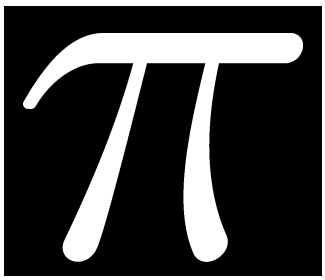


The
Pacific Institute
for the Mathematical Sciences

http://www.pims.math.ca
pims@pims.math.ca

Proceedings of the third

PIMS Industrial Problem Solving Workshop

PIMS IPSW 3

Co-sponsored by:

The Natural Sciences and Engineering Research Council of Canada

The Alberta Science and Research Authority

The British Columbia Information, Science and Technology Agency

Editor: C. Bose, University of Victoria

Proceedings of the Third Annual PIMS Industrial Problem Solving Workshop

Editor: C. Bose, University of Victoria

Co-sponsored by:
The Natural Sciences and Engineering Research Council of Canada
The Alberta Science and Research Authority
The British Columbia Information, Science and Technology Agency

January, 2000

Foreword by the PIMS Director

At the heart of the PIMS Industrial Outreach Program are the intensive week-long problem solving workshops initiated with the aim of creating a mutually beneficial link between researchers in industry and academic mathematicians. Researchers with industrial and commercial concerns are invited to present one of their current technical problems for study in working sessions with leading specialists from the academic community. As can be seen in these proceedings, problems may come from a wide variety of subject areas, but should be amenable to mathematical modeling and analysis.

PIMS held its first Industrial Problem Solving Workshop in August 1997 at UBC. Following the success of this workshop, it was decided to make IPSW an annual PIMS event and a second one –equally successful– was organized in May 1998 at the University of Calgary. This year, approximately 90 faculty, postdoctoral fellows, graduate students and industry scientists came together at the University of Victoria for the Third PIMS Industrial Problem Solving Workshop from May 31 to June 4, 1999.

Six problems were investigated at this year's workshop. The problems were contributed by *Searle Corporation (Chicago), RSI Technologies (Victoria), Charles Howard and Associates (Victoria), Chemex Labs (Vancouver), Merak (Calgary),* and *Enbridge (Calgary).* From the quality of these proceedings it is clear that significant progress and insight was made on each of the problems. Industry representatives have expressed satisfaction with the outcome of the workshop.

The success of the PIMS Industrial Problem Solving Program has been crucial to the development of the Network of Centers of Excellence MITACS: The Mathematics of Information Technology and Complex Systems Network which is currently the largest research network in Canada that comprises more than 185 researchers in 23 universities. Thanks to the PIMS outreach program, 22 private companies have already contributed more than $700,000 into the 9 MITACS projects led by the PIMS scientists. MITACS will benefit from a federal investment of close to $14.5 million over the next four years.

PIMS looks forward to continuing the Industrial Problem Solving Workshop Series, with the next one to be held at the University of Alberta in June 2000. More information about upcoming activities at PIMS may be found on the web page, *www.pims.math.ca.*

Workshops like this are only possible through the dedication and hard work of many PIMS scientists. Prof. Frank Ruskey, the PIMS Site Director at the University of Victoria, Kelly Choo, the PIMS Administrative Assistant at the University of Victoria and Dr. H. Huang, the former PIMS Industrial Facilitator for BC all worked tirelessly for the success of this workshop. Special thanks go to Prof. Chris Bose from the University of Victoria who took charge of compiling and editing these proceedings. I would also like to thank the Natural Sciences and Engineering Research Council of Canada, the Alberta Science and Research Authority, and the British Columbia Information, Science and Technology Agency for supporting PIMS and its many programmes.

Dr. Nassif Ghoussoub, Director
Pacific Institute for the Mathematical Sciences

Contents

List of Figures

List of Tables

Editor's Preface

From May 31 through June 4, 1999 approximately 90 mathematicians and industrial participants gathered at the University of Victoria for the Third Annual PIMS Industrial Problem Solving Workshop – the IPSW-3. These Proceedings document the work which was done during this intense week, as well as the significant effort put in after everyone went home by the report writers and their assistants. On the last day of the conference, each working group gave a presentation summarizing their findings. To a first-time IPSW participant such as myself, it was amazing to see the way in which the groups were able to meet the challenge of compiling the scientific and mathematical activities of their colleagues from the various working groups in such a cohesive and productive way and in such a short time. To some extent, these summary presentations formed the first step in the production of these Proceedings. However, the skill and effort required to take the summary presentations and from them produce the polished technical reports contained herein is not to be underestimated. I want to express my sincerest thanks to the writers who volunteered to handle this part of the job – I have only them to thank for the high standards of exposition and impressive scientific level of the reports. Of more interest to the editor than the reader, but important nonetheless, I also want to thank the writers for the prompt and professional way in which they completed their work.

The principal authors of the reports are as follows.

- John Stockie, Simon Fraser University (Charles Howard and Associates Limited)
- Luz Palacios and Rita Aggarwala, University of Calgary (Chemex Labs)
- Tim Myers, Cranfield University and Jim Brannan, Clemson University (Enbridge Incorporated)
- Min Tsao, University of Victoria (Merak Projects Limited)
- James Watmough, University of Victoria (RSI Technologies)
- Andrej Bona and Claude Laflamme, University of Calgary (Searle Corporation)

Each report was carefully proofread and 'refereed' by a colleague in our department, and I would like to thank my referees for their conscientious job.

I was also one of the principal organizers of IPSW-3, and this seems to be as good place as any to acknowledge the contributions of dozens of people who worked behind the scenes to make the event come together.

First, the PIMS industrial facilitators, Huaxiong Huang (UBC, now a faculty member at York University) and Mark Paulhus (Calgary) take the bulk of the credit for the scientific success of the workshop. The selection of the problems and of the 'experts' are critical steps, and they handled both with skill, diplomacy and flair.

The 'experts,' Chris Budd (Bath, UK), Tim Myers (Cranfield, UK) and Ellis Cumberbatch (Clairmont, USA) made it look easy, but I know that their jobs were a lot of work, requiring considerable skill to get the groups working and keep them moving in productive directions.

The company representatives were a great asset. While they may have been a bit unsure of what they were getting into when they came to give their presentations on Monday morning, standing in front of nearly 90 eager mathematicians, they held their own admirably, and in some cases became active members of the working groups. The industrial representatives this year were

- Doug Smith (Charles Howard and Associates Limited)
- Brenda Caughlin (Chemex Labs)
- Don Scott (Enbridge Incorporated)
- Jason McVean (Merak Projects Limited)
- David Lokhorst (RSI Technologies)
- Tom Doman (Searle)

Locally, we had a number of people helping to put this conference together. Kelly Choo (UVic PIMS Administrative Assistant) set up our computer network and handled the myriad of details about which the rest of us had no idea. Frank Ruskey (PIMS site director for UVic) kept us all moving forward and acted as go-between with the Vancouver office (as well as signing all the cheques). Rod Edwards did a fine job on the accommodation front, and Sean Bohun spent a long Saturday with Kelly setting up workstations. A number of other UVic faculty donated computers, volunteer time, or both.

Finally, it is a pleasure to thank the supporting institutions. UVic gave us rooms, equipment and computer services. The UVic Computer Science Department came through with much of the computing hardware, and most important of all, PIMS started the ball rolling three years ago with IPSW-1, and has kept the project going since then, devoting money, people and seemingly unlimited enthusiasm for the industrial-academic partnership.

Christopher J. Bose, Editor
Department of Mathematics and Statistics
University of Victoria

Chapter 1

PREFACE – Description of the Problems

Contaminant Transport in Municipal Water Systems – Charles Howard and Associates Limited

Problem Description

The problem is to determine water quality in a water distribution network as a function of time and location. The mathematical problem can be visualized (vaguely) as a system of equations that describe the spatial and temporal transport and dilution of a contaminant injected at a specific location.

Municipal water distribution systems are unsteady flow, nonlinear networks. In general, the unknowns are: the pressures (after eliminating elevation differences) at each node, the consumption at each node, and the frictional characteristics of each link. The flow in a link is a derived variable that depends on the difference in pressure across the link and the frictional characteristic of the link. In special cases some of the unknowns are known from field measurements, but in general the number of unknowns exceeds the number of equations that can be written (one equation of flow continuity for each node). In practice measurements are expensive so many of the unknowns are assumed to be known.

The main interest of the company is to determine the distribution of the contaminant volume, ie, its mass and concentration, as a function of time - who gets to drink it, for how long, and when. It is desirable to develop first, a closed form mathematical solution for an approximation to the real problem, and second, a numerical method that could be used to determine the range of applicability of the mathematical method. Then use the mathematical method to solve the inverse problem - given time varying measurements (of concentration and some of the pressures) determine the likely source(s) of the contaminant.

Presented by:
Doug Smith
Charles Howard and Associates Limited, Victoria, B.C.

Quality Control for Multi-variable Problems – Chemex Laboratories

Introduction

Analytical methods where many elements are determined simultaneously present special difficulties in quality control. The use of individual control charts for each variable (element) is generally not effective due to the combinative aspect of the variation. If there are n elements and the error on each is independent, the probability that all the results will fall within their respective 95% confidence limits is $(0.95)^n$. If $n = 30$ (multi-element analytical packages typically have 25 to 40 elements) this probability is equal to $(0.95)^{30}$ or 0.2146. In other words there will be one or more elements out of bounds in 78.53% of all batches, which on a simplistic interpretation suggests that about two-thirds of all batches will be rejected when nothing is actually wrong with the sample data.

In practice there will normally be both correlated error, (due to instrument variations that affect all elements or groups of elements) and independent error which is due to variations in the individual element measurement channels.

Some of the solutions that are being used or tested are listed in the next section.

Widen Control Limits

The general approach has been to widen the control limits for multi-element analyses to 3s. This has the effect of reducing the probability of false batch rejection for purely independent error to $1 - (0.9974)^n$ which is 0.075 (7.5%).

The disadvantages of this approach are a sacrifice in data quality for each individual element and as more and more elements are included in data packages (50+) the probability of false batch rejection is relatively high - greater than 10% for 50 elements.

Multivariate Quality Control

Principle component analysis has been examined. R.J. Howarth, M.H. Ramsey and B.J. Coles presented "The potential of multivariate quality control as a diagnostic tool in Geoanalysis" at Geoanalysis 97 (Vail, CO, June 97). Data from a 35 element multi-element analytical package was reduced to four principle components. They then suggested control be done by using the five control charts - one for each of the four components and a fifth for Hotelling's T - representing the multivariate distance.

This is a sophisticated approach that requires large training sets and significant calculations for each test. Ideally the QC control can be done 'on-line' during the analytical sequence (the instrument software recognizes a control sample and subroutine testing for in / out of control is down before the next sample is analyzed - in less than a minute - so that time is not wasted doing out of specification analysis). The calculations may be too complex for on-line control.

More problematic is the 'components' that are identified. Once a sample is identified as unacceptable the instrument operator needs to take corrective actions to resolve the problem before proceeding with further analyses. If the components are not closely related to physical properties (or are all combinations of a number of properties), identification of the 'fix' to the out of control problem may be difficult.

As well, multiple control charts, one for each component, are still being used.

Binomial Probabilities

Tables of the binomial probability of M out of N points falling above the 95th (90th or 99th as well) can be constructed. Control limits can then set for 95% (2s) and the number of elements failing versus the total number determined be evaluated. If the probability of this event is less than a predetermined level (i.e. 0.5) the sample can be determined to be a failure.

Presented by:
Brenda Caughlin
Chemex Laboratories, Vancouver, B.C.

Modelling Batch Interfaces – Enbridge Incorporated

Background Information

Petroleum products are most economically shipped in long distance pipelines. There can be significant differences between the characteristics of the various petroleum products that are shipped, so they are shipped in batches of fluids with like characteristics. At the limits of differing characteristics are 'clean' refined products (gasoline, diesel fuel, and aviation fuels) compared to heavy crude oils that contain sulfur compounds. The owners of the products insist that cross contamination of the batches be minimized. Separation of the batches by a mechanical devise such as a pipeline pig is not practically possible.

When the different petroleum products with different properties such as density and molecular viscosity are transported it is important to be able to distinguish the boundary between them so that upon arrival at the terminal consecutive products may be segregated into their respective tanks with a minimum of contamination. This process of segregation (or 'cutting' the stream) is made considerably more difficult when the interface between different products is tortuous with fingering of one component into the other. To reduce the effects of batch interfaces the product batches are made as large as possible so the batch interface becomes a smaller % of the entire batch. Some of the interfaces are considered as 'transmix' and must be re-refined (i.e. crude oil in gasoline). Other interfaces can be used by allowing the higher value product in the interface to be degraded (i.e. a premium gasoline/regular gasoline interface will all go to the regular gasoline tank).

Compounding the interface mixing are the mechanical components of a pipeline. A pipeline has bends and fittings (sectionalizing valves) that disturb the flow and repumping stations (centrifugal pumps) at which the pressure is boosted at approximately 50 mile intervals.

If it were possible to calculate batch interfaces, or even to better understand which parameters determine their growth rate it may be possible to optimize batch shipments by reducing the product degradation. This would have a considerable economic impact on petroleum pricing. As a side issue of this problem is the cost of shipping through a pipeline when the drag coefficient is larger than need be.

Statement of the Problems

The purpose of the workshop is to explore the mechanisms by which the bi-fluid vertical interface deforms with time when these two fluids flow through a pipe of cylindrical geometry under the driving force of a constant pressure gradient. These differing fluids may have different viscosity or different densities or both. Obviously the case of differing densities will be considerably more difficult because the flow will lose cylindrical symmetry due to the influence of gravity producing negative buoyancy effects.

Physics

The flow with both properties different will behave like a combination of a Poiseuille flow modified by gravity-current-like behaviour. That is the current will lose symmetry as the heavier fluid moves down the side of the container and under the lighter fluid.

Fluids Properties

Fluids Properties Consider that any combination of these fluids could be shipped or batched with any other (head to tail):

Empirical Information

Pipeline companies have developed means to cope with the problem. Firstly all the pipeline flows are fully turbulent, and pipeline start-up precedures have been developed that quickly take the flow through the laminar region. Typically the flow speed is 5 mph and Reynolds numbers are a minimum of 4000. Batch interfaces tend to stabilize in a 20" pipeline at about 250 m3. A 20" or 24" pipeline is a good size to consider.

MATERIAL	DENSITY (kg/cubic meter)	VISCOSITY (centistokes at 15 C)
NGL	560	0.1
Gasoline	700	0.4
Diesel Fuel	800	0.9
Synthetic Crude	585	5.5
Condensate	672	0.5
Sweet Lights	830	6.0
Sour Lights	852	15.0
Medium Crude	886	55.0
Heavy Crude	920	170.0

Table 1.1: Density and Viscosity of Petroleum Products

Presented by:
Don Scott
Enbridge Incorporated, Calgary, Alberta

Efficient Portfolio Selection – Merak Projects Limited

If portfolios made up of a selection of petroleum projects are plotted on a graph of expected value versus risk, there is an upper boundary above which no portfolios are found. This upper boundary is known as the efficient frontier.

Harry Markowitz revolutionized the field of portfolio theory with his pioneering work in the 1950s. His approach to efficient frontier analysis uses matrix algebra to determine an analytical expression for the line representing the efficient frontier.

In applying efficient frontier theory to the realm of the petroleum industry, Merak has taken a different approach. Merak's approach uses previously generated Monte Carlo results and a random portfolio generation scheme to plot thousands of valid portfolios on an efficient frontier graph. It quickly becomes apparent that there is an upper boundary, although a line is not explicitly drawn.

Both of these approaches have strengths and weaknesses. For instance, Merak's approach lacks the analytical certainty regarding the efficiency of promising portfolios that the Markowitz approach has. Even though there may seem to be no portfolios above a particular portfolio on the graph, it is always possible that the next randomly generated portfolio will be the best yet.

On the other hand, the Markowitz approach has some severe limitations when applied to the petroleum industry:

- The simplification of describing a risk profile with only a mean and variance will lead to inaccuracies in the efficient frontier.
- It may not be possible to participate in a project at an arbitrarily fine level of granularity. Some projects may be such that they require 100% investment or they cannot be done at all.
- The constraints that determine which portfolios are valid can be complicated in the petroleum industry. Constraints like "If A then also B, C, and D" or "If E then not F or G" or "At least 2 of H, I, J, and K" cannot be easily expressed as a linear equation, which is required for the Markowitz approach.

The goal of this project is to address the weaknesses of the two approaches, potentially by combining or partially combining them. A efficient frontier analysis method that combined the robustness of the Monte Carlo approach with the confidence of the Markowitz approach would be a powerful tool for any industry.

Presented by:
Jason McVean
Merak Projects Limited, Calgary, Alberta

Dynamics of Large Mining Excavators – RSI Technologies

Syncrude operates a large oil-sand mine in northern Alberta and owns several of the world's largest mining excavators. The boom and bucket, which look somewhat like those of a standard construction excavator, are driven by by large hydraulic cylinders operating at up to 30MPa (4350psi).

Operators of these large machines would benefit from real-time knowedge of (1) the payload in the bucket, and (2) the digging force at the bucket teeth. In order to be a viable product, this information must be related to the operator without the need to stop the motion of the machine or do anything that would hinder the production. To be useful, the information should be available to the operator with at most 1s latency and should be accurate to within 5% of the payload. Furthermore, a commercial product would be easily applied to any excavator without the need for extensive modeling.

Previous approaches to this problem have followed a traditional robotics/dynamics approach. A dynamic model of the boom and cylinders was developed. Readings from hydraulic pressure sensors and joint angle sensors were fed through the model to determine the payload. However, there are serious limitations of this approach since it is impractical to develop a detailed dynamic model of every machine to which the product would be applied.

The company is interested to know of another approach which can produce the required performance without the need to develop a more detailed dynamic model.

Presented by:
David Lokhorst
RSI Technologies, Victoria, B.C.

Classification of Chemical Compound Pharmacophore Structures – Searle Corporation

There are a staggeringly large number of "druglike" chemical compounds which can be made, and the number of unique molecular structures may very well be in the order of 10^{100}. To date, only on the order of 10^7 compounds have been made and characterized.

The main hope in a classification is the existence of a similarity pattern in the chemical structure, called "pharmacophore". Compounds matching the same pharmacophore interact with biological molecules (enzymes, receptors, etc) in a similar fashion, and a classification of these pharmacophores would be considered an enormous breakthrough in drug discovery, providing access to a "key" to every "lock".

An immediate question is to estimate the number of these pharmacophores. Mathematically, this is a question about how an object with a small, fixed number of pharmacophoric features fills a cavity, but also an exploration of the minimum set of objects required to display all pharmacophoric elements in all possible geometric arrangements.

Participants interested in this problem should familiarize themselves with the "pharmacophore" idea. That concept as well as others are discussed in the following papers. A number of references in these papers may be of interest as well for followup:

- Pickett, S.D.; Mason, J.S.; McLay, I.M. "Diversity Profiling and Design Using 3D Pharmacophores: Pharmacophore-Derived Queries (PDQ)" J. Chem. Inf. Comput. Sci. (1996), 36, 1214-1223.
- Ashton, M.J.; Jaye, M.C.; Mason, J.S. "New Perspectives in Lead Generation II: Evaluating Molecular Diversity" Drug Discovery Today (1996), vol. 1, no. 2, 71-78.
- Greene, J.; Kahn, S.; Savoj, H.; Sprague, P.; Teig, S. "Chemical Function Queries for 3D Database Search" J. Chem. Inf. Comput. Sci. (1994), 34, 1297-1308.
- Clark, D.E.; Willett, P.; Kenny, P.W. "Pharmacophoric Pattern Matching in Files of Three-Dimensional Chemical Structures: Implementation of Flexible Searching" J. Mol. Graphics (1993), vol. 11, 146-156.
- Kuhl, F.S.; Crippen, G.M.; Friesen, D.K. "A Combinatorial Algorithm for Calculating Ligand Binding" J. Comput. Chem. (1984), vol. 5, no. 1, 24-34.

