the Stage I conversion (307%) and secondary impurity composition (221%). It would appear that the large over prediction of the conversion in Stage 1, Figure 7.3 (b), contributes to the observed under prediction of the final solL to product ratio and over prediction of the Stage 1-15 total yield. A similar assessment may be made concerning the over prediction of the secondary impurity content in the final product.

Sensitivity Analysis is used to estimate the key contributions to the predicted output uncertainty with regard to the propagation of uncertainty in certain inter-stage process properties and the individual input uncertainties of the stochastic model. These indicate the relative importance of the uncertainty in the
model parameters characterising the current state of knowledge in the available process models and the associated phenomena.

Contributory process sub-sequences are defined by potentially viable data measurements in this case study. It is initially assumed that only the inter-stage criterion of reaction conversion is a measurement which can be used to define the sub-sequences for the endpoint total yield. For the impurities, it is assumed that the post reaction stream is a potential inter-stage measurement. The relative contributions of the sub-sequences associated to inter-stage criteria are shown in Table 7.7. These estimate the fraction of the total uncertainty (quantified as the width between the 5 and 95% fractiles) in the endpoint criterion which is attributed to each specified process sub-sequence. The initial indication is that the Stage 2 to 15 sub-sequence contributes the most uncertainty to the predicted uncertainty in the total yield (82% of the final uncertainty), while the Stage 1 reaction contributes the least (18% of the final uncertainty). The model parameter uncertainties in the models for Stages 2-15 provide a much larger contribution to the uncertainty in the endpoint total yield than the Stage 1 model parameter uncertainties. To reduce the uncertainty in the total yield, the sub-sequence contributors immediately indicate that further work should be focused on the models and uncertainties assumed in Stages 2-15. This is different to implying that the models for Stages 2-15 are the main cause for the deviation in total yield from the PPR plant data. The prediction of the key impurity in Stage 1 introduces a greater uncertainty in the final product content than all the following operations. The opposite is apparent for the secondary impurity.

Table 7.7. First generation model sub-sequence contributions to predicted uncertainty, Case Study II.

<table>
<thead>
<tr>
<th>Sub-sequence</th>
<th>Total yield</th>
<th>Key impurity</th>
<th>Secondary impurity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stage 1</td>
<td>0.18</td>
<td>0.62</td>
<td>0.34</td>
</tr>
<tr>
<td>Stage 2 to 15</td>
<td>0.82</td>
<td>0.38</td>
<td>0.66</td>
</tr>
<tr>
<td>Endpoint</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
</tbody>
</table>

If a sample of the pre-crystallisation crude were to be available for analysis, then the Stage 2-15 sub-sequence contribution of 0.82 for total yield uncertainty could be decomposed. The new contributions to the total yield uncertainty and impurity contents, are shown in Table 7.8. Now it is indicated that significant contributions from both the Stage 2-11 (0.35) and Stage 12-15 (0.47) sub-sequences are apparent. However, Stages 2-11 do not appear to introduce any additional uncertainty to the impurity compositions predicted from Stage 1. This is reasonable considering the realistic assumptions incorporated in the first generation reagent addition and layer separation models that any drug loss is independent of the concentrations of the other drug species and that there is no loss in the distillations.

Estimated values of the coefficient of determination close to unity for total yield, key and secondary impurity contents for unranked data (0.97, 0.98 and 0.99 respectively) and ranked data (0.97, 0.96 and
0.97 respectively) indicate that the linear input parameter contributors predicted by the Sensitivity Analysis should be reliable and rank transformation of the data is not required. Confirmation of this assumption is determined by examining the scatter plots between the stochastic inputs and outputs. Scatter plots between the Stage 12 crystal growth constant (k_p) and total yield and between key impurity content and the Stage 12 crystallisation key impurity 'solute loss' parameter (Cacc) are shown in Figure B13 (a) and (b), Appendix B. The relationships indicate that they are adequately measured using linear measures.

Table 7.8. Effect of more inter-stage measurements to sub-sequence contributions, Case Study II.

<table>
<thead>
<tr>
<th>Sub-sequence</th>
<th>Total yield</th>
<th>Key impurity</th>
<th>Secondary impurity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stage 1</td>
<td>0.18</td>
<td>0.62</td>
<td>0.34</td>
</tr>
<tr>
<td>Stage 2 to 11</td>
<td>0.35</td>
<td>-0.00</td>
<td>-0.00</td>
</tr>
<tr>
<td>Stages 12-15</td>
<td>0.47</td>
<td>0.38</td>
<td>0.66</td>
</tr>
<tr>
<td>Endpoint</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Figure 7.4. Uncertain parameter contributor measures for endpoint criteria, Case Study II.
The correlation coefficients (CC) and standardised regression coefficients (SRC) over all the uncertain parameters are shown in Figure 7.4 (a), where the parameter index numbers are specified in Table B2 (Appendix B). The key contributor parameters to the observed uncertainty in the endpoint yield as indicated by the CCs are \(k_6\) (0.74, index 16) then \(k_4\) (0.38, index 1) and \(k_2\) (-0.38, index 2). However, the induced correlation between the product formation reaction rate constant, \(k_1\), and the key impurity formation reaction rate constant, \(k_2\), in the sampling procedure results in a false estimation of the influence of \(k_2\) to total yield from the CCs. This is because the CCs do not measure the standardised global influence. The SRCs show that the influence of \(k_2\) (-0.05) is actually negligible compared to \(k_1\) (0.32), as would be expected. The parameters associated with product loss from the reagent addition and layer separation stages (indices 6 to 12) provide minor contributions (SRCs between -0.15 to -0.20). The key SRC contributors to the final impurity content are estimated as the key impurity ‘solute loss’ parameter, \(\zeta_{act}\) (0.78, index 21) then \(k_2\) (0.77, index 2) and then the wash efficiency, \(\eta_{wash}\) (-0.21, index 26), as shown in Figure 7.4 (b). A similar ranking is predicted for the secondary impurity but with a stronger influence on the secondary impurity ‘solute loss’ parameter, \(\zeta_{actE}\) (SRC of 0.96 compared to 0.32 for \(k_3\)) due to better fit of the reaction model for the parallel reaction to the bench scale data. The parameters identified as key contributors and priority do not provide any surprising results.

As may be expected the estimated key uncertain parameter contributors coincide with stages contained within the key sub-sequences. Whereas the sub-sequence contributors are useful in providing an initial idea to the key areas of the process sequence contributing uncertainty and provide a measure of the accumulation of uncertainty at specific points in the sequence based on certain (potentially measurable) inter-stage and endpoint outputs of the stochastic model, the CCs and SRCs provide a ranking of importance in the uncertain inputs which can differentiate between a large number of individual sources.

<table>
<thead>
<tr>
<th>Total yield</th>
<th>Key Stage</th>
<th>Key parameter</th>
<th>Characterised phenomena</th>
<th>Possible related phenomena</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>(k_6)</td>
<td>Growth kinetics</td>
<td>Nucleation kinetics, mixing etc.</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>(k_4)</td>
<td>Intrinsic kinetics</td>
<td>Intrinsic reagent kinetics, solids dissolution, mixing</td>
<td></td>
</tr>
<tr>
<td>2, 4, 7</td>
<td>(u_1)</td>
<td>Solubility loss</td>
<td>Mass transfer rate, equilibrium solubility</td>
<td></td>
</tr>
<tr>
<td>3, 6, 8</td>
<td>(u_2)</td>
<td>Imperfect phase cut</td>
<td>Phase dispersion band, drop entrapment</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Key and secondary impurity content</th>
<th>Key Stage</th>
<th>Key parameter</th>
<th>Characterised phenomena</th>
<th>Possible related phenomena</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>(\zeta_{act})</td>
<td>Impurity ‘solute loss’ rate</td>
<td>Various - very complex</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>(k_2)</td>
<td>Intrinsic kinetics</td>
<td>Intrinsic reagent kinetics, solids dissolution, mixing</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>(\eta_{wash})</td>
<td>Moisture displacement</td>
<td>Mass transfer rate</td>
<td></td>
</tr>
</tbody>
</table>
The identified key process sub-sequences and key parameter contributors of the predicted endpoint uncertainty in the whole process sequence can be used to provide a guide to the key phenomena which are not well characterised and introduce large amounts of uncertainty. Given the identified quantitative indicators for the predicted uncertainty, indicated in Table 7.9, a phenomenological knowledge priority list can be inferred, based on engineering intuition. With this information the data required to progress the model development can be ascertained, either to reduce the uncertainty associated with the parameters of the current model structures if the uncertain prediction encompasses the current data, or to develop models incorporating different phenomena and/or equations. The next step of the methodology is to determine how the effect of incoming data and knowledge can improve the current predictions (i.e. following Step 17 in Figure 5.2, Section 5.4).

7.5 Effect of new data

In the previous section, analysis of the process sequence under uncertainty indicated the key contributions to the predicted uncertainty observed in output criteria, given the model assumptions and uncertainties associated with the current data and level of knowledge. This comprised the first generation of process models (see Appendix B). In this section, the effect on the prediction of the key output criteria under uncertainty with developing models based on incoming data and observations, is investigated. Uncertainty and Sensitivity Analyses are used as before.

The order by which the new process data is incorporated into the system is not driven by the results of the methodology, since only a hypothetical situation of the availability of process development data is represented in this case study and only a limited amount of data is incorporated. Otherwise the methodology could provide a driving force for the collection of certain data to progress the model-based approach, where identified to be needed. The assumption made in this study is that new data is incorporated as and when it becomes available. The availability of new data is given in Table 7.10, in the assumed order of incorporation. The revised models and parameter uncertainties associated with each level of incorporated knowledge are given in Appendix C. The semi-empirical reactor model is developed to incorporate further limiting phenomena to account for the key responses. The layer separation and reagent addition models are transformed from speculative assumptions towards more mechanistic type models.

Tracking of the uncertainty (depicted by the predicted 5-95% fractiles) with sequential incorporation of knowledge into the system models, as data becomes available, is shown in Figure 7.5 (a) and Figure 7.6 (a) for endpoint yield and key impurity content predictions, respectively. The cumulative frequency plots for total yield and endpoint key impurity content given in Figures D1 and D2 (Appendix D) respectively, show the effect of new data on the distributions of these criteria. The respective relevant inter-stage criteria of conversion and post reactor key impurity composition (dashed fractile lines in Figure 7.5 (a) and Figure 7.6 (a)) indicate how much how uncertainty has accumulated and propagated between the Stage 1 and the Stage 15 predictions. Comparing the relative fractile interval widths between the Stage 1
and Stage 15 criteria it appears that while a large proportion of the uncertainty in the endpoint yield evolves after Stage 1, the uncertainty in the endpoint key impurity content is mainly due to that generated in Stage 1.

The addition of new data either leads to revised uncertainty characteristics for parameters in the same

![Graph](image)

(a) Predicted 5-95% fractiles relative to the plant data.

(b) Percent deviation of mean prediction relative to the plant data.

Figure 7.5. Effect of knowledge incorporation to total yield and conversion predictions, Case Study II.

Key: • = Total yield, o = Conversion.
model structures (knowledge levels 1 and 6) or to new models structures with different uncertain parameters (knowledge levels 2, 3, 4 and 5), as indicated in Table 7.10. In the former circumstance, new sets of data are required to re-estimate the uncertainty in the existing model parameters. This does not guarantee a reduction in the uncertainty of the predicted output if there is a wide spread in the new data sets or the model structure is inadequate. Hence, there is an increase in the predicted uncertainty in endpoint yield for knowledge levels 1 and 6 in Figure 7.5 (a). In the latter circumstance, while the predicted endpoint uncertainty may not decrease with the addition of new data, it is hoped that the prediction of the new model is closer to the data. This is indicated from the deviations in the mean prediction from the PPR plant data that are shown in Figure 7.5 (b) and Figure 7.6 (b) for total yield and key impurity content, respectively. For knowledge levels 2, 3, 4 and 5, the total yield predictions become closer to the data as the model structures are changed, and similarly for levels 3 and 6 for the key impurity content.

Consideration of these deviations with the associated uncertainties, provides an indication of the quality of the model system with respect to both the spread in the predicted distribution and accuracy relative to the data. The history of the Stage 1 conversion predictions, Figure 7.5 (a) and (b) shows that both the proximity of the mean prediction to the data, and the relative spread in the distribution are required to ascertain the quality of the model. At knowledge level 1, the mean prediction is closer to the data and under uncertainty the fractiles encompass the data. However, the large increase in the uncertainty of the prediction from level 0 to level 1 (an increase of over 800% in the fractile width) indicates a problem with the model. While the incorporation of more knowledge, in level 2, reduces the uncertainty, the prediction moves away from the data. Only at level 3, when the mixing phenomena are modelled in Stage 1, does the

![Figure 7.6](image_url)  
(a) Predicted 5-95% fractiles relative to the plant data.  
(b) Percent deviation of mean prediction relative to the plant data.

Figure 7.6. Effect of knowledge incorporation to key impurity composition predictions, Case Study II.  
Key: • = Dry crystal key impurity content, o = Post reaction crude key impurity composition.
Integrated design under uncertainty for pharmaceutical processes

prediction in the reaction conversion to the data improve without such a substantial increase in the uncertainty. The uncertainty in the predicted post reactor key impurity composition and endpoint content, Figure 7.6 (a), are also reduced at level 3, and the prediction accuracy to the data increases, Figure 7.6 (b).

The change in the accumulation of the uncertainty through the sequence as knowledge is incorporated can be analysed by the comparison between the relative magnitude of the inter-stage criteria fractile widths, as illustrated in Figure 7.5 (a) for the conversion and total yield criteria. A clearer representation is provided with the sub-sequence contributions.

The contributions to the total yield uncertainty for Stage 1 and Stage 2-15 sub-sequences are shown in Figure 7.7 (a). It is no surprise that the contribution of Stage 1 to the final uncertainty becomes larger than that of Stage 2-15 at level 1, when the parameters of the bench scale Stage 1 model are fitted to the larger scale data. This is redressed with the revised model, level 2. At level 3, the contribution of Stage 1 increases again due to the uncertainty in the mixing correlation employed. With the addition of the pre-crystallisation solL to product ratio as a measured inter-stage criterion, Figure 7.7 (b), it is indicated that the Stage 2-11 sub-sequence becomes an increasingly minor contributor compared to the Stage 12-15 sub-sequence with the incorporation of extra layer separation and drug solubility knowledge.

The importance of the deterministic model structure on the propagation of uncertainty in the stochastic model is demonstrated by the increase in Stage 2-15 sub-sequence contribution to the total yield at knowledge level 6, as shown in Figure 7.7 (b). At high values of the crystal growth rate constant (k_g) the Stage 12 crystalliser model predicts that the solute concentration approaches the saturation concentration.

Figure 7.7. Effect of incorporation of knowledge to the total yield uncertainty, Case Study II.
This suppresses the effect of uncertainty in \( k_g \) to the total yield. At knowledge level 6, the lower revised value of \( k_g \) results in a reduction in the suppression effect the model structure has on its uncertainty, despite no change in the relative level of uncertainty in \( k_g \).

The stages associated with the key contributing model parameters, identified using CCs and SRCs, do not necessarily coincide with the key contributing sub-sequences. For knowledge levels 0, 1 and 2, the unassociated variabilities in the Stage 12 key impurity 'solute loss' parameter (\( \zeta_{\text{solute}} \)) and the reaction rate constant for key impurity formation (\( k_2 \)) appear to each explain a similar fraction of the variability in the endpoint key impurity content (SRCs of 0.78, 0.61 and 0.72 for \( \zeta_{\text{solute}} \) compared to 0.77, 0.58 and 0.66 for \( k_2 \), Table D1, Appendix D). A scatter plot for the knowledge level 0 case, Figure D3 (Appendix D) does not indicate a greater relationship between either of these two inputs to the endpoint impurity composition. However, the Stage 1 sub-sequence appears to contribute a significantly greater proportion of the endpoint uncertainty than the Stage 2-15 sub-sequence as shown in Figure 7.8. Cumulative frequency plots, Figure D4 (Appendix D), show that the magnitude of the uncertainty in the endpoint composition relative to the uncertainty in the post reactor composition is not much greater. The propagation of minor uncertain inputs in Stage 1 provide an accumulation of uncertainty which overrides the single effect of the uncertainty in \( \zeta_{\text{solute}} \) in Stage 12. In this case it is important to differentiate between the key contributing sub-sequences and parameters. Focus on all the uncertainties in the Stage 1 sub-sequence would be more beneficial than on the Stage 2-15 sub-sequence, with regard to the uncertainty in the endpoint key impurity composition.

As the deterministic models are revised to accommodate different phenomena and different uncertain parameters are introduced the priority of the uncertainty contributors change. The final list of knowledge priorities at knowledge level 6 is given in Table 7.11. Compared to the priorities estimated at level 1 (Table 7.9) the main contributors to the prediction uncertainty in endpoint yield remain the crystallisation

![Figure 7.8. Effect of knowledge incorporation to the endpoint key impurity uncertainty. Case Study II.](image)

Key: o = Stage 1 contribution, • = Stage 2-15 contribution.
Table 7.11. List of key parameters and knowledge priorities in the final generation of models (knowledge level 6), in ascending order of priority, Case Study II.

<table>
<thead>
<tr>
<th>Total yield</th>
<th>Key Stage</th>
<th>Key parameter</th>
<th>Characterised phenomena</th>
<th>Possible related phenomena</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>$k_g$</td>
<td>Growth kinetics</td>
<td>Nucleation kinetics, mixing etc. various - complex</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>$k_1$</td>
<td>Intrinsic pseudo first order drug reaction kinetics</td>
<td>Intrinsic reagent-drug kinetics</td>
<td></td>
</tr>
<tr>
<td>2, 4, 7</td>
<td>$\sigma_{sl}$</td>
<td>Organic-aqueous phase drug solubility</td>
<td>Mass transfer rate</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>$\gamma_1, \gamma_2$</td>
<td>Rate limiting mixing case</td>
<td>Eddy formation and imperfect energy dissipation</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Key impurity content</th>
<th>Key Stage</th>
<th>Key parameter</th>
<th>Characterised phenomena</th>
<th>Possible related phenomena</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$k_2$</td>
<td>Intrinsic pseudo first order drug reaction kinetics</td>
<td>Intrinsic reagent-drug kinetics</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>$c_{actC}$</td>
<td>Impurity solute ‘loss’ rate</td>
<td>Various, very complex molecular scale phenomena</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>$\eta_{wash}$</td>
<td>Solution displacement</td>
<td>Mass transfer rate</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>$\gamma_1, \gamma_2$</td>
<td>Rate limiting mixing case</td>
<td>Eddy formation and imperfect energy dissipation</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Secondary impurity content</th>
<th>Key Stage</th>
<th>Key parameter</th>
<th>Characterised phenomena</th>
<th>Possible related phenomena</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$k_3$</td>
<td>Intrinsic pseudo first order drug reaction kinetics</td>
<td>Intrinsic reagent-drug kinetics</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>$c_{actE}$</td>
<td>Impurity solute ‘loss’ rate</td>
<td>Various, very complex molecular scale phenomena</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>$\eta_{wash}$</td>
<td>Solution displacement</td>
<td>Mass transfer rate</td>
<td></td>
</tr>
</tbody>
</table>

and reaction rate constants and drug-aqueous solubility parameters. Uncertainty in the layer separation parameters provide no significant contributions. Uncertainty in the Stage 1 reaction rate constants have become more important to the endpoint impurity than the Stage 12 impurity ‘solute loss’ parameters. At the expense of a more accurate deterministic Stage 1 model, the addition of the mixing effect (level 3) introduces minor contributions to the uncertainty in total yield and key impurity content. The SRC values of key parameter contributors for knowledge levels 0 to 6 are given in Table D1 (Appendix D).

### 7.6 Optimal uncertainty reduction

The key uncertain parameter contributors to the uncertainty in total yield, key and secondary impurity content predictions for knowledge levels 6 have already been identified in Table 7.11. The extent of the reduction in the uncertainty of these key contributors required to meet a specified reduction in the predicted output criteria can be quantified. An optimisation problem of the general formulation given in Problem P1 (Section 5.5.1) is solved for the knowledge level 6 process model system, Problem P8, objective function
Integrated design under uncertainty for pharmaceutical processes

\[
\max_{\delta_{\sigma_1}, \delta_{\sigma_2}, \delta_{\sigma_3}, \delta_{\sigma_{unc}}', \delta_{\eta_{wash}}} = \sum [\delta_{\sigma_1}, \delta_{\sigma_2}, \delta_{\sigma_3}, \delta_{\sigma_{unc}}', \delta_{\eta_{wash}}]
\]

subject to,

- deterministic process stage model equations \( \forall \ s = 1..S, \ m = 1..M \)

stochastic quality constraints

\[
FW_{5\%, 95\%, \gamma_{r}} \leq \alpha_{\gamma_{r}} FW'_{5\%, 95\%, \gamma_{r}}
\]
\[
FW_{5\%, 95\%, \text{wt}_{actC}} \leq \alpha_{\text{wt}_{actC}} FW'_{5\%, 95\%, \text{wt}_{actC}}
\]
\[
FW_{5\%, 95\%, \text{wt}_{actE}} \leq \alpha_{\text{wt}_{actE}} FW'_{5\%, 95\%, \text{wt}_{actE}}
\]

decision variable bounds

\[
0 < \delta_{\sigma_1} \leq 1, \ 0 < \delta_{\sigma_2} \leq 1, \ 0 < \delta_{\sigma_3} \leq 1, \ 0 < \delta_{\eta_{wash}} \leq 1
\]

uncertainty space

\[
\eta_{wash} = \left[ \eta_{wash} \bigg| U\left( \mu_{\eta_{wash}}', \delta_{\eta_{wash}}, \Delta\eta_{wash}, \mu_{\eta_{wash}}' + \delta_{\eta_{wash}}, \Delta\eta_{wash} \right) \right]
\]

\[
\Delta\eta_{wash} = \eta^{UB}_{wash} - \mu_{\eta_{wash}} = \mu_{\eta_{wash}} - \eta^{LB}_{wash}
\]

\[
k_1 = \left[ k_1 \big| N(\mu_k', \delta_k, \sigma_k') \right], \quad k_2 = \left[ k_2 \big| N(\mu_k', \delta_k, \sigma_k') \right]
\]

\[
\sigma_{\mu_{actC}} = \left[ \sigma_{\mu_{actC}} \big| N(\mu_{actC}', \delta_{actC}, \sigma_{actC}') \right], \quad \sigma_{\mu_{actE}} = \left[ \sigma_{\mu_{actE}} \big| N(\mu_{actE}', \delta_{actE}, \sigma_{actE}') \right]
\]

Stochastic inequality constraints for \( \alpha\% \) reduction in the width of the predicted 5-95\% fractile intervals for total yield, \( Y_T \), key and secondary impurity content, \( \text{wt}_{actC} \) and \( \text{wt}_{actE} \), are maintained by the minimum reduction in the standard deviations of the normal distributions of:

- the Stage 1 kinetic uncertain parameters (\( k_1 \) and \( k_3 \)),
- the Stage 12 crystal growth rate parameter (\( k_2 \)),
- the equilibrium drug-aqueous phase solubility parameters (\( \sigma_{\mu_{actC}} \) and \( \sigma_{\mu_{actE}} \)),
- the Stage 12 drug component impurity 'solute loss' parameters (\( \zeta_{actC} \) and \( \zeta_{actE} \)).
and by the tightening of the lower and upper bounds about the mean of the uniformly distributed Stage 14 wash efficiency parameter ($\eta_{\text{wash}}$). An equivalent reduction in the uncertainty in the Stage 1 consecutive reaction rate constant ($k_2$) is assumed to the reduction determined in the uncertainty of $k_1$. These inputs are the key contributing uncertainties to the output uncertainty, identified from the knowledge level 6 Sensitivity Analysis (Table D1, Appendix D). It is assumed that $\alpha_T, \alpha_{\text{actC}}, \alpha_{\text{actE}}$ are equal to each other in Problem P8.

Problem P8 was solved parametrically at different values of $\alpha$ to obtain a trade-off curve, Figure 7.9, between the required levels of total key input parameter uncertainty reduction (objective function in Problem P8) required to meet the desired uncertainty reductions in the output criteria. It appears that the required level of parameter reduction increases sharply (objective function decreases) after a desired combined output criteria uncertainty reduction of 50% (from the original levels). For desired reductions of 70% and greater, the problem was infeasible suggesting that other non-key parameter sources of uncertainty (which were not included as reducing decision variables) have become significant.

Individual relationships between key parameter uncertainty reductions and desired output criteria uncertainty reductions are shown in Figure 7.10 (a) for $k_1$, $k_3$, $k_p$, $\sigma_{\text{actC}}$, and Figure 7.10 (b) for $\zeta_{\text{actC}}$, $\zeta_{\text{actE}}$, $\eta_{\text{wash}}$. These graphs show that the uncertainty in $k_1$, $k_3$ and $k_p$ consistently needs to be reduced for all desired reductions in output uncertainty while reductions in uncertainty in the other considered parameters $\sigma_{\text{actC}}$, $\zeta_{\text{actC}}$, $\zeta_{\text{actE}}$ and $\eta_{\text{wash}}$ need only be obtained for desired output uncertainty reductions of greater than 30 or 40%. Associating the key parameters to the process stages indicates where research effort should be directed for different levels of desired output criteria uncertainty reduction. For the endpoint impurity contents, reduction in the uncertainty in the reaction kinetics ($k_3$ and $k_2$ via the correlation with $k_1$) is more important than the crystallisation parameter uncertainties ($\zeta_{\text{actC}}$ and $\zeta_{\text{actE}}$) until approximately 50-60%}

![Figure 7.9](image_url)
uncertainty reduction in the impurity contents. No such distinction between parameter uncertainties can be made regarding uncertainty reduction for the total yield. The optimisation results are tabulated in Table E1 (Appendix E) and computational statistics are presented in Table E9 (Appendix E).

The change in the sensitivities of the key uncertain parameters to the output criteria are shown in Figure 7.11 regarding absolute SRC measures ($R^2 > RR^2 > 0.90$ for all cases). Regarding the total yield, Figure 7.11 (a) indicates that as the uncertainty in the Stage 1 reaction rate constant for the product ($k_1$), Stage 12 crystal growth rate constant ($k_2$) and the drug-aqueous solubilities ($\sigma_d$) decrease as optimally determined (Figure 7.10 (a)), the SRC values measuring the contributions of the uncertainty in the Stage 1 time at which the initial rate limited period ends ($t''$, defined in Model C1, Appendix C) and the reaction mixing coefficient ($\gamma_1$, defined in Model C2, Appendix C) increase. A similar effect is shown in Figure 7.11 (b) regarding the endpoint key impurity content except that the Stage 1 reaction rate constant for the key impurity ($k_2$) and the Stage 14 wash efficiency ($\eta_{\text{wash}}$) replace $k_1$ and $\sigma_d$.

These plots indicate the desired levels of uncertainty reduction in the output criteria which may be achieved before the key contributor sensitivities change whereupon a change in the experimental and modelling effort would become necessary to provide further uncertainty reduction in the output criteria, due to the extent of the reductions in the original key parameter uncertainties. For this case study, beyond approximately 50-60% reduction in uncertainty in the total yield and impurity content output criteria it becomes more beneficial to reduce uncertainty in different uncertain parameters ($t''$ and $\gamma_1$), which would require a redirection of the experimental effort since these parameters are associated with different phenomena, Stage 1 initial rate limiting period and mixing regime (and may sometimes be associated with different process stages, though not in this case).

![Figure 7.10. Optimal degree of key input parameter uncertainty reduction (from original uncertain values) required to meet desired levels of output criteria uncertainty reduction, Case Study II.](image-url)
7.7 Conclusions

A comprehensive case study comprising of an integrated sequence of 15 process operations is introduced. In this chapter the integration of the proposed Risk Analysis methods with model development iterations as more process information becomes available and the knowledge incorporated, is demonstrated.

For Case Study II it is shown that as more information was incorporated into the process models the predicted distributions in the output criteria do compare more favourably with the independent pilot plant data. In general, the levels of uncertainty decrease with model development iterations but not in every instance. Stream variable uncertainties may be amplified or dampened as they propagate through the sequence. This is affected by the form of the model equations. An amplifying effect in the uncertainties in the Stage 12 solute concentration and total product yield was observed for lower values of the uncertain crystallisation growth rate constant ($k_g$) at knowledge level 6. The incorporation of three sets of data can be identified as key to the improvement of the distribution characteristics of the total yield prediction. The first and most critical is the 1000 US gallon Stage 1 reaction data at different agitation rates leading to the development of the mixing case model (knowledge level 3). The second is the incorporation of drug aqueous-organic phase solubility data in the development of the generic aqueous reagent addition model (knowledge level 5). The third is the incorporation of larger scale crystallisation yield data (knowledge level 6). The knowledge level 3 data is also critical in the improvement and reduction in the uncertainty of the prediction for the endpoint key impurity content (given the initial laboratory model, knowledge level 0).

Figure 7.11. Change in absolute SRC sensitivities between output criteria and key uncertain parameters with optimal uncertainty reductions, Case Study II.
Sensitivity Analysis showed that the relative contributions to uncertainty in the predicted process yield became approximately equivalent between the Stage I reaction and the downstream sequence. In the latter the contribution due to the uncertainty in the Stage 12 crystallisation grew while that due to the Stage 2-11 layer separation and solvent exchange operations receded. Uncertainty in the endpoint impurity content predictions were estimated to be mainly due to the uncertainty in the reaction model for all the knowledge level Risk Analysis iterations. In short and perhaps unsurprisingly, the key uncertain parameters (regarding the endpoint total yield and impurity contents) were associated with the intrinsic reaction kinetics and the crystallisation process, indicating the areas to which development efforts should be directed to increase the understanding and confidence in the process. For the final generation of models presented (knowledge level 6 models), optimal reduction in the uncertainties of the key uncertain parameter contributors is determined for increasing levels of desired uncertainty reduction in the output criteria. The indication is that uncertainty reduction in the reaction rate and crystal growth rate constants would be beneficial to obtain any degree of total yield and endpoint impurity content uncertainty reduction while additional reduction in the other less critical parameters only become necessary once a certain output uncertainty reduction threshold has been passed (Stage 12 crystallisation impurity 'solute loss' parameters and Stage 14 wash efficiency uncertainties become important at 50-60% output uncertainty reductions and Stage I transition time from initial rate limiting period and reaction mixing coefficient uncertainties become important beyond 60%).

As the Case Study II results show, the Risk Analysis methods applied in this chapter permit the quantification and tracking of the combined influence of parameter uncertainties contained in the entire sequence of process models as they are developed in the systematic model development procedures with the progression of process development. The Sensitivity Analysis methods allow the efficient estimation of the key uncertain parameters of the stochastic system and the internal sub-sequence contributions from the results of the Uncertainty Analysis. It is proposed that the information obtained may be used to ascertain levels of uncertainty and focus data collection and modelling effort towards those parts of the process sequence in which the uncertainty has the greatest influence on the output.
Chapter 8

CASE STUDY II: BASE CASE PROCESS FLOWSHEET OPTIMISATION UNDER UNCERTAINTY

8.1 Introduction

In the previous chapter the Risk Analysis approach was applied to the evolving models comprising the process sequence of Case Study II. The key uncertainty contributors were identified and the extent of input uncertainty reduction required to meet desired output levels was determined for the final generation of models. In this chapter the Case Study II investigation is extended to process flowsheet optimisation under uncertainty. The process described by the final generation of models in Chapter 7 (knowledge level 6 models), provides the Base Case flowsheet for this investigation. The general formulations for process flowsheet optimisation under uncertainty (Problem P2) and operating policy tolerance (Problem P3) are restated for the Base Case. The issues addressed in this chapter are:

- optimisation of an economic capacity of the Base Case flowsheet through manipulation of the operating policy variables under no uncertainty in comparison to stochastic optimisation accounting for model uncertainties,
- the effect on optimal solutions of different characterisations of the input uncertainties with respect to an increase in the size and a change in the shape of the stochastic input distributions,
- the value of perfect information with regard to the potential uncertainty in the purity of the feed API,
- the maximum tolerance to error in the implementation of the Base Case optimum operating policy under uncertainty.

8.2 Base Case problem formulations

The optimisation problems for the nominal flowsheet optimisation, the flowsheet optimisation under uncertainty and the operating policy tolerance optimisation are stated for the Base Case in this subsection.

8.2.1 Nominal flowsheet optimisation problem

Optimisation of the Base Case process flowsheet is based on a profitability objective function. The profitability, $P_{ty}$, is defined as the revenue from the end product, less the cost of the main solvents on a basis of the total feed mass of active pharmaceutical ingredient (API) charged to the reactor (Stage 1), $F_{act A,1}$, and the total batch processing time, $t_T$. The units for the profitability are dollars per kilogram of
API feed per hour of processing time. The values of the selling price, $C_{\text{drug}}$ and solvent costs, $C_{\text{solF}}$ and $C_{\text{solL}}$, are assumed at 2000 $\text{kg}^{-1}$ product, 5 $\text{kg}^{-1}$ solF solvent and 10 $\text{kg}^{-1}$ solL solvent, respectively.

It is assumed that six operating policy variables are available for the optimisation of the base process flowsheet, as follows:

- the Stage 1 agitation speed $N_1$ (rpm),
- the Stage 1 duration time, $t_{f,1}$ (min),
- the fraction of the total solL solvent used for the crystallisation (added over Stages 10 and 11), $AF_{10}$,
  added in Stage 10 with the remainder added in Stage 11,
- the fraction of the solL added in Stage 10 removed in the subsequent distillation, $RF_{10}$,
- the Stage 12 linear crystallisation cooling rate, $CR_{12}$ ($^\circ\text{C} \text{min}^{-1}$),
- the Stage 12 crystallisation holding period, $HP_{12}$ (min).

The optimisation problem under no uncertainty for maximum profitability with constraint limits on the endpoint key and secondary impurity contents and the pre-crystallisation solF composition, is shown in Problem P9. A fixed reboiler duty is assumed in the Stage 9, 10 and 11 batch distillation models.

objective function

$$\max_{t_{f,1}, N_1, AF_{10}, RF_{10}, CR_{12}, HP_{12}} Pr_y = \frac{Y_T}{100} \frac{1}{t_f} \left\{ C_{\text{drug}} - C_{\text{solF}} \sum_{s=1}^{15} \frac{F_{\text{solF},s}}{Z_{\text{drug},s}} - C_{\text{solL}} \sum_{s=1}^{15} \frac{F_{\text{solL},s}}{Z_{\text{drug},s}} \right\}$$

subject to:

process model equations

$$Y_T = \frac{Z_{\text{drug},15}}{F_{\text{actA,1}}} \times 100\%$$

inequality constraints

$$w_{\text{solF},11} \leq 0.5$$

$$w_{\text{actC},15} \leq 0.3$$

$$w_{\text{actE},15} \leq 2.0$$

decision bounds

$$200 \leq t_{f,1} \leq 400$$
where $Y_T$ is the product component (actB) yield over the entire process sequence based on the quantity of API feed.

### 8.2.2 Stochastic flowsheet optimisation problem

The proposed optimisation problem under uncertainty aims to maximise the expected potential profitability of the Base Case process flowsheet. However, certain realisations within the uncertainty space may result in poor process performances in the endpoint impurity contents. A potential loss in profitability is modelled as the average profitability which is lost due to violation of desired upper limits on either the key or secondary endpoint impurity contents (0.3 and 2.0 wt%, respectively). Some tolerance to these violations is allowed to reduce the tendency towards overly conservative solutions. This tolerance is quantified by the incorporation of a one-sided stochastic constraint allowing an average profitability loss of up to 3.0 $ \text{kg}_{\text{actA}}^{-1} \text{hr}^{-1}$.

The operating policy decisions are scenario independent, assuming the *a priori* 'here and now' mode of robust control where knowledge of particular realisations of the model parameter uncertainties is not assumed in the optimal operating policy solution. This results in six decision variables, as in the nominal optimisation given in Problem P9.

To solve this problem the stochastic optimisation formulation for the base flowsheet is given in Problem P10. This is derived from the general formulation, Problem P2, Section 5.5.2. The first stochastic inequality constraint tries to maintain an expected pre-crystallisation solF stream composition below 0.5 wt%. In this constraint the general continuous deviation function in Problem P10, $f_{\text{dev}}$, is replaced by the solF content at each scenario. The second stochastic inequality constraint maintains an average potential profitability loss below 3.0 $ \text{kg}_{\text{actA}}^{-1} \text{hr}^{-1}$. The impurity content binary variables, $\beta$, are one if the constraint thresholds are passed and zero otherwise. A profitability loss is returned if either the key or secondary content thresholds are violated. However, the profitability loss is not incorporated into the profitability objective since it is assumed only to be a potential loss which may be rectified with further purification iterations at further expense (not included in this problem). The associated general deviation functions in the general Problem P2, $d_1$ and $d_2$, are respectively zero and $Pty$ in Problem P10, i.e. the profitability loss is not a function of the extent of the impurity content violation. The resulting continuous deviation function, $f_{\text{dev}}$, is defined as the potential profitability loss.
Integrated design under uncertainty for pharmaceutical processes

The size of the problem is reduced by limiting the uncertainty space to sensitive inputs to the output criteria, as defined in the uncertainty space characterisation in Problem P10. A reduced convergence criterion of ± 2% deviation in the output distribution parameters is permitted to reduce the number of scenarios per objective function evaluation.

Objective function

$$\max_{f_{1.1}, N_1, AF_{10}, RF_{10}, CR_{12}, HP_{12}} \ E\{p_{y}\} = \frac{1}{M} \sum_{m=1}^{M} \left[ \frac{Y_{T,m}}{100} \left( C_{drug} - C_{sol} \sum_{i=1}^{15} F_{solF,i} - C_{sol} \sum_{i=1}^{15} F_{solL,i} \right) \right]$$

subject to:

process model equations

$$\forall \ s = 1..S, \ m = 1..M$$

binary variable approximations

$$\beta_{w_{act}C,15,m} = \frac{1}{2} \left[ \tanh \left( \frac{w_{act}C,15}{0.3} \right) + 1 \right] \quad \forall \ m = 1..M$$

$$\beta_{w_{act}E,15,m} = \frac{1}{2} \left[ \tanh \left( \frac{w_{act}E,15}{2.0} \right) + 1 \right] \quad \forall \ m = 1..M$$

stochastic inequality constraints

$$E\{w_{solF,1}\} = \frac{1}{M} \sum_{m=1}^{M} w_{solF,1,m} \leq 0.5$$

$$E\{p_{y,loss}\} = \frac{1}{M} \sum_{m=1}^{M} \left[ (1 - \beta_{w_{act}C,15,m} \beta_{w_{act}E,15,m}) p_{y,m} \right] \leq 3.0$$

decision bounds

$$200 \leq f_{1.1} \leq 400$$

$$60 \leq N_1 \leq 90$$

$$0.4 \leq AF_{10} \leq 0.7$$

$$0.3 \leq RF_{10} \leq 0.7$$

$$0.5 \leq CR_{12} \leq 4.0$$

$$40 \leq HP_{12} \leq 120$$

uncertainty space
The values for the constants of the hyperbolic smoothing functions for the endpoint key and secondary impurity contents, $\xi_{\text{wash},15}$ and $\xi_{\text{wash},15}$, were selected to be 1000 and 150, respectively. With these values the binary approximation smoothing functions calculate zero and one for criteria values to within approximately 1% of the threshold value.

8.2.3 Stochastic operating policy tolerance optimisation problem

The optimisation problem aims to maximise the tolerances (uncertainty) around the previous 'here and now' optimal operating policy variables subject to a stochastic constraint forcing the expected value of the potential profitability, $E\{P(t)\}$, to be at least 99% of the 'here and now' optimal solution, $E\{P(t)\}$. The total uncertainty space is expanded to include the operating policy variables tolerances, the extents of which are the decisions, $\delta_d^U$ and $\delta_d^L$, determined in the optimisation. The formulation for this optimisation is given in Problem P11.

Objective function

$$\max_{\delta_d^U, \delta_d^L} \frac{1}{D} \sum_{d=1}^{D} \frac{\delta_d^U (z_d^* - z_d^{LB}) + \delta_d^L (z_d^{UB} - z_d^*)}{(z_d^{UB} - z_d^{LB})}$$

where $d = t_f, N_1, AF_{10}, RF_{10}, CR_{12}, HP_{12}$

subject to:

- process model equations
- binary variable approximations
- stochastic inequality constraints

$$E\{w_{\text{tol},F,11}\} = \frac{1}{M} \sum_{m=1}^{M} w_{\text{tol},F,11,m} \leq 0.5$$
Integrated design under uncertainty for pharmaceutical processes

\[
E\{P_{\text{loss}}\} = \frac{1}{M} \sum_{m=1}^{M} \left[ (1 - \beta_{\text{waste},15.5,m} \beta_{\text{waste},15.5,m}) P_{\text{m}} \right] \leq 3.0
\]

\[
E\{P_{\text{y}}\} = \frac{1}{M} \sum_{m=1}^{M} \frac{Y_{\text{T},m}}{Y_{\text{T},m}} \left( C_{\text{drug}} - C_{\text{solF}} \frac{\sum_{i=1}^{15} F_{\text{solF,i}}}{Z_{\text{drug},15.5,m}} - C_{\text{solL}} \frac{\sum_{i=1}^{15} F_{\text{solL,i}}}{Z_{\text{drug},15.5,m}} \right) \leq 0.99 E\{P_{\text{y}}^*\}
\]

decision bounds

\[0 < \delta_{d}^{U} \leq 1, \quad 0 < \delta_{d}^{L} \leq 1\]

where \(d = t, N, AF_{10}, RF_{10}, CR_{12}, HP_{12}\)

uncertainty space

\[
\Theta_{U,d} = \left\{ \theta_{d} \left| \left( z_{d}^{*} - \delta_{d}^{L} (z_{d}^{*} - z_{d}^{LB}), z_{d}^{*} + \delta_{d}^{U} (z_{d}^{UB} - z_{d}^{*}) \right) \right\}
\]

where \(d = t, N, AF_{10}, RF_{10}, CR_{12}, HP_{12}\)

\[
k_{1,k_{2}} = \left( 1 + \frac{\kappa_{1} - \kappa_{2}}{k_{2} - \kappa_{1}} \right)^{T} \frac{\sigma_{k_{1}}^{2}}{\sigma_{k_{2}}^{2}} \frac{\rho \sigma_{k_{1}} \sigma_{k_{2}}}{\sigma_{k_{2}}^{2}} \left[ \frac{k_{1} - \kappa_{1}}{k_{2} - \kappa_{1}} \right]
\]

\[
k_{3,k_{t}} = \kappa_{t} + \kappa_{t}^{*} Y_{1}, Y_{2}, \sigma_{t}^{*} + \kappa_{g}^{*} \zeta_{\text{actA}}, \kappa_{g}^{*} \zeta_{\text{actB}}, \kappa_{g}^{*} \zeta_{\text{actC}}, \kappa_{g}^{*} \zeta_{\text{actD}}, \kappa_{g}^{*} \zeta_{\text{actE}}, FR, DR = \{ \theta_{t} | N(\mu_{t}, \sigma_{t}) \}
\]

\[
\eta_{\text{wash}} = \left\{ \eta_{\text{wash}} \left| \left( \eta_{\text{wash}}^{LB}, \eta_{\text{wash}}^{UB} \right) \right\}
\]

operating policy bounds

\[
\begin{align*}
\zeta_{i,1}^{LB} & = 200,400 \\
\zeta_{i,1}^{UB} & = 60,90 \end{align*}
\]

\[
\begin{align*}
\zeta_{N_{1}}^{LB} & = 60,90 \\
\zeta_{N_{1}}^{UB} & = 60,90 \end{align*}
\]

\[
\begin{align*}
\zeta_{AF_{10}}^{LB} & = 0.4,0.7 \\
\zeta_{AF_{10}}^{UB} & = 0.3,0.7 \end{align*}
\]

\[
\begin{align*}
\zeta_{CR_{12}}^{LB} & = 0.5,4.0 \\
\zeta_{CR_{12}}^{UB} & = 40,120 \end{align*}
\]

(Problem P11)

8.3 Base Case flowsheet optimisation results

The validated results of the optimisation with and without uncertainty are given in Table 8.1. It is immediately clear that when the optimal decisions obtained with no consideration of the uncertainty are
implemented in the uncertain process, the predicted potential profitability loss (15.08 $ kg_{act}^{-1}hr^{-1}) far exceeds the desired limit of 3.0 $ kg_{act}^{-1}hr^{-1}. This appears to be due to low probabilities of passing the thresholds on either the key or secondary impurity contents (0.67 and 0.75, respectively), resulting in a probability of passing the loss constraint (Pr_{pass}) of only 0.51. At the expense of a reduction in the expected profitability (30.97 from 31.13 $ kg_{act}^{-1}hr^{-1} due to a lower yield), the expected potential profitability loss limit is approximately maintained in the validated results of the robust optimisation.

The optimal decisions, shown in Table 8.2, clearly explain these results. The shorter Stage 1 reaction and Stage 12 crystallisation duration times determined in the uncertain optimisation, lead to a lower expected total yield (86.5%) but also restrict the formation of the impurities observed by the lower expected endpoint contents predicted (0.22 and 1.42 wt% for the key and secondary impurities, respectively). These comparisons are reflected in the cumulative frequency plots for the total yield and key impurity content, Figures 8.1 (a) and (b), respectively. The relative behaviours of the secondary impurity content predictions mirror those of the key impurity.

Table 8.1. Validated Base Case process flowsheet optimisation results under uncertainty, Case Study II.

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Nominal optimal operation</th>
<th>Uncertain optimal operation</th>
</tr>
</thead>
<tbody>
<tr>
<td>E(P_{ty}) ($ kg_{act}^{-1}hr^{-1}$)</td>
<td>31.13</td>
<td>30.97</td>
</tr>
<tr>
<td>E(P_{ty}loss) ($ kg_{act}^{-1}hr^{-1}$)</td>
<td>15.08</td>
<td>3.02</td>
</tr>
<tr>
<td>Pr_{pass}</td>
<td>0.514</td>
<td>0.894</td>
</tr>
<tr>
<td>[E(w_{act}), E(w_{actE}), E(w_{actF})] (%)</td>
<td>[0.28, 1.71, 0.43]</td>
<td>[0.22, 1.42, 0.43]</td>
</tr>
<tr>
<td>[FW(w_{act}), FW(w_{actE}), FW(w_{actF})]</td>
<td>[0.17, 1.42, 0.002]</td>
<td>[0.14, 1.19, 0.002]</td>
</tr>
<tr>
<td>E(Y_{T}) (%)</td>
<td>87.9</td>
<td>86.5</td>
</tr>
<tr>
<td>FW(Y_{T}) (%)</td>
<td>4.7</td>
<td>5.4</td>
</tr>
</tbody>
</table>

Table 8.2. Optimal decisions for the Base Case flowsheet, Case Study II.

<table>
<thead>
<tr>
<th>Decisions</th>
<th>Nominal optimisation</th>
<th>Uncertainty optimisation</th>
</tr>
</thead>
<tbody>
<tr>
<td>t_{f1} (min)</td>
<td>262</td>
<td>251</td>
</tr>
<tr>
<td>N_{1} (rpm)</td>
<td>90.0</td>
<td>90.0</td>
</tr>
<tr>
<td>AF_{10}</td>
<td>0.40</td>
<td>0.40</td>
</tr>
<tr>
<td>RF_{10}</td>
<td>0.70</td>
<td>0.70</td>
</tr>
<tr>
<td>CR_{12} (°C min^{-1})</td>
<td>0.57</td>
<td>0.50</td>
</tr>
<tr>
<td>HP_{12} (min)</td>
<td>67</td>
<td>54</td>
</tr>
</tbody>
</table>

The restricted formation of the key and secondary impurity are coupled with reductions in the uncertainties of the predicted endpoint contents (-16% and -17% in the respective 5-95% fractile widths).
Integrated design under uncertainty for pharmaceutical processes

However, the uncertainty in the total yield increases by 13%. The shorter crystallisation holding period has the effect that the final Stage 12 product drug concentration is further away from the equilibrium saturation which in turn has the effect that the propagation of the uncertainties entering Stage 12 are not suppressed as much in the output variables. The optimum Stage 1 agitation speed, $N_1$, and the solvent exchange decisions, $AF_{10}$ and $RF_{10}$, are unaffected by the incorporation of uncertainties.

The results of the Sensitivity Analysis for the optimised Base Case are given in Table 8.3 and Table 8.4. The drying time, $DR_{15}$, is clearly the highest ranked contributor to the uncertainty predicted in the profitability, followed by the Stage 12 crystallisation growth rate constant ($k_{d}$) and the Stage 1 reaction rate constant ($k_1$). For the total yield, $k_{yr}$, $k_1$, the aqueous-organic equilibrium drug solubilities ($\sigma_{in}$) and the delayed Stage 1 key impurity reaction start time, $t_{1''}$, exhibit the strongest relationships. For the key and secondary impurity contents, the Stage 1 reaction rate constants are the clearly the strongest contributors followed by the uncertainty in the Stage 12 crystallisation impurity ‘solute loss’ parameters ($\zeta_{sace}$ and $\zeta_{sact}$) and the Stage 14 wash efficiency ($\eta_{wash}$). The overwhelming importance of the uncertainty contained in the Stage 1 model to the impurity contents is corroborated by the sub-sequence contributions shown in Table 8.4. Implementation of the robust decisions compared to the nominal decisions did not qualitatively (nor significantly quantitatively) affect the outcomes of the Sensitivity Analysis.

From these results it can be inferred that with the importance of the time aspect incorporated in the profitability criterion, it would be most beneficial to try and reduce the uncertainty in the drying operation. However, since drying appears to be an extremely difficult process to model with any degree of accuracy it may be more realistic to direct action towards improving the current models (and parameter estimations) of the crystalliser and reactor models, and in particular the kinetics.
Table 8.3. SRC ranking of the key input uncertainty contributors to the Base Case flowsheet criteria, Case Study II.

<table>
<thead>
<tr>
<th>Criteria</th>
<th>R²</th>
<th>Key uncertainty contributors (SRC value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pty</td>
<td>0.99</td>
<td>DR₁₅ (-0.768), k₅ (0.405), k₁ (0.298), FR₁₃ (-0.184), σ₁₃’ (-0.146), t₁'' (-0.145)</td>
</tr>
<tr>
<td>Y₉</td>
<td>0.96</td>
<td>k₅ (0.681), k₁ (0.465), σ₁₃’ (-0.256), t₁'' (-0.231), X” (0.108), hband₃₆ (0.09)</td>
</tr>
<tr>
<td>wtactC</td>
<td>0.99</td>
<td>k₂ (0.848), ζactC (0.447), ηwash (-0.320), t₁'' (-0.240)</td>
</tr>
<tr>
<td>wtactE</td>
<td>0.99</td>
<td>k₃ (0.915), ζactE (0.331), ηwash (-0.186)</td>
</tr>
</tbody>
</table>

Table 8.4. Sub-sequence contributions to the uncertainty in the Base Case flowsheet, Case Study II.