Also, there is a web page which nicely describes a commercial implementation of pharmacophore analysis:

http://www.oxmol.co.uk/prods/chem-x/phtech.html

Presented by:
Tom Doman
Searle Corporation, Chicago, Illinois

Chapter 2

Contaminant Transport in Municipal Water Systems

Seema Ali[1], Jeremy Bell[2], Chris Budd[3], Adriana Davidova[2], Nathan Krislock[4], Margaret Liang[5], Yves Lucet[6], Scott MacLachlan[5], Marni Mishna[6], Shelly Pinder[5], Peilin Shi[2], Cristina Stoica[2], JF Williams[6], Zili Wu[2]

Report prepared by John Stockie[6]

2.1 Introduction.

The quality of municipal water supplies has become an issue of great importance for large, urban centres. Widespread industrialisation, pollution, and development of areas close to traditional water sources have all contributed to the degradation of our drinking water, and made the topic a subject of heated debate. Officials and planners at all levels of government have recognised the need to manage water distribution networks so as to maintain acceptable levels of water quality. Because of the complexity of even the simplest of flow networks, numerical computations have become an essential tool in the water quality management process, and simulations of networks with hundreds of junctions, pumping stations, tanks and reservoirs are commonplace [4].

The forward or dynamic simulation of water networks with known characteristics and contaminant inputs is fairly well-understood, and software packages such as EPANET [5], which is freely-available over the Internet, are widely used in the industry. However, a more difficult aspect of water networks is the "inverse" or "forensic" problem where, for example, one wants to know how much of a contaminant has been released given measurements of contaminant levels at certain locations[7]. This is the problem that Charles Howard & Associates Ltd. (CHAL), a local Victoria engineering consulting firm, brought to the PIMS problem-solving workshop.

Our first task from CHAL was to examine the dynamic (forward) algorithm used in EPANET, and ensure that it is computing physically reasonable solutions for networks that experience rapid "flow reversals," in which the flow in a pipe section changes direction over a short time intervals. This is a particular concern for EPANET, because the calculation of water flow in the pipe network is totally decoupled from the contaminant transport step. The discrete time step used for the flow is typically much larger than for the contaminant, and so EPANET may have difficulties handling flow reversals. CHAL is also concerned about whether or not the inverse problem is well-posed, and if it is possible to use existing forward solvers to construct an algorithm for solving the inverse problem.

[1]University of Calgary
[2]University of Victoria
[3]University of Bath
[4]University of Regina
[5]University of British Columbia
[6]Simon Fraser University

[7]The importance of forensic water quality analysis has received widespread attention with the recent release of the Hollywood movie "Civil Action" starring John Travolta.

We begin in the following sections with an overview of the physics of flow in water distribution networks, and the implementation used in the EPANET flow solver. The bulk of our results from the workshop are concerned with analysing simple flow networks and using the results to draw conclusions about the well-posedness of both the forward and inverse problems.

2.2 Flow in Water Distribution Networks.

A typical water distribution network, pictured in Figure 2.1, can be considered as a collection of *links* and *nodes*. The links come in two varieties[8]: *pipes* and *pumps*. The pipes convey water passively from one network

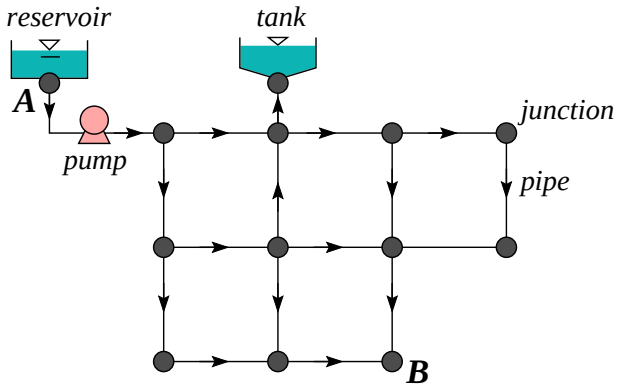


Figure 2.1: A typical water distribution network.

junction to another in the direction of decreasing pressure, except where forced by pumps which apply some external forcing to raise the hydraulic head of the water in the network. The nodes are either *junctions*, *tanks* or *reservoirs*. Junctions connect two pipe sections together and may also serve as points of water consumption (e.g., a residential or industrial customer) or external water sources (e.g., a water treatment plant). Tanks and reservoirs are external nodes that serve as water storage areas.

2.2.1 Network components.

Before moving on to a mathematical description of the physical processes in each of the network components, we need to define the basic variables that will be used to describe water transport in the network. Rather than the variables velocity and pressure that are typically used in fluid mechanics to describe flow, we will instead formulate the problem in terms of

- *hydraulic grade line* or *head* (H) which for 1D, steady, incompressible flow is given by Bernoulli's equation:

$$H = \frac{P}{\rho g} + \frac{V^2}{2g} + H^o,$$

  where P is the pressure, g is the acceleration due to gravity, V the fluid velocity, and H^o is the elevation;

- *volumetric flow rate* (Q) through a section of pipe, which is related to the velocity V by $Q = V \cdot A$, where A is the cross-sectional area of the pipe.

We now proceed to describe briefly the physics governing each of the network components implemented in EPANET, with complete details given in [5]:

[8]There are also several types of *valves* [5] which we will not consider here for the sake of simplicity.

Pipes: The flow in each pipe is assumed to be steady, uniform and characterised by fully-developed turbulence. As water flows along the pipe section joining node i to node j, there will be a loss in head, $\Delta H_{ij} = H_i - H_j$, due to friction between the water and the walls of the pipe. There are several semi-empirical equations used in hydraulics that relate this head loss to the flow rate, one of which is the *Hazen–Williams equation* [2, 8]

$$Q_{ij} = R_{ij} \frac{\Delta H_{ij}}{|\Delta H_{ij}|^{0.46}}, \tag{2.1}$$

where the constant R_{ij} (called the *conductance*) depends on the dimensions of the pipe and its frictional resistance. Notice from the form of (2.1) that Q_{ij} is always positive when the flow in pipe $i-j$ is directed from node i to node j (refer to Figure 2.2).

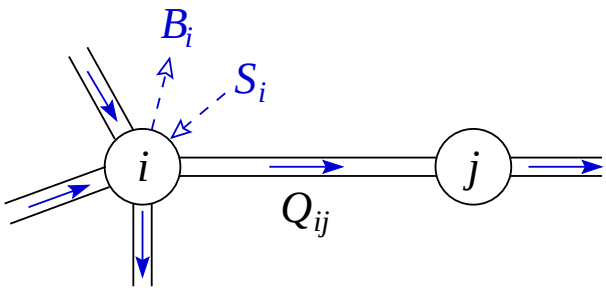


Figure 2.2: A typical pipe link in a network.

Pumps: Pumps are devices used to raise the hydraulic head of the water in the network. The flow rate in pump-driven links is related to the head gain by the following relationship:

$$Q_{ji} = \begin{cases} \left(\dfrac{H_{ij}^{off} - \Delta H_{ij}}{\alpha} \right)^{1/\beta} & \text{if } 0 \le \Delta H_{ij} \le H_{ij}^{off}, \\ 0 & \text{otherwise,} \end{cases} \tag{2.2}$$

where H_{ij}^{off} is the *shut-off head*, above which the pump is unable to operate. Notice that the pump only operates when the head at node i is greater than that at node j and hence always pumps water from node j to node i; i.e. pumps always work against the pressure gradient. The coefficients α and β are parameters specific to each pump that fit the flow rates to an operational *pump curve*.

Junctions: Pipe junctions may also be consumers or suppliers of water, with a given supply $S(t)$ or demand $B(t)$ flow rate, as pictured in Figure 2.2.

Tanks: A storage tank is a special type of node that has a free water surface where the head is simply given by the height above sea level. The time variation of the head H_s at a storage node connected by a single pipe to node s is described by the equation

$$\frac{\partial H_s}{\partial t} = \frac{Q_{is}}{A_s}, \tag{2.3}$$

where A_s is the cross-sectional area of the storage vessel.

Reservoirs: These are another type of storage node in which the water surface remains at the same level no matter what the flow rate is, and so the head is a constant. Reservoir nodes typically represent external sources of water such as lakes or rivers.

2.2.2 Hydraulic simulations.

The dynamic equations describing a flow network can be derived based on conservation principles. Suppose that a network contains N_j junctions, N_t tanks and N_r reservoirs. Within each node j (see Figure 2.2), the total of all flows into the node must equal that directed out of the node. Mathematically, this can be written as

$$S_j + \sum_{i=1}^{N} Q_{ij} = B_j, \tag{2.4}$$

where $N = N_j + N_t$ are the number of nodes with heads that vary (reservoirs are excluded here because they have a constant head). Since the flow rates satisfy $Q_{ij} = -Q_{ji}$, there is one additional conservation constraint that must be satisfied by the supply/demand input values; namely, that

$$\sum_{j=1}^{N} (S_j - B_j) = 0. \tag{2.5}$$

Based on the assumption that consumptions B_j, supplies S_j, and pump operations change slowly in time, we treat the system as one that can be modeled over time as a series of quasi–steady state simulations. By substituting into (2.4) the expressions for Q_{ij} from (2.1) and (2.2), and the discrete form $Q_{ij} = (H_j^{new} - H_j^{old}) \cdot (A_j/\Delta t)$ of the tank equation (2.3), we obtain a system of N nonlinear equations for the N unknown heads H_j at any given time.

2.2.3 Contaminant transport.

Imagine that a dissolved substance is introduced at a certain node (e.g., the reservoir node **A** in Figure 2.1), which is then transported through the network. The substance has no effect on the water flow rates we obtained in the hydraulic model and so its concentration can be determined independently.

Let $C_{ij}(x,t)$ represent the concentration of the substance in link $i-j$. While our assumption that the flow rate is constant along a given pipe link is reasonable, we cannot say the same for the contaminant, which must be taken as a function of time, t, and distance along the pipe, x. The equation of conservation for contaminant mass within each link then reads

$$\frac{\partial C_{ij}}{\partial t} + \frac{Q_{ij}}{A_{ij}} \frac{\partial C_{ij}}{\partial x} = -k_{ij}\, C_{ij}. \tag{2.6}$$

A_{ij} is the pipe's cross-sectional radius and k_{ij} is a rate constant that describes the decay of contaminant due to reaction with the pipe walls, and bulk reactions with the water[9].

The quantity (Q_{ij}/A_{ij}) represents a known velocity with which the substance is convected in the pipe and therefore, (2.6) is simply a linear advection equation. In order to solve this equation we need only determine the value $C_i^o(t)$ of the concentration at $x_{ij} = 0$, where we have used the convention that the coordinate x along each pipe is chosen so that the origin is at the upstream end (i.e., at node i if $Q_{ij} > 0$ and node j if $Q_{ij} < 0$). This situation is pictured in Figure 2.3.

The boundary values $C_i^o(t)$ can be found for each node i by assuming that the turbulent flow into the node is well-mixed and then applying the following conservation argument. We require that the total mass of contaminant transported into junction i

$$M_i + \sum_{k, Q_{ki} > 0} C_{ki}(L_{ki}, t)\, Q_{ki}$$

(where M_i is the mass inflow rate of contaminant and L_{ki} is the length of the pipe k–i) must be balanced by the total mass out

$$C_i^o \cdot \left[B_i + \sum_{k, Q_{ki} < 0} (-Q_{ki}) \right].$$

[9]The overall rate constant takes the form $k = k_b + 2k_w\, k_f/[R(k_w + k_f)]$ where k_b is the first-order bulk reaction rate constant, k_f is the mass transfer coefficient between wall and pipe, k_w is the reaction rate at the wall, and R is the pipe radius.

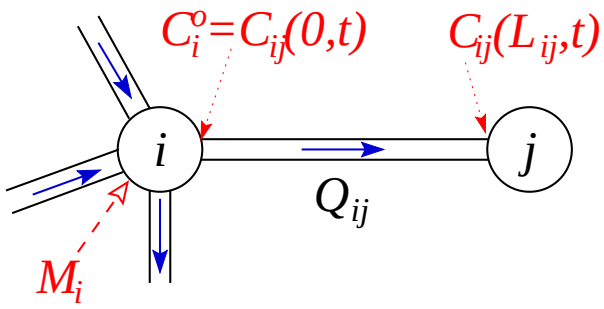


Figure 2.3: The boundary condition on the concentration in a link, $C_i^o(t)$, is obtained by combining the mass inflow of contaminant, M_i, and the mixing of inflows into node i.

Equating these two expressions and using (2.4), we can solve for the boundary value

$$C_i^o(t) = \frac{M_i + \sum_{k,Q_{ki}>0} C_{ki}(L_{ki},t)\,Q_{ki}}{S_i + \sum_{k,Q_{ki}>0} Q_{ki}} \tag{2.7}$$

Equation (2.6) can now be solved in each link using (2.7) as a boundary condition, and with the flow rate Q held constant over the time interval.

2.3 The EPANET solution algorithm.

The solution procedure within the EPANET water quality simulation package proceeds as follows in each hydraulic time step:

1. Calculate the supply and demand at each node, and the status of the pump links using user-specified time "patterns."
2. Solve the non-linear system of equations (2.1)–(2.4) for the head values H_i using a full Newton solver. The Jacobian matrix is constructed explicitly and the resulting linear system is solved using a Cholesky factorisation which takes advantage of the sparse structure of the matrix.
3. Perform the water quality calculations using the *Discrete Volume Element Method* (or DVEM), which proceeds as follows:
   (a) Select the water quality time step $\tau = \min_{ij}(L_{ij}/V_{ij})$, where $V_{ij} = Q_{ij}/A_{ij}$ is the flow velocity in pipe $i-j$. Consequently, τ is equal to the shortest travel time through all pipes in the system, which ensures that water is not transported beyond the downstream node in any link within a single time step.
   (b) Interpolate the concentrations from the discrete volume elements at the beginning of the hydraulic time step onto their current locations (since the position of the elements changes in each hydraulic time step).
   (c) Each pipe $i-j$ is then re-partitioned into η_{ij} discrete elements of equal volume, with the number of elements chosen so that

$$\eta_{ij} = \left\lfloor \frac{L_{ij}}{\tau\, V_{ij}} \right\rfloor .$$

   This ensures that the transport in all links is resolved adequately over the time step τ.

(d) Propagate the concentrations through the network using equation (2.6) and boundary condition (2.7). The water quality time step τ is typically much smaller than the hydraulic time step, and so this will require several integration steps.

Details of the Newton solver for the network flow are given in [8] and the DVEM contaminant transport algorithm is detailed in [6].

2.3.1 EPANET's limitations: reverse flow.

One of the major limitations of EPANET is the total decoupling of the hydraulic and water quality computations, so that the time steps used in each component need not coincide. The default hydraulic time step in EPANET is $\Delta t_h = 1\ hour$, whereas the water quality computations are typically done on a much finer time scale of $\Delta t_q \lessapprox 0.1$ h. Solving the network flow problem can be expensive because EPANET uses a full Newton solver, and so the choice of a larger Δt_h is justified in terms of efficiency. However, holding the flow rates constant over a large hydraulic time step can lead to inaccurate or even non-physical results. For example, if the actual flow in a pipe link reverses in the middle of a hydraulic time step, then the contaminant in that pipe may be transported in the wrong direction!

To illustrate the non-physical results that can occur when choosing a Δt_h that is too large, consider a pipe section of length 1 unit wherein the flow velocity varies periodically as shown in Figure 2.4(a). Suppose that a

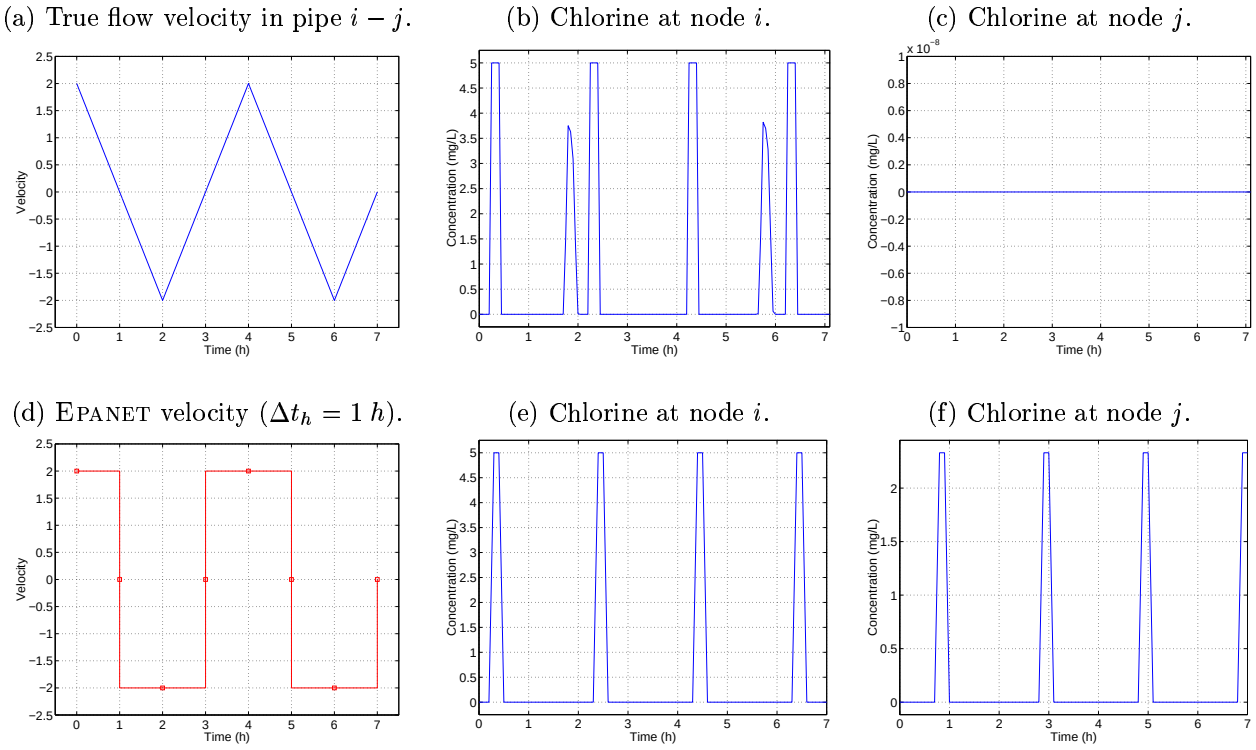


Figure 2.4: Computed chlorine concentrations at the upstream node for periodically-reversed flow: (a,b,c): "true" triangular velocity profile, computed with $\Delta t_h = 0.025\ h$; (d,e,f): typical EPANET calculation, with $\Delta t_h = 1\ h$.

pulse of contaminant is released at time $t = \frac{1}{4}\ h$ in the upstream node (i) of the pipe section $i-j$. The pulse will travel only $x = \frac{9}{16}$ of the distance down the pipe (which is simply the area under the velocity curve between $t = \frac{1}{4}$ and 1), at which time the pulse will reverse direction. Therefore, the pulse returns eventually to node i and never gets transported into the rest of the network through node j.

If this same problem were to be computed using EPANET with a hydraulic time step of 1 h, then the velocity variation would be represented as a step function, as pictured in Figure 2.4(d). A pulse of contaminant released

at time $t = \frac{1}{4}$ would reach the downstream node at time $t = \frac{3}{4}$ and be transported into the rest of the network before the pipe flow experiences a reversal. Consequently, taking a hydraulic time step that is too large to resolve variations in the water flow or contaminant transport can lead to incorrect water quality solutions.

To see that this is in fact what happens in EPANET, Figures 2.4(a,b,c) depict a simulation of the "true flow" conditions with a triangular velocity profile and a small hydraulic time step. The upstream node (i) is given periodic pulses of contaminant, and the return pulse is clearly superimposed on the input pulse. The downstream node (j) remains at a contaminant level of zero throughout the simulation. Figures 2.4(d,e,f) on the other hand, show the results for a hydraulic time step of 1 hour corresponding to the step function velocity profile. Contaminant is released at the same frequency, and the pulses are clearly seen to exit the loop via the downstream node before the flow reverses.

There is no reason in principle that Δt_h cannot be chosen much smaller so that it is equal to Δt_q in order to minimise this phase error due to flow reversal. However, the calculations become much more expensive. Not only this, the format of the input files requires that the user provide a "pattern" for the time variation of nodal supply and demand and pump operations, which must be given at each hydraulic time step (i.e., EPANET does no interpolation). In reality, demand patterns are known only at 1– or 2–hour intervals, and so manual interpolation is required if a smaller hydraulic time step is to be taken.

Nonetheless there is a practical limit to how small the time step can be taken. Within the EPANET algorithm is embedded a Newton solver that inverts a full Jacobian matrix in each iteration. This is not much of a concern when the hydraulic time step is on the order of hours, but could be prohibitively expensive if the time step is taken small enough to accurately resolve frequent flow reversals. An obvious alternative is the class of *quasi–Newton* methods, which never explicitly construct the Jacobian matrix, and solve the system using iterative techniques rather than direct methods. Many public–domain codes are available that could be incorporated in the EPANET software, one of which is DNSQE[10]. We did not have the time during the Workshop to implement a non-linear solver in EPANET, but we do believe that the modifications required to the code for such an addition are straightforward.