<table>
<thead>
<tr>
<th>Sub-sequence</th>
<th>Total yield</th>
<th>Key impurity</th>
<th>Secondary impurity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stage 1</td>
<td>0.47</td>
<td>0.80</td>
<td>0.93</td>
</tr>
<tr>
<td>Stage 2 to 15</td>
<td>0.53</td>
<td>0.20</td>
<td>0.07</td>
</tr>
<tr>
<td>Endpoint</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
</tbody>
</table>

The fact that the nominal optimal decisions, which maintain the deterministic impurity constraints, perform so poorly when extended to the uncertain process system, underlines the importance of the consideration of the main uncertainties in the optimisation determining these decisions. This importance has been quantified through the stochastic criteria estimated in the optimisation under uncertainty. The effect of the optimal decisions on the predicted output criteria distributions can be explained with respect to the propagation of uncertainty in the stream variables through the process sequence due to the deterministic structures of the process models.

8.3.1 Robustness Analysis of the input uncertainties

The importance of the state of knowledge of the input source uncertainties to the results of the process flowsheet optimisation under uncertainty (Problem PIO) is investigated and quantified in what Kleijnen (1997) terms a Robustness Analysis. Since Problem PIO is constrained by the potential profitability loss due to failure of some portion of the upper distribution tails of the endpoint impurity contents, the optimal solution and the corresponding decisions may be sensitive to the size and form of the input uncertainties (the location of the means are not varied). Two variations to the Base Case problem uncertainties are considered for the re-optimisation of the process flowsheet under uncertainty,

- Case 1 considers the sensitivity to the size of the input uncertainties, where the uncertainties in the distributions of the original problem are increased by 50%,

- Case 2 considers the sensitivity to the form of the input uncertainties, where uniform distributions replace the normal distributions assumed in the original problem. The upper and lower bounds are fixed at two standard deviations (of the original normal distributions) from the mean value.

It can be immediately seen from the bar chart showing percent deviations in the expected values and 5-95% fractile widths of the key criteria from the Base Case, Figure 8.2, that the assumption of uniform
input uncertainties (Case 2) imposes very little effect compared with a 50% increase in the input uncertainties (Case 1), relative to the original normal input distributions of the Base Case. A deviation of only -0.3% in the expected profitability and virtually identical optimal decisions to those of the Base Case are predicted for Case 2. A deviation of -1.2% in the profitability for Case 1 is predicted, resulting from reduced expected yield due to the shorter Stage 1 reaction time (246 minutes) and Stage 12 crystallisation time (44 minutes). As expected for Case 1, the significant increases in uncertainty in the impurity contents (+36.5% and +37.1% in the fractile width of the key and secondary impurity contents, respectively) results in the distributions being shifted to the left in order to satisfy the potential profitability loss constraint and a decrease in the expected values (see Figure 8.2).

These observations are reflected in the cumulative frequency plots for total yield and key impurity content comparing the three input uncertainty cases, shown in Figure E1 (a) and (b) (Appendix E), respectively. The behaviour of the secondary impurity content predictions mirror those of the key impurity.

This investigation underlines the importance that the state of knowledge of the input uncertainties can have in the Uncertainty Analysis approach to process flowsheet optimisation under uncertainty. Clearly this importance can depend on whether the optimisations are concerned with averages or the tails of distributions. Within the confines of good assumptions for the input uncertainty bounds, the form of the distributions appears to be of negligible significance to the optimal solutions determined. On the other hand, a good estimate to the magnitude of the input uncertainties is essential to the results and decisions obtained using an Uncertainty Analysis approach to process flowsheet optimisation. This is particularly true when the tails of output distributions are important.

![Cumulative Frequency Plots](image)

(a) Expected value. (b) Fractile width.

Figure 8.2. Deviation in optimal results with input uncertainty variations, from the Base Case, Case Study II.

Key: ■ = Case 1, ○ = Case 2.
8.3.2 The importance of process input specification

The concept of the value of perfect information (VPI) is applied to the Base Case flowsheet with regard to potential feed purity knowledge of the API, $p_{f,1}$. The VPI is defined in Equation 8.1, and approximated in Equation 8.2 as the expected gain in the potential profitability when using an informed 'wait and see' optimal approach as opposed to the uninformed 'here and now' approach, with penalisation of violations in an acceptable profitability loss constraint in the latter. A linear function in the extent of violation in the loss constraint penalises the value of information expected with the 'here and now' approach, $V_{l,here}$.

\[
VPI = E_{\theta, \phi} \left\{ V_{l,wait} - V_{l,here} \right\}
\]

\[
= E_{\rho_{f,1}} \left\{ \left( E_{\theta, \phi, \rho_{f,1}} \left\{ P_{t, wait} \right\} \right)_{m'} - \left[ \left( E_{\theta, \phi, \rho_{f,1}} \left\{ P_{t, here} \right\} \right)_{m'} - \beta_{P_{t,loss,here,m'}} \left( \left( E_{\theta, \phi, \rho_{f,1}} \left\{ P_{t,loss,here} \right\} \right)_{m'} - 3.0 \right) \right] \right\}
\]

where $m'$ are the observations in $p_{f,1}$ space. Under potential uncertainty in the feed purity of 74 ± 3 wt%, the VPI with knowledge of this uncertainty is a profitability of 11.46 $kg_{act}^{-1} hr^{-1}$. It is clear from the increasing value of the VPI with increasing feed purity, Figure 8.3 (a), that at low values of $p_{f,1}$ (below 72.5 wt%) the main contribution to the VPI is incurred. As indicated in Figure 8.3 (b) the $V_{l,here}$ is greatly reduced for these values of feed purity. The relatively large violation of the profitability loss constraint is shown by the dashed line representing the $V_{l,here}$ without any penalisation of the constraint violation. A shallow optimum in the value of the feed purity (~74 wt%) is observed in the $V_{l,wait}$ solutions, Figure 8.3.

![Figure 8.3](image-url)

(a) Value of Perfect Information.  
(b) Value of Information.  
Key: ● = $V_{l,wait}$, o = $V_{l,here}$,  
--- = unpenalised $V_{l,here}$

Figure 8.3. Value of feed purity information to potential profitability, Case Study II.
Integrated design under uncertainty for pharmaceutical processes

(b). This could be useful knowledge if the feed purity can be more closely specified. At lower feed purities optimal solutions are constrained by the increased propensity for formation of the secondary impurity in the Stage 1 reaction leading to high values in the endpoint secondary impurity content. At higher feed purities the greater amount of product formed results in increasing proportions of the key impurity.

The importance of certain a priori knowledge to the process has been quantified in the form of the VPI. The results show that this value can be very sensitive within certain ranges of the available information.

8.4 Base case operating policy tolerance optimisation results

Solution of Problem P11 gives an optimum value for the average tolerance of 0.088, shown in Table 8.5. This is the maximum average fraction of the space of all the operating policy variables (defined by the upper and lower policy bounds) within which a feasible solution is permitted under error or uncertainty in the implementation of the ‘here and now’ optimal policy. As shown in Table 8.5, the solution is tightly constrained by the stochastic constraints on the expected potential profitability (30.66 \( \text{kg actA}^{-1} \text{hr}^{-1} \)), at 99% of the ‘here and now’ optimal solution), the expected potential profitability loss (3.00 \( \text{kg actA}^{-1} \text{hr}^{-1} \)) and the pre-crystallisation solF content (0.50 wt%).

The resulting tolerance limits around the ‘here and now’ optimal operating policy variables (given previously in Table 8.2) on the operating policy are given in Table 8.6. The tolerances permitted on the reaction time \((t_{r1})\) and the crystallisation time \((HP_{12})\) indicate that the expected profitability criteria is relatively insensitive to the total yield due to a strong dependence on the overall process time to which the available optimisation operating policy variables provide a relatively small contribution. Virtually no tolerance is acceptable in any of the other operating policy variables, hence the relatively low value for the average tolerance. The values of the optimal decisions \((\delta^U_1\) and \(\delta^L_1\)) are given in Table E5 (Appendix E).

| Table 8.5. Results for tolerance optimisation of the Base Case flowsheet operating policy, Case Study II. |
|----------------|-----------------|----------------|
| Objective       | Average tolerance | 0.088          |
| Stochastic constraints | E(Pty)(\( \text{kg actA}^{-1} \text{hr}^{-1} \)) | 30.66          |
|                  | E(Ptyloss)(\( \text{kg actA}^{-1} \text{hr}^{-1} \)) | 3.00           |
|                  | E(\%wt solF) (%) | 0.49           |

| Table 8.6. Tolerance limits for Base Case optimal operating policy, Case Study II. |
|----------------|-----------------|----------------|----------------|----------------|----------------|
|                | \(z^U_d\) (min) | \(z^L_d\) (rpm) | \(AF_{10}\) | \(RF_{10}\) | \(CR_{12}\) \(\degree C \text{min}^{-1}\) | \(HP_{12}\) (min) |
| \(t_{r1}\)   | 267             | 90.0            | 0.41          | 0.70          | 0.52           | 64             |
| \(z_{r1}\)   | 235             | 89.5            | 0.40          | 0.68          | 0.50           | 43             |
8.5 Conclusions

In this Chapter the use of a multiscenario stochastic optimisation approach for the optimisation of integrated flowsheets under model parameter uncertainty is demonstrated with respect to the set of models comprising the Base Case process flowsheet of Case Study II. Its application allowed the selection of 'here and now' operating policy decisions to optimise and manage certain aspects of the distributions of the uncertain output predictions. An expected profitability objective function is optimised within a threshold on an expected profitability loss due to failure in endpoint impurity contents. This action corresponds to the second management response identified in Chapter 5. Under no uncertainty, the nominal optimal decisions result in poor behaviour of the endpoint impurity content distributions and a high potential profitability loss when the quantified model parameter uncertainty is considered in Uncertainty Analysis. This highlights the importance of the consideration of uncertainty in the process optimisation.

It is also shown that the state of knowledge of the model uncertainties is important with respect to the magnitude but not so much to the characteristic distribution. The value of perfect information in potentially uncertain process stream inputs is considered with respect to the feed purity. It was found that below a certain value in the feed purity (72.5%), perfect knowledge had a very significant impact and that a shallow optimum value existed (~74%).

The key uncertainties under the optimal policy decisions are identified in the Sensitivity Analysis with the conclusion that realistic efforts to reduce the current levels of uncertainty in the profitability and yield should be primarily directed towards improving the confidence in the crystallisation kinetics followed by the intrinsic rate constant for the product reaction. The ability to provide a better prediction for the time taken for the drying is the key factor in reducing the uncertainty in the profitability due to its dependency on the total batch process time. However, the provision of a useful mechanistic model for drying may be unrealistic considering the complexity of the physical phenomena associated with drying. To reduce the uncertainty predicted in the endpoint impurity contents the analysis (not unexpectedly) strongly indicates that efforts aimed at reducing the uncertainty in the Stage I intrinsic reaction kinetics would be the most rewarding. This information, based on systematic and quantitative procedures, provides focus to the actions which may be invested towards the development of more reliable models for the first management response (Chapter 5).

A further stochastic optimisation determined the maximum uncertainty allowed around the optimal 'here and now' operating policy values of the Base Case flowsheet. The upper and lower tolerance limits are constrained to meet a small relaxation of the 'here and now' profitability together with the original stochastic constraints. This quantifies the relaxation permitted to the optimal operating policy actions determined in association to the second management response.

The third management response, considering alternative process flowsheets, is investigated in Chapter 9.
Chapter 9

CASE STUDY II: ASSESSMENT OF ALTERNATIVE PROCESS FLOWSHEET OPTIONS

9.1 Introduction

In the previous chapter stochastic optimisation of the Base Case flowsheet under uncertainty in the model parameters was shown to provide a better solution than the equivalent optimal solution under no uncertainty. In addition, the maximum permitted tolerance around the values of the optimal operating policy was determined, to achieve an expected profitability within 99% of the 'here and now' optimal solution. In this chapter three alternative process flowsheet options are assessed with respect to optimal performance and permitted operating policy tolerance.

9.2 Process flowsheet alternatives

Four alternative flowsheet options are considered:

- Base Case, the original process flowsheet and associated uncertainties,
- Alternative 1, back extraction of post-reaction aqueous phase with solF,
- Alternative 2, recycle of prior solL mother liquor into the crystallisation,
- Alternative 3, removal of the crystal wash operation.

Alternatives 1, 2, and 3 are briefly discussed in the following sections. The stochastic optimisations for expected profitability, Problem P10, and operating policy tolerance, Problem P11, are solved for each flowsheet option under their respective uncertainties.

9.2.1 Alternative 1: Post-reactor back extraction

The total product yield may be increased with the solF back extraction of residing organic components in the Stage 3 layer separation aqueous waste stream. The additional Stage 3A solF back extraction and Stage 3B layer separation are shown in Figure 9.1 in comparison with the Base Case flowsheet shown in Figure 7.1 (Section 7.1). A compromise is incurred at the expense of the extra solF solvent required, the increase in the total batch process time for the additional extraction and layer separation operations and the increase in the distillation time during the removal of solF. The number of uncertain inputs to the stochastic process is increased by the additional uncertainties in the drug solubilities associated with the Stage 3A model, and the uncertainties associated with the subsequent layer separation model, Stage 3B.
2. DILUTION
3. LAYER
4. rH
5. DESTRUCTION
6% basel
50% basel

Distilled water
Post reactor stream [3]

Figure 9.1. Process flowsheet showing the additional back extraction in Alternative 1, Case Study II.

The quantity of pure solF used in the back extraction is set as a fraction of that added in the prior reaction, \( R_{\text{solF-back}} \). It is an additional decision variable in the optimisation under uncertainty, between bounds of 0.01 and 0.5.

9.2.2 Alternative 2: solL mother liquor recycle

The recycle of solL mother liquor from a prior crystallisation may be desirable to obtain higher yields if the desired impurity levels can be maintained. The addition of recycled drug product in the mother liquor may allow shorter reaction and crystallisation times reducing the possibility of impurity formation and crystallisation take-up. The reduction in the quantity of pure solL solvent required is another benefit. However, the possibility of the build-up of impurities may be critical with uncertainty assumed in the composition of the mother liquor recycle. The process flowsheet for Alternative 2 is similar to the Base Case but with the addition of the solL mother liquor recycle stream to the pure solL solvent feeds in the Stage 10 and 11 distillations, shown in Figure 9.2. The ratio of the solL mother liquor recycle to the pure solvent (at constant total solL mass), \( R_{\text{solL-recycle}} \), is set as a decision variable in the optimisation under uncertainty, between bounds of 0.01 and 0.5, with the total quantity of solL remaining the same as for the Base Case.
9.2.3 Alternative 3: No post-filtration crystal wash

The crystal wash operation is removed in Alternative 3. The removal of the crystal wash operation in Alternative 3 saves time and the quantity of pure soIL solvent required but the lack of residual trace impurity removal may limit the reaction and crystallisation operations. The pertinent part of the process flowsheet of Alternative 3 is shown in Figure 9.3. The only change in the stochastic system definition is the removal of the uncertainty assumed in the efficiency of the wash.

9.3 Comparison of optimal process flowsheets under uncertainty

Comparison between the optimal alternate flowsheets are made with respect to the criteria presented by subsequent Uncertainty and Sensitivity Analyses,
• expected values for profitability, total yield, and key and secondary impurity content,

• uncertainty in the process variable predictions for total yield, and key and secondary impurity content,

• the major contributions to the uncertainty in the profitability, total yield, and key and secondary impurity content.

Comparison of the deviations in the optimal expected values of the potential profitability from the Base Case value (30.97 $ kg\text{-acta}^{-1} \text{hr}^{-1}, \text{Table 8.1, Section 8.1}) shows that Alternative 2 is the best with a 4.2% improvement over the Base Case, as indicated by the deviations in the bar chart, Figure 9.4 (a). On an annualised basis, the 4.2% improvement in profitability (32.26 from 30.97 $ kg\text{-acta}^{-1} \text{hr}^{-1}) becomes a significant annual improvement of $5.65 million, assuming back-to-back batch cycles of 500 kg API feed. The estimated annualised improvement for Alternative 1 is $0.70 million and Alternative 3 is $2.98 million (over the Base Case). The process flowsheet stochastic optimisation results are summarised in Table E2 and E3 (Appendix E).

Back extraction of the post reactor aqueous stream in Alternative 1 results in the best flowsheet prediction for expected yield (89.2%) and the lowest expected endpoint impurity contents. The assumption of independent drug component solubilities leads to a higher transfer of the relatively high concentration of drug product to the organic phase but very little transfer of low composition impurities. The decreases in expected impurity content may not be that significant since the potential profitability loss constraint (based on the impurity content thresholds) is still maintained. It does indicate that the spread in the impurity content distributions is increased, as shown by the deviation in fractile width in Figure 9.4 (b), but only significantly so for that part of the distribution below the impurity threshold (see Figure E3 (b),

Figure 9.4. Deviations in optimal results under uncertainty for alternative flowsheet options, from the Base Case, Case Study II.

Key: ■ = Alternative 1, ● = Alternative 2, □ = Alternative 3.
Appendix E). Despite the significant yield increase (89.2% compared to 86.5% in the Base case), the expected profitability is constrained by the extra solF solvent cost and particularly the increase in the total batch process time due to the back extraction, layer separation and increased distillation time required for solF removal, Stage 9.

Recycle of the solL mother liquor in Alternative 2 provides the best profitability (32.26 $ kg\text{act}^{-1} \text{hr}^{-1}$ or $141$ million per year with the annualised assumptions mentioned above) due to the reduced solL solvent cost and lower reaction and crystallisation times, see Table E3 (Appendix E) for the optimal decisions. The lower impurity formation in the reaction is offset by the introduction of impurities in the recycle stream. At the optimal operating conditions the slight increase in the expected impurity contents shown in Figure 9.4 (a) are fairly insignificant.

Although removal of the crystal wash in Alternative 3, results in a slightly lower yield expected due to the shorter reaction time which is required for the maintenance of the expected profitability loss constraint, this reduction is less significant to the profitability than the lower total process time and lower solL solvent usage. The cumulative frequency plots for profitability, total yield and key impurity content, see Figure E2 and E3 (a) and (b), Appendix E, show the distributions predicted with each flowsheet option.

Without any uncertainty assumed in the model parameters, the results for optimisation of profitability under no uncertainty leads to the same conclusion that Alternative 2 has the potential to provide the best profitability. However, as observed previously with the Base Case flowsheet, the nominal optimal decisions for the alternative flowsheets provide very poor performance with respect to the potential profitability loss under the uncertainties considered.

9.3.1 Comparison of uncertainty in output criteria predictions

The uncertainty in the predictions associated with each flowsheet alternative is also an important criterion for comparison when using the model-based approach. From Figure 9.4 (b), Alternative 2 predicts the greatest uncertainty in the total yield by a significant margin (16% more than the Base Case). However, by far the greatest uncertainty in the endpoint key and secondary impurity content are predicted for Alternative 1 (18% and 15% more than the Base Case, respectively). If the uncertainty associated with the prediction is an important issue in the flowsheet selection then the increased prediction uncertainty and the possibility of the further modelling effort required to reduce these uncertainties should be considered in the cases of the Alternative 2 and Alternative 1 options. However, the increased uncertainty in the predicted impurity contents for Alternative 1 is mainly due to the portion of the distributions below the upper threshold limits.

9.3.2 Comparison of uncertainty contributions

Sensitivity Analysis is used to compare how the alternative process options with their associated levels of knowledge in the integrated flowsheets can affect the key contributions to the uncertainty in the
predictions. The main results are given in Table 9.1 and Figure 9.5. More complete results for the uncertainty contributors are given in Table E4 (Appendix E). With the consideration of time importance, the assumed long duration of the drying operation and the relatively large amount of associated uncertainty, make the drying time, \(DR_{15}\), the key uncertainty contributor to the potential profitability for all the flowsheet alternatives.

While no change in the top ranked SRC contributors are observed following Sensitivity Analysis of Alternative 1, the addition of the back extraction results in a small decrease in the Stage 2-15 sub-sequence contribution to the total yield uncertainty and increases to the key and secondary impurity content uncertainties, as shown in the bar charts shown in Figure 9.5. Although the actual observed changes in the uncertainties of the key impurity compositions through the process between Alternative 1 and the Base Case are small (fractile width, FW, of 0.434 wt% for the post-reactor key impurity composition and 0.172 wt% for the endpoint impurity content in Alternative 1, compared to the FW of 0.444 wt% and 0.145 wt% in the Base Case), the significance lies in the shift in the contribution to the

![Sub-sequence contribution to predicted uncertainty, for flowsheet alternatives, Case Study II.](image)

(a) Total yield.  (b) Key impurity content.  (c) Secondary impurity content.

Figure 9.5. Sub-sequence contributions to predicted uncertainty, for flowsheet alternatives, Case Study II. Key: ■ = Stage 1 sub-sequence, ◯ = Stage 2-15 sub-sequence.
Integrated design under uncertainty for pharmaceutical processes

sub-sequence within which the uncertainty is being introduced and propagated. This shows that with the use of a solF back extraction, the contribution to the uncertainty in the endpoint impurity content is increased over the Stage 2-15 sub-sequence in Alternative 1 (compared to the Base Case) and that if greater quantities of solF are used in the extraction (the optimal value is only 5.2% of the quantity used in the reaction) that the Stage 2-15 sub-sequence contribution to the uncertainty may increase in a more significant manner. The opposite effect is observed with regard to the uncertainty in the total yield. These observations may have important connotations with respect to the state of knowledge that can be ascertained over certain parts of the process sequence and what can be done to improve it.

In Alternative 2, the uncertainty in the product composition of the recycled solL mother liquor, wt\text{solL}_{ml2,10} becomes the fifth ranked contributor to profitability uncertainty and the fourth ranked to the total yield uncertainty. In contrast the uncertainties in the drug impurity compositions of the recycle appear to have negligible effect on the endpoint impurity contents. With elimination of the wash operation in Alternative 3, the uncertainty in the level of dampness of the post-filtration solids stream, LOD_{13}, becomes an important factor to the key impurity content. This would be expected since any quantities of the key impurity contained in the residue moisture is no longer removed from the crystal product.

With the introduction of different contributors or changes in the relative contributions of the common sources of uncertainty associated with alternative flowsheets, qualitative judgement regarding the possibility of the relevant process stage model improvement and uncertainty reduction becomes important if the levels of predicted output uncertainty are desired to be limited.

9.4 Comparison between flowsheet operating policy tolerances

The basis for this comparison between each flowsheet alternative is the maximum average tolerances achieved under a percentage based relaxation in the ‘here and now’ optimal expected potential profitability and an examination of the trade-offs between the key operating policy tolerance windows. The former quantity represents the average fraction of the total operating policy space available in all the considered variables (defined by the upper and lower policy bounds), within which policy uncertainty permits a feasible solution. It can take values between zero for no tolerance to operating policy uncertainty about the optimal values, to one for complete tolerance within the specified policy bounds.

Comparison of the average operating policy tolerances with a 99% achievement of the ‘here and now’

<table>
<thead>
<tr>
<th>Operating policy variables</th>
<th>Average tolerance</th>
<th>E[Pty] $ kg_{actA}^{-1} hr^{-1}</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base Case</td>
<td>0.088</td>
<td>30.66</td>
</tr>
<tr>
<td>Alternative 1</td>
<td>0.080 (0.090)</td>
<td>30.81</td>
</tr>
<tr>
<td>Alternative 2</td>
<td>0.083 (0.081)</td>
<td>31.94</td>
</tr>
<tr>
<td>Alternative 3</td>
<td>0.055</td>
<td>31.33</td>
</tr>
</tbody>
</table>

Table 9.2. Alternative flowsheet operating policy tolerance optimisation results, Case Study II.
optimal expected profitability is shown in Table 9.2. This basis measures the relative degree of relaxation for each flowsheet in an unbiased manner. Otherwise, the use of a constant expected profitability constraint threshold value would give a greater relaxation advantage to the flowsheet with the greatest 'here and now' optimal profitability. The policy variables considered are the same as those used in the flowsheet optimisations. Table 9.2 indicates that the Base Case flowsheet is the most tolerant to error in the 'here and now' optimal operating policy, although the values for Alternatives 1 and 2 are very close.

The average tolerance values for all the flowsheets are relatively small since negligible tolerances are permitted in the Stage 1 agitation speed (N1), the Stage 10 solL addition and removal fractions (AF10 and RF10), the Stage 12 crystallisation cooling rate (CR12) nor in the fraction (of the Stage 1 quantity) of solF for the back extraction (R_{solF,back}) in Alternative 2, or the fraction of solL mother liquor recycle employed in Stage 10 and 11 (R_{solL,recycle}) in Alternative 3. The values of the optimisation decision variables and the tolerance limits for Alternatives 1, 2 and 3 are given in Tables E6, E7 and E8 (Appendix E), respectively.