2.4 Analysis of Simple Networks.

We now proceed to analyse some very simple distribution networks, in order to make conclusions about the solvability of the forward and inverse problems for contaminant transport. We assume for the remainder that we have a given set of flow rates Q_{ij} for the links in the network.

2.4.1 Exact solution for contaminant flow

We consider the simplest possible network consisting of two junctions connected by a single pipe. Letting $V(t) = Q(t)/A$ denote the flow velocity, we are reduced to solving the equation

$$\begin{aligned} C_t + V(t)\, C_x &= -k\, C, \\ C(0,t) &= C^o(t) \qquad \text{(given)}, \end{aligned}$$

where the boundary condition is given at the upstream node. If we define

$$\xi(t) = \int_0^t V(\tau)\, d\tau$$

to be the distance traveled by fluid particles from time 0 to t, then the solution to this problem can be written explicitly as

$$C(x,t) = C^o(\xi^{-1}(\xi - x))\, e^{-k(t-\xi^{-1}(\xi-x))}. \tag{2.8}$$

[10]DNSQE is part of the SLATEC library, which is freely available on the Internet through Netlib (`http://www.netlib.org/bib/gams.html`).

A very important special case is the situation of constant velocity, $V(t) \equiv V_o$, for which

$$C(x,t) = C^o(\underbrace{t - x/V_o}_{delay}) \quad \underbrace{e^{-kx/V_o}}_{attenuation} . \tag{2.9}$$

From this form of the solution, it is clear that flow in the pipe is a combination of *delay* (due to the finite propagation velocity of contaminant in the system) and *attenuation* (from the decay of contaminant due to reactions at the pipe wall and in the bulk flow).

An important consequence of the form of (2.9) is that even though the transport of contaminant is a linear process (i.e., $C(x,t) \propto C^o$), it is *not* a convolution (because of the exponential decay term). Therefore, we expect that there may be some difficulties with the solution of the inverse problem.

Nevertheless, let us continue on and investigate the solution of the inverse problem for a simple "loop" network, pictured in Figure 2.5. Two nodes, A and B, are connected by a pair of pipes with different lengths (L_1

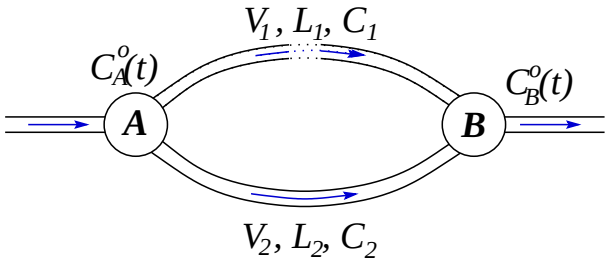


Figure 2.5: A simple reduced network consisting of two nodes connected in a loop, with both flows in the same direction. The dotted lines in pipe 1 represent the fact that there cannot be two pipes joining a single pair of junctions in EPANET; rather there must be some collection of other nodes and links in between for which we take "effective" variables V_1, L_1 and C_1.

and L_2) and flow velocities (V_1 and V_2). Supposing that contaminant enters node A with a variation given by $C^{in} = C^o_A(t)$, we begin by determining how the outflow concentration C^o_B measured at node B depends on C^o_A. Assuming the velocities are positive and constant, equation (2.9) gives us concentrations of the contaminant at the end of the pipe section:

$$C_1(L_1,t) = C^o_A(t - L_1/V_1)\, e^{-kL_1/V_1},$$
$$C_2(L_2,t) = C^o_A(t - L_2/V_2)\, e^{-kL_2/V_2}.$$

Once the inflows from the two pipes meet in node B and mix, the outflow from the node is given by

$$\begin{aligned} C^o_B &= \frac{V_1C_1 + V_2C_2}{V_1 + V_2} \\ &= \frac{V_1C^o_A(t - L_1/V_1)\, e^{-kL_1/V_1} + V_2C^o_A(t - L_2/V_2)\, e^{-kL_2/V_2}}{V_1 + V_2}. \end{aligned} \tag{2.10}$$

2.4.2 Ill-posedness of inverse problem for simple loops.

We will now use the exact solution (2.10) to answer the following questions:

> *Is the inverse problem well-posed? Or, in other words, is there a unique inflow of contaminant $C^o_A(t)$ corresponding to each measured outflow $C^o_B(t)$?*

In fact, we will show that a constant outflow $C^o_B = 0$ can be generated by more than one input for a given set of network parameter and flow rates.

Consider the single-loop network in Figure 2.5, and assume that $V_2 > V_1 > 0$. To simplify the situation, we look for input concentrations of the form $C_A^o(t) = \overline{C}\, e^{\theta t}$, where θ is a complex number. Substituting this expression into (2.10) and setting $C_B^o = 0$, we obtain

$$e^{\delta\theta} = -\frac{V_2}{V_1}\, e^{-\delta k},$$

where $\delta = (\frac{L_2}{V_2} - \frac{L_1}{V_1})$ is the difference in time it takes for the concentration pulse to travel through pipe 2 versus pipe 1. Taking the logarithm of this expression, we can solve for θ:

$$\theta = -k + \frac{1}{\delta}\log\left(\frac{V_2}{V_1}\right) + \frac{(2n+1)\pi}{\delta}\sqrt{-1}, \tag{2.11}$$

where n is any integer. Therefore, the inflows we are interested in take the form

$$C_A^o(t) = P(t)\, e^{[-k+\log(V_2/V_1)/\delta]\, t}, \tag{2.12}$$

where $P(t)$ is *any* periodic function with period $2\pi/\delta$.

Consequently, there are an *infinite* number of possible inflows for which the measured concentration at node B will be zero (or constant). A second such inflow is pictured in Figure 2.6, where it is clear that after an initial transient, the measured concentration is identically constant, even though the contaminant continues to vary at node A.

(a) Contaminant input in node A. (b) Contaminant measured at node B.

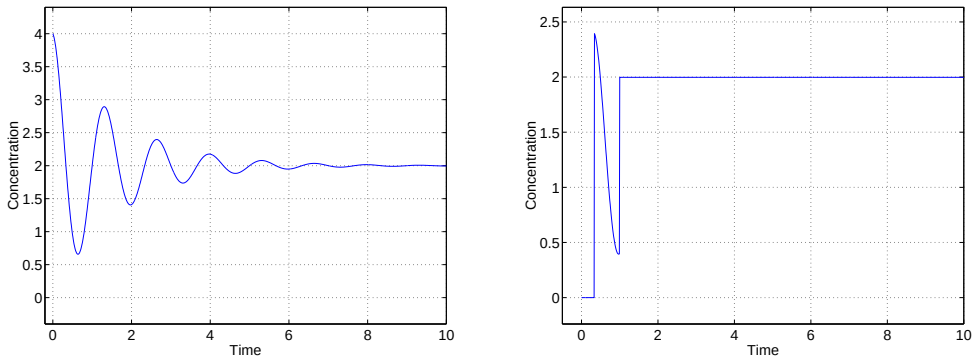


Figure 2.6: An exact solution of (2.10) with a decaying sinusoidal input concentration, and $L_1 = 10$, $V_1 = 10$, $L_2 = 5$, $V_2 = 15$, $k = 0.002$ (so that $\delta = -2/3$ and $\theta \approx -0.610198$).

In real water distribution networks, the sources do not decay in time in an exponential fashion. Instead, it is much more typical for sources to vary periodically *without* decay, corresponding, for example, to 24–hour cycles of demand and supply. In order for the solution in (2.12) to be undamped, the term in the exponent must be zero, so that

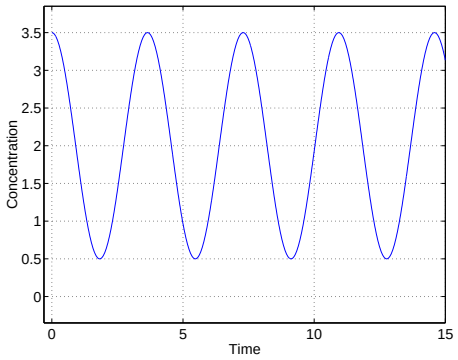
$$\delta = \frac{1}{k}\log\left(\frac{V_2}{V_1}\right),$$

which can also be written as a restriction on network parameters

$$L_2 = V_2\left[\frac{L_1}{V_1} + \frac{1}{k}\log\left(\frac{V_2}{V_1}\right)\right].$$

Figure 2.7 presents an undamped solution for another network which again clearly exhibits a constant concentration at node B after the initial transients have died out.

(a) Contaminant input at node A.



(b) Contaminant measured at node B.

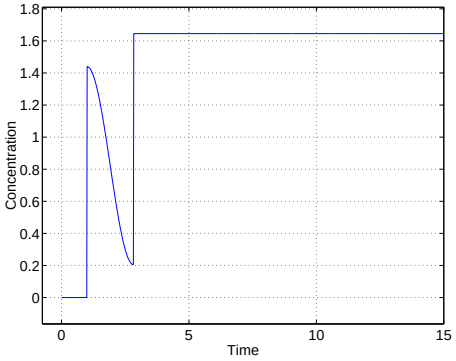


Figure 2.7: An exact solution of (2.10) for an undamped periodic input, with $L_1 = 10$, $V_1 = 10$, $L_2 \approx 33.8786$, $V_2 = 12$, $k = 0.1$ ($\delta \approx 1.823216$, $\theta = 0.0$).

All of our efforts in this section have so far been theoretical and so it is worthwhile for us to verify that the same type of behaviour is observed in actual EPANET computations. We constructed a simple 3-node network consisting of a loop, as pictured in Figure 2.8(a) (the extra junction is needed for the lower pipe since EPANET has a restriction that only one pipe can connect any two junctions). We applied a periodic, square-wave input of contaminant to node A, as shown in Figure 2.8(b), and then measured the concentration at node B. The measured chlorine levels at node B in Figure 2.8(c) again show that the contaminant reaches a constant level after the initial transients have died out.

2.4.3 Ill-posedness of forward problem.

We will now show that even the forward contaminant problem can be ill-posed. For this purpose, we will consider a similar loop network to that just considered, except that the velocity $V_1 < 0$ so that there is a *return loop*. The situation is pictured in Figure 2.9. The important distinction here is that we are not attempting to find C^{in}, the concentration flowing into node A. Instead, C^{in} is a given function.

Using the same mixing arguments from the previous section, the concentration of the boundary value at the head of pipe 2, C_A^o, is a linear combination of C^{in} and the inflow from pipe 1:

$$C_A^o(t) = aC^{in}(t) + bC_1(L_1, t),$$

(where a and b are constants depending on the flow velocities). However, if we apply mixing to the inflows at node B, then $C_1(L_1, t)$ is an attenuated and time-shifted version of $C_A^o(t)$:

$$C_1(L_1, t) = C_A^o(t - \delta)\, e^{-k\delta},$$

with $\delta = \frac{L_2}{V_2} + \frac{L_1}{V_1}$ (notice the change in sign from the expression for δ in Section2.4.3!). If we then assume $C_A^o = \overline{C}e^{\theta t}$ (and take $C^{in} = 0$ for simplicity), then we can solve these equations as before to obtain

$$C_A^o(t) = P(t)\, e^{(-k + \log b/\delta)t},$$

where $P(t)$ is any periodic function with period $\frac{2\pi}{\delta}$. Physically, this behaviour can be interpreted in terms of a given contaminant pulse generating *echoes* or *ringing* in networks containing a return loop.

Consequently, for a *given* contaminant inflow $C^{in}(t)$, there can be more than one outflow C_A^o into pipe 2. This suggests that the forward problem is also ill-posed. Whether or not such a situation is actually encountered in practice is not clear. However, at the very least it does suggest that we proceed with caution when attempting to compute solutions to the network flow problem.

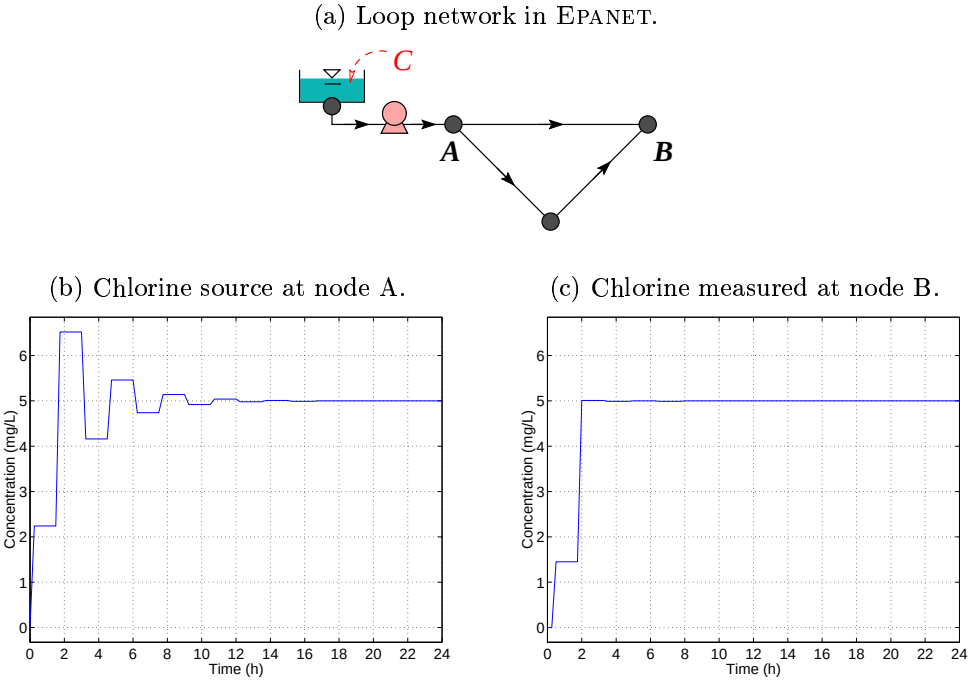


Figure 2.8: An EPANET computation with a decaying, periodic, square wave input that produces a constant output at node B.

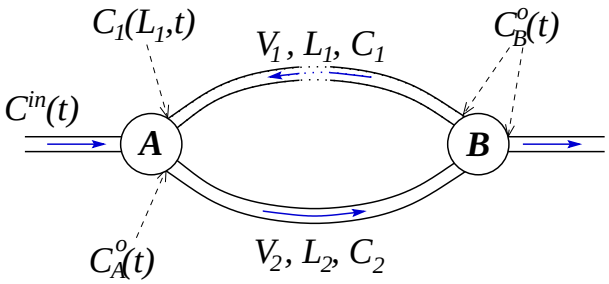


Figure 2.9: A second network with a "feedback" loop.

2.4.4 Solving the Inverse Problem.

There are many possible approaches to solving the inverse water quality problem, one of which is described in [1]. During the workshop, we were not able to implement any inverse solution algorithms or to investigate in detail the merits of the possible approaches. However, with an efficient forward solver in hand, there are many public domain software packages available for non-linear optimisation that use the forward solution as a "black box" in solving the inverse problem. One package which would be worth considering is DASOPT [3], or its successor COOPT [7], which are based on the well-known nonlinear differential-algebraic system solver DASSL. Provided that a forward time step of the water quality calculation could be incorporated into this package, then the contaminant concentrations or positions could be treated as free parameters and optimised using these algorithms.

2.5 Summary and Future Directions.

Our main conclusion is that the EPANET code is calculating the correct solution, provided that the hydraulic time step is taken small enough to resolve flow reversals. As a result, we make the following suggestions:

- First, the hydraulic time step and the water quality time step should be explicitly chosen equal to each other, instead of letting EPANET default to a hydraulic step of one hour.
- Second, patterns for the time variation of consumption, supply and pump/valve operations must be interpolated onto the hydraulic time steps from known (measured) data. Two options for dealing with this are:
  - preprocessing the input data files to interpolate the known patterns onto a denser series of time points. This requires no modifications to EPANET and can easily be made transparent to the user.
  - incorporating an interpolation algorithm into EPANET, which is much more problematic since one must deal with the internal data structures used to store the flow information.
- Because the flow equations are now being solved at a much larger number of time steps, faster algorithms are required for solving the nonlinear system. Our suggestion is to use a quasi-Newton algorithm, such as DNSQE.
- We have found concrete examples where the inverse problem for contaminant transport is ill-posed when there are "loops" in the network, and so it may be difficult or impossible to solve in practice. Nevertheless, we have no evidence that this is true in general and so it may still be worthwhile to pursue the implementation of an inverse solver. Our advice: *Proceed with caution!*
- Once an efficient forward solution algorithm is working, there are many free software packages available on the Internet for solving the inverse problem, in which the forward solver is used as a simple "black box."

Bibliography

[1] Timothy R. Ginn and Ellyn M. Murphy. A transient flux model for convective infiltration: Forward and inverse solutions for chloride mass balance studies. *Water Resources Research*, 33(9):2065–2079, 1997.

[2] Horace W. King, Chester O. Wisler, and James G. Woodburn. *Hydraulics*. John Wiley & Sons, New York, 5th edition, 1948.

[3] Linda Petzold, J. Ben Rosen, Philip E. Gill, Laurent O. Jay and Kihong Park. Numerical optimal control of parabolic PDEs using DASOPT, in *Proceedings of the IMA Workshop on Large-Scale Optimization*, 1996. Available at `http://www.engineering.ucsb.edu/~cse/dasopt.ps`.

[4] L. S. Reddy, Lindell E. Ormsbee, and Don J. Wood. Time-averaging water quality assessment. *Journal of the American Water Works Association*, pages 64–73, July 1995.

[5] Lewis A. Rossman. *EPANET Users Manual (Version 1.1)*. Risk Reduction Engineering Laboratory, U.S. Environmental Protection Agency, Cincinnati, Ohio, January 1994. Code and documentation are available from `http://www.epa.gov/docs/RREL/`.

[6] Lewis A. Rossman, Paul F. Boulos, and Tom Altman. Discrete volume-element method for network water-quality models. *Journal of Water Resources Planning and Management*, 119(5):505–517, 1993.

[7] Radu Serban. COOPT – Control and optimization of dynamic systems: User's guide. Technical Report UCSB-ME-99-1, University of California at Santa Barbara, 1999. Available at `http://www.engineering.ucsb.edu/~cse/coopt.ps`.

[8] Uri Shamir and Charles D. D. Howard. Water distribution systems analysis. *Journal of the Hydraulics Division, Proceedings of the American Society of Civil Engineers*, 94(HY1):219–234, January 1968.

Chapter 3

Quality Control for Multi-variable Problems

Eric Agyekum[2], Andreea Amariei[3], Brenda Caughlin[7], Jie Cheng[3], Ella Huszti[3], Veselin Jungic[4], Emmanuel Ngembo[5], Grace So[6], Paul Wiebe[1],

Report prepared by Luz Palacios[1]

Introduction, Discussion and Editing by Rita Aggarwala [1]

3.1 Introduction

Chemex Laboratories measures concentrations of various minerals in land samples provided by their clients. The problem presented by Chemex is that they would like to be able to detect when their measurement apparatus needs re-calibration due to a shift in process parameters, so as to minimize rework.

Currently, calibration of the measuring machine is monitored by measuring mineral concentrations in control samples about every 40 minutes or 20 samples. Chemex would ideally like to make a decision about the calibration at every control sample assessment, that is, in real time, with a low error rate. Therefore the tool implemented must be quite simple, powerful and readily interpretable. This is complicated by the fact that multiple dependent measurements are made on each control sample, and only one set measurements is made before the decision is made. At present, the re-calibration decision is made using traditional, univariate control charts for a few key mineral measurements.

Briefly, the measurement process can be described as follows: Some compound, whose composition is of interest is prepared and is passed to a machine, the ICP, which sprays it through a plasma flame exciting the atoms. Photons of characteristic wavelengths are emitted as the atoms return to their ground states. These are counted to determine the concentrations of the atoms present in the compound.

There are many sources of error in this process but, the three with the largest contributions are: the sample is susceptible to contamination at the preparation stage, the spray nozzles often get plugged and the plasma flame may not heat to precisely the correct temperature.

Each sample contains more than 20 elements whose atom concentrations are to be measured. In consultation with the industrial mentor, 16 of these elements were identified to be of interest for this problem: Ag, Al, As, Ca, Co, Cr, Cu, Fe, Mg, Mn, Ni, P, Pb, Sr, V and Zn.