It is immediately clear that Alternative 3 has a rather lower average tolerance (0.055) than the other flowsheets. The removal of the wash operation acts to remove some of the tolerance to error in those operating policy variables controlling the formation and crystallisation of the impurities, the Stage 1 reaction time (t_{r1}) and the Stage 12 crystallisation holding period (HP_{12}), respectively. This may be reasonable to expect.

The bracketed values for Alternative 1 and Alternative 2 in Table 9.2, are the average tolerances over only the operating policy variables common to all the flowsheet alternatives (t_{r1}, N1, AF10, RF10, CR12, HP_{12}). Comparing these average tolerances removes the bias in the averages associated with additional policy variables (solF back extraction ratio and solL mother liquor recycle ratio operating policy variables, R_{solF,back} and R_{solL,recycle}, available in Alternative 1 and Alternative 2, respectively). Under this comparison
the average tolerance of Alternative 1 (0.090) becomes comparable to that of the Base Case (0.088).

The tolerance limits on the Stage 1 reaction time ($t_{f1}$) and the Stage 12 crystallisation holding period ($HP_{12}$) are compared between the flowsheets, in Figure 9.6 (a) and (b). These policy variables are the only two in which any significant tolerance is permitted regarding the expected profitability objective and profitability loss associated with the total yield, key and secondary impurity content criteria. The dots encompassed by the tolerance limits are the ‘here and now’ optimal values and the dashed lines are the bounds on the decision variable. The large tolerance in Alternative 1 $t_{f1}$ (69 min) is offset by a small tolerance in $HP_{12}$ (8 min). For the Base Case the opposite is true, where the $HP_{12}$ tolerance (25 min) is offset by that in $t_{f1}$ (32 min). The tolerance limits on $N_1$ and $CR_{12}$ are very narrow and on $AF_{10}$ and $RF_{10}$ they are the same for each flowsheet. The limited tolerance in both $t_{f1}$ and $HP_{12}$ for the removal of the wash operation in Alternative 3, is observed in Figure 9.6.

9.5 Conclusions

In this Chapter the third management response to process uncertainty regarding the assessment of alternative flowsheet options (see Section 5.3) is implemented for Case Study II for three alternative flowsheets. Stochastic optimisation approaches are used for the quantified assessment of the alternative process flowsheet options with respect to the expected profitability and loss criteria under model parameter uncertainty and the maximum permitted tolerance in the subsequent optimal operating policy. This permitted the selection of the best flowsheet according to these different criteria under their respective model uncertainties.

It was shown that the Alternative 2 flowsheet achieves the best expected profitability under a soft potential profitability loss constraint due to endpoint purity violations. The recycle of soil mother liquor saves solvent cost. In addition, the introduction of recycled product permits a shorter reaction time (reducing the formation of impurities) and a lower crystallisation holding period. The drawback is the observed increase in uncertainty in the total yield, resulting from the introduction of uncertainty in the mother liquor product composition.

The Base Case flowsheet appears to be the most tolerant to uncertainty in its ‘here and now’ optimal operating policy variables, although there is little difference between the Base Case and Alternatives 1 and 2. The Stage 1 reaction time and the Stage 12 crystallisation holding period were the only policy variables exhibiting tolerance with respect to the profitability criterion. The average tolerance associated with Alternative 3 is significantly lower than for the other flowsheets due to the removal of the wash purification operation. This provides a compromise to the improved profitability predicted for this option.

The overall results presented in this chapter show that although the Alternative 2 flowsheet (soil mother liquor recycle) predicts the greatest profitability (annualised $2.67 million greater than the nearest rival option, Alternative 3) the flowsheet with the greatest degree of operating policy tolerance is the Base Case. However, the Base Case exhibited the lowest expected profitability.
Chapter 10

CONCLUSIONS AND FUTURE WORK

10.1 Conclusions

In this thesis the subject of uncertainty in a model-based approach for the design of pharmaceutical processes is addressed. Quantitative uncertainty management considerations are systematically applied to model building procedures and to the optimisation of complete integrated process sequences.

Chapter 1 introduces the motivation for the work within the context of the pharmaceutical industry. A background to the process development challenges faced by pharmaceutical industry together with the current role of process modelling is given in Chapter 2. It is acknowledged that significant benefits could be achieved with a more comprehensive use of model-based approaches. The opportunities to integrate systematic model development procedures with current process development cultures are apparent but the necessity not to increase process development cycle times further may be seen to provide a key stumbling block in the short term. Chapter 2 ends with a long term future vision for the integration of model-based approaches with process development cultures in the pharmaceutical industry.

Aspects of uncertainty in the context of model-based applications are discussed in Chapter 3. A range of Uncertainty and Sensitivity Analysis methods are discussed. In a stochastic system framework, the major emphasis of the problem considered in this work is defined to support deterministic process model parameter uncertainties as opposed to inherent system variabilities which are irreducible. Sample based approaches for the quantification of the combined influence of parameter uncertainties propagating to the process performance variables are identified. A variety of Sensitivity Analysis techniques and measures are defined and compared in terms of the limiting criteria of efficient computation and information they contain. These aim to identify the key uncertain input sources contributing to the predicted uncertainty in the output performance criteria. It was concluded that methods which are computationally efficient and flexible to manipulation in different analysis situations would be the most amenable for this work.

Methods for design optimisation under uncertainty are reviewed in Chapter 4. In general these problems involve multiple scenarios, either stated a priori or generated from bounded ranges of parameter uncertainty or from probability distributions. Methods involving equipment design and control variables as decisions were discussed. These were categorised into three classes: approaches using bounded range characterisations of uncertainty and which aim to guarantee a fixed degree of feasibility, approaches using probabilistic information on the uncertainties and which aim to determine more accurately the feasible region, and approaches using probabilistic information on the uncertainties and which consider the entire space of the uncertainties and permit levels of partial feasibility. It was concluded that the robust partial feasibility approaches were the most suitable for the problem considered in this thesis.
Integrated design under uncertainty for pharmaceutical processes

Three management responses to the presence of large amounts of uncertainty in the available models are considered in this work: identification of the key sources of uncertainty, manipulation of the available operating policy decisions and consideration of alternate process topology. In Chapter 5 a methodology based on a Risk Analysis approach for model uncertainty is proposed for the provision of these responses. This approach is integrated with systematic model development procedures so that the uncertainty in the important process output criteria are quantified and the key sources of input uncertainty are identified.

A stochastic description of the process system is generated with probability distributions for input parameter uncertainties based on the process data and parameter estimation methods where possible. This is approximated with the simulation of a set of deterministic models at discrete realisations in the defined input uncertainty space in an efficient Hammersley sampling based approach. In this way the combined influence of a large number of uncertainties over a complete process system can be measured. Efficient Sensitivity Analysis techniques which directly use the sample results of the Uncertainty Analysis are used to estimate linear, monotonic and/or non-monotonic nonlinear relationships between the uncertain inputs and the outputs. An estimation method for process sub-sequence contributions to uncertainty in related inter-stage stream properties is used to determine the cumulative and propagative effects on the uncertainty in inter-stage process qualities within the process sequence.

A stochastic optimisation algorithm is used for the solution of four different stochastic programming problems. The key components of the stochastic system generated in the Risk Analysis approach provide the basis for these multiscenario optimisation problems. General formulations are stated for the estimation of the smallest reduction in the stochastic input uncertainties required to meet certain output requirements, process flowsheet optimisation under uncertainty for a desired process or economic objective, for the estimation of the value of perfect information in potentially certain process properties and for the determination of maximum operating policy error tolerances maintaining desired levels of performance under uncertainty. These problems take into account the entire space of the uncertainties and do not rely on worst case scenarios.

Two case studies have been investigated to verify the methodology. In Chapter 6 the basic components of the Risk Analysis approach and stochastic process optimisation under uncertainty are demonstrated for a multiphase reaction process model, Case Study I. Uncertainty Analysis showed that model uncertainty was important and the optimum operating policy determined with no consideration of uncertainty resulted in very poor performances in some of the important criteria. Optimisation under uncertainty was able to provide significantly more desirable distributions in these poor performance aspects. Sensitivity Analysis methods were able to identify the key parameters for which reductions in uncertainty would be most beneficial. It is concluded that the methods implemented in Case Study I verify that this approach can provide useful information to the process development engineer. The significant (adverse) effect on optimised processes that uncertainty can impose was shown and that although better solutions can be achieved using optimisation under uncertainty techniques, this is not always guaranteed and it may be more beneficial to consider alternative process options or operating strategies.
A comprehensive case study of an integrated sequence of 15 process operations is introduced in Chapter 7. Here the Risk Analysis methods proposed in Chapter 5 are demonstrated in a more rigorous manner as they are applied within the confines of systematic model development procedures. The combined influence of all the potentially significant parameter uncertainties in the complete process sequence is measured and the uncertainty in the important output criteria is tracked as more process information is incorporated into the model development. Key process sub-sequences were identified as contributing relative amounts of uncertainty to the output criteria predictions. Specific key uncertain parameters were identified over the entire process sequence and the lack of understanding in the associated physico-chemical phenomena was inferred. The endpoint process criteria of yield and impurity content were most sensitive to the uncertainty in the models for the reaction and crystallisation operations. This may be a reasonable expectation from engineering intuition. Trade-off levels between the minimum reductions in the key uncertain parameters to meet target levels of uncertainty in endpoint yield and impurity contents. These indicated the relative levels of uncertainty reduction required in different sources, for decisions concerning further data collection and resourcing levels.

Optimisation under model uncertainty of the complete process flowsheet is addressed in Chapter 8. An optimal operating policy which takes into account the combined influence of the important uncertainties in the sequence provides better performance than the nominal policy, in terms of a profitability objective. Specified tolerance to potential losses in profitability (due to violation of desired impurity content limits) is included through the use of one-sided stochastic constraints. Tolerance to error or uncertainty in this optimal operating policy is determined for a specified relaxation of the optimum objective function value. For this case study the key policy variables exhibiting tolerance to error were associated with the reaction and crystallisation operations. The value of perfect information in potentially measurable properties is determined for uncertainty in the purity of the active pharmaceutical ingredient (API) feed stream to the reactor. This knowledge was indicated to be very important for purity levels below a certain level beyond which significant drops in profitability could otherwise be incurred.

The third management response is addressed in Chapter 9. Implementation of the Risk Analysis approach in the stochastic optimisation techniques provide useful information for comparison between alternative process options based on the associated levels of process knowledge. For Case Study II this comparison is made between four alternative options with respect to the expected profitability objective and the associated uncertainty and the degree of operating policy error tolerance permitted. This allows a desired trade-off to be made at the discretion of the decision maker.

The main conclusion of this thesis is that the systematic incorporation of quantitative Risk Analysis methods for uncertainty inclusion in a model-based approach is able to provide useful information to aid the integrated design of complete pharmaceutical processes with even simple models. The results of the case studies provide evidence that the management of uncertainty in a systematic manner is not only prudent but also has the potential to improve the quality of the process. In the second case study, tracking of the uncertainty with incoming development data indicated the availability of which data had the most beneficial impact on the uncertainty contained in the models. With the indication of the key uncertainties
in the current process sequence models, engineering judgement can be used to deduce the presence of associated phenomena which are significant but unconsidered. This information could be used to focus efforts within the development process.

Since the results of the case studies appear reasonable and could be explained logically, some confidence can be drawn from the conclusions made. In the second case study, the relative importance of the reaction and crystallisation operations with regard to uncertainty in endpoint yield and impurity content levels provides confidence to the prior expectation that these would be the critical stages of the integrated process sequence. With this in mind it can be postulated that with the availability of more realistic models and more accurate quantification of the model uncertainty the results and conclusions would still be useful and valid (within the limiting assumptions).

The evidence provided in this work should provide encouragement for the incorporation of uncertainty management and model-based support tools into process development cultures. What is shown in this work is a basis for a systematic framework where process development procedures and process modelling tools are complementary. Referring back to Figure 2.1 (Chapter 2) indicating the typical roles of the key stake-holders within pharmaceutical process development, it can be postulated how such a combined framework can be used to link the different data collected and the resulting information determined by each stake-holder. The chemists provide a model of the intrinsic chemistry and reaction pathways. The development engineers then combine this knowledge with engineering phenomena associated with scale-up and pilot plant equipment within the paradigm of the framework. Though the same data may be collected and a rational basis for the collection of additional data provided, it is how this data is processed and how the information is used which is the essence of the framework. The integration of specific process operation knowledge with that of the other processes in the sequence is also seen as an essential element of process decision-making. The output is an estimate of the level of risk associated with the prediction of process performance given the current structured knowledge for the complete sequence and an evaluation of risk management response options, in an underlying (computational) framework where relevant information from all the stake-holders has been incorporated in a systematic and iterative manner. In this way the process characterisation passed to the pilot plant and ultimately the production engineers has a more rational and structured basis, which can be used to help determine better plant operating conditions and what measurements to take to improve the rationality. As Basu et al. (1999) pointed out, ultimately less scale-up iterations and higher quality processes and shorter development times could be expected.

However, it is acknowledged that a limiting factor depends on how systematic model development procedures can be successfully implemented without substantially increasing the strain on resources and in particular the time observed with current practices. While it is also acknowledged that some processes remain too difficult to model reliably and in other cases empirical models may still need to be incorporated, ultimately the aim is that with advances in the understanding of the processes a shift towards more mechanistic modelling approaches will follow. This is why the potentially important uncertainties should be systematically incorporated into the structured process knowledge to identify areas where improved mechanistic understanding would be beneficial. An important feature of the methodology
presented in this thesis is the flexibility to easily implement deterministic process models ranging from very simple mass balances to more complex models, so that useful information can be obtained from its application in early stages to the later stages of process development. The efficiency of the Hammersley sampling scheme and that of the employed Sensitivity Analysis measure estimations should permit the incorporation of more complex models with a greater number of parameters without creating an impractical computational demand. While it is proposed that the information and results determined in this work provide a new and useful contribution in this field, there certainly remains further issues which should be addressed in order to advance the case.

10.2 Future work

One of the main assumptions of the combined model development and Risk Analysis methodology is that the uncertainty is contained in the process model parameters and the model structure is correct. A further step would be to incorporate procedures which also determine the adequacy of the proposed model structures in the face of parameter uncertainties. Implementation of techniques for this concept would be strongly dependent on the quality of the data made available for modelling. One approach incorporates the fuzzy set theory for non-random uncertainties (Kubic and Stein, 1986). This was used to provide measures of model structure confidence which incorporates expert knowledge. They are expressed as membership functions and ignorance factors which test the validity of structural assumptions and the need for experimental data to validate the model. Zhang and Qian (1998) combined fuzzy and deterministic models into an integrated hybrid flowsheet so that poorly understood operations could be better expressed using expert knowledge. This could have application in the analysis of complete pharmaceutical process sequences where the solid phase operations are particularly poorly understood, but the objective remains to develop mechanistic models where justifiable.

Since uncertainty in process knowledge and model development are inherently dependent on the available data, further involvement in the design of laboratory experiments/pilot plant data collection is a logical advance for the methodology proposed in this thesis. Work by Pinto (1998), Martinez (2000) and Asprey and Macchietto (2000 and 2002) are examples where the link between experimental design and model discrimination and model parameter uncertainty is addressed. It is important that knowledge obtained from experimental designs be transformed into knowledge contained in deterministic models. With an increased emphasis on obtaining experimental data, methods like those of Vasquez and Whiting (1999) and Vachhani et al. (2001) considering systematic and random errors in the experimental data used to estimate model parameters, would add an important aspect to the analysis of the uncertainty. In the case of highly non-linear models, methods which determine more accurate estimations of joint correlated parameter uncertainties than is captured assuming linearisation methods employing Taylor series expansion based covariance matrices (as assumed in this thesis) would be more appropriate. The work of Vasquez and Whiting (2000) and Rooney and Biegler (2001) provide recent advances in this area.
Integrated design under uncertainty for pharmaceutical processes

The applicability of the approach proposed in this work to the difficulties of scale-up is worth consideration. While energy considerations were largely neglected in the case studies presented in this thesis for simplicity, future work should certainly consider uncertainty and optimisation in more rigorous models which describe heat transfer and mixing phenomena as are often problematically encountered.

Application of process modelling towards the examination of critical parameters, safe determination of practical achievable operating ranges and process specification failure limits in Sensitivity Analysis and process validation and possibly even in conjunction with the stipulations of the regulatory authorities has already drawn the attention of Wright and Bramfitt (1997) and Basu et al. (1999). The work presented in this thesis aims to provide one step towards the progress required to meet these visions.

The formulation of the stochastic optimisation of the process flowsheet could be improved by including information concerning state variable measurements by using feedback correction policy methods (as proposed by Terwiesch et al. (1994) and Bhatia and Biegler (1997). This would reduce the degree of conservatism of the 'here and now' approach but avoid the optimistic and unrealistic assumptions inherent in the 'wait and see' approach.

Since the formulation of the stochastic programming problem with one-sided stochastic constraints is approximated as a stochastic optimisation with non-convex binary approximations, this results in a non-convex NLP for which gradient information may not be of good quality and the global optimality of solutions is not guaranteed. Alternative non-gradient based optimisation methods (e.g. genetic algorithms, simulated annealing) to the SQP method used in this work are designed to be more reliable in obtaining non-local optima and less dependent on the initial points. However, consideration of the computational demand such methods require is important. The efficient Hammersley sequence sampling method as used in this work to approximate uncertain quantities still required more than a hundred deterministic model evaluations per objective function evaluation. The incorporation of more complex models may prove to be computationally prohibitive. The cubature integration method introduced by Bernardo et al. (1999) improves efficiency in some cases but a more generic method would be useful. Consideration should be given to other low discrepancy number sequences which exhibit greater efficiency in P-dimensional space.
REFERENCES


[22]. Bollyn M., A. van den Bergh and A. Wright, Accelerated scale-up: Reaction calorimetry and reactor simulation - combined techniques accelerate scale-up of a highly exothermic reaction, 1995, 7th RC User Forum, Interlaken, Switzerland.


[111]. Shah N., N.J. Samsatli and M. Sharif, Modelling and optimisation for pharmaceutical and fine chemical process development, 1999, FOCAPD, Colorado, USA.


Integrated design under uncertainty for pharmaceutical processes

**NOTATION**

\(a\) = Taguchi loss proportionality constant or radix-R integer coefficients

\(A\) = Fourier cosine coefficient or Arrhenius coefficient (dm\(^3\) mol\(^{-1}\) hr\(^{-1}\) second order reaction)

\(AF\) = fraction of a total solvent feed quantity added to an operation

\(b\) = least squares point estimate of a multi-linear regression coefficient, or vector of these estimates

\(B\) = Fourier sine coefficient

\(c\) = sequence of points generated during computation of the Hammersley sequence points

\(C\) = general cost function or desired correlation matrix of an uncertain factor matrix or material price ($ kg\(^{-1}\) material)

\(C^*\) = desired rank correlation matrix of an uncertain factor matrix

\(C\text{_{Tag}}\) = cost incurred for a Taguchi performance tolerance violation

\(\hat{C}\) = parameter correlation matrix

\(\text{Cov}\) = covariance

\(\text{CC}\) = correlation coefficient

\(\text{CR}\) = correlation ratio or crystallisation cooling rate (°C min\(^{-1}\))

\(d\) = (equipment) design variables

\(d_1, d_2\) = general one-sided deviation functions

\(\text{det}\) = determinant

\(\text{DC}\) = differential coefficient

\(\text{DR}\) = drying rate (min kg\(^{-1}\) solids)

\(e\) = Hammersley sequence points

\(E\) = expected value or correlation matrix of a substitution matrix of the uncertain factor matrix

\(E_a\) = activation energy (dm\(^3\) mol\(^{-1}\) hr\(^{-1}\), second order reaction)

\(E_{\text{viol}}\) = expected value of the extent of a deterministic constraint violation

\(f\) = general deterministic model equations

\(f_{\text{dev}}\) = general two-sided deviation function
Integrated design under uncertainty for pharmaceutical processes

$F$ = deterministic flexibility index or general stochastic objective function or feed mass (kg)

$F_{\alpha}$ = value of the F-distribution with an area $\alpha$ to its right

$FCV$ = fractional contribution to output performance variance

$FR$ = filtration rate ($\text{min kg}^{-1} \text{solids}$)

$FW_{\alpha_1,\alpha_2}$ = width between $\alpha_1$ % and $\alpha_2$ % fractiles

$g$ = general deterministic inequality constraints

$G$ = Fourier transformation function

$h$ = general deterministic equality constraints

$h_{\text{band}}$ = height of a two-phase dispersion band (m)

$H$ = Hessian matrix

$HP$ = crystallisation process isothermal holding period (min)

$J$ = Jacobian matrix of model parameter estimation predictions or total number of inequality constraints

$J_0$ = general deterministic initial conditions

$k$ = constant in Chebyshev's rule or reaction rate constant (units dependent on reaction order)

$k_g$ = crystal growth constant ($\text{m min}^{-1} (\text{kg kg}^{-1} \text{solvent})^{1/2}$)

$K$ = substitution matrix of the uncertain factor matrix

$K'$ = rearranged substitution matrix of the uncertain factor matrix

$L$ = Taguchi's cost associated with a quality loss or a lower triangular matrix of a desired correlation matrix

$\text{LOD}$ = level of dampness of solids (%)

$m_a$ = number of preceding sample observations used for convergence tolerance test

$m_b$ = current sample observation iteration number

$M$ = total number of scenarios, periods or sample observations

$M'$ = total number of sample observations (scenarios) in potentially measurable uncertainty space

$N$ = number of measurement data points used in a model parameter estimation problem or normal distribution or agitation speed (rpm)
<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n_i$</td>
<td>sample size required to evaluate individual variance integrals in measures of importance</td>
</tr>
<tr>
<td>$p$</td>
<td>purity (wt%)</td>
</tr>
<tr>
<td>$P$</td>
<td>total number of uncertain factors (dimensions) or number of parameters in a model parameter estimation problem</td>
</tr>
<tr>
<td>$P_{ft}$</td>
<td>profit</td>
</tr>
<tr>
<td>$Pr$</td>
<td>probability</td>
</tr>
<tr>
<td>$Pr_{pass}$</td>
<td>probability of passing a constraint</td>
</tr>
<tr>
<td>$Pr_{viol}$</td>
<td>probability of deterministic constraint violation</td>
</tr>
<tr>
<td>$PC$</td>
<td>perturbation coefficient</td>
</tr>
<tr>
<td>$PCC$</td>
<td>partial correlation coefficient</td>
</tr>
<tr>
<td>$Pty$</td>
<td>profitability ($\text{kg API}^{-1} \text{ hr}^{-1}$)</td>
</tr>
<tr>
<td>$Pty_{loss}$</td>
<td>potential profitability loss ($\text{kg API}^{-1} \text{ hr}^{-1}$)</td>
</tr>
<tr>
<td>$Q$</td>
<td>inter-stage sub-sequence quality criterion or lower triangular matrix of an non-identity correlation matrix</td>
</tr>
<tr>
<td>$R$</td>
<td>feasible region or prime number</td>
</tr>
<tr>
<td>$R^2$</td>
<td>coefficient of determination</td>
</tr>
<tr>
<td>$R_{sol_back}$</td>
<td>fraction of a feed quantity of solvent used in a back extraction</td>
</tr>
<tr>
<td>$R_{sol_recycle}$</td>
<td>fraction of a feed quantity of solvent which is from a recycle</td>
</tr>
<tr>
<td>$RF$</td>
<td>fraction of a solvent feed quantity removed in an operation</td>
</tr>
<tr>
<td>$RSC$</td>
<td>response surface coefficient</td>
</tr>
<tr>
<td>$RSS$</td>
<td>residual sum of squares between the vector of deterministic model performance outputs and the vector of regression model performance outputs</td>
</tr>
<tr>
<td>$s$</td>
<td>sample standard deviation or Fourier space</td>
</tr>
<tr>
<td>$s^2$</td>
<td>residual variance of a model parameter estimation problem</td>
</tr>
<tr>
<td>$S^{\text{FAST}}$</td>
<td>Fourier amplitude sensitivity test index</td>
</tr>
<tr>
<td>$So$</td>
<td>Sobol' sensitivity index</td>
</tr>
<tr>
<td>$SF$</td>
<td>stochastic flexibility index</td>
</tr>
<tr>
<td>$SRC$</td>
<td>standardised regression coefficient</td>
</tr>
<tr>
<td>Symbol</td>
<td>Definition</td>
</tr>
<tr>
<td>--------</td>
<td>------------</td>
</tr>
<tr>
<td>$SS_{\phi\phi}$</td>
<td>sum of squares of the distances from the mean value of the deterministic model performance output</td>
</tr>
<tr>
<td>$SS_{\phi\theta}$</td>
<td>sum of squares of the distances from the mean value of the uncertain factor</td>
</tr>
<tr>
<td>$SS_{\phi\theta}$</td>
<td>sum of products of distances from the deterministic model performance output and uncertain factor values from their respective means</td>
</tr>
<tr>
<td>$SSC$</td>
<td>sub-sequence contribution to endpoint performance uncertainty</td>
</tr>
<tr>
<td>$t$</td>
<td>time (hr, Case Study I or min, Case Study II - units analogous with available case study data)</td>
</tr>
<tr>
<td>$t''$</td>
<td>time at which initial rate limited period ends, Case Study II (min)</td>
</tr>
<tr>
<td>$t_{\alpha}$</td>
<td>value of the Student t-distribution with an area $\alpha$ to its right</td>
</tr>
<tr>
<td>$T$</td>
<td>bounded region of deterministic uncertainties or temperature (K, Case Study I or °C, Case Study II - units analogous with available case study data)</td>
</tr>
<tr>
<td>$TSS$</td>
<td>total sum of squares of the distances between deterministic model performance outputs and its mean</td>
</tr>
<tr>
<td>$u_1$</td>
<td>fraction of drugs originally in the organic phase which are soluble in the aqueous phase (reagent addition model, Case Study II)</td>
</tr>
<tr>
<td>$u_2$</td>
<td>fraction of drugs originally in the organic phase which is cut into the aqueous phase (layer separation model, Case Study II)</td>
</tr>
<tr>
<td>$U$</td>
<td>uniform distribution</td>
</tr>
<tr>
<td>$\hat{V}$</td>
<td>covariance matrix</td>
</tr>
<tr>
<td>$\text{Var}$</td>
<td>variance</td>
</tr>
<tr>
<td>$VI$</td>
<td>value of information</td>
</tr>
<tr>
<td>$\text{VPI}$</td>
<td>value of perfect information</td>
</tr>
<tr>
<td>$w$</td>
<td>distribution parameter (mean or standard deviation)</td>
</tr>
<tr>
<td>$wf$</td>
<td>weighting factor</td>
</tr>
<tr>
<td>$wt$</td>
<td>weight percent (%)</td>
</tr>
<tr>
<td>$W$</td>
<td>level of dampness (% LOD)</td>
</tr>
<tr>
<td>$\dot{x}$</td>
<td>derivatives of differential variables with respect to time</td>
</tr>
<tr>
<td>$x$</td>
<td>state variables or differential variables</td>
</tr>
<tr>
<td>$X$</td>
<td>matrix of uncertain factor values or conversion</td>
</tr>
</tbody>
</table>
$X^*$ = matrix of uncertain factor values with a desired correlation structure

$y$ = algebraic variables

$Y$ = yield, based on active pharmaceutical feed (%)

$z$ = control (operating policy) variables

$Z$ = mass (kg)

Greek letters

$\alpha$ = confidence

$\beta$ = binary variable

$\chi$ = objective value of the max-min-max sub-problem equivalent to the feasibility test

$\delta$ = fraction of the original spread of the uncertainty

$\delta^2$ = fraction of the variability of the deterministic model performance output that is explained by the variability in the uncertain factors

$\Delta$ = deviation or tolerance

$\varepsilon$ = error or tolerance limit

$\phi$ = deterministic quality criterion

$\Phi$ = deterministic objective quality criterion (or a vector of these values)

$\hat{\phi}$ = predicted value of a quality criterion from a regression model (or a vector of these values)

$\bar{\Phi}$ = sample mean of deterministic model performance output

$\gamma$ = desired constraint target

$\gamma_1$ = power characterising the mixing regime on the rates of reaction

$\gamma_1$ = power characterising the effect of mixing on the parameter characterising the initial rate limiting period

$\eta_{\text{wash}}$ = crystal wash operation efficiency

$\varphi$ = inverse radix number

$\kappa$ = mean-variance weighting parameter

$\mu$ = population mean or nominal value (or vector of these values)

$\nu$ = time dependent operating variables or gaussian quadrature collocation point
Integrated design under uncertainty for pharmaceutical processes

\[ \Theta \] = uncertain factor (or vector of these values)
\[ \overline{\Theta} \] = sample mean of uncertainty factor
\[ \Theta \] = total uncertainty space
\[ \Theta_U \] = input uncertainty space over uniformly distributed space
\[ \Theta_N \] = input uncertainty space over joint normal distributed space
\[ \rho \] = Pearson product moment population correlation coefficient
\[ \sigma \] = population standard deviation
\[ \sigma_{eq} \] = equilibrium solubility of solute in organic-aqueous phase mixture (wt%)
\[ \nu \] = time invariant operating decisions
\[ \omega \] = angular frequency
\[ \xi \] = binary approximation parameter
\[ \Psi \] = penalty function for the penalisation of stochastic constraint violations
\[ \zeta \] = first order rate constant for loss of impurity concentration (min\(^{-1}\))

Subscripts

\[ \text{add} \] = addition
\[ \text{d} \] = index for operating policy variables, of total D
\[ \text{diss} \] = dissolution
\[ \text{dt} \] = index for uniformly distributed input uncertainties, of total DT
\[ \text{f} \] = final or feed
\[ \text{here} \] = 'here and now' optimal solution
\[ \text{i} \] = index for iteration number
\[ \text{iso} \] = isothermal
\[ \text{j} \] = index for Fourier coefficients or index for inequality constraints
\[ \text{k} \] = index for radix-R integer coefficients
\[ \text{m} \] = index for sample observations (scenarios) in uncertainty space, of total M
\[ \text{m'} \] = index for sample observations (scenarios) in potentially measurable uncertainty space, of total M'
\[ \text{n} \] = index for data points used in a model parameter estimation problem, of total N
nominal = nominal optimum operating policy
non-iso = non-isothermal
o = value at time zero
p = index for parameters estimated in a model parameter estimation problem or in the entire stochastic problem, of total P
pc = index for potentially measurable uncertainties, of total PC
pe = index for time period in optimal control problem
q = index for output performance criteria, of total Q
r = index for uncertain factors in a stochastic problem, where r \neq p
robust = robust optimum operating policy
s = index for process stages, of total S
ss = index for process sub-sequences, of total SS
st = index for normally distributed input uncertainties, of total ST
std = standardised
T = endpoint
v = index associated with radix-R integer coefficients
wait = 'wait and see' optimal solution

Superscripts

c = critical value
L = lower range
LB = lower bound
N = nominal value
th = constraint threshold value
U = upper range
UB = upper bound
* = optimal value
' = original value (associated with a prior problem solution)
Appendix A

A. CASE STUDY I

The deterministic model equations, uncertain parameter models and additional results for Case Study I are stated in Appendix A.

A.1 Deterministic process model for semi-batch reactor

Component mole balance,

\[
\frac{dm_A}{dt} = -k_1 \frac{m_A m_B}{V_T} - k_2 \frac{m_A m_C}{V_T}
\]

\[
\frac{dm_B}{dt} = v \frac{m_{b0}}{V_{b0}} - k_1 \frac{m_A m_B}{V_T}
\]

\[
\frac{dm_C}{dt} = k_1 \frac{m_A m_B}{V_T} - k_2 \frac{m_A m_C}{V_T}
\]

\[
\frac{dm_D}{dt} = k_2 \frac{m_A m_C}{V_T}
\]

Arrhenius equations,

\[
k_1 = \left[ (1 - \beta_{diss}) A_1 \exp\left( \frac{-Ea_1}{T_{iso}} \right) \right] + \left[ \beta_{diss} A_{diss} \exp\left( \frac{-Ea_{diss}}{T_{iso}} \right) \right]
\]

\[
k_2 = A_2 \exp\left( \frac{-Ea_2}{T_{iso}} \right)
\]

\[
\beta_{diss} = \begin{cases} 
1 & \text{if } X_A \leq X_{diss} \\
0 & \text{if } X_A > X_{diss}
\end{cases}
\]

Volume change,

\[
\frac{dV_T}{dt} = v \bigg|_{v=0} \bigg|_{t_{add} \rightarrow 0} = \frac{V_{b0}}{t_{add}}
\]

Termination condition,
$X_A = \frac{m_{A0} - m_A}{m_{A0}}$

$$\frac{dX_A}{dt} = \left( \frac{k_1 m_A m_B}{V_T m_{A0}} + \frac{k_2 m_A m_C}{V_T m_{A0}} \right) \leq 0.001 \text{ hr}^{-1}$$

Cooling capacity constraint on operating policy variables,

$$T_{iso,max} \geq 7.06(t_{add})^2 - 43.50(t_{add}) + 352.67$$

Output criterion,

$$Y_D = \frac{m_D}{m_{A0}} \times 100\%$$

Initial conditions (inside reactor)

$m_{A0} = 1.075 \text{ moles}$

$m_{B0} = 0$

$m_{C0} = 0$

$m_{D0} = 0$

$V_T = 0.7 \text{ dm}^3$

$X_{A0} = 0$  \hspace{1cm} (Model A1)

where

$A = \text{ Arrhenius coefficient, dm}^3 \text{ mol}^{-1} \text{ hr}^{-1} \text{ (second order elementary reaction)}$

$E_a = \text{ activation energy, kJ mol}^{-1}$

$k = \text{ rate constant, dm}^3 \text{ mol}^{-1} \text{ hr}^{-1} \text{ (second order elementary reaction)}$

$m = \text{ moles, mol}$

$t = \text{ time, hr}$

$T = \text{ temperature, K}$

$V = \text{ volume, dm}^3$

$X = \text{ conversion}$

$Y = \text{ yield, } \%$

$\nu = \text{ volumetric flowrate, dm}^3 \text{ hr}^{-1}$

Subscripts
The assumptions made in this model are:

- elementary reactions,
- perfect mixing,
- duration of the initial dissolution period is only dependent on conversion,
- constant rate of feed B addition,
- isothermal operation,
- use of a quadratic approximation for the cooling capacity limitation on the operating policy variables (see Figure A1).

The available decision variables are the addition time (of a predetermined quantity of B), $t_{\text{add}}$, and the isothermal temperature, $T_{\text{iso}}$. The total moles of B added is 3.01 moles, the value of $X_{\text{diss}}$ is 0.55 and the estimated values of $E_{\text{adi}}$, $E_{A_1}$, and $E_{A_2}$ are 78.98, 90.46 and 65.97 kJ mol$^{-1}$, respectively, and $A_{\text{adi}}$, $A_{A_1}$, and $A_{A_2}$ are $1.76 \times 10^{13}$, $4.68 \times 10^{15}$ and $1.00 \times 10^{10}$ dm$^3$ mol$^{-1}$ hr$^{-1}$, respectively. The kinetic parameters were determined through plots of $\ln(k)$ versus $1/T$, where the experimental data points have been taken from Sano et al. (1998).
A.2 Uncertainty and Sensitivity Analysis

The probability distribution characterisations for the parameter uncertainties considered in the Uncertainty Analysis for Case Study I are assumed to be uncorrelated and are given in Table A1. Figure A2 shows the evolution of the mean and variance parameters for the predicted impurity yield. The scatter plot between normalised values of impurity Yield D and the sub-reaction activation energy, Figure A3 (a), shows the effect of rank transformation for the computation of monotonic but non-linear relationships using linear
Table A1. Uncertainty characterisation in the parameters of Case Study I.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Normal distribution, N(μ, σ)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Uniform distribution, U(min, max)</td>
</tr>
<tr>
<td>$E_a_1$</td>
<td>N(90.46 × 10^3, 10% nominal)</td>
</tr>
<tr>
<td>$A_1$</td>
<td>N(4.68 × 10^15, 10% nominal)</td>
</tr>
<tr>
<td>$E_a_2$</td>
<td>N(65.97 × 10^3, 10% nominal)</td>
</tr>
<tr>
<td>$A_2$</td>
<td>N(1.00 × 10^10, 10% nominal)</td>
</tr>
<tr>
<td>$E_{adiss}$</td>
<td>N(78.98 × 10^3, 10% nominal)</td>
</tr>
<tr>
<td>$A_{diss}$</td>
<td>N(1.76 × 10^13, 10% nominal)</td>
</tr>
<tr>
<td>$X_{diss}$</td>
<td>N(0.55, 10% nominal)</td>
</tr>
<tr>
<td>$T_{iso}$</td>
<td>U(-1°C(μ_{Tiso}), +1°C(μ_{Tiso}))</td>
</tr>
</tbody>
</table>

(a) Effect of rank transformation (nominal policy). Key: ● = unranked data (SRC = -0.59), ○ = ranked data (SRRC = -0.76)

(b) Effect of different optimal operating policies on rank transformed data. Key: ● = robust policy (SRRC = -0.88), ○ = nominal policy (SRRC = -0.76).

Figure A3. Normalised scatter plots between Yield D and sub-reaction activation energy ($E_{a_2}$) sample results, Case Study I.

measures. Figure A3 (b) shows the increased strength in the monotonic relationship between Yield D and sub-reaction activation energy under the robust optimum operating policy.

A.3 Accuracy of the hyperbolic smoothing function

The binary variable approximation curves in Figure A4 (a) and (b) show the hyperbolic smoothing function for values of 160 and 40 for the impurity yield and final time criteria, respectively. At these
Integrated design under uncertainty for pharmaceutical processes

(a) Impurity yield.

(b) Final time.

Figure A4. Binary variable approximation hyperbolic smoothing curves, Case Study I.

values of the constants, the binary approximations calculate zero or one for predictions that fall within 1\% (dotted lines) of the desired threshold value (dashed lines).

A.4 Computation statistics

The computational statistics for the optimisation problems solved for Case Study I are listed in Table A2. They were solved using the MATLAB (Version 6.0) programming software and the Optimisation Toolbox function for non-linear constrained optimisation algorithm based on SQP. They were performed on a RS6000 IBM workstation. The CPU times for the stochastic optimisation problems were approximately two orders of magnitude greater than the CPU times for the respective nominal optimisation problems, indicating that the direct solution of a stochastic problem poses a substantial computational burden even with the use of an efficient sampling technique.

Table A2. Computational statistics for optimisation problems, Case Study I.

<table>
<thead>
<tr>
<th>Problem formulation</th>
<th>Problem formulation</th>
<th>CPU time (seconds)</th>
<th>Optimisation iterations</th>
<th>Function evaluations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nominal isothermal process optimisation</td>
<td>P4</td>
<td>5.3</td>
<td>20</td>
<td>85</td>
</tr>
<tr>
<td>Stochastic isothermal process optimisation</td>
<td>P5</td>
<td>339</td>
<td>13</td>
<td>55</td>
</tr>
<tr>
<td>Nominal non-isothermal process optimisation</td>
<td>P6</td>
<td>27.3</td>
<td>36</td>
<td>361</td>
</tr>
<tr>
<td>Stochastic non-isothermal process optimisation (over a number of starting points)</td>
<td>P7</td>
<td>1×10^3 to 8×10^3</td>
<td>50 to 85</td>
<td>500 to 800</td>
</tr>
</tbody>
</table>
Appendix B

B. CASE STUDY II: FIRST GENERATION PROCESS MODELS

The deterministic models, associated modelling assumptions and uncertain parameter characterisations for the processes comprising the complete process sequence investigated in Case Study II are stated in Appendix B. In addition, an expression is given which is used to account for additional levels of parameter uncertainty due to the violation of desired ratios or ranges in inter-stage state variable measurements. Some additional Uncertainty Analysis results for Case Study II are presented.

B.1 Reaction model

The available bench scale data for the reaction comprises concentration-time profiles of each drug species, reG and reH. The following experimental results and analysis are obtained from private communication with a pharmaceutical company. These are discussed to understand the assumptions in the model. With this data it is assumed an intrinsic kinetic model for the reaction process can be developed. Since the fates of aqueous reagent, reH, solid organic reagent, reG, and the resulting oxidant, oxG, are complex and not well understood, it is not possible to develop a rigorous kinetic model which accounts for these species. Under the conditions used oxG does not appear to be a limiting factor in the kinetics of the drug reactions. Simplified pseudo-first order reaction kinetics based on the stoichiometry shown in Equations B1 and B2, are assumed in the organic solvent phase. The first generation model for the Stage 1 reaction is given in Model B1 and the process diagram is shown in Figure B1.

\[
\begin{align*}
\text{actA}_{\text{org}} & \rightarrow \text{actB}_{\text{org}} \rightarrow \text{actC}_{\text{org}} & (\text{B1}) \\
\text{actD}_{\text{org}} & \rightarrow \text{actE}_{\text{org}} & (\text{B2})
\end{align*}
\]

![Figure B1. Stage 1 reaction, Case Study II.](image)

Feed specification,

\[
f_{\text{drug},i} = \frac{P_f}{100} F_{\text{drug}}
\]
Integrated design under uncertainty for pharmaceutical processes

\[ F_{relH} = \text{molratio}_{relH} \frac{RMM_{relH}}{RMM_{drug,i}} \times \frac{f_{drug,i}}{\text{wtag}_{relH}} \times 100 \]

\[ V_{relH} = \frac{F_{relH}}{\rho_{relH}} \]

Drug component mole balance,

\[ \frac{dm_{1,org}}{dt} = -k_1 m_{1,org} \]

\[ \frac{dm_{2,org}}{dt} = k_1 m_{1,org} - k_2 m_{2,org} \]

\[ \frac{dm_{3,org}}{dt} = k_2 m_{2,org} \]

\[ \begin{align*}
\frac{dm_{4,org}}{dt} &= 0 \\
I_{i',i'} &= -k_3 m_{4,org} \\
\frac{dm_{5,org}}{dt} &= 0 \\
I_{i',i'} &= k_3 m_{4,org}
\end{align*} \]

Solvent balance,

\[ z_{1,6,org} = F_{solF} \]

\[ z_{1,7,org} = 0 \]

\[ z_{1,8,org} = F_{relH} \]

End point criteria,

\[ z_{1,i,org} = \frac{RMM_{i,org} m_{1,org}}{1000} \quad \text{for } i = 1..5 \]

\[ X_1 = \frac{m_{0,1,org} - m_{1,org}}{m_{0,1,org}} \]

\[ \text{comp}_i = \frac{z_{1,i,org}}{\sum_{i=1}^{5} z_{1,i,org}} \times 100 \quad \text{for } i = 1..5 \]

Initial conditions,

\[ m_{0,i} = \frac{f_{drug,i}}{RMM_{drug,i}} \times 1000 \quad \text{for } i = 1..5 \]

(Model B1)

where
Integrated design under uncertainty for pharmaceutical processes

comp = composition by weight, wt%
f = component feed mass, kg
F = total stream feed mass, kg
k_1 = first order rate constant for main reaction, min^{-1}
k_2 = first order rate constant for consecutive reaction, min^{-1}
k_3 = first order rate constant for sub-reaction, min^{-1}
m = moles
molratio = mole ratio of reagent to moles of active pharmaceutical ingredient in feed
ρ_r = purity of active pharmaceutical ingredient feed, wt%
RMM = relative molecular mass (in aqueous streams, refers to RMM of solute compound)
V = volume, m^3
wtaq = reagent strength in aqueous solution, wt%
X = conversion
z = mass, kg
ρ = density, kg m^{-3}

subscripts
aq = aqueous phase
i = component species \{actA, actB, actC, actD, actE, solF, solL, aq\}
o = initial condition
org = organic phase
t_o = reaction starting time (zero), min
t' = time at which sub-reaction starts, min
t_f = total time of Stage I operation, min

Possible degrees of freedom include \( F_{drug}, \) \( molratio_{act} \), \( wtaq_{act} \) and \( t_f \). The main assumptions made in this model include:

- pseudo first order stoichiometry and elementary reaction kinetics (oxG in excess throughout),
- observation of intrinsic kinetics in the assumed reactions (perfect mixing throughout),
- the sub-reaction for the secondary impurity (actE) starts after 60 minutes,
- no feed solids dissolution effects.
• instantaneous addition of reH feed,
• reG and reH species are not explicitly modelled, since their fate is not understood,
• isothermal operation and no other limiting heat transfer effects,
• no mass transfer of drug species from the organic phase to the aqueous phase.

The fitted parameter values for the bench scale model are given in Table B1. Since the predicted drug component profiles exhibit a reasonably good fit to the bench scale data, see Figure B2, the assumptions of pseudo-first order kinetics and perfect mixing appear satisfactory for this system.

<table>
<thead>
<tr>
<th>Table B1. Parameters for the bench scale Stage I model, Case Study II.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fitted model parameter values</td>
</tr>
<tr>
<td>$k_1 = 0.0169 \text{ min}^{-1}$</td>
</tr>
<tr>
<td>$k_2 = 7.14 \times 10^{-5} \text{ min}^{-1}$</td>
</tr>
<tr>
<td>$k_3 = 1.66 \times 10^{-3} \text{ min}^{-1}$</td>
</tr>
<tr>
<td>$\tau' = 60 \text{ min}$</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

(a) Key to data points: $\circ = \text{actA}$, $\ast = \text{actB}$, $\times = \text{actD}$. (b) Key to data points: $\ast = \text{actC}$, $+$ = actE.

Figure B2. Bench scale drug profile predictions for the first generation Stage I model, Case Study II.

**B.2 Reagent addition model**

A mass balance comprises the first generation model used to describe the addition of aqueous reagent operations. The first generation model for the Stage 2, 4, 5 and 7 reagent addition operations is given in Model B2 and the process diagram is shown in Figure B3.
Feed specification,

\[ F = V_F \rho_F \quad \text{for distilled water} \]

or

\[ F = \text{molratio}_F \frac{RMM}{RMM_1} \frac{100}{\text{wtaq}_F} \quad \text{for aqueous solution} \]

Drug component balance,

\[ z_{2,i,\text{org}} = (1 - u_1) z_{1,i,\text{org}} \quad \text{for } i = 1..5 \]

\[ z_{2,i,\text{aq}} = (u_1 z_{1,i,\text{org}}) + z_{1,i,\text{aq}} \quad \text{for } i = 1..5 \]

Solvent balance

\[ z_{2,i,\text{org}} = z_{1,i,\text{org}} \quad \text{for } i = 6..8 \]

\[ z_{2,8,\text{aq}} = z_{1,8,\text{aq}} + F \]

\[ z_{2,i,\text{aq}} = z_{1,i,\text{aq}} \quad \text{for } i = 6,7 \quad \text{(Model B2)} \]

where \( u_1 \) is a parameter representing the fraction of drugs in the organic phase of the input stream, which are soluble in the aqueous phase and the component species, \( i \), are \{actA, actB, actC, actD, actE, solF, solL, aq\}. An additional parameter, \( u_o \), not explicitly expressed in Model B2, is used specifically for the Stage 2 dilution model to represent the time in minutes, after the desired Stage 1 termination point (\( t_{1,1} \), whereupon the diluent is added), that passes before the reaction has fully terminated. Degrees of freedom could be either \( V_F \) for the addition of distilled water or \( \text{molratio}_F \) and \( \text{wtaq}_F \) for the addition of an aqueous reagent solution. The main assumptions of this model include:

- instantaneous addition of feed stream, \( F \),
- instantaneous reactions, no mixing effects,
- no limiting heat transfer effects,
- any mass transfer of drugs to the aqueous phase, \( z_{2,i,\text{aq}} \) due to solubility, is represented by parameter \( u_1 \), and is assumed instantaneous and the same proportion for each drug species.
B.3 Layer separation model

A simple mass balance comprises the first generation model used to describe layer separations. The first generation model for the Stage 3, 6 and 8 layer separation operations is given in Model B3 and the process diagram is shown in Figure B4.

```
Component mass balance

\[ z_{2,i,\text{org}} = (1 - u_2)z_{1,i,\text{org}} \quad \text{for } i = 1..8 \]
\[ z_{3,i,\text{org}} = u_2 z_{1,i,\text{org}} \quad \text{for } i = 1..8 \]
\[ z_{3,i,aq} = z_{1,i,aq} \quad \text{for } i = 1..8 \]  

(Model B3)
```

where \( u_2 \) is a parameter representing the fraction of organic phase of the input stream, which is an undesired cut in the aqueous waste stream phase and the component species, \( i \), are \{actA, actB, actC, actD, actE, solF, solL, aq\}. There are no degrees of freedom in this model. The main assumptions made in this model include:

- light aqueous phase (reH) is disperse, heavy organic phase (solF) is continuous,
- no mixing or time-dependent effects,
- no aqueous phase hold-up in the output organic stream, \( z_2 \), so that the efficiency of subsequent chemical destruction or solvent exchange operations is maintained (i.e. instead a small amount of organic phase loss, \( z_{3,\text{org}} \), is incurred in the aqueous phase cut, \( z_3 \)),
- any organic phase hold up (within a dispersion band) retained in the aqueous waste stream, \( z_3 \), is characterised by the parameter fraction \( u_2 \),
- no drug solubility in the aqueous phase.
B.4 Batch distillation model

The model used here is a batch distillation from a reboiler, with an energy balance to evaluate the vapour flowrate to the top product. A total condenser is not explicitly modelled but its operation is assumed. The first generation model for the Stage 9, 10 and 11 solvent exchange operations is given in Model B4 and the process diagram is shown in Figure B5.

\[
\frac{dm_{L,i}}{dt} = -y_i V_{\text{flow}} 
\]

for \( i = 1 \ldots 8 \)

\[ z_{2,i} = m_{L,i} RMM_i 
\]

for \( i = 1 \ldots 8 \)

\[ z_{3,i} = \left( F_i \cdot \frac{\text{comp} F_i}{100} \right) - z_{2,i} 
\]

for \( i = 1 \ldots 8 \)

Energy balance

\[
\frac{d(M_L H_L)}{dt} = -H_V V_{\text{flow}} + Q_r
\]

Vapour-liquid equilibrium

\[ y_i = x_i K_i 
\]

for \( i = 1 \ldots 8 \)

\[ K_i = \frac{P_{o,i}}{P} 
\]

for \( i = 1 \ldots 8 \)

Total liquid mole balance,

\[ M_L = \sum_{i=1}^{8} m_{L,i} \]

Product volume,

\[ V_{\text{top}} = \frac{\sum_{i=1}^{8} z_{3,i}}{\sum_{i=1}^{8} y_i P_i} \]
Integrated design under uncertainty for pharmaceutical processes

\[ V_{\text{tot}} = \left( \sum_{i=6}^{8} z_{2,i} \right) + \left( \sum_{i=6}^{8} z_{2,i} \right) \left( 1 - \sum_{i=6}^{8} x_{i} \text{grad}_i \right) \]

\[ \sum_{i=6}^{8} x_{i} \rho_i \]

Enthalpy,

\[ H_L = \sum_{i=1}^{8} x_i C_p_{L,i} \]

\[ H_V = \sum_{i=1}^{8} x_i \left( C_p_{L,i} + dH_{pV_0,i} \right) \]

Liquid mole fraction,

\[ m_{L,i} = x_i M_L \quad \text{for} \quad i = 1..8 \]

Normalisation equations,

\[ \sum_{i=1}^{8} y_i = 1 \]

\[ \sum_{i=1}^{8} x_i = 1 \]

Initial conditions,

\[ m_{\text{v},i} = \left( F_i \frac{\text{wt}\% \text{ feed}, i + z_{1,i}}{100} \right) \left( \frac{1}{RMM_i} \right) \quad \text{for} \quad i = 1..8 \quad (\text{Model B4}) \]

where

\( C_p \) = pure component heat capacity, J/kmol/K

\( dH_{pV_0} \) = pure component heat of vaporisation, J/kmol

\( F \) = mass solvent feed, kg

\( H \) = enthalpy, J/kmol

\( K \) = vapour liquid equilibrium K-value

\( m \) = moles, kmol

\( M \) = total moles

\( p_o \) = pure component pressure, Pa

\( P \) = total pressure, Pa

\( Q_r \) = reboiler duty J/h
The assumptions made in this model include:

- distillation time is an important variable otherwise a constant vaporisation rate model could be used,

- the condenser is not explicitly modelled (not assumed to be a limiting factor), but total condensation is assumed in the equation for top volume ($V_{\text{top}}$),

- the physical properties, $dH_{pv}$, and $C_{pl}$, for the drug components are unknown, so the properties for dioctylphthalate (ChemCAD V database, Chemstations, Inc., USA), $C_{12}H_{38}O_4$ (RMM = 390), are assumed due to a similarity in RMM (its properties predict no vaporisation under the range of operating conditions considered in this case study),

- the physical property methods used are: ideal VLE from K-value model, the generic physical property equations for pure component vapour pressure, heat of vaporisation for pure liquid components and pure liquid component heat capacities, as specified in the physical property library of the ChemCAD V simulation software,

- operation at zero reflux and 1 bar pressure,

- the assumption that an estimation for the maximum available reboiler duty per hour is available,

- the equation for bottom volume ($V_{\text{bot}}$) is derived from an assumed linear function between drug solute concentration and solution density when solF is the only solvent present (the bottom volume estimation is only required for the solF single solvent case), where the gradient ($\text{grad}_i$) is assumed to be 0.5 for solF ($i=6$) and the intercept is the pure solvent density ($\rho$).
Possible degrees of freedom in this model could include the reboiler duty, $Q_R$, and the initial quantity of mixture, $z_i$, and pure solvent feed, $F$.