[1]University of Calgary
[2]University of Victoria
[3]University of Alberta
[4]Simon Fraser University
[5]University of Montreal
[6]University of Toronto
[7]Chemex Laboratories

Chemex has considered univariate quality control analyses, that is, they have considered errors in measurement of each individual characteristic one at a time, but have not considered all of the characteristics together, thereby potentially losing information about dependencies between mineral measurements and inflating the chances of false alarms. In addition, the large number of measurements involved renders this a cumbersome procedure. The decision rule used by Chemex at present is that recalibration is performed if more than 3 of the 16 mineral concentrations for the control sample are outside the traditional univariate Shewhart "two-sigma control limits."

Different multivariate approaches are presented here for the 16 elements of interest, considering the characteristics together and taking advantage of computing capability available on even the most basic systems in order to arrive at effective decision rules which accurately and precisely flag truly out-of-control measurements on the process. Three very contrasting approaches to the problem were taken during the PIMS 1999 Industrial Problem Solving Workshop. They are: considering the process as a Markov Chain; use of Bayesian Belief Networks; and Multivariate control chart implementation. Preliminary data analysis for data provided by Chemex was also performed in order to initially assess some of the models considered.

Approaches

3.2 The Markov Chain Model

Paul Wiebe and Eric Agyekum

3.2.1 Introduction

The Chemex Labs multivariate quality control problem is to decide the 'best' rules for controlling their process of determining atomic concentrations. However before a quality tool can be implemented the mathematical foundations of the process of interest must be uncovered and compared with the mathematical assumptions associated with the tool. The suitability of a vector Markov Chain model for the Chemex measurement process is studied in this context. It is a good model for a simulator which is easy to implement and on which potential control rules can be tested without tying up the machine.

3.2.2 The Model

The physics of the interactions in the plasma flame are not well described and the other two sources of error (contamination at the preparation stage and spray nozzle plugging) seem to be purely stochastic in nature except that the plugging should effect a trend. This led us to a Markov chain description.

Let,

$$X : \Omega \times N \to S$$

be the discrete–time process

$$X := (X^{\mathtt{Pb}}, X^{\mathtt{Au}}, \cdots, X^{\mathtt{Zn}})$$

with state space

$$S := S^{\mathtt{Pb}} \times S^{\mathtt{Au}} \times \cdots \times S^{\mathtt{An}}.$$

The decision to 'exert control', or the **control rule**, on the ICP is modelled as a stopping time (with respect to the filtration generated by X),

$$T : \Omega \to N.$$

This is a function on the sample paths of the process X into discrete time: each sample path is associated to a time T, the time to 'exert control'. It may be argued that T is a stopping time and that X has the Markov property.

Let $\sigma(X(\cdot,1),\cdots,X(\cdot,n))$ be the element of the filtration generated by X up to time n.

Since it is impossible to exercise a control rule which depends on future states of the process (if it were then this problem would not exist), it must be the case that $1_{\{T\leq n\}} \notin \cup_{m>n}\sigma(X_1,\cdots,X_m)$. Hence a stopping time is a good model for the control rule.

The Markov property,

$$P(X(\cdot,n+1)\in B|\sigma(X(\cdot,1),\cdots,X(\cdot,n)) = P(X(\cdot,n+1)\in B|\sigma(X(\cdot,n))$$

where $B \in \sigma(S)$, is interpreted as follows: to determine the probable state of the process at time $n+1$ the information about the complete history of the process is no more (or less) useful than the information about its current state (at time n). So it is argued that the transition from one state to another (possibly the same state) in one time step does not depend on states other than the current state. It is conceivable that the transition does depend on previous states; for example if the nozzle is plugged there may be a time after which it becomes unplugged. If this time is short than we can ignore the plugging. In the event that this isn't the case, plugging will arrest the process. In any event the Markov property remains intact. Such arguments can be prolonged *ad infinitum* but the real reason for considering the Markov chain model is its well–developed theory, and the fact that Markov chain models are often useful in approximating physical processes.

Immediately, two aspects of Markov chain theory are useful: the transition probability matrix and the chain's equilibrium distribution. The transition probability matrix describes the dynamics of the process. For any given current state the matrix gives the conditional distribution of advancing to 'next' states. This is important because it is information on the future of the process. For example, if the conditional distribution is 'nearly degenerate' then it can be acted upon with confidence. The equilibrium distribution of the Markov chain is the long-term proportion of time each element is in each state. So, if an element has far exceeded its expected time in a certain state, the operator may suspect that a shift has occurred in the process parameters.

In the multivariate case, if a subset of elements are correlated, these should be grouped. This means, at worst, taking the Cartesian product of their respective state spaces; if the structure of the multi–correlation can be determined then a smaller state space is feasible. Overall we would like to have a partition of the full set of elements into statistically unrelated groups whose dynamics can be determined independently.

The Markov Chain model is also an excellent candidate for simulation. One only needs to determine the transition matrices of each of the groups of related elements. These can be estimated most simply (but there are probably more elegant and robust ways) by counting the state transitions. To simulate, begin the process in a known state (the expected state); the transition matrix gives the probability distribution of 'next' states given the current one, so draw from it. Now the process is in another known state and the transition matrix gives the (new) distribution of 'next' states, so draw from it, etc. Based on the history of the simulated process, control rules can be tested and the future states of the chain compared with other future states based on different control rules.

3.2.3 Postscript

A Martingale model was also considered but discarded since the Martingale property implies that the expectation of the process at any time is the same as it was in the beginning. In fact, this appears to be a part of the notion of a process being 'in control'. So it seems initially that a control rule monitoring the mean of a Martingale should never demand that control be exerted. A similar statement could be made about the (time) cross–sectional variance of the process: as soon as it ceases to be constant (probabilistically) but starts to expand (perhaps like a Brownian motion) current Shewhart rules object and demand that control be exerted. This idea brings to mind the notion of an **ergodic theorem**.

An ergodic theorem is a tool which tells the practitioner that statistical properties of a process, under some stationarity [8] condition, usually can be approximated by the statistics of a single sample path. It seems that control rules are dictated by ergodic theorems, or in particular by their contrapositives. That is, given an ergodic theorem on some statistic `S` of the process, as soon as it can be determined with some certainty that the estimate of `S` based on the sample path is deviating then it is time to exert control.

[8] A process Y is called **stationary** if its finite dimensional distributions are invariant under shifts of time ; it is called **weakly stationary** if its mean and autocorrelation are constant in time.

3.3 Using Bayesian Belief Networks for Quality Control

Jie Cheng

The Bayesian belief network (BN) is powerful tool for knowledge representation and reasoning under conditions of uncertainty. BNs have been used successfully in fault diagnosis, decision support and classification. Unlike rule-based knowledge base systems, the BN is based on the solid foundation of probability theory and graph theory - it can express uncertainties in a natural way. Unlike a neural network, which is often interpreted as a "black box", the nodes in a BN are domain variables and the links are causal connections among the variables. Therefore, human experts can create a BN by hand. In the case that the BN is developed from training data, human experts can easily understand it and make modifications. A BN usually requires less training than a neural network does. Evidence shows that BNs often outperform other tools like decision trees, neural networks and statistical programs in classification tasks. For each node in a BN, there is an associated probability table that specifies the probability distribution of the variable given its parents.

For the Chemex problem, we can create a group of BNs, one for each particular physical problem of the instrument. After getting the readings from the instrument as input, each BN will give the probability that the physical problem occurs. (This computation will take less than 0.01 second on an average PC.) Then we can set (and fine tune) a rule that will trigger the alarm when a probability appears to be too high.

There are three ways to get these BNs.

1. Learn both the structure and the parameters (probability tables) of the BNs from data.
2. Let human experts specify the structure of BNs and determine the parameters from data.
3. Let human experts specify both the structure and the parameters of the BNs.

Based on the information from the workshop, we can probably construct the BNs by hand, i.e., the diagnostic (problem) node is the parent of all elements. For example, in the power failure BN, the power supply node is the parent of all elements and the elements are independent given the status of the power supply. For the power supply node, the probability might be P(failure)=0.05, P(normal)=0.95. For an element node, the probability might be:

$$
\begin{aligned}
P(-s < Ag < s \mid failure) &= 0.2 \\
P(-2s < Ag < 2s \ and \ s > Ag > -s \mid failure) &= 0.3 \\
P(Ag < -3s \ or \ Ag > 3s \mid failure) &= 0.5
\end{aligned}
$$

$$
\begin{aligned}
P(-s < Ag < s \mid normal) &= 0.8 \\
P(-2s < Ag < 2s \ and \ s > Ag > -s \mid normal) &= 0.15 \\
P(Ag < -3s \ or \ Ag > 3s \mid normal) &= 0.05
\end{aligned}
$$

(Strictly speaking, this way of constructing BNs assumes that each problem occurs independently and the problems do not occur at the same time. If the assumption is unacceptable, we need a more complex BN structure instead of a group of simple BNs.)

If such probability distributions cannot be given, data collection and experiments will be necessary. The data can be collected through everyday operation by recording the readings of the normal condition and the readings when something is wrong. (The diagnostic results should also be recorded.) Alternatively, the data can be collected by doing experiments - create a physical problem intentionally and record the readings.

3.4 Multivariate Control Charts

Chemex provided data containing 109 'in-control' data points (that is, sample measurements which appeared to be within control limits as determined by Chemex's current rules), estimates of population mean concentrations of the various substances of interest when the process is operating within natural variation, and 348 'out-of-control' samples (that is, measurements which included both in-control and out-of-control points as per Chemex's current rules). It should be noted that good, valid data collection, although not the focus of this report, is another issue

which should be addressed before implementing any control tool. The data provided by Chemex were used in order to assess two multivariate algorithms, which are discussed in the following two subsections. The final subsection on multidimensional Shewhart control charts explores the mathematical generalization of univariate control rules to the multidimensional case. The use of two or more of the tools presented in this section can potentially provide a very strong, practical tool for both pinpointing when and assessing why the process of interest appears to be out of control.

3.4.1 Principal Component Analysis

Grace So

The key idea of principle components is dimension reduction. Our goal is to reduce the dimension of the problem to be smaller than 16, with enough explanation of variability.

The correlation matrix of the in-control data was used to carry out the principal components analysis.

The first 8 eigenvectors were chosen, which explain about 85% of the variation.

1. A matrix U consisting of the eigenvectors $u_1, u_2, \cdots, u_8$ was formed. Also, $U'SU = L$ was obtained, where L is an 8×8 diagonal matrix consisting of the eigenvalues $l_1, l_2, \cdots, l_8$ of S. The principle components of x can be formed by the following transformation:

$$z = U'(x - \bar{x}) \tag{3.1}$$

2. We use the Hotelling T^2 for the (8 dimensional) principal components to determine whether each observation is out of control. If a particular observation is out of control, we should be able to check which of the 8 Principal Components caused the problem in further analysis. We are now ready to calculate the Hotelling T^2. Before doing this, the principle components were required to be scaled. The formula used for scaling the principle components is

$$w_i = \frac{u_i}{\sqrt{(l_i)}} \tag{3.2}$$

such that a scaled matrix, W was formed and

$$W'W = L^{-1} \tag{3.3}$$

The Hotelling T^2 is defined as

$$T^2 = z'L^{-1}z \tag{3.4}$$

The 109 values of T^2 were obtained from the in-control data. The 95% quantile of the distribution of the T^2 statistic can be easily found, since[9]

$$T^2 \sim \frac{p(n-1)}{(n-p)} F_{p,n-p} \tag{3.5}$$

In this example, $p = 8,\ n = 109$, and $\alpha = .05$. $F_{8,109,0.05} = 2.0314$ so $\frac{8(109-1)}{(109-8)} F_{8,109,0.05} = 17.3774$. *All* of the 109 observed T^2 values were greater than this value, indicating a problem either in the data or in the application of the methodology. As the problem was further pondered, the difficulty of interpretation was discussed, since the individual entries of the main principal components were similar. This method was therefore discarded as the best approach this problem, and another approach was considered, as presented in the following subsection.

[9] Assumptions of multivariate normality and independence between samples must be checked and verified.

3.4.2 Confidence Region and Simultaneous Bounds

Luz Palacios and Emmanuel Ngembo

Since the dimension of the problem could not be reduced efficiently through principal component analysis, further analyses were done with the 16 characteristics, using Confidence Regions and Simultaneous Bounds. [10]

In order to compare the assessment of which points are classified as in-control by Chemex's current rules and the Hotelling multivariate approach considered here, Hotelling's T^2 for the original (16-dimensional) data was calculated for each of the 109 Chemex in-control samples, using the means and covariance matrix from this data.

$$T^2 = (\mathbf{X} - \mu)^T S^{-1} (\mathbf{X} - \mu)$$

The upper control limit is

$$\chi^2_{16} = 26.296.$$

The lower control limit is 0.

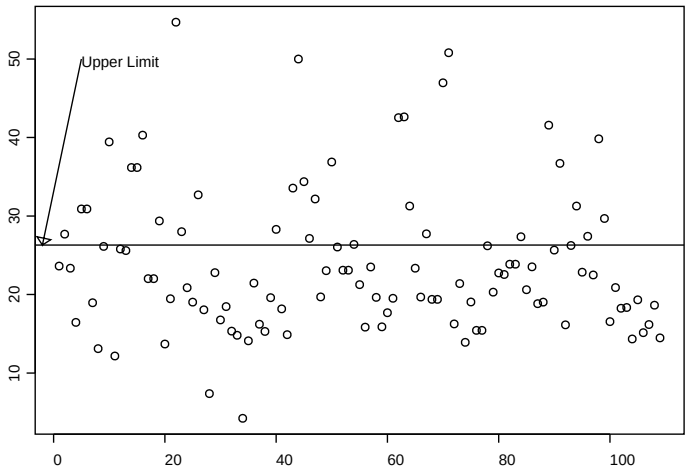


Figure 3.1: T^2 in-control data

As seen on Figure 3.1, many points fall outside of the limits. This is due to the fact that Chemex used a different method for determining in-control samples.

Hotelling's T^2 was also calculated for each of the out-of-control samples, using the mean and covariance matrix mentioned above from the in-control data. The upper and lower limits are the same as before.

As seen on Figure 3.2, many samples fall outside the limit.

At this point, we would like to be able to determine the cause (particular element measurements) contributing to the process falling outside control limits, if the process parameters have actually shifted. This would be done for each point falling outside the control limits.

In order to illustrate the rest of the analysis, sample 18 was chosen as an example. The T^2 value for this sample is approximately 80, which is greater than 26.296, which means that a problem might exist in this sample, so simultaneous bounds for each characteristic were calculated. These are based on Bonferroni Simultaneous Confidence Intervals, and have an overall confidence level of about 100(1-α)% provided associated assumptions are met. (Here, p is the dimension of the problem, in this case, 16.)

[10] Assumptions of multivariate normality and independence between samples must be checked and verified.

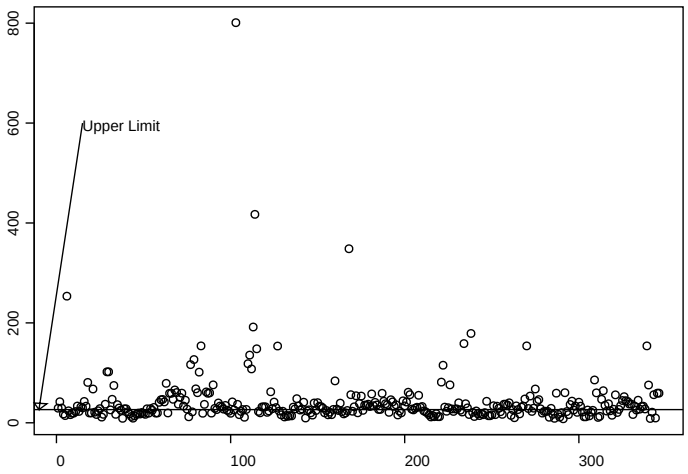


Figure 3.2: T^2 out-control data

$$\mu_i \pm Z_{\frac{\alpha}{2p}}\sigma_i \qquad i = 1, ..., 16$$

A table is presented showing the Silver (Ag) and Manganese (Mn) characteristics using $\alpha = 0.05$.

Characteristic	Upper limit	Lower limit	Observation
Ag	5.877	2.922	6.8
Mn	1,079	775	1,020

Figure 3.3 and Figure 3.4 are the graphs for Silver and Manganese.

The value for Silver falls outside of the limits, so we can say that a problem might exist in this characteristic. The value for Manganese does not fall outside of the limits, which means there is no signal for saying that a problem exists in this characteristic.

These analyses can be done with each sample associated with a point falling outside the control limits, and each characteristic, while maintaining the overall confidence level for each point.

3.4.3 Multidimensional Shewhart Control Charts

Veselin Jungic, Ella Huszti, and Andreea Amariei

This note is inspired by the fact that Shewhart control charts work well in the univariate (one characteristic) case. The effectiveness of Shewhart's idea in years of practice and its well defined and simple constraints give another reason for this exercise. Also, a motive for this note is that one can easily imagine a situation that if an instrument that measures n characteristics simultaneously fails, then it rarely does so for only one characteristic.

Over the years, the more rules for out-of-control detection have been implemented in analyzing control charts. For example, in addition to the decision rule based on a point falling outside the 3-sigma control limits, the following rules are used for univariate control charts in many industrial settings to alert the operator that the system appears to be operating in an out-of-control state:

7 successive points up or down

2 successive points outside the 2-sigma control limits

4 successive points outside the 1-sigma control limits

The mathematical formulation for extending these rules to the multivariate situation follows.

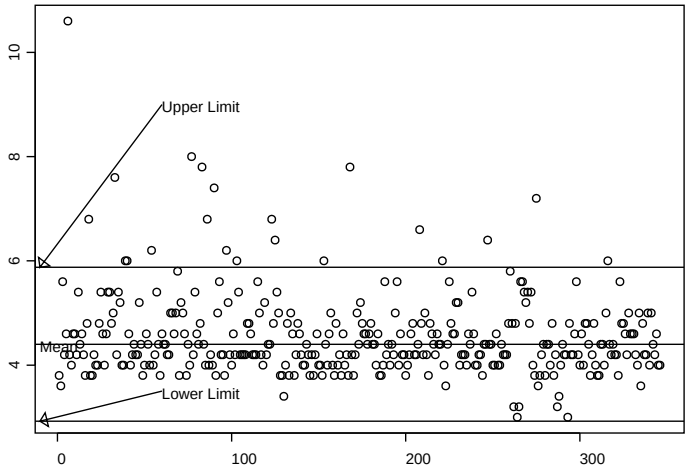


Figure 3.3: Silver

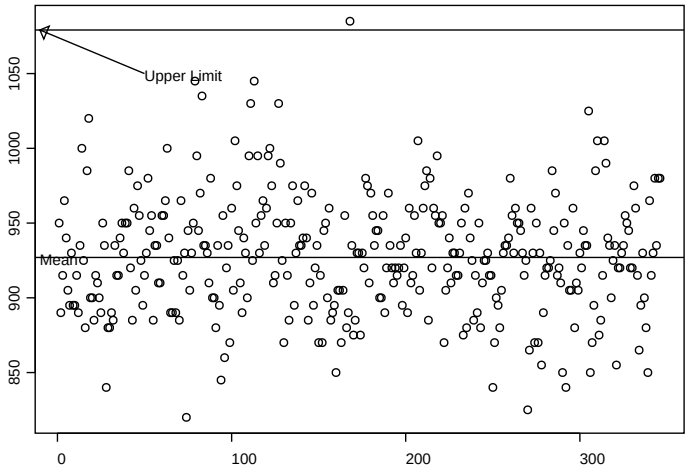


Figure 3.4: Manganese

Let $m, n \in \mathbf{N}$. Let, for $i \in [1,3]$, $(a_1^{(i)}, \ldots, a_n^{(i)}) \in \mathbf{R}^n$ be such that

$$(\forall j \in [1,n])\ 0 < a_j^{(1)} < a_j^{(2)} < a_j^{(3)} \ .$$

For the polyhedra defined by

$$P_i = \{(x_1, \ldots, x_n) : |x_j| \leq a_j^{(i)}\},\ i \in [1,3]$$

we have that

$$P_1 \subseteq P_2 \subseteq P_3 \ .$$

Suppose that a function

$$f : [1,m] \to \mathbf{R}^n$$

is given and let us write

$$f(l) = (x_1^{(l)}, \ldots, x_n^{(l)}) \ .$$

Also, for $l \in [1,m]$

$$S_l = \{j \in [1,n] : |x_j^{(l)}| > a_j^{(2)}\}$$

and

$$T_l = \{j \in [1,n] : |x_j^{(l)}| > a_j^{(1)}\} \ .$$

Note that

$$S_l \subseteq T_l \ .$$

Let k be the maximum of all $j \in [1,m]$ such that

$$(\forall i \in [1,j])\ f(i) \in P_3$$

(this is analogous to the univariate rule of 1 point outside the 3-sigma control limits) and with the following three properties.