### B.5 Cooling batch crystallisation model

The cooling batch crystalliser model used in this case study incorporates conventional growth kinetics for the product drug component, in which the method of moments is used to solve the population balance. Hulbert and Katz (1964). The first generation model for the Stage 12 crystallisation is given in Model B5 and the process diagram is shown in Figure B6.

![Figure B6. Crystalliser, Case Study II.](image-url)

Seeded moments,

$$\frac{dN_i}{dt} = 0$$

$$\frac{dN}{dt} = G N_i$$

$$\frac{dA_i}{dt} = 2G L_i$$

$$\frac{dV_i}{dt} = 3G A_i$$

Growth kinetics,

$$G = k_g \Delta c_{2,2,liq}^g$$

Solute concentration balance,

$$\Delta c_{2,2,liq} = c_{2,2,liq} - c_{2,2,liq}(T)^*$$

$$c_{2,2,liq} = c_{v,2,2,liq} - Z_{pc}$$

$$Z_{pc} = V \rho_v f_v$$

Component mass balance,
\[ z_{2,2,\text{crys}} = Z_{pC} z_{2,7,\text{liq}} \]

\[ z_{2,2,\text{liq}} = z_{1,2} - z_{2,2,\text{crys}} \]

\[ z_{2,i,\text{crys}} = z_{2,7,\text{liq}} \left( c_{o,2,i,\text{liq}} - c_{2,i,\text{liq}} \right) \quad \text{for } i = 1, 3, 4, 5 \]

\[ z_{2,i,\text{liq}} = z_{1,i} - z_{2,i,\text{crys}} \quad \text{for } i = 1, 3, 4, 5 \]

\[ z_{2,i,\text{crys}} = z_{1,i} \quad \text{for } i = 6, 7, 8 \]

\[ z_{2,i,\text{crys}} = 0 \quad \text{for } i = 6, 7, 8 \]

Impurity growth,
\[ \frac{dc_{i,\text{liq}}}{dt} = \zeta_i c_{2,i,\text{liq}} \quad \text{for } i = 1, 3, 4, 5 \]

Initial conditions,
\[ N_{o,s} = \frac{Z_{o,s}}{f_c \rho_c L_{o,s} z_{1,7,\text{liq}}} \]
\[ L_{o,s} \]
\[ A_{o,s} = \frac{Z_{o,s} F}{\rho_c L_{o,s} z_{1,7,\text{liq}}} \]
\[ V_{o,s} = \frac{Z_{o,s}}{\rho_c z_{1,7,\text{liq}}} \]
\[ N_{o,n} = L_{o,n} = A_{o,n} = V_{o,n} = 0 \]
\[ c_{o,2,i,\text{liq}} = \frac{z_{1,i}}{z_{1,7}} \quad \text{for } i = 1, 3, 4, 5 \]

Operating policy,
\[ \frac{dT}{dt} = CR \left|_{t',t'} \right. = 0 \left|_{t',t'} \right. \]
\[ t_f = t' + HT \quad \text{(Model B5)} \]

where
\[ A \quad = \quad \text{total crystal surface area, } m^2 \text{ kg}^{-1} \text{ solvent} \]
\[ c \quad = \quad \text{solute concentration, } kg \text{ kg}^{-1} \text{ solvent} \]
\[ c^* \quad = \quad \text{equilibrium solubility solute concentration, } kg \text{ kg}^{-1} \text{ solvent} \]
### Integrated design under uncertainty for pharmaceutical processes

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>CR</td>
<td>cooling rate, °C min⁻¹</td>
</tr>
<tr>
<td>f_v</td>
<td>volumetric shape factor</td>
</tr>
<tr>
<td>F</td>
<td>overall shape factor</td>
</tr>
<tr>
<td>g</td>
<td>kinetic order of growth</td>
</tr>
<tr>
<td>G</td>
<td>growth rate, m min⁻¹</td>
</tr>
<tr>
<td>HT</td>
<td>holding time, min</td>
</tr>
<tr>
<td>k_g</td>
<td>kinetic growth rate constant, m min⁻¹ (kg kg⁻¹ solvent)¹/z</td>
</tr>
<tr>
<td>L</td>
<td>total crystal length, m kg⁻¹ solvent</td>
</tr>
<tr>
<td>N</td>
<td>number of crystals, kg⁻¹ solvent</td>
</tr>
<tr>
<td>t</td>
<td>time, min</td>
</tr>
<tr>
<td>t'</td>
<td>time at which holding temperature is achieved, min</td>
</tr>
<tr>
<td>t_f</td>
<td>termination time, min</td>
</tr>
<tr>
<td>T</td>
<td>temperature, °C</td>
</tr>
<tr>
<td>V</td>
<td>total crystal volume, m³ kg⁻¹ solvent</td>
</tr>
<tr>
<td>Z</td>
<td>stream mass, kg kg⁻¹ solvent</td>
</tr>
<tr>
<td>ρ</td>
<td>crystal density, kg m⁻³</td>
</tr>
<tr>
<td>ζ</td>
<td>first order rate constant for loss of impurity concentration, min⁻¹</td>
</tr>
</tbody>
</table>

**subscripts**

- **crys** = solid phase
- i = component species \{actA, actB, actC, actD, actE, solF, solL, aq\}
- **liq** = liquid phase
- o = initial value
- **pc** = product crystal
- s = seeds

Possible degrees of freedom for the batch crystallisation model could include: CR, HT, \(L_{o,s}\), \(Z_{o,s}\). The assumptions in the crystalliser model include:

- seeded operation, with \(Z_{o,s}\) kg of crystals of size \(L_{o,s}\), and no nucleation,
- a power law function is suitable to describe growth kinetics for the crystallisation of the product drug.
since a lower holding temperature is believed to lead to increased crystal impurity content, but no data or mechanistic knowledge is available, the holding temperature is not considered a degree of freedom.

due to the lack of understanding regarding the drug impurity effects, their crystalline presence is explained using first order solute loss functions of liquid phase drug impurity concentration (as opposed to an alternative assumption of linear impurity concentration loss which is not sensitive to changes in the initial value), and independent to temperature (the holding temperature remains constant),

since data is not available concerning the presence of crystalline impurities other than the drug components (i.e. reG, soLF, soLL) no characterisation for these effects is portrayed in the model, although the presence of soLF in the pre-crystallisation stream is considered an important criterion in this case study,

estimation of growth rate constant, \( k_p \), is based on solubility data for an alternative high relative molecular mass organic compound in pure soLL solvent (based on three temperature data points and fitted with a 2\(^{nd}\) order polynomial, RMM = 354, Crossfire Beilstein Database, Beilstein Chemiedaten und Software GmbH) since solubility data for the drug is not available in this study,

the limitation in process understanding precludes crystal size distribution (CSD) prediction,

an assumed fixed value for growth rate order (\( g = 1.2 \)) due to lack of profile data points,

perfect cooling control at a constant rate and associated heat transfer effects are not limiting,

size independent growth,

perfect spheres assumed for overall shape factor (\( F \)) and volumetric shape factor (\( f_v \)),

the initial mixture is at the composition boiling point predicted by the ideal VLE batch distillation model with the physical properties of dioctylphthalate (ChemCAD V) used to represent the unknown drug properties.

**B.6 Filtration model**

The first generation filtration operation is described with a simple mass balance. The lack of available data precludes the use of conventional filtration/centrifugation models found in chemical engineering literature. The first generation model for the Stage 13 washing operation is given in Model B6 and the process diagram is shown in Figure B7.
Component mass balance,

\[ z_{2,i,\text{crys}} = z_{1,i,\text{crys}} \quad \text{for } i = 1..8 \]

\[
LOD = \frac{\sum_{i=1}^{8} z_{2,i,\text{liq}}}{\sum_{i=1}^{8} z_{1,i,\text{liq}} + \sum_{i=1}^{8} z_{2,i,\text{crys}}} \times 100\%
\]

\[
\frac{\sum_{i=1}^{8} z_{1,i,\text{liq}}}{\sum_{i=1}^{8} z_{2,i,\text{liq}}} = \frac{\sum_{i=1}^{8} z_{2,i,\text{liq}}}{\sum_{i=1}^{8} z_{3,i,\text{liq}}} \quad \text{for } i = 1..8
\]

Batch filtration time,

\[ t_f = FR \sum_{i=1}^{8} z_{i,\text{crys}} \quad \text{(Model B6)} \]

where

\begin{align*}
FR &= \text{filtration rate, min kg}^{-1} \text{ solids} \\
i &= \text{component species \{actA, actB, actC, actD, actE, solF, solL, aq\}} \\
LOD &= \text{level of dampness in solids, } \% \\
t_f &= \text{operation time, min}
\end{align*}

No degrees of freedom are associated with this model, since not enough information is available to justify a model which relates an operating policy to performance. The assumptions for the filtration model include:

- no change in the slurry liquid composition such that the composition of drugs in entrained in the damp solids, \( z_{2,i,\text{liq}} \) is the same as in the filtrate, \( z_{3,i,\text{liq}} \).
- no change in dry solids composition or mass,
- a pre-determined desired value of the level of dampness (LOD) is achieved,
- a fixed processing rate per mass of solids, independent of scale, LOD and CSD.
B.7 Washing model

The first generation washing model consists of a mass balance with displacement of residual moisture with wash solvent. The first generation model for the Stage 14 washing operation is given in Model B7 and the process diagram is shown in Figure B8.

![Process diagram](image)

**Figure B8. Washing unit, Case Study II.**

Non-wash solvent component mass balance,

\[
LOD = \frac{\sum_{i=1}^{8} z_{2,i,\text{liq}}}{\sum_{i=1}^{8} z_{2,i,\text{liq}} + \sum_{i=1}^{8} z_{2,i,\text{crys}}} \times 100%
\]

\[z_{2,i,\text{crys}} = z_{1,i,\text{crys}} \quad \text{for } i = 1..8\]

\[z_{2,i,\text{liq}} = \left(1 - \eta_{\text{wash}}\right) z_{1,i,\text{liq}} \quad \text{for } i = 1..8, \quad i \neq 7\]

\[z_{3,i} = \eta_{\text{wash}} z_{1,i,\text{liq}} \quad \text{for } i = 1..6,8\]

Wash solvent component (solL, i = 7) mass balance,

\[z_{2,7,\text{liq}} = \left(1 - \eta_{\text{wash}}\right) z_{1,7,\text{liq}} + \left(\eta_{\text{wash}} \sum_{i=1,i \neq 7}^{8} z_{1,i,\text{liq}}\right)\]

\[z_{3,7} = \eta_{\text{wash}} z_{1,7,\text{liq}} + \left(F - \eta_{\text{wash}} \sum_{i=1,i \neq 7}^{8} z_{1,i,\text{liq}}\right)\]

(Model B7)

where

- \(F\) = mass of wash solvent feed, kg
- \(LOD\) = level of dampness in solids, %
- \(t_f\) = operation time, min
- \(\eta_{\text{wash}}\) = wash efficiency, representing the split fraction of initial residual moisture which is replaced with pure wash solvent
and subscript i represents the component species (actA, actB, actC, actD, actE, solF, solL, aq). No degrees of freedom are associated with this model, since not enough information is available to justify a model which relates an operating policy to performance. Assumptions for the washing model include:

- a final LOD equal to the initial LOD is achieved,
- a fractional displacement of the initial residual moisture \((z_{1,i,\text{aq}})\) with pure wash solvent \((F)\) represented with an assumed wash efficiency \((\eta_{\text{wash}})\),
- negligible dissolution of crystalline drug components in the pure solL wash solvent (operated at ambient temperature),
- the composition of the displaced residual moisture is equal to the composition of the initial moisture,
- no change in dry solids composition or mass,
- a fixed wash time, independent of scale.

B.8 Dryer model

Mass and heat transfer effects are likely to be complex and a lack of data and general understanding of the drying process permits only a simple mass balance model based on an efficiency measure in drying rate. The first generation model for the Stage 15 drying operation is given in Model B8 and the process diagram is shown in Figure B9.

![Drying Unit Diagram](image)

Figure B9. Drying unit, Case Study II.

Drug component mass balance,

\[
z_{2,i,\text{crys}} = z_{1,i,\text{crys}} \quad \text{for } i = 1..8
\]

\[
z_{2,i,\text{aq}} = z_{1,i,\text{aq}} \quad \text{for } i = 1..5
\]

Solvent component mass balance

\[
\text{LOD} = \frac{\sum_{i=1}^{8} z_{2,i,\text{aq}}}{\sum_{i=1}^{8} z_{2,i,\text{aq}} + \sum_{i=1}^{8} z_{2,i,\text{crys}}} \times 100\%
\]
Integrated design under uncertainty for pharmaceutical processes

\[
\sum_{i=6}^{8} z_{2,i,liq} + \sum_{i=1}^{5} z_{2,i,liq} = \left( \sum_{i=1}^{8} z_{2,i,cr} \right) \left( \frac{\sum_{i=1}^{8} z_{2,i,cr}}{100 - LOD} \right)
\]

\[
\frac{z_{2,i,liq}}{z_{1,i,liq}} = \frac{\sum_{i=6}^{8} z_{2,i,liq}}{\sum_{i=6}^{8} z_{1,i,liq}} \quad \text{for } i = 6..8
\]

\[
z_{3,i} = z_{1,i,liq} - z_{2,i,liq} \quad \text{for } i = 6..8
\]

\[
P_{drycr} = 100\% \left( \frac{z_{2,i,cr} + z_{2,i,liq}}{\sum_{i=1}^{5} z_{2,i,cr} + \sum_{i=1}^{5} z_{2,i,liq}} \right) \quad \text{for } i = 1..5
\]

\[
t_f = DR \sum_{i=1}^{5} z_{1,i,cr} \quad \text{(Model B8)}
\]

where

**DR** = drying rate, min kg\(^{-1}\) solids

**LOD** = level of dampness in solids, %

**P_{drycr}** = purity of final crystals, dry weight percent % (excluding solvent moisture)

and subscript **i** represents the component species \{actA, actB, actC, actD, actE, solF, solL, aq\}. No degrees of freedom are associated with this model, since not enough information is available to justify a model which relates an operating policy to performance. Assumptions for the drying model include:

- no change in dry solids composition or mass,
- any drug components in the initial residual moisture \((z_{1,i,liq})\) is retained and does not leave in the evaporate \((z_{3,i})\),
- a pre-determined desired value of the LOD is achieved,
- a fixed processing rate per mass of solids (DR), independent of scale, LOD, CSD and ambient temperature.
### B.9 Uncertainty in first generation models

Table B2. Uncertainty characterisation in the parameters of the first generation models, Case Study II.

<table>
<thead>
<tr>
<th>Stage</th>
<th>Parameter (stochastic model parameter index)</th>
<th>Normal distribution, ( \mathcal{N}(\mu, \sigma) )</th>
<th>Uniform distribution, ( \mathcal{U}(\text{min}, \text{max}) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>( k_1 (1), k_2 (2) )</td>
<td>( \begin{cases} \mathcal{N}(\mu, \sigma) \ \mathcal{U}(\text{min}, \text{max}) \end{cases} )</td>
<td>( \begin{cases} 1.69 \times 10^{-2} \ 7.14 \times 10^{-5} \end{cases} )</td>
</tr>
<tr>
<td></td>
<td>( k_3 (3) )</td>
<td>( \mathcal{N}(1.67 \times 10^{-5}, 1.51 \times 10^{-6}) )</td>
<td>( \mathcal{N}(60, 10% \text{ nominal}) )</td>
</tr>
<tr>
<td>2</td>
<td>( u_o (5) )</td>
<td>( \mathcal{U}(3, 6) )</td>
<td>( \mathcal{U}(0, 0.01) )</td>
</tr>
<tr>
<td>3</td>
<td>( u_1 (6) )</td>
<td>( \mathcal{U}(0, 0.01) )</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>( u_2 (7) )</td>
<td>( \mathcal{U}(0, 0.01) )</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>( u_3 (8) )</td>
<td>( \mathcal{U}(0, 0.01) )</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>( u_4 (9) )</td>
<td>( \mathcal{U}(0, 0.01) )</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>( u_5 (10) )</td>
<td>( \mathcal{U}(0, 0.01) )</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>( u_6 (11) )</td>
<td>( \mathcal{U}(0, 0.01) )</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>( v_{p, \text{solF,A}} (13) )</td>
<td>( \mathcal{N}(101.6, 0.25% \text{ nominal}) )</td>
<td></td>
</tr>
<tr>
<td></td>
<td>( v_{p, \text{solF,C}} (14) )</td>
<td>( \mathcal{N}(-12.2, 0.25% \text{ nominal}) )</td>
<td></td>
</tr>
<tr>
<td></td>
<td>( \text{grad}_{\text{solF}} (15) )</td>
<td>( \mathcal{N}(0.49, 10% \text{ nominal}) )</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>( v_{p, \text{solF,A}} (13) )</td>
<td>( \mathcal{N}(101.6, 0.25% \text{ nominal}) )</td>
<td></td>
</tr>
<tr>
<td></td>
<td>( v_{p, \text{solF,C}} (14) )</td>
<td>( \mathcal{N}(-12.2, 0.25% \text{ nominal}) )</td>
<td></td>
</tr>
<tr>
<td></td>
<td>( \text{grad}_{\text{solF}} (15) )</td>
<td>( \mathcal{N}(0.49, 10% \text{ nominal}) )</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>( v_{p, \text{solF,A}} (13) )</td>
<td>( \mathcal{N}(101.6, 0.25% \text{ nominal}) )</td>
<td></td>
</tr>
<tr>
<td></td>
<td>( v_{p, \text{solF,C}} (14) )</td>
<td>( \mathcal{N}(-12.2, 0.25% \text{ nominal}) )</td>
<td></td>
</tr>
<tr>
<td></td>
<td>( \text{grad}_{\text{solF}} (15) )</td>
<td>( \mathcal{N}(0.49, 10% \text{ nominal}) )</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>( k_v (16) )</td>
<td>( \mathcal{N}(6.61 \times 10^{-3}, 2.31 \times 10^{-3}) )</td>
<td></td>
</tr>
<tr>
<td></td>
<td>( C_{210C}^* (17) )</td>
<td>( \mathcal{N}(0.0056, 5% \text{ nominal}) )</td>
<td></td>
</tr>
<tr>
<td></td>
<td>( C_{30C}^* (18) )</td>
<td>( \mathcal{N}(0.0110, 5% \text{ nominal}) )</td>
<td></td>
</tr>
<tr>
<td></td>
<td>( C_{78C}^* (19) )</td>
<td>( \mathcal{N}(0.2530, 5% \text{ nominal}) )</td>
<td></td>
</tr>
<tr>
<td></td>
<td>( \zeta_{\text{sacA}} (20) )</td>
<td>( \mathcal{N}(0.0025, \sigma_{\text{sacA}}) )</td>
<td></td>
</tr>
<tr>
<td></td>
<td>( \zeta_{\text{sacC}} (21) )</td>
<td>( \mathcal{N}(0.0058, \sigma_{\text{sacC}}) )</td>
<td></td>
</tr>
<tr>
<td></td>
<td>( \zeta_{\text{sacD}} (22) )</td>
<td>( \mathcal{N}(0.0021, \sigma_{\text{sacD}}) )</td>
<td></td>
</tr>
<tr>
<td></td>
<td>( \zeta_{\text{sacE}} (23) )</td>
<td>( \mathcal{N}(0.0034, \sigma_{\text{sacE}}) )</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>( \text{FR} (24) )</td>
<td>( \mathcal{N}(0.5, 10% \text{ nominal}) )</td>
<td></td>
</tr>
<tr>
<td></td>
<td>( \text{LOD} (25) )</td>
<td>( \mathcal{N}(25, 10% \text{ nominal}) )</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>( \eta_{\text{wash}} (26) )</td>
<td>( \mathcal{U}(0, 1) )</td>
<td></td>
</tr>
<tr>
<td></td>
<td>( \text{LOD} (27) )</td>
<td>( \mathcal{N}(25, 10% \text{ nominal}) )</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>( \text{DR} (28) )</td>
<td>( \mathcal{N}(2.0, 10% \text{ nominal}) )</td>
<td></td>
</tr>
<tr>
<td></td>
<td>( \text{LOD} (29) )</td>
<td>( \mathcal{N}(6, 10% \text{ nominal}) )</td>
<td></td>
</tr>
</tbody>
</table>
B.10 Uncertainty from violation of predetermined operating ranges

To accommodate the lack of understanding and mechanistic knowledge regarding possible consequences due to deviations from desired operating conditions obtained from design of experiment analyses, an extra degree of uncertainty is incorporated. This introduces a form of the stochastic system model where the uncertainty is dynamic, dependent on future decisions or knowledge. In these instances it is assumed that a deviation of a particular measured criterion from the desired value or range results in an increase in the prior uncertainty for a parameter characterising the possible consequence,

\[
\sigma_{\theta,i} = \begin{cases} 
\sigma_{\theta,0} 
& \text{if } \left| Q_{j,i} - Q_{j}^{UB} + Q_{j}^{LB} \right| \leq \frac{Q_{j}^{UB} - Q_{j}^{LB}}{2} \\
\sigma_{\theta,0} \times \left( 1 + \sum_{j} g_{j} \left( \left| Q_{j,i} - Q_{j}^{UB} + Q_{j}^{LB} \right| - \frac{Q_{j}^{UB} - Q_{j}^{LB}}{2} \right) \right) 
& \text{if } \left| Q_{j,i} - Q_{j}^{UB} + Q_{j}^{LB} \right| \geq \frac{Q_{j}^{UB} - Q_{j}^{LB}}{2} 
\end{cases}
\]

\[
\text{for } j = 1 \ldots J \\
\text{and } i = 1 \ldots I
\]

where \(\sigma_{\theta,i}\) is the standard deviation used to generate the \(i\)th parameter scenario \(\theta_i\), \(\sigma_{\theta,0}\) is the standard deviation of uncertain parameter \(\theta\) used if no range violation occurs, \(Q_j\) is the value of the \(j\)th criterion and \(g\) is the factor by which \(\sigma_{\theta,i}\) increases linearly from \(\sigma_{\theta,0}\) with deviation of \(Q_j\) outside the desired criterion range \(Q_j^{UB}\) and \(Q_j^{LB}\). In a conservative assumption, the standard deviation for a given uncertain parameter increases additively for deviations from multiple criteria ranges (which may be associated with the uncertain parameter in question), as shown in Equation B3. In the event of a criterion deviation outside the desired range, the uncertain parameter scenario is re-sampled from the newly characterised distribution.