1. For all $i \in [7,j]$,

$$(\exists l \in [i-6, i-1]) \min_{p \in [1,n]} (a_p^{(1)} - x_p^{(l)}) \leq \min_{p \in [1,n]} (a_p^{(1)} - x_p^{(l+1)}) \ .$$

(this is analogous to the univariate rule of 7 successive points up or down)

2. For all $i \in [2,j]$,

$$\{f(i), f(i-1)\} \subseteq P_3 \backslash P_2 \ \Rightarrow\ S_i \cap S_{i-1} = \phi \ .$$

(this is analogous to the univariate rule of 2 successive points outside the 2-sigma control limits)

3. For all $i \in [4,j]$,

$$\{f(i-p) : p \in [0,3]\} \subseteq P_3 \backslash P_1 \ \Rightarrow\ \cap_{l=0}^{3} T_{i-l} = \phi \ .$$

(this is analogous to the univariate rule of 4 successive points outside the 1-sigma control limits)

If $k = m$, we are done.

Otherwise, do the adjustment, i.e., form a new function, call it f again, so that, for $i \leq k$, $f(i)$ is the same as before and that $f(k+1) \in P_1$ with

$$\min_{p \in [1,n]} (a_p^{(1)} - x_p^{(k)}) \leq \min_{p \in [1,n]} (a_p^{(1)} - x_p^{(k+1)}) \ .$$

Continue till $k = m$.

Other univariate control rules may be similarly generalized.

3.5 Discussion

The problem posed by Chemex laboratories can be attacked in a variety of ways. In all the models considered, the common approach is to define a model, and discuss its suitability to the Chemex problem. The next step is then to determine a suitable measure on which to base a decision on whether or not to recalibrate. A measure which is often used in multiple-component systems and which has not been touched upon here is 'defects per unit'. This has the nice property of a very natural dimension reduction, but may not be a very powerful technique since it is an attributes (specifically, a count) measure. Once a measure has been chosen, real data may be used to determine whether or not the model actually performs satisfactorily. Model and distributional assumptions, data dependence and forms of data dependence are often factors which are overlooked and may render a tool useless or needing some adjustments. Natural data dependence such as running samples in batches, as is done by Chemex, can often be used to the advantage of the practitioner in specifying a more precise model. Actual data collection may differ from model requirement and therefore, if the study is a retrospective one, it is very important to know exactly how the data being analyzed have been collected. For example, in the Chemex problem, the 'in-control' data may or may not have been representative of the entire process, since operators and times were not known. The 'out-of-control' data may have been a series of points over time or a gathering of points from wherever they were available. These data may have been such that adjustments had been made by operators, if they were collected in real time. If the study requires data to be collected, they should be collected in a way such that a suitable model can be built or used. Since no mathematical model mirrors reality, simulation is a commonly used technique which allows experimenters to compare theoretical confidence levels with actual obtained confidence levels.

Bibliography

[1] Alt, F. B. (1985). Multivariate Quality Control, *Encyclopedia of Statistical Sciences*, **6** edited by N. L. Johnson & S. Kotz, John Wiley and Sons, New York.

[2] Caughlin, B. (Chemex Laboratories) - Dialogue and Correspondence during and following PIMS 1999 Industrial Problem Solving Workshop.

[3] Montgomery, D. C. (1996). *Introduction to Statistical Quality Control*, 3rd ed. John Wiley and Sons, New York, NY.

Chapter 4

Density Driven Turbulent Mixing at Batch Interfaces

Jim Brannan[1], Ellis Cumberbatch[2], Bryant Moodie[3], Tim Myers[4], Gordon O'Connell[5], Douglas Pickering[6].
Report prepared by Tim Myers & Jim Brannan

4.1 Introduction

Pipelines provide an economically efficient method of transporting liquid hydrocarbon products over long distances. A pipeline may be batched, that is, it may contain a number of different types of liquids, each batch having its own fluid characteristics of density and viscosity. Liquid hydrocarbons could be fluids such as crude oil (heavy, medium, light or synthetic) or refined products (gasoline diesel fuel, aviation fuel) or volatile liquids such as Natural Gas Liquids (NGL). It is economically desirable, especially with refined products, that cross contamination between adjacent batches be minimized. If the densities of the fluids in adjacent batches are different and the flow is laminar, buoyancy forces lead to an undesirable horizontal mixing process between the two fluids with the heavier fluid forcing itself under the lighter fluid as they are both convected along the pipe by a constant pressure gradient. This type of mixing is primarily driven by density gradients and gravity, but differences in viscosity will also affect the mixing process. In the laminar flow regime, the amount of cross contamination between batches is roughly proportional to the distance that the fluids are transported. Since the average fluid velocity is constant this means that the width of the interface, W, grows proportional to time t. Pipeline companies minimize mixing by operating the pipeline in the fully turbulent flow region. Reynolds numbers of greater than 2000 are considered to be fully turbulent but most companies use a lower limit of 4000. This assures that the flow will be fully turbulent. Upon start-up and shutdown of a pipeline the flow does pass through the laminar region but this is typically a short interval. Besides, starts and stops of pipeline are generally not frequent. Experimental measurements show that this results in an interface growth rate $W \propto t^{0.56}$.

In the following analysis models are developed for the turbulent mixing and growth at a batch interface. In §4.2 the mathematical model is defined, this depends crucially on the choice of diffusion coefficient D. In §4.3 a model is analysed where D is the harmonic average of the mixing coefficients of the two pure fluids. This is likely to be a good approximation when the density difference between the fluids is small.

When the density difference is large, in the laminar flow regime fingering will occur and there will be a relatively sharp interface between the fluids. However, in the turbulent case, as gravity drives the denser fluid into the less dense one the invading fluid is immediately mixed by turbulent diffusion. This means that sharp interfaces will not exist. Instead there will be a finite mixing region where the volume fraction of each fluid

[1]Clemson University
[2]Clairmont Graduate University
[3]University of Alberta
[4]Cranfield University
[5]University of Victoria
[6]Brandon University

changes from 1 to 0. In this case D will depend upon the relative concentration of the fluids. This approach leads to a degenerate diffusion problem which is analysed in §4.4.

4.2 Governing equations

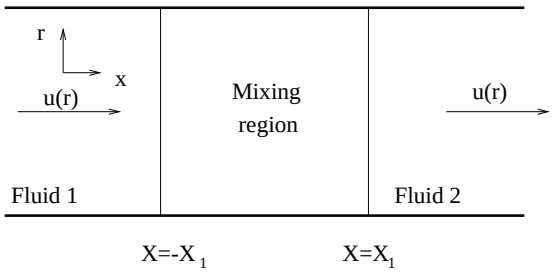


Figure 4.1: Problem configuration for $t > 0$

At an initial time $t = 0$ we consider fluid 1 occupying the section of pipe $x < 0$ and fluid 2 occupying the section of pipe $x > 0$, as depicted in see Figure 4.1. Let $c_1(x,t)$ and $c_2(x,t)$ denote concentrations (number per unit volume) of virtual particles of types 1 and 2, respectively. The concentrations of the two populations of virtual particles are assumed to be proxies or indicators for the mixing of the two fluids. The two types of particles are assumed to be distinguishable from one another and respond to the forces of the adjacent fluid elements just as if they were themselves characteristic volume elements of the fluid in its turbulent state. Assume that the fluids are incompressible so that

$$c_1(x,t) + c_2(x,t) = \bar{c}, \qquad t > 0\ . \tag{4.1}$$

The turbulent diffusion equations for the two sets of particles are:

$$\frac{\partial c_i}{\partial t} + \mathbf{u}\cdot\nabla c_i = \nabla\cdot(D_i\nabla c_i)\ , \tag{4.2}$$

where D_i is the turbulent diffusion coefficient in fluid i and $\mathbf{u} = (u,w)$ is the fluid velocity. Finally, we assume initial concentrations

$$c_1(x,0) = \begin{cases} \bar{c}, & x < 0 \\ 0, & x > 0 \end{cases} \tag{4.3}$$

and

$$c_2(x,0) = \begin{cases} \bar{c}, & x > 0 \\ 0, & x < 0 \end{cases}\ . \tag{4.4}$$

If $\alpha_1(x,t)$ and $\alpha_2(x,t)$ represent the volume fractions of fluid 1 and fluid 2 respectively, because of the identification discussed above, we set

$$\alpha_1 = \frac{c_1}{c_1+c_2} = \frac{c_1}{\bar{c}} \qquad\qquad \alpha_2 = \frac{c_2}{c_1+c_2} = \frac{c_2}{\bar{c}}. \tag{4.5}$$

Since $\bar{c}$ is a constant the turbulent diffusion equations may be rewritten as

$$\frac{\partial \alpha_i}{\partial t} + \mathbf{u}\cdot\nabla\alpha_i = \nabla\cdot(D_i\nabla\alpha_i)\ . \tag{4.6}$$

The turbulent diffusion coefficient in (4.6) must be determined experimentally for the fluids in question. The turbulent velocity $\mathbf{u}$ may be estimated from available existing results. This is discussed below.

With a 20" (0.508m) pipe, an average velocity of 5mph (2.235 m/s) and the fluid dynamic viscosity range supplied by Enbridge $\nu \in [0.1, 170]$cS ($\nu \in [10^{-7}, 1.7 \times 10^{-4}]$m^2/s) the Reynolds number for the flow is in the range $R_e \in [6.68 \times 10^3, 1.1 \times 10^7]$. The flow is therefore turbulent for all cases of interest. Experimental work shows that for turbulent flow, the velocity component along the pipe is closely approximated by

$$u = U_{max} \left(1 - \frac{r}{R}\right)^{1/n} , \tag{4.7}$$

where $n = 6, 7, 10$ for Reynolds numbers $R_e \approx 4 \times 10^3, 1.1 \times 10^5, 3.2 \times 10^6$ [1]. The wide variation calculated for the Reynolds number in the present case indicates that the appropriate value for n will depend strongly upon the fluids in question.

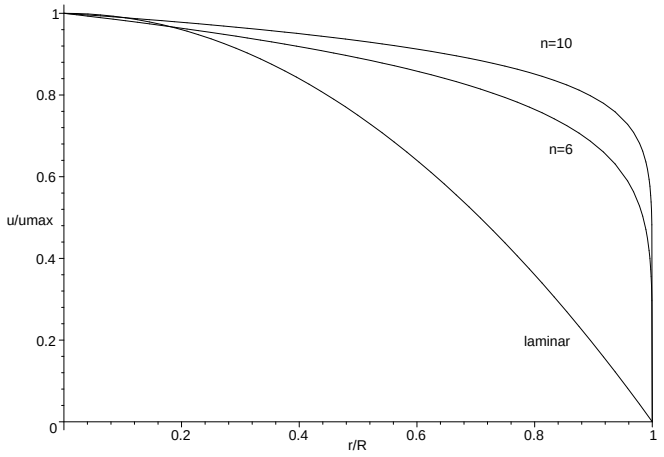


Figure 4.2: Velocity profile from $r = 0$ to R

The velocity profiles corresponding to values of $n = 6$ and 10 are shown in Figure 4.2. The laminar, parabolic profile is also plotted. When the flow is turbulent, in the vicinity of $r = R$, there is a boundary layer, away from this region the velocity varies only very slowly. For this reason u is well approximated by a constant value U_a, except for in the vicinity of $r = R$. However, since this boundary layer occupies only a small proportion of the cross-section it will not have a significant effect on the bulk properties and so can be neglected. The experience of pipeline companies indicates that for pipes of diameter greater than 12" the boundary layer is not hydraulically significant.

Since the flow is turbulent, mixing across the pipe happens over a very short time-scale. This means that any radial variations are quickly smoothed out and so radial derivatives will be negligible compared to derivatives along the pipe. Further, the (time) average velocity in the radial direction, w, must be zero. The diffusion equations may therefore now be written in a considerably simpler form

$$\frac{\partial \alpha_i}{\partial t} + U_a \frac{\partial \alpha_i}{\partial x} = \frac{\partial}{\partial x}\left(D_i \frac{\partial \alpha_i}{\partial x}\right) , \tag{4.8}$$

where radial variations have been neglected.

The heart of the current problem lies in determining the turbulent diffusion coefficients D_i. In the following section a standard diffusion model will be employed. The diffusion coefficient will be chosen as the harmonic average of the two diffusion coefficients for the pure fluids. At either end of the interface the harmonic average reduces to the value for a single fluid. As with standard diffusion problems this will imply an infinite speed of propagation and the width of the mixing zone will be infinite. This is the essence of the approach adopted in [2]. In this case the mixing zone must be defined by the region where the concentration of each component is above some minimum value. This model will be most appropriate when there is very little variation in the

densities of the fluids in question. In §4.4, it will be assumed that the amount of mixing depends on the relative concentration of the components. At the centre of the mixing zone, the mixing will be greatest, due to the high concentration of each component, away from the centre the mixing will decrease and reach zero at the edges of the mixing zone.

4.3 Mixing of fluids with a similar density

In this section the diffusion coefficient will be chosen as the harmonic average of the two diffusion coefficients of the pure fluids:

$$D = \frac{1}{\frac{1}{D_1}\alpha_1 + \frac{1}{D_2}\alpha_2} \,. \tag{4.9}$$

Note that $D \to D_1$ as $\alpha_2 \to 0$ and $D \to D_2$ as $\alpha_1 \to 0$. The rationale for using the harmonic average comes from an analogy with diffusion in homogenized media. The effective diffusion coefficient in this case is well known to be the harmonic average of the diffusion coefficients over a period length of the microscopic media.

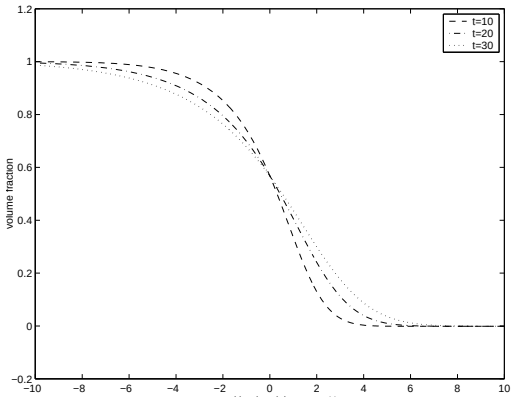


Figure 4.3: Volume fractions, α_1, when $t = 10, 20, 30$; D_1=1/2 and D_2=1/10

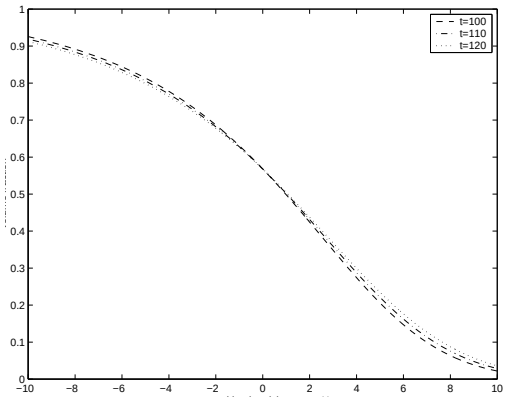


Figure 4.4: Volume fractions, α_1, $t = 100, 110, 120$; D_1=1/2 and D_2=1/10

Shifting co-ordinate system to move with the fluid $X = x - U_a t$ the problem is now defined by

$$\frac{\partial \alpha_1}{\partial t} = \frac{\partial}{\partial X}\left(D\frac{\partial \alpha_1}{\partial X}\right) \tag{4.10}$$

$$\alpha_1(X,0) = \begin{cases} 1, & X<0 \\ 0, & X>0 \end{cases} \tag{4.11}$$

$$\alpha_1(X,t)\to 1 \quad \text{as} \quad X\to-\infty \qquad , \qquad \alpha_1(X,t)\to 0 \quad \text{as} \quad X\to\infty \tag{4.12}$$

where

$$D=\frac{1}{\frac{1}{D_1}\alpha_1+\frac{1}{D_2}(1-\alpha_1)}=\frac{D_1D_2}{D_2\alpha_1+D_1(1-\alpha_1)}\ . \tag{4.13}$$

4.3.1 A similarity solution

If we set $\alpha_1(x,t)=f(\eta)$ where $\eta=\frac{X}{\sqrt{t}}$ we find that f must satisfy the two point boundary value problem

$$\left\{\frac{D_1D_2}{D_2f+D_1(1-f)}f'\right\}'+\frac{1}{2}\eta f'=0, \qquad \lim_{\eta\to-\infty}f(\eta)=1, \qquad \lim_{\eta\to\infty}f(\eta)=0. \tag{4.14}$$

The solution of this system will be a good approximation to that of (4.10)-(4.12) for large times.

With a constant diffusion coefficient a similarity solution exists for all time [3]

$$f=\frac{1}{2}\mathrm{erfc}\left(\frac{\eta}{\sqrt{2D}}\right)\ . \tag{4.15}$$

When the diffusion coefficient is defined by (4.13) the problem must be solved numerically. With values of $D_1=0.5$, $D_2=0.1$ solutions for times $t=10,20,30$ are shown in Figure 4.3, for $t=100,110,120$ solutions are shown in Figure 4.4. Two solutions are compared in Figure 4.5 to demonstrate the difference in concentrations when the diffusion coefficients are swapped. This highlights the fact that for any two fluids the concentration profile in the mixing zone depends on the order of the fluids, *i.e.* which is driving and which is being driven.

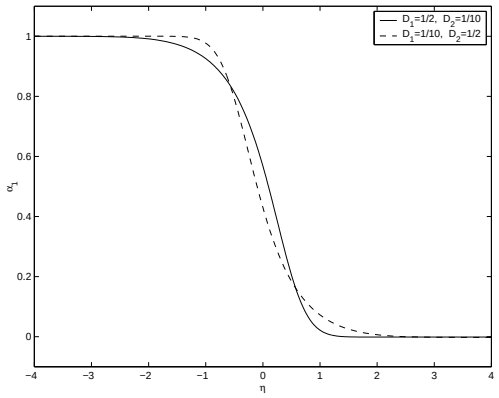


Figure 4.5: Solution to (4.14), two cases: D1=1/2 and D2=1/10 and vice-versa

If we arbitrarily define the position of the trailing edge, $x_t(t)$, and the position of the leading edge, $x_l(t)$, of the mixing zone to be locations where the composition is 90% fluid 1 and 90% fluid 2, respectively, then

$$x_t(t)=\eta_1\sqrt{t}+U_at$$

and the leading edge

$$x_l(t)=\eta_2\sqrt{t}+U_at$$

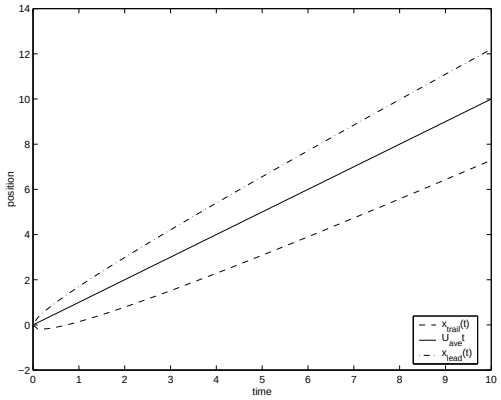


Figure 4.6: $x_t = U_a t + \eta_1\sqrt{t}, x_l = U_a t + \eta_2\sqrt{t}, x_c = U_a t; D_1 = 1/2, D_2 = 1/10$

where η_1 and η_2 are determined from $f(\eta_1) = 0.9$ and $f(\eta_2) = 0.1$. Graphs of $x_t(t)$ and $x_l(t)$ are illustrated in Figure 4.6 for the case $D_1 = 0.5$ and $D_2 = 0.1$ (so that the efficiency of the turbulent mixing is greater in fluid 1 than it is in fluid 2). The effect of this is shown by the greater thickness of the mixing zone below the line $x = U_a t$ than above.

In the case that the leading fluid is more efficient at mixing than the trailing fluid ($D_1 = 0.1$ and $D_2 = 0.5$) the situation is reversed with the thickness of the mixing zone above the line $x = U_a t$ greater than the thickness below the line.

4.4 Density driven mixing

The model of the previous section is based on turbulent diffusion. It has infinite speed of propagation and so cannot describe a finite interface. An alternative approach, discussed in §1, is a density driven mixing model. This requires a turbulent mixing coefficient which is zero at either end of the interface, reflecting the fact that no mixing occurs beyond this point. Close to an interface, the mixing will be small due to the small amount of one of the fluids. The mixing then increases to a maximum at the centre and decreases again as the other interface is approached. The simplest model which captures these features is $D_i = D_0\alpha_1(1-\alpha_1)$; because the fluids are well-mixed the diffusion coefficient is the same in each phase. The value of D_0 will be constant for each pair of fluids, it is expected that D_0 will vary with the viscosity and density difference between the fluids, in particular it must increase as the density difference increases. It seems reasonable to assume $D_0 = D_0(R_{e_1}, R_{e_2})$. As highlighted by the results of the previous section D must be calculated for each pair of fluids and with every fluid driving and being driven. The problem is now governed by

$$\frac{\partial \alpha_1}{\partial t} = \frac{\partial}{\partial X}\left(D_0\alpha_1(1-\alpha_1)\frac{\partial \alpha_1}{\partial X}\right) \ , \tag{4.16}$$

where, once again $X = x - U_a t$. Another simple model for the mixing coefficient is $D = D_0\partial\alpha_1/\partial x$. However, this leads to an interface which grows with $t^{1/3}$, whereas experiments indicate the growth is approximately $t^{0.56}$ so this model will not be discussed further.

Equation (4.16) may be solved numerically, alternatively a lot of useful information may be gained from the similarity solutions of the following section. Note, equation (4.16) is invariant under the transformation $\alpha_1 \to 1 - \alpha_1$, so there is symmetry about the origin. The problem therefore only requires solving on one side of the line $X = 0$.

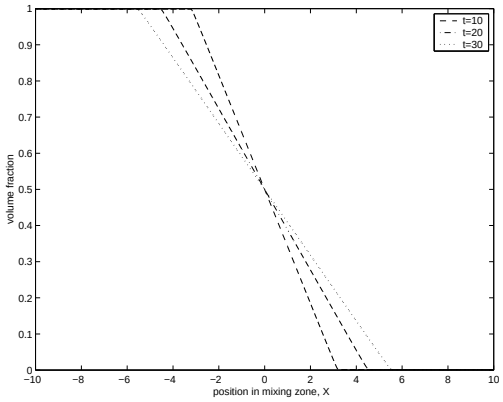


Figure 4.7: Concentration, $\alpha_1(x,t)$, when t = 10,20,30; D_0=1

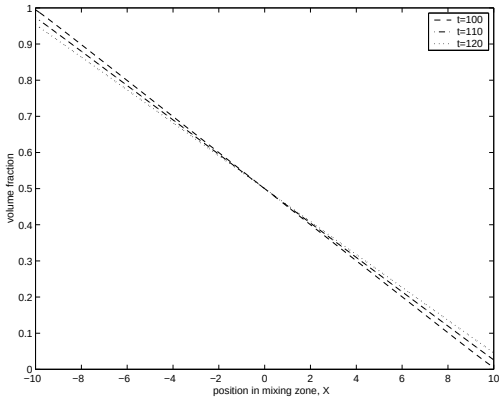


Figure 4.8: Concentration, α_1, when t = 100,110,120; D_0=1

4.4.1 Similarity solution

Employing the similarity substitution $\eta = X/\sqrt{D_0 t}$, equation (4.16) becomes

$$-\frac{1}{2}\eta\alpha_1' = \frac{\partial}{\partial\eta}\left(\alpha_1(1-\alpha_1)\alpha_1'\right) \ . \tag{4.17}$$

Appropriate boundary conditions are

$$\alpha_1(0) = \frac{1}{2} \qquad \alpha_1''(0) = 0 \qquad \alpha_1(\eta_1) = 0 \ , \tag{4.18}$$

where η_1 is the unknown position of the edge of the mixing zone. The first two conditions arise from symmetry considerations.

An exact solution to this problem is $\alpha_1 = (1-\eta)/2$. That this is the appropriate solution is confirmed by the numerical solutions of (4.17)-(4.18) shown in Figures 4.7, 4.8. Note, the ordinate in these figures is X, when η is taken as the ordinate all the curves collapse onto a single line, as shown in Figure 4.9. From this figure it is clear that the width of the mixing zone is $2\eta_1 = 2$.

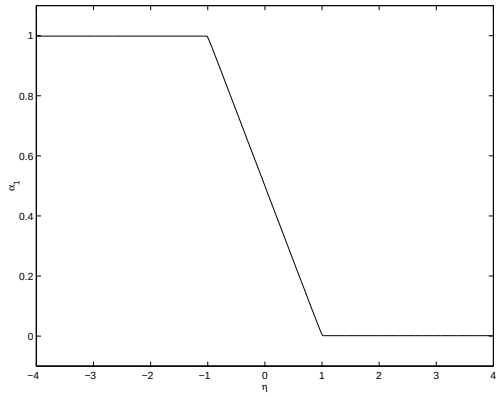


Figure 4.9: Solution to (4.17)-(4.18)

4.4.2 Width of interface

The width of the interface is simply given by $W = 2X_1$, where X_1 is the position of the leading edge, $X_1 = \eta_1\sqrt{D_0 t}$. The exact solution of the previous section shows $\eta_1 = 1$, and the interface width is therefore

$$W = 2\sqrt{D_0 t}\ . \tag{4.19}$$

This indicates that the interface grows proportional to $t^{0.5}$ and the constant of proportionality is $2\sqrt{D_0}$. The numerical results of Figures 4.7, 4.8, where D_0 is set to 1, have an interface width $2\sqrt{t}$. Experimental results given in [2] indicate $W \sim t^{0.56}$. A plot of the growth of the interface with time is shown in Figure 4.10.

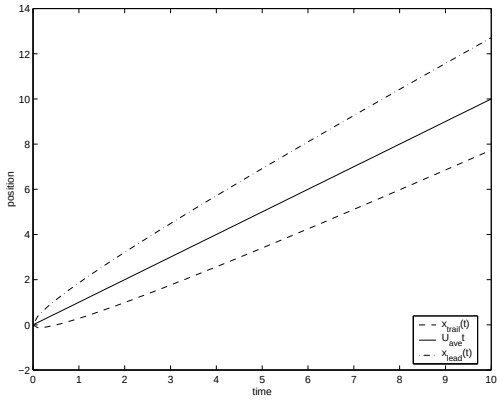


Figure 4.10: $x_t = U_a t + \eta_1\sqrt{t}, x_l = U_a t + \eta_2\sqrt{t}, x_c = U_a t; D_0 = 1/2$

4.5 Conclusions

Two models have been discussed in this paper. The first model will be most relevant to flows where the density difference between the fluids is small. In this case the diffusion/mixing coefficient is related to the diffusion coefficients in the pure fluids, D_i. For the analysis of §4.3 the coefficient is chosen as the harmonic average of D_1 and D_2. The second model will be relevant to flows where the density difference is significant. In this case gravity drives the mixing and the diffusion coefficient is not necessarily related to D_1 and D_2.

Both models studied lead to a growth of the interface with $t^{0.5}$, which is in close agreement with the experimental results given in [2]. These indicate a growth rate proportional to $t^{0.56}$. The problem description provided

by Enbridge stated that the width of the mixing zone stabilizes after a sufficiently long time. The rate of change in width of the interface in both of the current models is proportional to $1/\sqrt{t}$ (this rate would also be predicted by a constant diffusion coefficient model). As indicated by Figures 3, 4, 7, and 8 this decrease in the rate of change of mixing zone width may be considered as consistent with the stabilization observed by Enbridge within the limits of experimental error

The interface width also depends crucially upon the value of the diffusion coefficient, which must be determined experimentally. In the turbulent mixing section a very simple formula was obtained for the interface width, $W = 2\sqrt{D_0 t}$. So, the only parameter controlling the width is the diffusion coefficient. The best choice of fluids to travel together will therefore be the pair with the minimum value of D_0. The results of §4.3 highlight the fact that the mixing between any two fluids depends upon which fluid is driving and which is being driven. In light of this fact it is recommended that when determining D_0 tests are carried out on each pair of fluids and with every fluid driving and being driven. It is likely that the value of D_0 is a function of the Reynolds numbers of the two fluids, hence scaled down tests may be carried out provided the Reynolds numbers are scaled appropriately.

Bibliography

[1] R.B. Bird, W.E. Stewart & E.N. Lightfoot. *Transport Phenomena*, John Wiley & Sons 1960.

[2] G.I.Taylor. The dispersion of matter in turbulent flow through a pipe, *Proceedings of the Royal Society of London*, 223A, pp446-468, 1954.

[3] J.D. Crank, *Free and moving boundary value problems*, Oxford Science Publications, 1984.

Chapter 5

Efficient Portfolio Selection

Benyounes Amjoun[2], Dan Calistrate[1], Myriam Caprioglio[3], Brenda Hawkins[4], Mounia Kjiri[3], Tamara Koziak[4], Vincent Lemaire[3], Jason McVean[5], Miro Powojowski[1], Bill Reed[6], Satoshi Tomoda[1], Julie Zhou[6].

Report prepared by Min Tsao[6],

with assistance from Rita Aggarwala[1], Hassan Aurag[3], and Marc Paulhus[1].

5.1 Introduction

The problem described in this report deals with finding an efficient or optimal portfolio from many possible portfolios each of which consists of a subset of a large set of projects. The problem was brought to the PIMS workshop by Merak Project Ltd. of Calgary, a developer of oil and gas software for the petroleum industry. If portfolios made up of a selection of petroleum projects are plotted on a graph of expected value versus risk, there is an upper boundary above which no portfolios are found. The portfolios on this boundary are said to be efficient. The collection of these efficient portfolios is known as the efficient frontier.

For a portfolio to be efficient, it must be the case that no other portfolio has more value while having the same or less risk, and there is no other portfolio that has less risk while having the same or more value. Harry Markowitz revolutionized the field of portfolio theory with his pioneering work in the 1950s (Markowitz, 1952, 1959). His approach to efficient frontier analysis uses matrix algebra to determine an analytical expression for the curve representing the efficient frontier.

Merak develops economic software for the petroleum industry. In applying efficient frontier theory to the realm of the petroleum industry, Merak has taken a different approach from the traditional Markowitz technique. Using previously generated Monte Carlo results randomly selected portfolios are generated and plotted on an efficient frontier graph. As more portfolios are plotted, it quickly becomes apparent that there is an upper boundary. The efficient frontier is, therefore, implied by the upper boundary, but a curve is not explicitly drawn.

Both of these approaches have strengths and weaknesses. For instance, Merak's approach lacks the analytical certainty regarding the efficiency of promising portfolios that the Markowitz approach has. Even though there may seem to be no portfolios above a particular portfolio on the graph, it is always possible that the next randomly generated portfolio will be better.

On the other hand, the Markowitz approach has some severe limitations when applied to the petroleum industry:

[1]University of Calgary
[2]Ocean and Coastal Environmental Sensing Inc.
[3]University of Montreal
[4]University of Alberta
[5]Merak Project Ltd.
[6]Univerity of Victoria

- The simplification of describing a risk profile with only a mean and variance will lead to inaccuracies in the efficient frontier.
- It may not be possible to participate in a project at an arbitrarily fine level of granularity. Some projects may be such that they require 100% investment or they cannot be done at all.
- The constraints that determine which portfolios are valid can be complicated in the petroleum industry. Constraints like "If A then also B, C, and D" or "If E then not F or " or "At least 2 of H, I, J, and K" cannot be easily expressed as a linear equation, which is required for the Markowitz approach.

Merak brought this problem to the workshop in the hope that we could find a way to address the weaknesses of the two approaches, potentially by combining or partially combining them. An efficient frontier analysis method that combines the robustness of the Monte Carlo approach with the confidence of the Markowitz approach would indeed be a powerful tool for any industry. However, it soon became clear to us that there are other ways to address the problem which do not require a Monte Carlo component. Members of our group formed three subgroups and each subgroup developed a different approach for solving the problem. The first is the Portfolio Selection Algorithm Approach where we try to develop a practical searching algorithm which will lead us to the efficient portfolio without having to examine each and every possible portfolio. The second approach is the Statistical Inference Approach where we discuss statistical estimation and inference of the efficient portfolio. This approach provides a solution to the weakness of the Monte Carlo method of Merak. The third approach is the Integer Programming Approach where we try to find the exact efficient portfolio by setting up the problem as an integer programming problem and then solving it.

The rest of this report is organized as follows: Sections 5.2, 5.3 and 5.4 cover the three approaches, respectively. Section 5.5 contains a short summary.

5.2 The Portfolio Selection Algorithm Approach

Members of the subgroup which developed this approach are Hassan Aurag, Myriam Caprioglio, Mounia Kjiri and Vincent Lemaire, all of the University of Montreal. This approach is motivated by the fact that exhaustive computation of portfolios cannot be undertaken when the number of projects exceeds 30. Thus we have to propose a selective method. In Merak's Monte Carlo method, portfolios are generated using a Monte Carlo technique. Since the portfolios are selected randomly, the best ones may not appear in the graph. Can we find an algorithm that would eventually select a representative set of portfolios such that none of the best portfolios are missed? In the following, we describe a portfolio selection algorithm for this purpose. We will explain the algorithm using an example data set.

5.2.1 The Data

The data consists of a set of 10 projects. For each project, we had two sets of numbers (of length 200 each) corresponding to Net Present Value (NPV) and Capital Investment (CI). The expected value and risk of a project are respectively the mean and standard deviation of the NPV's. The expected value and risk of a portfolio are respectively the mean and standard deviation of the sum of its component projects.

5.2.2 The Constraints

The first constraint is on total capital investment. Then one has to deal with some petroleum specific constraints, e.g., constraints 2-4 below. In all our tests, we used the following set of rules:

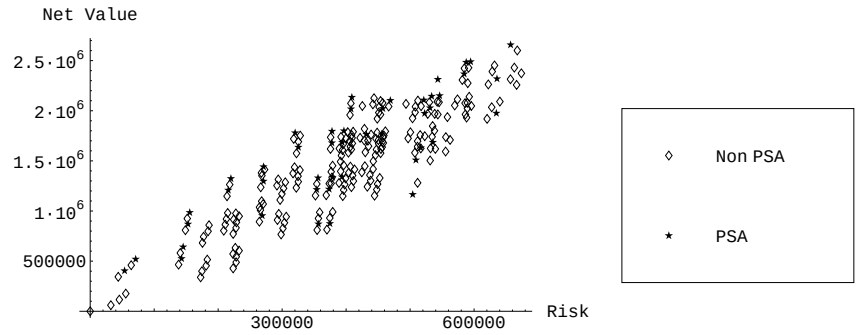
1. The maximum investment is $2,000,000.
2. Projects 1 and 4 may not belong to the same portfolio.
3. If a portfolio contains project 9, it must also contain one and only one of projects 7 and 10.
4. If a portfolio contains project 4 then it must contain project 8.

5.2.3 Portfolio Selection Algorithm (PSA)

The idea is to classify the set of all portfolios by the number of projects they contain. One then starts by picking a random project among the subset of best portfolios. At the next stage, you consider a sampling of portfolios containing two projects with the constraint that one of them must be the one we started with. In all subsequent stages, we keep adding a sampling of projects. Moreover, at each stage, we will require our portfolios to be valid and we only keep those considered best.

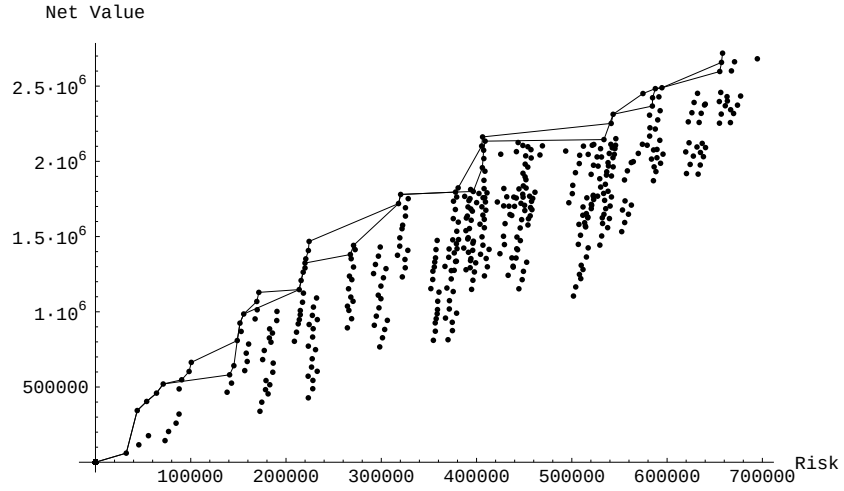
Notice that we cannot always compare two portfolios. Hence, in selecting the best portfolios we will only keep those portfolios that satisfy the following: Its NPV is greater than the NPV of all portfolios that have less or equal risk.

5.2.4 PSA Results



5.2.5 Results Analysis

Initial results show that adding boolean type constraints does not affect the efficient frontier very much. Using our approach, we have noticed that we didn't miss the best portfolios. The plot below further illustrates this point.



In the plot the dots represent all portfolios not all of which satisfy the boolean type constraints. One of the two lines represents the efficient frontier for all portfolios and the other represents that for those that satisfy the boolean type constraints. The two lines are very close to each other.

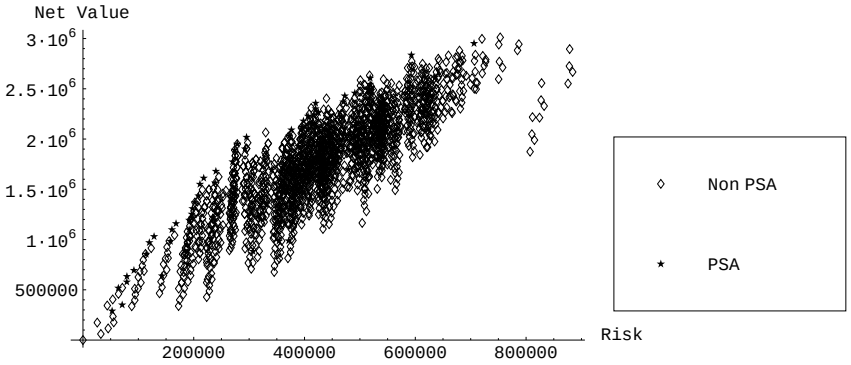
It should be noted that the data given to us by Merak was not strongly correlated.

5.2.6 Second Case: a 15 projects database

Here we used the original 10 projects and made up another 5 of the original ones using linear combinations, in order to obtain higher correlation between some of the projects. The correlation matrix of the set of 15 projects is a 15 by 15 matrix, too large to be shown in full here. The first 10 columns are:

$$\begin{pmatrix}
1.0000 & -0.0210 & 0.0290 & -0.0360 & 0.0610 & -0.1100 & -0.0067 & 0.0051 & 0.1200 & 0.0370 & \dots \\
-0.0210 & 1.0000 & -0.1200 & -0.0620 & 0.0022 & -0.0320 & 0.0190 & -0.0570 & 0.0640 & 0.0770 & \dots \\
0.0290 & -0.1200 & 1.0000 & 0.0720 & -0.0620 & 0.0940 & -0.0270 & -0.0980 & 0.0730 & -0.0860 & \dots \\
-0.0360 & -0.0620 & 0.0720 & 1.0000 & -0.1300 & 0.0066 & -0.1000 & -0.1700 & 0.0130 & -0.1300 & \dots \\
0.0610 & 0.0022 & -0.0620 & -0.1300 & 1.0000 & -0.1100 & -0.0340 & 0.1200 & -0.0071 & 0.2400 & \dots \\
-0.1100 & -0.0320 & 0.0940 & 0.0066 & -0.1100 & 1.0000 & 0.0290 & 0.0150 & -0.0460 & -0.0460 & \dots \\
-0.0067 & 0.0190 & -0.0270 & -0.1000 & -0.0340 & 0.0290 & 1.0000 & 0.0730 & 0.0420 & 0.1000 & \dots \\
0.0051 & -0.0570 & -0.0980 & -0.1700 & 0.1200 & 0.0150 & 0.0730 & 1.0000 & 0.0150 & 0.0047 & \dots \\
0.1200 & 0.0640 & 0.0730 & 0.0130 & -0.0071 & -0.0460 & 0.0420 & 0.0150 & 1.0000 & -0.0100 & \dots \\
0.0370 & 0.0770 & -0.0860 & -0.1300 & 0.2400 & -0.0460 & 0.1000 & 0.0047 & -0.0100 & 1.0000 & \dots \\
0.1400 & -0.0590 & -0.0930 & -0.1700 & 0.1300 & 0.0004 & 0.0720 & 0.9900 & 0.0320 & 0.0096 & \dots \\
-0.0220 & 0.9900 & -0.1200 & -0.0760 & -0.0024 & -0.0270 & 0.1500 & -0.0460 & 0.0680 & 0.0900 & \dots \\
-0.0960 & -0.0550 & 0.2900 & 0.0210 & -0.1200 & 0.9800 & 0.0230 & -0.0056 & -0.0300 & -0.0620 & \dots \\
0.0500 & -0.0150 & -0.0410 & 0.1500 & 0.9600 & -0.1100 & -0.0630 & 0.0740 & -0.0033 & 0.2000 & \dots \\
0.0340 & -0.0340 & -0.0140 & 0.4500 & 0.8200 & -0.0960 & -0.0890 & 0.0140 & 0.0013 & 0.1400 & \dots
\end{pmatrix}$$

Again, one obtains similar results as in the case of the 10 original projects. The graph below represents the results of our algorithm.