For the violation of the total initial and pre-crystallisation desired solv. solvent volume to product mass ratio operating ranges (14-15 and 7-8, respectively) it is assumed that the uncertainty (standard deviation) in the Stage 12 crystallisation parameters characterising the crystal impurity content, \(\zeta\), increase linearly at a rate of unity with extent of the (additive) deviations from the limits of the initial and final solvent ratio ranges.
B.11 Uncertainty Analysis results for the first generation of models, Case Study II.

Figure B10 shows the quantitative effect of employing the expression for additional uncertainty (Equation B3) due to the violation of the desired initial and final range in the predicted solvent to product ratio (latter shown in Figure B10 (b)) on the crystallisation key impurity 'solute loss' parameter ($\gamma_{actC}$). Figure

(a) Key impurity 'solute loss' parameter.
Key: $\bullet =$ with additional uncertainty,
$\circ =$ without additional uncertainty.

(b) Pre-crystallisation solvent volume to product mass ratio.

(c) Key impurity content.
Key: $\bullet =$ with additional uncertainty,
$\circ =$ without additional uncertainty.

Figure B10. Effect of additional uncertainty in crystallisation 'solute loss' parameter for the key impurity due to violation of desired solvent volume to product mass operating range, Case Study II.
B10 (a), and on the endpoint impurity content (wt ecc). Figure B10 (c).

Figure B11 shows the arrival of sampling convergence in the evolution of the mean and variance parameters for the predicted total yield. The effect of inducing rank correlation using Iman and Conover (1982) technique, as expressed in Figure B12, is discussed. The parameter regression covariance matrix for the knowledge level 0 Stage I reaction rate constants, \(k_1\) and \(k_2\), was determined assuming linearisation around the optimal estimates from which the associated correlation matrix is obtained (as stated in Section 5.4),

\[
\hat{\Sigma} = \begin{bmatrix}
2.27 \times 10^{-6} & -1.80 \times 10^{-8} \\
-1.80 \times 10^{-8} & 1.48 \times 10^{-10}
\end{bmatrix}, \quad \hat{C} = \begin{bmatrix}
1.0 & -0.9820 \\
-0.9820 & 1.0
\end{bmatrix}
\]

(B4)

\(\hat{C}\) is set as the desired correlation matrix for the sample generated \(k_1\) and \(k_2\) vectors. A matrix \(K\), obtained from an independently generated Hammersley sequence sample, has a correlation matrix \(E\),

\[
E = \begin{bmatrix}
1.0 & -0.0024 \\
-0.0024 & 1.0
\end{bmatrix}
\]

(B5)

Since \(E\) is close to the identity matrix (the correction for \(K^*\) is not required) and the correlation and rank correlation (\(E^*\) and \(E_{rk}^*\)) matrices of \(K^*\) are close to each other,

\[
E^* = \begin{bmatrix}
1.0 & -0.9821 \\
-0.9821 & 1.0
\end{bmatrix}, \quad E_{rk}^* = \begin{bmatrix}
1.0 & -0.9829 \\
-0.9829 & 1.0
\end{bmatrix}
\]

(B6)

then the desired rank correlation can be induced into \(k_1\) and \(k_2\) sample vectors by rearranging the elements according to the rank order of \(K^*\). The resulting sample correlation matrix for the 431 sample of \(k_1\) and \(k_2\) is,

\[
\hat{C}_s = \begin{bmatrix}
1.0 & -0.9866 \\
-0.9866 & 1.0
\end{bmatrix}
\]

(B6)

which is close to the desired correlation matrix, \(\hat{C}\). Figure B12 (a) shows the contrast of the rearranged observations of the independent Hammersley sample matrix for a desired correlation of -0.9820 (dots) to the unit hyper-cube sample matrix (circles), before inversion over the standard normal cumulative distribution. \(X^*\) is expressed for the normally distributed \(k_1\) and \(k_2\) parameters with the same desired correlation in Figure B12 (b). The circles represent the distributed observations before the induced correlation and the dots represent the observations after the rearrangement. The 95% confidence regions assume linearisation of the model about the optimal parameter estimates (see Section 5.4). The inducement of the correlation in the sample appears to be reasonable compared to the 95% confidence region for the correlated parameters (the solid ellipse). The scatter plots in Figure B13 shows that the
Linear Sensitivity Analysis measures (CC and SRC) based on unranked data are adequate measures of the key uncertain parameter contributions for the performance criteria shown.

Figure B11. Evolution of distribution parameter estimates with sample observations for the total yield of the first generation of models (knowledge level 0), Case Study II.

(a) The first two dimensions of the unit hypercube sample (before inversion over the standard normal distribution) Key: o = initial Hammersley sample, * = rearranged sample.

(b) Normally distributed k₁ and k₂ reaction rate constants. Key: o = initial Hammersley sample, * = rearranged sample, --- = 95% confidence region for uncorrelated sample, ----- = 95% region for rank correlated sample.

Figure B12. Scatter plots showing the effect of induced rank correlation ($\rho = -0.98$) in the Stage 1 rate constant parameters for the first generation of models (knowledge level 0), Case Study II.
Integrated design under uncertainty for pharmaceutical processes

Figure B13. Scatter plots of the key uncertain parameters with endpoint criteria for the first generation of models (knowledge level 0), Case Study II.

(a) Stage 12 crystal growth constant and total yield (SRC = 0.82).

(b) Stage 12 crystallisation key impurity 'solute loss' parameter and key impurity content (SRC = 0.78).
Appendix C

C. CASE STUDY II: PROCESS MODEL REVISIONS INCORPORATING LEVELS OF KNOWLEDGE

Appendix C contains revised process models and uncertainty characterisations for Case Study II. These are associated with the employment of additional experimental data or pilot plant measurements into systematic model development procedures.

C.1 Knowledge level 1. First generation Stage 1 reactor model fitted to new data

Profile data from a 1000 US gallon reaction process operated at 90 rpm (the maximum agitation possible in the equipment), is available. Fitting the parameters of the first generation Stage 1 model to this new data introduces an increase in the amount of uncertainty in the confidence region for the $k_1$ and $k_2$ parameters (Table C1) compared to the uncertainty in the parameter fit to the bench scale data (Table B2, Appendix B).

<table>
<thead>
<tr>
<th>Stage</th>
<th>Parameter</th>
<th>Normal distribution, $N(\mu, \sigma)$</th>
<th>Increase in $\sigma^2$ over fit to bench scale data (Table B2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$k_1$</td>
<td>$N(1.46 \times 10^{-2}, 1.68 \times 10^{-8})$</td>
<td>7.4</td>
</tr>
<tr>
<td></td>
<td>$k_2$</td>
<td>$N(3.3 \times 10^{-3}, 7.10 \times 10^{-9})$</td>
<td>2.2</td>
</tr>
<tr>
<td></td>
<td>$k_3$</td>
<td>$N(9.92 \times 10^{-4}, 2.56 \times 10^{-6})$</td>
<td>1.7</td>
</tr>
<tr>
<td></td>
<td>$t'$</td>
<td>$N(60, 10%$ nominal)</td>
<td>-</td>
</tr>
</tbody>
</table>

C.2 Knowledge level 2. Second generation Stage 1 reactor model

A second generation Stage 1 model is proposed. From the new profile data at the larger scale (1000 US gallon, 90 rpm), an initial rate limiting period is observed in the product reaction. The possibility of non-instantaneous addition of reH and/or the dissolution of solid reG is postulated in Table 7.2 (Section 7.3). Since the role of the reagent species on the observed drug reactions is unclear and it not possible to differentiate between these two phenomena without further investigation or data, the rate of reaction is assumed to proceed at a constant rate, $k_1''$, during the initial rate limited period. This leads to a second generation Stage 1 model, comprising:
the first generation Stage 1 model (Model B, Appendix B), with the exception of the following drug mole balances,

\[
\frac{dm_{1,\text{org}}}{dt} = -k_1^n \bigg|_{t^n} = -k_1 m_{1,\text{org}} \bigg|_{t^n}\]

\[
\frac{dm_{2,\text{org}}}{dt} = k_2^n \bigg|_{t^n} = k_1 m_{1,\text{org}} - k_2 m_{2,\text{org}} \bigg|_{t^n}\]

\[
\frac{dm_{3,\text{org}}}{dt} = 0 \bigg|_{t^n} = k_2 m_{1,\text{org}} \bigg|_{t^n}\]

\[k_1^n = \frac{m_{0,\text{I}} X''}{t''}\] (Model C1)

where the notation is defined in Model B (Appendix B) and additionally

\[k_1''\] = zero order rate constant during initial rate limited period, mol min\(^{-1}\)

\[t''\] = time at which initial rate limited period ends, min

\[X''\] = conversion at which initial rate limited period ends

The same assumptions apply to this model as for the first generation Stage 1 model (Model B, Appendix B), with the exception of:

- a zero order (linear profile) initial reaction rate limiting period, assumed to be independent of scale (since there are not enough data points to justify any other assumption), in which \(t''\) and \(X''\) remain constant given the same operational reH addition time (1 hour),

- the reaction for the formation of the key impurity starts after the end of the initial rate of product formation limiting period.

The fitted parameter values for the second generation model are given in Table C2 and the uncertainty characterisations are given in Table C3. The second and first generation model profile predictions for the new 1000 US gallon reactor data, the solid and dashed curves in Figure C1, respectively, compare the difference due to the structural revision. There is a clear difference in the actA and actB drug component profiles when a rate limiting phenomena is introduced. Very similar values for the \(k_1\) rate constant are obtained (0.0168 min\(^{-1}\) compared to the first generation model value of 0.0169 min\(^{-1}\) based on bench scale data) leads to the conclusion that at 90 rpm agitation, close to fully disperse conditions are achieved in the 1000 US gallon vessel, indicating intrinsic kinetic control. Despite keeping the same first order equation structure for the rate of secondary impurity formation, the change in the fitted value of \(k_3\) (9.92 \(\times\) 10\(^{-4}\) compared to 1.67 \(\times\) 10\(^{-3}\) min\(^{-1}\)) with the 1000 US gallon data has a large effect on the prediction for the secondary impurity (actE), Figure C1 (b).
Table C2. Parameters for the second generation Stage 1 model, Case Study II.

<table>
<thead>
<tr>
<th>Fitted model parameter values</th>
<th>Imposed operating conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td>$k_1 = 0.0168$ min$^{-1}$</td>
<td>$t_f = 230$ min</td>
</tr>
<tr>
<td>$k_2 = 7.98 \times 10^{-3}$ min$^{-1}$</td>
<td>$\text{wt}%<em>{F</em>{drug}} = [79.1, 0, 0, 20.1, 0]$</td>
</tr>
<tr>
<td>$k_3 = 9.92 \times 10^{-4}$ min$^{-1}$</td>
<td>$\Gamma_{\text{drug}} = 195.3$ kg</td>
</tr>
<tr>
<td>$t' = 60$ min</td>
<td>$\text{molratio}_{\text{eff}} = 10.4$</td>
</tr>
<tr>
<td>$X''' = 0.37$</td>
<td>$\text{wt}%<em>{F</em>{\text{eff}}} = 30%$</td>
</tr>
<tr>
<td>$t'' = 75$ min</td>
<td></td>
</tr>
</tbody>
</table>

Table C3. Uncertainty characterisation in the second generation Stage 1 model parameters, Case Study II.

<table>
<thead>
<tr>
<th>Stage</th>
<th>Parameter</th>
<th>Normal distribution, $N(\mu, \sigma)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$k_1, k_2$</td>
<td>$\mathcal{N}\left(1.68 \times 10^{-2}, \begin{bmatrix} 1.46 \times 10^{-6} &amp; -1.61 \times 10^{-6} \ -1.61 \times 10^{-6} &amp; 1.77 \times 10^{-10} \end{bmatrix}\right)$</td>
</tr>
<tr>
<td></td>
<td>$k_3$</td>
<td>$\mathcal{N}(9.92 \times 10^{-4}, 2.56 \times 10^{-4})$</td>
</tr>
<tr>
<td></td>
<td>$t'$</td>
<td>$\mathcal{N}(60, 10%$ nominal$)$</td>
</tr>
<tr>
<td></td>
<td>$X''$</td>
<td>$\mathcal{N}(0.37, 10%$ nominal$)$</td>
</tr>
<tr>
<td></td>
<td>$t''$</td>
<td>$\mathcal{N}(75, 10%$ nominal$)$</td>
</tr>
</tbody>
</table>

(a) Key: Pilot scale data points  
$\circ = \text{actA}, \ast = \text{actB}, \bigcirc = \text{actD}.$

(b) Key: Pilot scale data points  
$\bigcirc = \text{actC}, \bigast = \text{actE}.$

Figure C1. Pilot scale (1000 US gallon, 90 rpm) drug profile, Case Study II.
Key: $-$ = Second generation model prediction, $---$ = first generation model prediction.
C.3 Knowledge level 3. Third generation Stage 1 reactor model

Profile data becomes available at 60 and 75 rpm, at the 1000 US gallon scale. With this data, in addition to the 90 rpm data, it is possible to develop a third generation Stage 1 model, to account for mixing effects. The intention is to use this model to predict the agitation required to maintain a desired performance of the reaction. The development of the mixing case is stated from private communication with a pharmaceutical company. The parameters and associated uncertainties for the assumed mixing model incorporated in the third generation Stage 1 model, are identified in this section.

The reaction data obtained to develop the second generation Stage 1 model indicated that close to fully disperse conditions (intrinsic kinetics observed) were obtained at 90 rpm in the 1000 US gallon reactor. It cannot be assumed that the 78 rpm agitation conditions in the PPR plant vessel achieves fully disperse conditions. If the relative importance of the Stage 1 model to the overall process sequence is high, then any effect of mixing on the reaction kinetics is likely to be a key factor. The third generation model comprises:

the second generation Stage 1 model (Model C1),

\[
\frac{k_j}{k_{j,\text{disp}}} = \left[ \frac{N}{N_{\text{disp}}} \right]^{\gamma_1}
\]

\[
\frac{k''}{k''_{\text{disp}}} = \left[ \frac{N}{N_{\text{disp}}} \right]^{\gamma_2}
\]

(Model C2)

where the notation is defined in Model C1 (Appendix C) and Model B1 (Appendix B) and additionally

\[N\] = agitation speed, rpm

\[\gamma_1\] = a power which characterises the mixing regime on the intrinsic rates of reaction

\[\gamma_2\] = a power which characterises the effect of mixing on the parameter \(k''\) characterising the initial rate limiting period

subscripts

\[\text{disp}\] = fully disperse

\[j\] = reaction rate constant index \(\{1, 2, 3\}\)

Assumptions for the third generation Stage 1 model include:

- the predicted mixing regime is valid at both the 1000 US gallon and PPR plant scales, and all the drug reactions are similarly affected,

- an agitation of ~90 rpm gives close to fully disperse conditions in the PPR plant vessel, i.e. a similar power input per unit volume is achieved to that in the 1000 US gallon scale 90rpm run,
the power, $\gamma_2$, adequately describes the effect of agitation on $k_{1,\text{disp}}''$ with respect to the assumed disperse value, $k_{1,\text{disp}}''$ (based on $m_{0,1}$, $X''$ and $t''$ at 90 rpm in Model C2).

- $t''$ is independent of agitation (i.e. $t_{\text{disp}}'' = t''$, based on comparison of concentration profile data).

The parameter values for the third generation Stage 1 model are given in Table C4. The uncertainty these parameters are characterised in Table C5.

Table C4. Parameters for the third generation Stage 1 model, Case Study II.

<table>
<thead>
<tr>
<th>Fitted model parameter values</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t'' = 75$ min</td>
</tr>
<tr>
<td>$X_{\text{disp}}'' = 0.37$</td>
</tr>
<tr>
<td>$k_{1,\text{disp}} = 0.0168$</td>
</tr>
<tr>
<td>$k_{2,\text{disp}} = 7.98 \times 10^{-5}$</td>
</tr>
<tr>
<td>$k_{3,\text{disp}} = 9.92 \times 10^{-4}$</td>
</tr>
<tr>
<td>$t' = 60$ min</td>
</tr>
<tr>
<td>$\gamma_1 = 1.2$</td>
</tr>
<tr>
<td>$\gamma_2 = 2.0$</td>
</tr>
</tbody>
</table>

Table C5. Uncertainty in the parameters of the third generation Stage 1 model, Case Study II.

<table>
<thead>
<tr>
<th>Stage</th>
<th>Parameter</th>
<th>Distribution, $N(\mu, \sigma), U(\text{min, max})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$k_1, k_2$</td>
<td>$\mathcal{N}\left(k_1' = 168 \times 10^{-2}, k_2' = 7.98 \times 10^{-5}\right)$</td>
</tr>
<tr>
<td></td>
<td>$k_3$</td>
<td>$\mathcal{N}(9.92 \times 10^4, 2.56 \times 10^4)$</td>
</tr>
<tr>
<td></td>
<td>$t'$</td>
<td>$\mathcal{N}(60, 10% \text{ nominal})$</td>
</tr>
<tr>
<td></td>
<td>$X_{\text{disp}}''$</td>
<td>$\mathcal{N}(0.37, 10% \text{ nominal})$</td>
</tr>
<tr>
<td></td>
<td>$t''$</td>
<td>$\mathcal{N}(75, 10% \text{ nominal})$</td>
</tr>
<tr>
<td></td>
<td>$\gamma_1$</td>
<td>$U(1.2-0.37, 1.2+0.37)$</td>
</tr>
<tr>
<td></td>
<td>$\gamma_2$</td>
<td>$U(2.0-0.92, 2.0+0.92)$</td>
</tr>
</tbody>
</table>

Development of the mixing case using a mixing correlation (Godfrey et al. 1989), based on the actA profiles under different agitation rates, indicated that a dispersion limited mixing regime (no mass transfer limitation) is observed for this system. This was determined from a plot of the observed $k''/k_{1,\text{disp}}''$, against $N$, when compared with the correlation relationship of $(N/N_{\text{disp}})^\gamma$ and $k_{1,\text{disp}}$ and $N_{\text{disp}}$ are assumed to be the values from the 90 rpm data. A value of the power $\gamma_1$ of 1.2 (predicting a dispersion limited regime) appears to fit the observations reasonably well. A similar relationship was assumed to predict the effect of mixing on the $k_{1,\text{disp}}''$ parameter and a value of 2 for the power $\gamma_2$ appears to fit the observations. Since these are linear relationships the uncertainty in the values of the log plot gradients, $\gamma_1$ and $\gamma_2$, are
estimated from the standard equation used to predict a confidence interval for the true slope of a single variable linear regression model. The uncertainties for \( \gamma_1 \) and \( \gamma_2 \), shown in Table C5, are estimated at a 90% confidence level given three data points.

C.4 Knowledge level 4. Second generation layer separation model

A simple development for the layer separation model assumes observations of the possible presence of a durable dispersion band of uncertain height in the phase cut following a pre-determined settling time. The quantity of aqueous droplets contained in continuous organic phase band depth is described, as opposed to the assumption of a total drug loss fraction \( (u_2) \) used in the first generation layer separation model (Model B3, Appendix B). The second generation layer separation model is given in Model C3 where the stream number and phase indices refer to Figure B4 (Appendix B).

Organic phase volume loss,

\[
V_{3,org} = h_{band} A_{vessel} \left(1 - \varepsilon^*\right)
\]

Component mass balance,

\[
\sum_{i=1}^{8} z_{1,i,org} V_{3,org} \rho_{org} = 1000 \sum_{i=1}^{8} z_{l,org} Z_{l,org}
\]

\[
z_{3,i,org} = z_{1,i,org} - z_{3,i,org}
\]

for \( i = 1 \ldots 8 \) (Model C3)

where the notation is defined in Model B3 (Appendix B) and additionally

\[
A_{vessel} = \text{cross-sectional area of vessel, m}^2
\]

\[
h_{band} = \text{height of a two phase dispersion band, m}
\]

\[
V_{3,org} = \text{volume of organic phase within dispersion band, m}^3
\]

\[
\varepsilon^* = \text{equilibrium disperse phase hold-up fraction in dispersion band}
\]

\[
\rho_{org} = \text{density of organic phase, kg dm}^{-3}
\]

There are no degrees of freedom in this model since past observations do not appear to justify detailed investigations to model the complex dependence of dispersion band height on settling and coalescence rates, relating time and agitation effects. It is assumed that the diameter of the PPR plant vessel is 2.13 metres (with a height to diameter aspect ratio of 1:1, in the 2000 US gallon PPR vessel). The main assumptions made in this model include those made in the first generation layer separation model (Model B3, Appendix B) with the exception that:
• an equilibrium dispersion band height, $h_{\text{band}}$, is assumed to be attained, independent of vessel geometry, the relative volumes of aqueous and organic phases, mixing and dynamics (since settling and coalescence rate data is unavailable),

• the density of the organic phase is estimated from the linear function of drug concentration also used in the first generation batch distillation model (Model B4),

• an approximate value within the range of 0.5 to 0.75 is assumed for the equilibrium value of the disperse phase band hold-up (Godfrey and Slater, 1994), $\varepsilon^*$.

The assumed parameter values for the second generation layer separation model are given in Table C6, and the associated uncertainties in Stages 3, 6 and 8 are given in Table C7.

Table C6. Parameters for the second generation layer separation model, Case Study II.

<table>
<thead>
<tr>
<th>Model parameter values</th>
</tr>
</thead>
<tbody>
<tr>
<td>$h_{\text{band}} = 0.01$ m</td>
</tr>
<tr>
<td>$\varepsilon^* = 0.63$</td>
</tr>
</tbody>
</table>

Table C7. Uncertainty characterisation in the second generation layer separation models, Case Study II.

<table>
<thead>
<tr>
<th>Stage</th>
<th>Parameter</th>
<th>Uniform distribution, $U(\text{min}, \text{max})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>$h_{\text{band}}$, $\varepsilon^*$</td>
<td>$U(0, 0.02)$, $U(0.5, 0.75)$</td>
</tr>
<tr>
<td>6</td>
<td>$h_{\text{band}}$, $\varepsilon^*$</td>
<td>$U(0, 0.02)$, $U(0.5, 0.75)$</td>
</tr>
<tr>
<td>8</td>
<td>$h_{\text{band}}$, $\varepsilon^*$</td>
<td>$U(0, 0.02)$, $U(0.5, 0.75)$</td>
</tr>
</tbody>
</table>

C.5 Knowledge level 5. Second generation reagent addition model

Incorporation of liquid-liquid equilibrium tie-lines for loss of drugs through aqueous phase solubility effects provides a prediction for drug loss based on an estimation of physical property behaviour as opposed to a fractional loss assumption, $u_1$, used in the first generation reagent addition model (Model B2, Appendix B). The model for the second generation reagent addition model comprises:

the first generation reagent addition model (Model B2, Appendix B), with the exception of the following drug mole balances,

$$Z_{Sl} = z_{1,\text{org}} + z_{1,\text{aq}} \quad \text{for} \quad i = 1..5$$
Integrated design under uncertainty for pharmaceutical processes

\[ Z_{orgsv} = z_{1,org} \]

\[ z_{aqsv} = v_f \rho_f \left( \frac{100 - wt\%_F}{100} \right) + v_{1,aq} \rho_{aq} \left( \frac{100 - wt\%_{aq}}{100} \right) \]

\[ z_{2,org} \cdot z_{2,aq} = f(Z_{sl} \cdot Z_{orgsv} \cdot Z_{aqsv} \cdot \sigma^*_{sl,orgsv,aqsv}) \quad \text{for} \quad i = 1..5 \quad (Model C4) \]

where the stream number and phase indices refer to Figure B3 (Appendix B) and the notation is defined in Model B2 (Appendix B) and additionally,

\[ v \quad = \quad \text{volume, m}^3 \]
\[ f \quad = \quad \text{function determining solute split based on physical property equilibrium tie-line solubility data and phase mass,} \]
\[ Z \quad = \quad \text{stream mass, kg} \]
\[ \sigma^* \quad = \quad \text{equilibrium solubility of solute in organic-aqueous phase mixtures, wt\%} \]

subscripts

\[ aqsv = \text{pure aqueous solvent} \]
\[ orgsv = \text{pure organic solvent} \]
\[ sl = \text{drug solute} \]
\[ F = \text{reagent feed} \]
\[ aq = \text{aqueous phase} \]

The possible degrees of freedom are either the volume of distilled water added in the dilution operation or the addition reagent to active drug ingredient feed mole ratio governing the aqueous reagent strength. The uncertainty characterisations of the parameters in the second stage reagent addition model are given in Table C8.

Table C8. Uncertainty characterisation in the second generation reagent addition models, Case Study II.

<table>
<thead>
<tr>
<th>Stage</th>
<th>Parameter</th>
<th>Normal distribution at 95% confidence, N(\mu, \sigma)</th>
<th>Uniform distribution, U(min, max)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>(u_0)</td>
<td>(\sigma_{u*})</td>
<td>(U(3, 6))</td>
</tr>
<tr>
<td></td>
<td>(\sigma_{u*})</td>
<td>N(equilibrium tie-line data point, 10% nominal)</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>(\sigma_{u*})</td>
<td>N(equilibrium tie-line data point, 10% nominal)</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>(\sigma_{u*})</td>
<td>N(equilibrium tie-line data point, 10% nominal)</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>(\sigma_{u*})</td>
<td>N(equilibrium tie-line data point, 10% nominal)</td>
<td></td>
</tr>
</tbody>
</table>
Uncertainty is considered in the aqueous phase solubility equilibria data. The same assumptions apply to the second generation reagent addition model (Model B2, Appendix B) with the following exceptions:

- equilibrium solubility is observed at the endpoint of the operation, replacing the fractional loss parameter, $u_1$,
- the solubility behaviour of each drug component species is the same with the effect that the same tie-line data can be used,
- the solubility behaviour of each drug component species is independent of each other and the presence of any other impurities.