5.2.7 Conclusions

We would like, as a conclusion, to raise some questions and give our views regarding their answers.

1. When we have a large number of projects, we need to start choosing a random number of portfolios at each step. Do we still get the accuracy obtained in our case?

2. It is not necessary to start with 1-project portfolios. Can we for example start with r-projects portfolios where $r < N$, and N is the number of projects in the company's database?

3. Is it possible to merge this approach with Merak's current Monte-Carlo approach?

We believe the answers to these questions might be of great help in achieving better accuracy for the portfolio analysis software. In the short time we had for this study, we could not make all the necessary tests. Thus, we must say that the answer to the first question is unknown to us.

Depending on our goal, the answer to question 2 can be yes, if an approach such as ours is used to refine Merak's current model. We can easily imagine a situation where a *manager* would be able to select certain portfolios and ask for a refinement of the result using our PSA. In this case, we would just start from the given r.

Concerning our last point, if the answer to the first is negative, we believe we could still use our approach as an optional path for the end-user (e.g., to refine results), as proposed above. The user would supply the projects and we would build the portfolios using our approach. In this case, the user should be aware that no more than a certain number of projects can be selected; this number depending on the computational cost of our approach.

Finally, we would like to mention that all our simulations were done using Mathematica software. Even though Mathematica is a good software for simulation, it is certainly true that much faster code can be produced using C for instance. Mathematica was able to handle, exhaustively, up to 15 projects in 30 minutes without any optimization. In general, this means that one can go up to 20 projects, exhaustively, in less than 2 minutes, without counting expert programmer's optimizations.

5.3 The Statistical Inference Approach

Members of the subgroup which developed this approach are Rita Aggarwala of the University of Calgary, Brenda Hawkins and Tamara Koziak of the University of Alberta, and Bill Reed, Min Tsao and Julie Zhou of the University of Victoria. This approach explores the question of what constitutes *adequate sampling* of portfolio values in the sense that the client's *desired confidence* in obtaining a solution *close to* the true optimal portfolio value can be specified and satisfied if adequate sampling is carried out. The ideas of confidence and close to the true optimal will also be discussed more thoroughly.

Every portfolio is composed of a number of projects, each of which may be associated with various distributional and constraint assumptions. For example, a particular project may involve drilling an oil well. The results (which an attempt is usually made to quantify, for example, the net return) of such a drill will not be known prior to the drilling, however an idea of the various possibilities or result outcomes, may be adequately expressed by applying a distributional assumption to these possibilities. In addition, legal issues may dictate constraints on the drilling of the well. The total portfolio itself may also be subject to assumptions and constraints, such as budgetary and resource considerations, and/or restrictive relationships between possible projects.

Since each portfolio is usually composed of projects with uncertain returns, the return on any particular portfolio itself is also uncertain and therefore associated with some probability distribution. Present practice assigns each portfolio a *value*, based on the *expected* return of the portfolio, and a *risk* in choosing the portfolio, based on the differences between the *possible* returns for the portfolio and the expected return. These two quantities are calculated using the distributional assumptions and constraints discussed above, which are supplied by a client wishing to decide upon a suitable portfolio. As a general rule, portfolios with higher values tend to have higher corresponding risk.

As can be gathered from this preliminary discussion, there are quite a few assumptions and approximations involved in computing any single portfolio value and risk. Therefore, although it is very easy to become absorbed in the search for the theoretically absolute optimal portfolio, it should be kept in mind that even if the theoretical optimal were to be defined as in the following discussions, and found (which in most cases is either not possible, not feasible or both at present), it would be based on some approximated, albeit experience-based inputs given by the client. The conclusion that this portfolio is, in reality, *the* optimal one is almost certainly incorrect. However, given the information available is the best that we have, the ability to find the theoretically optimal portfolio, and perhaps others in that region, would be useful in making an intelligent decision. Based on this reasoning, being in the *region* of the theoretical optimum with high confidence is a very reasonable target for a client.

Let us suppose that there are M possible portfolios meeting the required constraints, and therefore M associated probability distributions, giving M value and risk pairs. Quite often, M is unknown, however we are usually able to put at least some rough constraints on M. In the simple case where there are k possible projects, and each project is either opted in to or opted out of, we can say that $M \leq 2^k$. In practice, some projects may also be opted in to at less than full commitment, for example some projects may allow commitment levels of 0%, 50%, or 100%. Constraints added to the portfolio will determine how many portfolios are eliminated from the collection represented by the upper bound on M. In situations where the bound on M is not too large for exhaustive enumeration of portfolio values and risks (and simulated probability densities for returns), exhaustive enumeration is the method taken, and statistical sampling and inference techniques need not be applied.

Notice that the collection of values (expected returns) itself can be seen as M measurements on the population of all M possible portfolios. Since each portfolio value is the expected value of some distribution which is arrived at through different (possibly continuous) distributional assumptions, constraints, etc, it is highly unlikely that any two of the M possible values will be the same. The standard procedure, in the case where the number of possible portfolios is too large to enumerate exhaustively and too complicated to attack analytically is to

randomly sample a pre-determined number m of unique portfolios from the population of possible portfolios and obtain their m corresponding values, that is, to obtain a random sample (without replacement) of size m from the probability mass function for portfolio values. The portfolio decided upon could then be the largest of the m sampled. In addition to the cardinality M of the probability mass function being unknown, the support of this probability mass function (the range of allowable portfolio values) is also unknown. Similar reasoning can be applied towards the collection of portfolio risks.

Our approach was to define an optimal portfolio as that which gave the maximum value or expected return for a specified *risk bin*. A risk bin is taken to be a range of quantified risk which the client is comfortable with and indifferent to. We will assume that N is the number of allowable portfolios in the specified risk bin, and that a *random* sample of n values can be attained from this risk bin. We will also assume that this random sampling approach will be applied in cases where N is prohibitively large to enumerate all portfolio values and risks in the specified risk bin exhaustively, and an analytic solution for the maximum value is not attainable. Thus, a statistical approach to deciding upon a suitable value of n given a desired confidence and range of risk is required. It should be noted that unless we are able to (internally) fit a suitable probability distribution to individual clients' samples of n portfolio values and use parametric inference from that point on, it is important to develop robust procedures.

The subgroup of investigators explored the following specific questions from a number of angles:

- What is the confidence associated with being in the top $100\alpha\%$ of possible portfolios values, if n portfolios are randomly sampled?
- Can a confidence interval be specified for the estimation of the theoretically optimal (highest possible) portfolio value?

5.3.1 Percentile Estimation

Robust techniques for percentile estimation have been explored in many books on order statistics, see for examples, David (1981) and Arnold, Balakrishnan and Nagaraja (1992). The application of these techniques will apply to this problem, as well, provided we assume that the random sample of size n of portfolio values from a risk bin is much smaller than the risk bin population N. (This is in order that we may justify the assumption of approximately sampling from an infinite population). This assumption is actually a conservative one, in that if it does not hold, the results of this subsection are generally stronger than stated.

A simple and robust probabilistic argument may be used to determine the probability β that the r^{th} largest value in the sample will be in the top $100\alpha\%$ of possible portfolio values based on a sample of size n, where α is generally taken to be a small proportion. Once the sample has been taken, the term *probability* must be replaced by *confidence level*. Conversely, for fixed r, α, and desired confidence level β, the required sample size n can be determined. The expressions given hold for continuous distributions of portfolio values. Since we are actually dealing with a discrete distribution of portfolio values, the probabilities β obtained here will be approximations. They will be very close approximations if it may be assumed that (at least the top $100\alpha\%$ of) the distribution of portfolio values can be approximated by a continuous distribution. Notice that this implies that the theoretical max does not "stand alone." This will be important in our discussion of approximating the actual maximum possible portfolio value in the next section. From examination of typical plots and discussions with the industrial mentor, these are not unreasonable assumptions. The probability β is given as follows. We will denote the i^{th} order statistic (ordered value) in the random sample of size n by $Y_{i:n}$.

$$\begin{aligned}
\beta &= P\left(\begin{array}{c}\text{at least } r \text{ of the } n \text{ observed values are in}\\ \text{the top } 100\alpha\% \text{ of possible portfolio values}\end{array}\right)\\
&= \sum_{i=r}^{n}\binom{n}{i}\alpha^i(1-\alpha)^{n-i}\\
&= 1-\sum_{i=0}^{r-1}\binom{n}{i}\alpha^i(1-\alpha)^{n-i}.
\end{aligned}$$

It is easily seen that for $r = 1$, where we are only interested in the probability of the largest observed value being in the top $100\alpha\%$ of possible values, the above expression for β reduces to $1 - (1 - \alpha)^n$. By specifying β and α it is also easy to determine the required sample size n in the case $r = 1$. For larger values of r, solving for n will involve polynomial root finding, which is handled with ease using any simple mathematical software tool. Recall that, in view that all calculated portfolio values are based on approximations themselves, it may be of greater practical interest to consider a few feasible portfolios in the top percentiles of possible values rather than a single one, and perhaps base the final selection of portfolios on other considerations such as convenience.

A table displaying selected values of n, r, α and β follows.

n	100	100	100	200	200	200	500	500	500
r	1	1	3	1	1	5	1	3	10
α	.01	.05	.05	.01	.05	.05	.01	.01	.05
β	.634	.994	.882	.866	1.00	.974	.993	.877	1.00

Thus, for example, if 200 points are sampled from the desired risk bin, we will have approximately 97% confidence that the top 5 observed portfolio values are within the top 5% of all possible portfolio values, and almost 100% confident that the highest observed portfolio value is within the top 5% of all possible portfolio values.

5.3.2 Efficient Frontier Estimation

The previous section on percentile estimation discussed robust methods for determining the probability of observing one or more portfolio values in the top $100\alpha\%$ of all possible values of portfolios, for a specified risk bin. However, the true maximum possible portfolio value was never assumed nor estimated. In fact, to use the techniques of the previous section, no true maximum needs to exist. In the present context, there is a theoretical true maximum portfolio value θ which the client desires to be at or close to. In this section we will explore point and interval estimation of the value θ based on a random sample of size n of the N possible portfolio values in a risk bin.

We will again assume that n is much smaller than N, for if this were not the case, all possible portfolios would be enumerated. We will also assume that the distribution of portfolio values can be approximated by some (perhaps piecewise disjoint) continuous distribution. Recall that this implies that the theoretical maximum does not "stand alone." Again, this seems to be a reasonable assumption, however, for this section, the concept of not standing alone is clarified and quantified.

Specified Intervals

In this approach, the client would want to know the probability that their maximum observed portfolio value $Y_{n:n}$ is within δ of the true maximum possible portfolio value, θ, where δ is a number (perhaps a percentage of the largest observed value) specified by the client. It should be noted that an appropriate sample size may be chosen by methods in the section on percentile estimation. The probability which we will estimate here will be easily computed after the sample has been observed, as a post-hoc analysis, and therefore will be viewed as a confidence level.

We would like to estimate $P(Y_{n:n} \in [\theta - \delta, \theta])$. Since we have a random sample of observations, the probability that any observed value lies in the interval $[\theta - \delta, \theta]$ is the same, say α. Therefore,

$$P(\text{at least one observation is in } [\theta - \delta, \theta]) = 1 - (1 - \alpha)^n .$$

Notice that this expression looks very similar to the binomial sum expression for $r = 1$ in the previous section on percentile estimation. However, in this case, α is not specified by the client, rather it is something which must be estimated from the data in order to arrive at an approximate probability. Assuming the distribution of values behaves similarly in $[Y_{n:n} - \delta, Y_{n:n}]$ and $[\theta - \delta, \theta]$, we may estimate α by

$$\frac{\text{number of observations in } [Y_{n:n} - \delta, Y_{n:n}]}{n}.$$

This approximation would be reasonable if the distribution of portfolio values were assumed to behave uniformly in $[Y_{n:n} - \delta, \theta]$. Choosing a sample size large enough that there is a reasonable probability of sampling a few values in the higher percentiles by using the techniques discussed in the previous section should elicit a good idea of the shape of the distribution, even in the tails, since the sampling is random.

Likelihood Intervals

The idea of likelihood intervals is explored, for example, in Kalbfleisch (Volume 2, 1985) and Royall (1997). In the present situation, likelihood point estimation is very intuitive, and therefore likelihood intervals are a natural way to estimate θ. The interpretation of likelihood intervals can be compared to that of traditional confidence intervals for large sample sizes, for example, the authors mentioned above discuss the near equivalence of a 14.7% likelihood interval with a 95% confidence interval, and a 3.6% likelihood interval with a 99% confidence interval for regular distributions from which large samples have been drawn. As a general rule, points inside a 10% likelihood interval are labeled as "plausible values" for the parameter θ, and points outside a 1% likelihood interval as "very implausible values." It is important to realize here that lower percentage likelihood intervals are more desirable, whereas with traditional confidence intervals, higher percentage confidence intervals are desired. This is simply due to the construction of the intervals.

In general, the Maximum Likelihood Estimate of a parameter is the value of the parameter which maximizes the likelihood function of the observed data. The likelihood function is simply the product of individual probability functions for the observed data when the data are a random sample. We will seek robust estimates and intervals, with minimal assumptions made on probability functions. Specifically, we will assume that the distribution of portfolio values can be approximated by a truncated (possibly interval piecewise) continuous distribution. Following discussions with the industrial mentor and examination of some sample data, it seems that it is quite common for the density of portfolio values to appear to be truncated at θ. This may be due to one or more of the constraints associated with a problem.

We assume the truncated distribution takes the following form:

$$\begin{aligned} f(x) &= \frac{g(x)}{G(\theta)},\ 0 \le x \le \theta \\ &= 0,\ \text{otherwise,} \end{aligned}$$

where $g(x) = \frac{d}{dx}G(x)$, and $G(x)$ is a valid, differentiable cumulative distribution function. The lower bound on the support need not be 0.

The likelihood function for a random sample of n values from this distribution is then

$$\begin{aligned} L(\theta) &= \frac{\prod_{i=1}^{n} g(y_i)}{[G(\theta)]^n},\ \theta \ge y_{n:n} \\ &= 0,\ \theta < y_{n:n}. \end{aligned}$$

Since $G(\theta)$ must be an increasing function, this likelihood will take its maximum value at $\theta = y_{n:n}$, and thus $y_{n:n}$ is the maximum likelihood estimate of θ here.

In general a $100\gamma\%$ likelihood interval for a parameter θ is given by

$$\{\theta : L(\theta) > \gamma L(\theta^*)\}$$

where θ^* is the maximum likelihood estimate of θ. A $100\gamma\%$ likelihood interval for θ in this problem is therefore

$$\left\{\theta : G(y_{n:n}) \le G(\theta) \le \gamma^{-1/n} G(y_{n:n})\right\}.$$

If one can assume a parametric form for $G(\theta)$ (for example, by internally fitting a truncated distribution to the observed random sample), one can solve this inequality. For example if the distribution of portfolio values can be assumed to be uniformly distributed on $(0, \theta)$, then the likelihood interval for θ is

$$\left\{\theta : y_{n:n} \le \theta \le \gamma^{-1/n} y_{n:n}\right\},$$

whereas if the distribution of portfolio values can be assumed to come from a truncated exponential distribution with mean $1/\lambda$, the likelihood interval is

$$\left\{\theta : y_{n:n} \leq \theta \leq -\ln\left[1-\gamma^{-1/n}\left(1-\exp\left(-\lambda y_{n:n}\right)\right)\right]/\lambda\right\}.$$

Notice that each of these examples assumes a tidy parametric form for the density $f(x) = g(x)/G(\theta)$ over the entire range of feasible portfolio values, $0 \leq x \leq \theta$. It is possible that the data would allow such a density to be fit to portfolio values, however, we will give a few ideas in the likely event that this is not the case.

Firstly, a Taylor's series expansion of $\ln G(\theta)$ about $\theta = y_{n:n}$ will give us the interval

$$\left\{\theta : \frac{g(y_{n:n})}{G(y_{n:n})}(\theta - y_{n:n}) + \cdots \leq -\frac{1}{n}\ln\gamma\right\}.$$

One may proceed to obtain the first order likelihood interval for θ

$$\left\{\theta : y_{n:n} \leq \theta \leq y_{n:n} - \frac{\ln\gamma}{nf(y_{n:n})}\right\}$$

where a non-parametric estimate of the density $f(\cdot)$ at $y_{n:n}$ may be substituted for $f(y_{n:n})$ provided n is large enough to obtain a good estimate of this value. Similarly, a second order likelihood interval may be obtained by solving the quadratic equation arising from the Taylor's series expansion above. This will involve non-parametric estimates of $f(\cdot)$ and $f'(\cdot)$ at $y_{n:n}$.

Another possible approach is to argue that since the likelihood function is non-zero only for $\theta \geq y_{n:n}$, this is the region in which a parametric form for $G(\cdot)$ (and as a result, $f(\cdot)$) will be needed in obtaining the desired interval. Unfortunately, we have no data in this interval! If we consider using a few, say r of the upper observed ordered values in order to estimate the shape of $G(\theta)$ in that region, it will be sufficient to assume that the *upper tail* of the distribution, from $Y_{n-r:n}$ to θ can be approximated by a continuous distribution on one interval segment, and the sample size of observed values should be chosen large enough that at least r values of the n sampled will be in this upper tail with high probability (the more robust techniques of the previous section may be used to determine a large enough sample size for this latter condition; If, for example, it is felt that it is safe to assume the top $100\alpha\%$ percent of values can be approximated by a continuous distribution, a sample size may be chosen so that the corresponding probability β discussed in the previous section on percentile estimation is high for some reasonable value of r). Then, the assumed parametric form for $G(\cdot)$ in this tail region, which includes $\theta \geq y_{n:n}$ may be used to solve the likelihood interval

$$\left\{\theta : G(y_{n:n}) \leq G(\theta) \leq \gamma^{-1/n}G(y_{n:n})\right\}.$$

This approach may be most reasonable if, for example, the sample of portfolio values does not seem to be continuous on one interval, but rather over interval segments. The industrial mentor did indicate that quite often, data is observed in "clumps".

Finally, it should be noted that if parametric forms are assumed for the probability density of portfolio values, classical confidence intervals may also be explored for θ. Care should be taken that these confidence intervals make sense, in that they do not include values of $\theta < y_{n:n}$. As one can see from the likelihood interval above, this is not a concern for the likelihood intervals discussed here, since $G(\cdot)$ is an increasing function.

We close with the following example: suppose it is reasonable to assume that $f(\cdot)$ can be approximated by a uniform distribution on $[Y_{n-r:n}, \theta]$ with high probability. Then $G(\cdot)$ will be linear in this interval. This could be quite a reasonable assumption for large enough n. Thus, $G(x) = ax + b$ and the likelihood interval is given by

$$\left\{\theta : \theta \leq \alpha^{-1/n}y_{n:n} + \frac{b}{a}\left(\alpha^{-1/n} - 1\right)\right\}.$$

The quantity $\frac{b}{a}$ may be approximated by drawing a line of best fit through the largest r values of the empirical cumulative distribution function for $f(\cdot)$. Thus, since $f(x) = g(x)/G(\theta)$, and $F(x) = G(x)/G(\theta)$, the slope

of the line will be an estimate for $a/G(\theta)$ and the intercept will be an estimate for $b/G(\theta)$, and an estimate of b/a can be obtained using the ratio of the fitted intercept to the fitted slope.