C.6 Knowledge level 6. First generation Stage 12 crystallisation model fitted to new data

Input and output drug product concentration data for the Stage 12 crystallisation process from a process sequence run of similar scale to that referred in Appendix C.2 (1000 US gallons), is available. However, without further profile data, revision of the first generation model structure (Model B5, Appendix B) is unjustified. This data is used to revise the value of $k_e$ to $6.1 \times 10^{-5}$ compared to $7.94 \times 10^{-5}$ m min$^{-1}$ (kg kgsolvent$^{-1}$)$^{1/2}$ obtained from the bench scale data. The uncertainty in $k_e$ is assumed to be normally distributed, $N(6.1 \times 10^{-5}, 10\%$ nominal). No data is available to revise the values of the other parameters in the first generation Stage 12 model.
D. CASE STUDY II: RESULTS FOR THE EFFECT OF INCOMING DATA

Additional results for Case Study II are presented in Appendix D, which have been referred to in Chapter 7. The cumulative frequency plots for the predicted total yield at different levels of knowledge are shown in Figure D1. Knowledge level 6 shows evidence that it provides the best prediction to the plant data since the distribution clearly encompasses the data point value (solid vertical line). Figure D2 shows the corresponding plots for the predicted key impurity endpoint content and that knowledge levels 3, 4, 5, and 6 exhibit the least uncertainty for this criterion. The main Sensitivity Analysis results for the different knowledge levels are shown in Table D1. The scatter plots for knowledge level 0 data, Figure D3, verify that the scatter of the key impurity reaction rate constant, $k_2$, and the key impurity 'solute loss' parameter, $\beta_{actC}$, with the endpoint key impurity content, are comparable (SRC of 0.78 and 0.77, Table D1). The relative uncertainty for the post-reactor key impurity composition and the endpoint content, shown in the cumulative frequency plots, Figure D4, indicate that a large proportion of the uncertainty accumulated in the endpoint criterion appears to be already present after Stage I (corresponding to a Stage I sub-sequence contribution of 0.62 compared to Stage 2-15 contribution of 0.38).

![Cumulative frequency plots](image)

Figure D1. Cumulative frequency plots of total yield prediction under uncertainty, Case Study II.
Figure D2. Cumulative frequency plots of key impurity content prediction, Case Study II.
Key: * = Knowledge level 0, o = Knowledge level 1, * = Knowledge level 2,
+ = Knowledge level 3, 4, 5, and 6.

Figure D3. Scatter plots of the key parameters with the endpoint key impurity content, Case Study II.
Key: * = Key impurity reaction rate constant, $k_2$, o = Key impurity 'solute loss' parameter, $\zeta_{\text{solute}}$. 
Figure D4. Cumulative frequency plots of inter-stage key impurity composition under uncertainty, Case Study II.

Key: * = Post-reaction key impurity composition, wt%, o = Endpoint key impurity content, wt%.
Table D1. Main SRC Sensitivity Analysis results for levels of knowledge incorporation, Case Study II.

<table>
<thead>
<tr>
<th>Knowledge level</th>
<th>Total yield</th>
<th>Key impurity content</th>
<th>Secondary impurity content</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>$R^2$</td>
<td>$R^2$</td>
<td>$R^2$</td>
</tr>
<tr>
<td>$k_2$</td>
<td>0.82</td>
<td>$\zeta_{actC}$ 0.78</td>
<td>$\zeta_{actE}$ 0.96</td>
</tr>
<tr>
<td>$k_1$</td>
<td>0.32</td>
<td>$k_2$ 0.77</td>
<td>$k_3$ 0.32</td>
</tr>
<tr>
<td>$u_1$</td>
<td>-0.17 to -0.20</td>
<td>$\eta_{wash}$ -0.21</td>
<td>$\eta_{wash}$ -0.19</td>
</tr>
<tr>
<td>$u_2$</td>
<td>-0.15 to -0.18</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>$R^2$</td>
<td>$R^2$</td>
<td>$R^2$</td>
</tr>
<tr>
<td>$k_3$</td>
<td>0.80</td>
<td>$\zeta_{actC}$ 0.61</td>
<td>$\zeta_{actE}$ 0.76</td>
</tr>
<tr>
<td>$k_2$</td>
<td>0.40</td>
<td>$k_3$ 0.58</td>
<td>$k_3$ 0.60</td>
</tr>
<tr>
<td>$u_1$</td>
<td>-0.12 to -0.13</td>
<td>$\eta_{wash}$ -0.18</td>
<td>$\eta_{wash}$ -0.10</td>
</tr>
<tr>
<td>$u_2$</td>
<td>-0.10 to -0.12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>$R^2$</td>
<td>$R^2$</td>
<td>$R^2$</td>
</tr>
<tr>
<td>$k_3$</td>
<td>0.79</td>
<td>$\zeta_{actC}$ 0.72</td>
<td>$\zeta_{actE}$ 0.73</td>
</tr>
<tr>
<td>$k_2$</td>
<td>0.30</td>
<td>$k_3$ 0.66</td>
<td>$k_3$ 0.67</td>
</tr>
<tr>
<td>$u_1$</td>
<td>-0.15 to -0.22</td>
<td>$\eta_{wash}$ -0.20</td>
<td>$\eta_{wash}$ -0.14</td>
</tr>
<tr>
<td>$u_2$</td>
<td>-0.16 to -0.21</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>$R^2$</td>
<td>$R^2$</td>
<td>$R^2$</td>
</tr>
<tr>
<td>$k_3$</td>
<td>0.67</td>
<td>$k_3$ 0.89</td>
<td>$k_3$ 0.89</td>
</tr>
<tr>
<td>$k_2$</td>
<td>0.58</td>
<td>$\zeta_{actC}$ 0.54</td>
<td>$\zeta_{actE}$ 0.44</td>
</tr>
<tr>
<td>$u_2$</td>
<td>-0.13 to -0.21</td>
<td>$\eta_{wash}$ -0.25</td>
<td>$\eta_{wash}$ -0.15</td>
</tr>
<tr>
<td>$\gamma_1$</td>
<td>-0.16</td>
<td>$\gamma_1$ -0.14</td>
<td></td>
</tr>
<tr>
<td>$\gamma_2$</td>
<td>-0.16</td>
<td>$\gamma_2$ -0.13</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>$R^2$</td>
<td>$R^2$</td>
<td>$R^2$</td>
</tr>
<tr>
<td>$k_2$</td>
<td>0.60</td>
<td>$k_2$ 0.89</td>
<td>$k_3$ 0.87</td>
</tr>
<tr>
<td>$k_1$</td>
<td>0.58</td>
<td>$\zeta_{actC}$ 0.55</td>
<td>$\zeta_{actE}$ 0.45</td>
</tr>
<tr>
<td>$t$</td>
<td>-0.30</td>
<td>$t$ -0.26</td>
<td>$\eta_{wash}$ -0.15</td>
</tr>
<tr>
<td>$u_1$</td>
<td>-0.14 to -0.17</td>
<td>$\eta_{wash}$ -0.25</td>
<td></td>
</tr>
<tr>
<td>$\gamma_1$</td>
<td>-0.15</td>
<td>$\gamma_1$ -0.14</td>
<td></td>
</tr>
<tr>
<td>$\gamma_2$</td>
<td>-0.15</td>
<td>$\gamma_2$ -0.12</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>$R^2$</td>
<td>$R^2$</td>
<td>$R^2$</td>
</tr>
<tr>
<td>$k_3$</td>
<td>0.64</td>
<td>$k_2$ 0.75</td>
<td>$k_3$ 0.88</td>
</tr>
<tr>
<td>$k_2$</td>
<td>0.47</td>
<td>$\zeta_{actC}$ 0.63</td>
<td>$\zeta_{actE}$ 0.44</td>
</tr>
<tr>
<td>$\sigma_{st}$</td>
<td>-0.23 to 0.44</td>
<td>$\tau$ -0.27</td>
<td>$\eta_{wash}$ -0.14</td>
</tr>
<tr>
<td>$\gamma_1$</td>
<td>-0.17</td>
<td>$\gamma_1$ -0.15</td>
<td></td>
</tr>
<tr>
<td>$\gamma_2$</td>
<td>-0.15</td>
<td>$\gamma_2$ -0.12</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>$R^2$</td>
<td>$R^2$</td>
<td>$R^2$</td>
</tr>
<tr>
<td>$k_3$</td>
<td>0.77</td>
<td>$k_2$ 0.75</td>
<td>$k_3$ 0.88</td>
</tr>
<tr>
<td>$k_2$</td>
<td>0.38</td>
<td>$\zeta_{actC}$ 0.63</td>
<td>$\zeta_{actE}$ 0.44</td>
</tr>
<tr>
<td>$\sigma_{st}$</td>
<td>-0.20 to 0.36</td>
<td>$t$ -0.26</td>
<td>$\eta_{wash}$ -0.13</td>
</tr>
<tr>
<td>$\gamma_1$</td>
<td>-0.14</td>
<td>$\gamma_1$ -0.15</td>
<td></td>
</tr>
<tr>
<td>$\gamma_2$</td>
<td>-0.12</td>
<td>$\gamma_2$ -0.12</td>
<td></td>
</tr>
</tbody>
</table>
Appendix E

E. CASE STUDY II: OPTIMISATION RESULTS

Appendix E contains additional optimisation results for Case Study II which have been referred to in Chapter 7, 8 and 9. These include input uncertainty reduction optimisation results, Robustness Analysis on parameter uncertainty characterisation results and results presented for the analysis of alternative process flowsheets regarding flowsheet optimisation under uncertainty, Uncertainty and Sensitivity Analysis and optimal operating policy error tolerance results. Computational statistics are presented for the Case Study II optimisation problems.

E.1 Optimisation results for uncertainty reduction.

The results in Table E1 show how the optimal key input parameter uncertainties vary with parametric increases in the desired reduction in the output criteria (total yield, key and secondary endpoint impurity contents). No feasible solution was found for desired uncertainty reduction levels above 60% (of the original uncertainty in the knowledge level 6 model). Computational statistics for this problem are given in Table E6.

Table E1. Optimisation results for optimal key input parameter uncertainty reductions for knowledge level 6 Base Case flowsheet stochastic model, Case Study II.

<table>
<thead>
<tr>
<th>Desired uncertainty reduction in output performance, %</th>
<th>0</th>
<th>10</th>
<th>20</th>
<th>30</th>
<th>40</th>
<th>50</th>
<th>60</th>
<th>70, 80, 90</th>
</tr>
</thead>
<tbody>
<tr>
<td>Objective function, $\sum \delta_i$</td>
<td>7.00</td>
<td>6.30</td>
<td>5.94</td>
<td>5.56</td>
<td>4.83</td>
<td>3.93</td>
<td>1.72</td>
<td>-</td>
</tr>
<tr>
<td>Decisions</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\delta_{\sigma_{y1}}$</td>
<td>1.00</td>
<td>0.93</td>
<td>0.73</td>
<td>0.64</td>
<td>0.47</td>
<td>0.28</td>
<td>1.0x10^2</td>
<td></td>
</tr>
<tr>
<td>$\delta_{\sigma_{y2}}$</td>
<td>1.00</td>
<td>0.82</td>
<td>0.81</td>
<td>0.56</td>
<td>0.50</td>
<td>0.40</td>
<td>9.1x10^2</td>
<td></td>
</tr>
<tr>
<td>$\delta_{\sigma_{n1}}$</td>
<td>1.00</td>
<td>1.00</td>
<td>0.72</td>
<td>0.98</td>
<td>0.58</td>
<td>0.39</td>
<td>0.34</td>
<td></td>
</tr>
<tr>
<td>$\delta_{\sigma_{n2}}$</td>
<td>1.00</td>
<td>0.82</td>
<td>0.68</td>
<td>0.55</td>
<td>0.49</td>
<td>0.40</td>
<td>0.27</td>
<td></td>
</tr>
<tr>
<td>$\delta_{\sigma_{n3}}$</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.83</td>
<td>0.92</td>
<td>0.82</td>
<td>6.3x10^2</td>
<td></td>
</tr>
<tr>
<td>$\delta_{\sigma_{w1}}$</td>
<td>1.00</td>
<td>0.97</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.93</td>
<td>0.44</td>
<td></td>
</tr>
<tr>
<td>$\delta_{\sigma_{w2}}$</td>
<td>1.00</td>
<td>0.77</td>
<td>1.00</td>
<td>1.00</td>
<td>0.86</td>
<td>0.69</td>
<td>1.1x10^2</td>
<td></td>
</tr>
<tr>
<td>5-95% Fractile width</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$Y_f$, %</td>
<td>7.63</td>
<td>6.87</td>
<td>6.10</td>
<td>5.34</td>
<td>4.58</td>
<td>3.82</td>
<td>3.05</td>
<td></td>
</tr>
<tr>
<td>$w_{f_{\text{solC}}}$, %</td>
<td>0.17</td>
<td>0.15</td>
<td>0.13</td>
<td>0.12</td>
<td>0.10</td>
<td>0.08</td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td>$w_{f_{\text{solE}}}$, %</td>
<td>1.43</td>
<td>0.98</td>
<td>1.15</td>
<td>1.00</td>
<td>0.86</td>
<td>0.71</td>
<td>0.40</td>
<td></td>
</tr>
</tbody>
</table>
E.2 The importance of the state of knowledge of the uncertainties results

The results of the Robustness Analysis to input uncertainty distribution, Figure E1, shows only a limited difference between output criterion distributions predicted using either normally or uniformly distributed input uncertainties (Case 2), but as expected, a significant difference when the input uncertainty spread is increased (Case 1).

![Cumulative frequency plots for validated optimisation results under input uncertainty variations, Case Study II.](image)

(a) Total yield. (b) Key impurity content.

Figure E1. Cumulative frequency plots for validated optimisation results under input uncertainty variations, Case Study II.

Key: o = Base Case (normally distributed), • = Case 1 (50% increase in spread, normally distributed), x = Case 2 (uniformly distributed).

E.3 Alternate flowsheet optimisation under uncertainty results.

Tables E2 and E3 and Figures E2 and E3 show the process flowsheet stochastic optimisation results for the four alternative flowsheets considered in Case Study II (Chapter 9). Alternative 2 returns the best profitability predictions under uncertainty. Only the Stage 1 reaction time ($t_r$) and Stage 12 crystallisation holding period (HP$_{12}$) decision variables appear sensitive to the different flowsheet configurations. Table E4 shows the main results of the Sensitivity Analysis where only SRC measures were required.
Table E2. Validated alternate process flowsheet optimisation under uncertainty results, Case Study II.

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Base Case</th>
<th>Alternative 1</th>
<th>Alternative 2</th>
<th>Alternative 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E[\text{P}<em>{\text{ty}}](\text{S kg}</em>{\text{cat}}^{-1} \text{hr}^{-1})$</td>
<td>30.97</td>
<td>31.13</td>
<td>32.26</td>
<td>31.65</td>
</tr>
<tr>
<td>$E[\text{P}<em>{\text{ty}} \text{loss}](\text{S kg}</em>{\text{cat}}^{-1} \text{hr}^{-1})$</td>
<td>3.02</td>
<td>2.97</td>
<td>3.04</td>
<td>3.10</td>
</tr>
<tr>
<td>$\text{P}_{\text{pass}}$</td>
<td>0.894</td>
<td>0.904</td>
<td>0.905</td>
<td>0.901</td>
</tr>
<tr>
<td>$<a href="%25">E(\text{wt}<em>{\text{act},C}), E(\text{wt}</em>{\text{act},E})</a>$</td>
<td>[0.22, 1.42]</td>
<td>[0.21, 1.36]</td>
<td>[0.23, 1.42]</td>
<td>[0.23, 1.42]</td>
</tr>
<tr>
<td>$<a href="%25">\text{FW}(\text{wt}<em>{\text{act},C}), \text{FW}(\text{wt}</em>{\text{act},E})</a>$</td>
<td>[0.14, 1.19]</td>
<td>[0.17, 1.36]</td>
<td>[0.15, 1.14]</td>
<td>[0.14, 1.19]</td>
</tr>
<tr>
<td>$E(\text{Y}_{\text{T}})(%)$</td>
<td>86.5</td>
<td>89.2</td>
<td>86.9</td>
<td>85.6</td>
</tr>
<tr>
<td>$\text{FW}(\text{Y}_{\text{T}})(%)$</td>
<td>5.4</td>
<td>5.4</td>
<td>6.3</td>
<td>5.7</td>
</tr>
</tbody>
</table>

Table E3. Optimal decisions for alternate process flowsheet optimisation under uncertainty, Case Study II.

<table>
<thead>
<tr>
<th>Decisions</th>
<th>Base Case</th>
<th>Alternative 1</th>
<th>Alternative 2</th>
<th>Alternative 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t_{\text{c},1}$ (min)</td>
<td>251</td>
<td>247</td>
<td>247</td>
<td>241</td>
</tr>
<tr>
<td>$N_1$ (rpm)</td>
<td>90.0</td>
<td>90.0</td>
<td>90.0</td>
<td>90.0</td>
</tr>
<tr>
<td>$\text{AF}_{10}$</td>
<td>0.40</td>
<td>0.40</td>
<td>0.40</td>
<td>0.40</td>
</tr>
<tr>
<td>$\text{RF}_{10}$</td>
<td>0.70</td>
<td>0.70</td>
<td>0.70</td>
<td>0.70</td>
</tr>
<tr>
<td>$\text{CR}_{12}$ ($^\circ{\text{C}} \text{min}^{-1}$)</td>
<td>0.50</td>
<td>0.50</td>
<td>0.50</td>
<td>0.50</td>
</tr>
<tr>
<td>$\text{HP}_{12}$ (min)</td>
<td>54</td>
<td>52</td>
<td>47</td>
<td>50</td>
</tr>
</tbody>
</table>

$R_{\text{redF-back}} = 0.05$ $R_{\text{roll-recycle}} = 0.5$ 

Figure E2. Cumulative frequency plots for potential profitability, Case Study II.
Key: + = Base Case. * = Alternative 1. o = Alternative 2. x = Alternative 3.
Figure E3. Cumulative frequency plots for validated optimisation results under input uncertainty variations, Case Study II.

Key: + = Base Case, * = Alternative 1, o = Alternative 2, x = Alternative 3.
E.5 Alternate flowsheet operating policy tolerance optimisation results.

The optimal control variable range tolerance fraction decisions for the four process flowsheet alternatives based on 99% achievement of their respective ‘here and now’ optimal mean profitabilities, are given in Tables E5, E6, E7 and E8. Zero δ values indicate no tolerance in a particular policy variable deviation direction.

Table E5. Optimal decisions for the Base Case operating policy tolerance, Case Study II.
Table E6. Optimal decisions and tolerance limits for Alternative 1 operating policy tolerance, Case Study II.

| \( \delta_{a+}^{U} \) | \( \delta_{a+}^{L} \) | \( z_{d+}^{UB} \) | \( z_{d+}^{LB} \) | \( t_{F,1} \) | \( N_{1} \) | \( A_{F10} \) | \( R_{F10} \) | \( CR_{12} \) | \( HP_{12} \) | \( R_{\text{vol}F_{\text{back}}} \) |
|----------------|----------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|
| 0.281          | 0.544          | 290         | 221         | 0           | 0.042       | 0           | 0           | 0.001       | 0.001       | 0.027       |

Table E7. Optimal decisions and tolerance limits for Alternative 2 operating policy tolerance, Case Study II.

| \( \delta_{a+}^{U} \) | \( \delta_{a+}^{L} \) | \( z_{d+}^{UB} \) | \( z_{d+}^{LB} \) | \( t_{F,1} \) | \( N_{1} \) | \( A_{F10} \) | \( R_{F10} \) | \( CR_{12} \) | \( HP_{12} \) | \( R_{\text{vol}F_{\text{recycle}}} \) |
|----------------|----------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|
| 0.127          | 0.488          | 266         | 224         | 0           | 0.042       | 0           | 0           | 0.016       | 0.065       | 0           |

Table E8. Optimal decisions and tolerance limits for Alternative 3 operating policy tolerance, Case Study II.

| \( \delta_{a+}^{U} \) | \( \delta_{a+}^{L} \) | \( z_{d+}^{UB} \) | \( z_{d+}^{LB} \) | \( t_{F,1} \) | \( N_{1} \) | \( A_{F10} \) | \( R_{F10} \) | \( CR_{12} \) | \( HP_{12} \) |
|----------------|----------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|
| 0.110          | 0.059          | 258         | 238         | 0           | 0.042       | 0           | 0           | 0.004       | 0.042       |

E.6 Computation statistics

These stochastic optimisation problems were solved using the MATLAB (Version 6.0) programming software and the Optimisation Toolbox function for non-linear constrained optimisation algorithm based on SQP. They were performed on a RS6000 IBM workstation. The stochastic optimisation problems involving decision variables which redefine the input uncertainty space in some way (optimal input uncertainty reduction - Problem type P1, and maximum operating policy tolerance - Problem type P3) required CPU times of an order of magnitude greater than the stochastic flowsheet optimisations of fixed
space definition (Problem type P2). The latter required CPU times two orders of magnitude greater than their respective nominal optimisation problems.

Table E9. Computational statistics for optimisation problems, Case Study II.

<table>
<thead>
<tr>
<th>Input parameter uncertainty reduction:</th>
<th>Problem formulation</th>
<th>CPU time (seconds)</th>
<th>Optimisation iterations</th>
<th>Function evaluations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base case 10% red'n</td>
<td>P8 (type P1)</td>
<td>2.46×10^4</td>
<td>53</td>
<td>548</td>
</tr>
<tr>
<td>20% red'n</td>
<td>5.02×10^4</td>
<td>120</td>
<td>1094</td>
<td></td>
</tr>
<tr>
<td>30% red'n</td>
<td>3.04×10^4</td>
<td>78</td>
<td>753</td>
<td></td>
</tr>
<tr>
<td>40% red'n</td>
<td>5.45×10^4</td>
<td>123</td>
<td>1129</td>
<td></td>
</tr>
<tr>
<td>50% red'n</td>
<td>6.12×10^4</td>
<td>139</td>
<td>1312</td>
<td></td>
</tr>
<tr>
<td>60% red'n</td>
<td>1.01×10^5</td>
<td>24</td>
<td>221</td>
<td></td>
</tr>
<tr>
<td>Nominal flowsheet optimisation:</td>
<td>P9</td>
<td>67</td>
<td>11</td>
<td>102</td>
</tr>
<tr>
<td>Uncertain flowsheet optimisation:</td>
<td>P10 (type P2)</td>
<td>3256</td>
<td>11</td>
<td>66</td>
</tr>
<tr>
<td>Uncertain flowsheet optimisation with input uncertainty variation:</td>
<td>Type P10 (type P2)</td>
<td>3719</td>
<td>12</td>
<td>77</td>
</tr>
<tr>
<td>Case 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Uncertain flowsheet optimisation with input uncertainty variation:</td>
<td>Type P10 (type P2)</td>
<td>4149</td>
<td>13</td>
<td>85</td>
</tr>
<tr>
<td>Case 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Uncertain flowsheet optimisation with uncertain feed purity:</td>
<td>Type P10 (type P2)</td>
<td>3213</td>
<td>11</td>
<td>66</td>
</tr>
<tr>
<td>Base Case</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nominal flowsheet optimisation:</td>
<td>Type P9</td>
<td>50</td>
<td>27</td>
<td>197</td>
</tr>
<tr>
<td>Alternative 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Uncertain flowsheet optimisation:</td>
<td>Type P10 (type P2)</td>
<td>5365</td>
<td>10</td>
<td>85</td>
</tr>
<tr>
<td>Alternative 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nominal flowsheet optimisation:</td>
<td>Type P9</td>
<td>19</td>
<td>10</td>
<td>77</td>
</tr>
<tr>
<td>Alternative 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Uncertain flowsheet optimisation:</td>
<td>Type P10 (type P2)</td>
<td>2387</td>
<td>6</td>
<td>41</td>
</tr>
<tr>
<td>Alternative 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nominal flowsheet optimisation:</td>
<td>Type P9</td>
<td>19</td>
<td>12</td>
<td>79</td>
</tr>
<tr>
<td>Alternative 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Uncertain flowsheet optimisation:</td>
<td>Type P10</td>
<td>3028</td>
<td>11</td>
<td>80</td>
</tr>
<tr>
<td>Alternative 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Operating policy tolerance optimisation:</td>
<td>P11 (type P3)</td>
<td>1.93×10^4</td>
<td>52</td>
<td>411</td>
</tr>
<tr>
<td>Base Case</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Operating policy tolerance optimisation:</td>
<td>Type P3</td>
<td>4.08×10^4</td>
<td>77</td>
<td>798</td>
</tr>
<tr>
<td>Alternative 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Operating policy tolerance optimisation:</td>
<td>Type P3</td>
<td>3.85×10^4</td>
<td>80</td>
<td>806</td>
</tr>
<tr>
<td>Alternative 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Operating policy tolerance optimisation:</td>
<td>Type P3</td>
<td>2.72×10^4</td>
<td>73</td>
<td>605</td>
</tr>
<tr>
<td>Alternative 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>