An interesting observation arising from this example can be made. If we do assume a uniform (or even some other) distribution of portfolio values in the upper tail of the density, we may consider the following: a useful property of order statistics is that given $Y_{n-r:n} = y_{n-r:n}$, the remaining larger order statistics $Y_{n-r+1:n}, \ldots, Y_{n:n}$ behave as a random sample of size r from the same distribution (uniform or some other) left truncated at $y_{n-r:n}$. A likelihood interval may then be determined directly for θ from first principles, but only using the r values in the tail as discussed earlier. Notice that this approach does not require the entire distribution of portfolio values to be of a truncated form. In the case of the uniform assumption for the upper tail, the likelihood interval for θ becomes

$$\left\{\theta : y_{n:n} \leq \theta \leq \frac{y_{n-m:n} - \alpha^{-1/r} y_{n:n}}{1 - \alpha^{-1/r}}\right\}.$$

For the uniform distribution tail assumption, a traditional confidence interval for θ is also easily obtained by using the fact that order statistics from uniform distributions behave as Beta random variables. A $100\,(1-\alpha)\,\%$ upper confidence interval for θ in this case is given by

$$\left\{\theta : y_{n:n} \leq \theta \leq y_{n-r:n} + \frac{y_{n:n} - y_{n-r:n}}{(1-\alpha)^{1/r}}\right\}.$$

The advantage in both of these cases is that the form of the upper bound on θ is very simple. A disadvantage is that the number of values r is very important in the resulting width of the interval, and a larger sample size n will be necessary to ensure observation of enough of these tail values with high probability.

Deciding upon the best method to employ (or if other methods should be sought) should involve experiments and simulations with typical projects and portfolios.

5.4 The Integer Programming Approach

Members of the subgroup which developed this approach are Benyounes Amjoun of the Ocean and Coastal Environmental Sensing Inc. and Marc Paulhus, Miro Powojowski and Satoshi Tomoda of the University of Calgary. This approach provides a rigorous method for solving a large class of portfolio selection problems. We now give a detailed description of this approach, beginning with some background material.

The problem is to find an efficient portfolio of projects under some constraints. If risk is measured by a single parameter, and the company is assumed to be *rational* (prefer more wealth to less wealth and less risk to more risk) then "efficient" is easy to define. If we plot all available portfolios on a standard expected return versus risk graph, then a portfolio P is *efficient* if there are no portfolios both above and to the left of P. See Figure 5.1.

We will assume that there are n projects which are available to a company, which we will label $0 \ldots n-1$. The company is not free to enter these projects at any level of granularity, indeed, we shall assume that a company is either invested in a project or not invested in a project (in Section 5.4.1 we explain how to relax this slightly). Hence, for project $i \in \{0, \ldots, n-1\}$ we can associate the variable x_i, such that

$$x_i = \begin{cases} 0 & \text{if project } i \text{ is not included} \\ 1 & \text{if project } i \text{ is included.} \end{cases}$$

Thus the vector $X = (x_0, \ldots, x_{n-1})'$ is a vector of zeros and ones which defines a portfolio.

Further associate with project i:

- a dollar cost c_i (let $C = (c_0, \ldots, c_{n-1})'$),
- an expected net return μ_i (let $\boldsymbol{\mu} = (\mu_0, \ldots, \mu_{n-1})'$).

In our approach the project costs are taken to be fixed. The generalization of our approach where costs are stochastic might be found in the literature on *stochastic programming*. The interested reader could start with [4].

Let $\Sigma = (\sigma_{ij})$ be the covariance matrix, that is σ_{ij} is the covariance of projects i and j and σ_{ii} is the variance (not the standard deviation) of the return of project i. We will make use of the fact that Σ is non-negative definite.

Our method requires two assumptions:

1. The measure for risk is the standard deviation,
2. All the constraints are expressible as linear or parabolic equations.

The next section will show how the second assumption might not be too restrictive.

5.4.1 Constraints

In the last section we mentioned that for our approach to work all the constraints would have to be expressible as linear or parabolic equations.

Naturally there will be a budgetary constraint, perhaps the company must spend less than M_h dollars and more than M_l on this portfolio. Hence

$$M_l \leq X'C \leq M_h.$$

There also might be other constraints such as the choice of one project forbids the option to choose another project. Many of these types of constraints can be written linearly. For example:

- "If project a then also projects b, c and d" can be described as

$$3x_a \leq x_b + x_c + x_d.$$

- "If project e then not projects f or g" can be described as

$$x_f + x_g \leq 2 - 2x_e.$$

- "At least two of h, i, j and k" can be described as

$$x_h + x_i + x_j + x_k \geq 2.$$

- "If project l then not both projects m and n" can be described as

$$x_m + x_n \leq 2 - x_l.$$

Of course it is possible to construct constraints which cannot be written as a linear equation, for example

- "Exactly one or exactly three of projects a, b and c."

But it is simple enough to consider the two different feasible regions corresponding to

- "Exactly one of projects a, b and c" which can be described as

$$x_a + x_b + x_c = 1,$$

- "Exactly three of projects a, b and c" which can be described as

$$x_a + x_b + x_c = 3.$$

Then solve the problem over each of these regions and compare the solutions.

In the introduction we stated that the company can either be fully invested in a project or not invested at all. In reality this might not be true. Perhaps, for a particular project A for example, not only can the company be either completely in the project or completely out of the project, but they might be able to invest in $1/2$, $1/3$ or $2/3$ of the project. In this case we define four new projects:

- Project a: invest in $1/2$ of project A.
- Project b: invest in $1/3$ of project A.
- Project c: invest in $1/3$ of project A.
- Project d: invest in $1/3$ of project A.

Note that projects b,c and d are identical. The constraint is

$$x_b + x_c + x_d \leq 3 - 3x_a.$$

Hence, if we choose to invest in project a ($1/2$ investment in A) then we are forbidden to invest in any of projects b, c and d. Otherwise we are free to invest exactly one of projects b, c and d ($1/3$ investment in A), exactly two of projects b, c and d ($2/3$ investment in A) or invest in all three of projects b, c and d (100% investment in A). Perhaps an interesting generalization would be the corresponding mixed-integer problem where some projects can be included at a continuum of levels.

Define S to be the set of all the constraints. In what follows we assume that S consists of only linear and parabolic equations.

A portfolio X will be called *feasible* if it satisfies all of the constraints in S. Plotting all the feasible portfolios on a graph of expected return versus risk we would get a graph similar to Figure 5.1. Note that due to the finite granularity constraint there will only be a finite number of possible portfolios. Hence, unlike the traditional Markowitz portfolio problems [9] where a continuous "efficient frontier" is expected, the efficient and feasible portfolios from this problem will form a set of "efficient fenceposts". In Figure 5.1 the efficient portfolios (fenceposts) are shown as x's and the dominated portfolios are shown as o's. One portfolio, labeled P_0, is the global minimum for risk and is an efficient fencepost.

Under the assumption that the standard deviation is the measure of risk we can locate point P_0 simply by minimizing the objective function $X'\Sigma X$ under the constraint set S. The next section will describe how to do this. Section 5.4.3 will describe how to locate the other fenceposts.

5.4.2 The Integer Programming Problem

The problem is to find the X which minimizes $X'\Sigma X$ subject to a set S of linear and parabolic equations. Following the outline of a method described in [5] we will describe how to transform this quadratic problem into an equivalent problem which can be solved.

Definition 5.4.1 *A parabolic constraint of rank k is one which can be put into the form*

$$a_{00} - L_0(X) - b_1(L_1(X))^2 - \cdots - b_k(L_k(X))^2 \geq 0,$$

where

$$L_s(X) = a_{s1}x_1 + \cdots + a_{sn}x_n, \qquad s = 0, 1, \cdots, k$$

are a set of $k+1$ linearly independent homogeneous linear forms of n variables and

$$b_i \geq 0, \qquad i = 1, \cdots, k.$$

What we do is transform the objective function $X'\Sigma X$ into a new objective function z and add a parabolic constraint

$$z - X'\Sigma X \geq 0$$

to S. The variable z is called a slack variable. This means that we have to express $z - X'\Sigma X \geq 0$ in the form

$$a_{00} - L_0(X) - b_1(L_1(X))^2 - \cdots - b_k(L_k(X))^2 \geq 0$$

as in Definition 5.4.1. To this end, consider

$$A = \sigma_{00}x_0^2 + 2\sigma_{01}x_0x_1 + \cdots + 2\sigma_{0n-1}x_0x_{n-1}$$

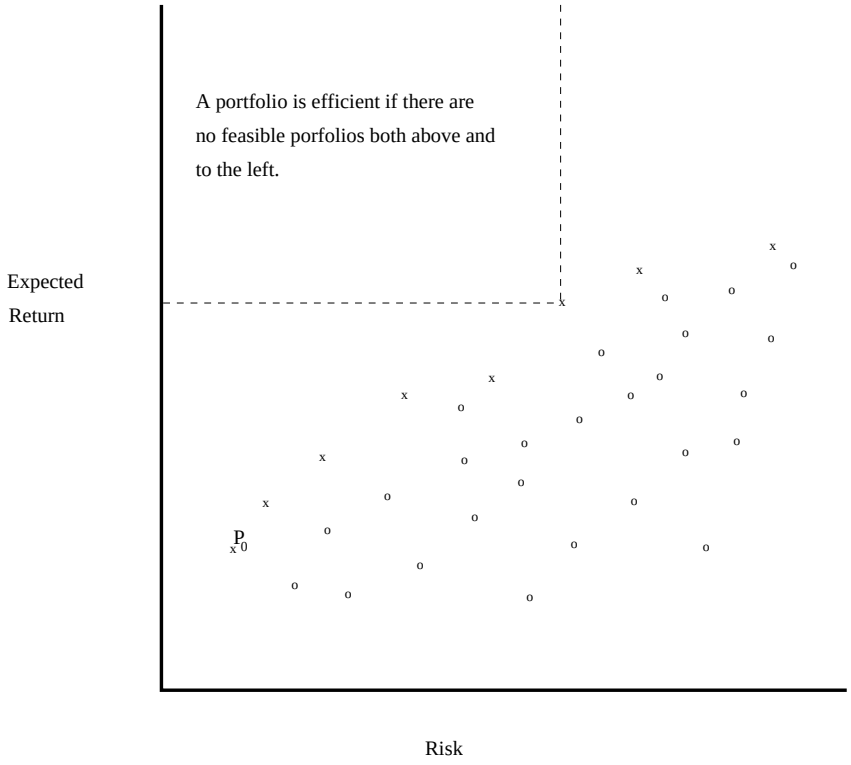


Figure 5.1: A plot of all feasible solutions. Efficient fenceposts are shown as x's. Dominated portfolios are shown as o's. The fencepost corresponding to the globally minimal risk is labeled P_0.

which is all the terms with the variable x_0 in the expansion of $X'\Sigma X$ (Note that the covariance matrix Σ is symmetric). We can rewrite A as

$$A = \frac{1}{\sigma_{00}}(\alpha - (\beta + \gamma))$$

by completing the square, where

$$\begin{aligned} \alpha &= (\sum_{j=0}^{n-1} \sigma_{0j} x_j)^2, \\ \beta &= \sum_{j=1}^{n-1} (\sigma_{0j} x_j)^2, \text{ and} \\ \gamma &= \sum_{1 \le i < j}^{n-1} 2\sigma_{0i}\sigma_{0j} x_i x_j. \end{aligned}$$

Note that $X'\Sigma X - A$, β and γ have no x_0 terms. Thus, since Σ is non-negative definite, $X'\Sigma X$ is a non-negative definite quadratic form (let k be its rank) and therefore the resulting expression $X'\Sigma X - A - \frac{1}{\sigma_{00}}(\beta + \gamma)$ is a non-negative definite quadratic form (with rank $k-1$). We can repeat this process of completing the squares for each variable. Clearly, the resulting expression $z - (\sum_{0 \le i < j}^{n-1} a_{ij} x_j)^2 \ge 0$ satisfies the conditions of the parabolic constraint defined in Definition 5.4.1. After this transformation, our problem can be stated as follows:

$$\min z$$

subject to the original constraint set, S, described in Section 5.4.1 together with the parabolic constraint

$$z - (\sum_{0 \le i < j}^{n-1} a_{ij} x_j)^2 \ge 0$$

where a_{ij} is an appropriate coefficient derived from completing the square for the variable x_j.

The algorithm to solve this transformed problem is rather lengthy and shall be omitted. The interested reader is invited to explore [5], page 277. A number of commercial software packages exist which should solve the transformed problem. An example might be the "Professional Linear Programming System" available from Sunset Software Technology[7] [11].

5.4.3 Finding the Efficient Fenceposts

In the last section we described how to locate the efficient fencepost P_0. We are left with the task of locating the other fenceposts. Associated with each fencepost $P_0 \dots P_m$ there exists an associated expected return $r_0 \dots r_m$ as shown in Figure 5.2. Once we have located P_0 we know the value of r_0. Thus, if we can minimize the objective function $X'\Sigma X$ under the constraint set S plus a further constraint $X'\boldsymbol{\mu} \ge r_0 + \epsilon$ (ϵ is less than a penny and is included to insure our feasible space is closed), then the solution will be P_1. Since the new constraint is linear, the process described in Section 5.4.2 can be used. Iterating this process (until our algorithm returns a "no solution" result) will locate all the efficient fenceposts and hence solve the problem given.

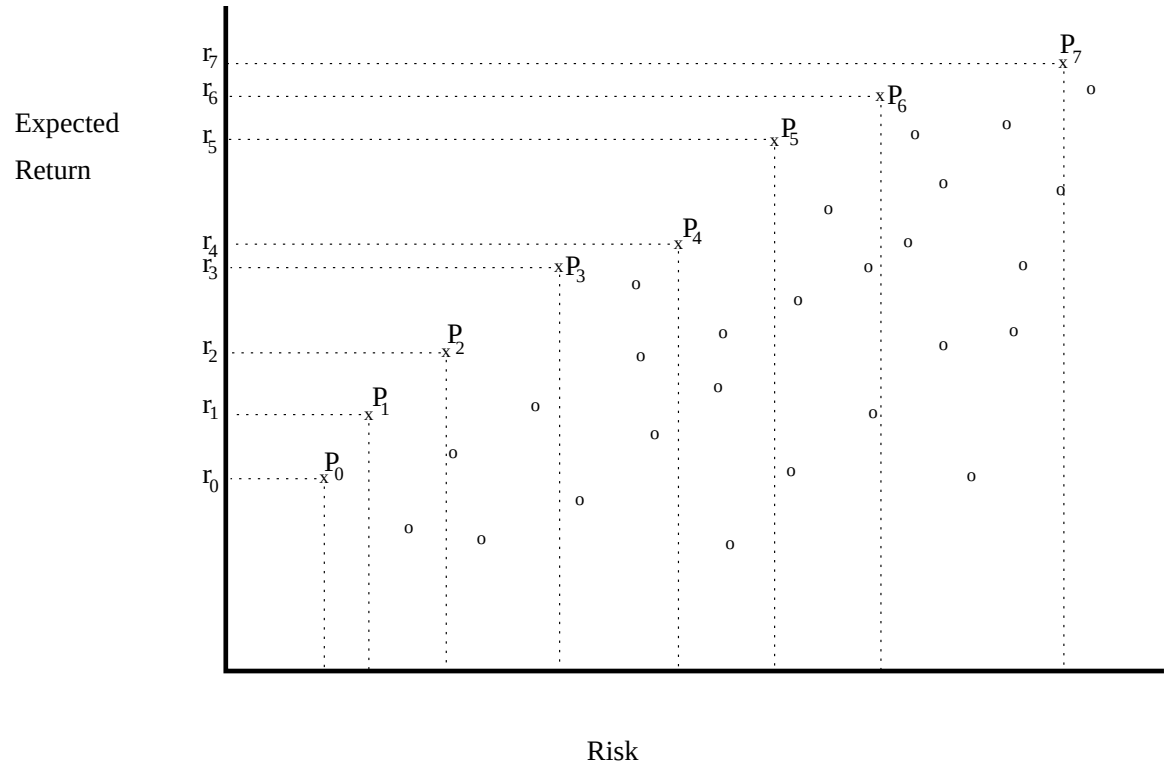


Figure 5.2: Associated with each efficient fencepost $P_0 \dots P_7$ is an expected return $r_0 \dots r_7$.

5.4.4 Conclusion

We have presented a rigorous method for solving the portfolio problem under some assumptions.

[7] www.sunsetsoft.com

The assumption that risk is measured by standard deviation can not be easily relaxed. Although the standard deviation is a common measure for risk, for some reasonable utility functions and non-normal payoff distributions the standard deviation might not be a good measure for comparing portfolios. There are many advantages to using a single parameter risk measure, but, in the case where standard deviation is thought to be inadequate perhaps the concept of *stochastic dominance* should be used. Stochastic dominance is a partial-order relation used to compare payoff distributions. Second-degree stochastic dominance is exactly the concept that an investor is rational. We refer the reader to any graduate text in finance, for example [6], or to [2] for some results which simplify the computations necessary to compare distributions.

We argue in Section 5.4.1 that the assumption that all constraints must be expressible as linear or parabolic equations might not be too restrictive.

The final point we have not yet addressed is the question of computational efficiency. The method described requires multiple solutions to potentially large integer programming problems. These problems can be solved very efficiently but it is possible, especially when there are many projects and many constraints, that the method might be too computationally demanding for commercial applications.

Note that the problem has some nice properties: the variables are binary and the covariance matrix is symmetric. It might be possible to exploit these properties to make the method even more efficient.

In any case, this method can be used to solve a large class of interesting and non-trivial portfolio problems. If the general case does not satisfy the assumptions required by this method or is too computationally demanding for commercial applications, then perhaps this method can be used to test and benchmark more heuristic approaches.

5.5 Summary

Each of the three approaches has its advantages and disadvantages. The Portfolio Selection Algorithm approach is based on the appealing idea of trying to find the efficient portfolio with the minimum amount of effort. It is easy to implement and it works well on examples we have considered. If problems raised in Section 5.2.7 can be successfully resolved, this approach will provide a valuable solution to Merak's problem.

The Statistical Inference Approach is also easy to use. Its main disadvantage is that it does not give the (exact) efficient portfolio. On the other hand, it has the advantage that it can be used in any situation, regardless of the number of projects and the nature of the constraints. In the absence of a universally applicable method which will always find the efficient portfolio, this approach provides a practical solution to Merak's problem.

The practicality of the Integer Programming Approach depends on the number of projects and the nature of the constraints. Nevertheless, it has the advantage that it gives the efficient portfolio for problems where it is applicable. Since most clients of Merak will likely want to know the efficient portfolio if it can be found, for problems where this method is applicable, it is the most preferred approach.

Bibliography

[1] Arnold, B. C., Balakrishnan, N. and Nagaraja, H. N. (1992). *A First Course in Order Statistics*, John Wiley & Sons, New York.

[2] Calistrate, D., Paulhus, M. and Sick, G. (1998). "Using Real Options to Manage Risk", *Proceedings of The 2nd Real Options Conference, Ernst & Young*, Northwestern University.

[3] David, H. A. (1981). *Order Statistics*, Second Edition, John Wiley & Sons, New York.

[4] Dempster, M. A. H. (1980). *Stochastic Programming*, Academic Press.

[5] Hu, T. C. (1969). *Integer Programming and Network Flows*, Addison-Wesley Publishing Company, Inc.

[6] Ingersoll, J. E. (1987). *Theory of Financial Decision Making*, Rowman & Littlefield.

[7] Kalbfleisch, J. G. (1985). *Probability and Statistical Inference Volume 2: Statistical Inference*, Springer-Verlag, New York.

[8] Markowitz, H. (1952). Portfolio Selection. *The Journal of Finance*, Vol. VII, No. 1, pp. 77-91.

[9] Markowitz, H. (1959). *Portfolio Selection: Efficient Diversification of Investments*, Blackwell Publishers, Cambridge MA and Oxford UK, second ed., 1997.

[10] Royall, R. (1997). *Statistical Evidence*, Chapman and Hall, New York.

[11] Sunset Software Technology (1987). "Professional Linear Programming System", 1613 Chelsea Road, Suite 153. San Marino, CA 91108, USA.

Chapter 6

Dynamics of Large Mining Excavators

Chris Budd[1], Huaxiong Huang[2], David Lokhorst[3], Keith Promislow[4], Rex Westbrook[6], Brian Wetton[2]

Report prepared by James Watmough[5]

6.1 Introduction

Operators of large mining excavators could improve their efficiency if they were provided with a real-time knowledge of both the mass in the bucket and the force on the digging teeth of the machine. Although shaped roughly like a standard construction excavator, mining excavators are many times larger. Their payload can be over 40 tonnes, with the hydraulic cylinders operating at up to 30 MPa. Obviously, the intuition and feel that an operator may have for a smaller machine is lost on these goliaths, and any information about the forces on the bucket is important to the operator. During normal operation, one truck is loaded on one side of the machine while a second truck moves into position on the other side. The machine loads vehicles continuously, spending approximately one minute with each truck. Hence, to be of use to the operator, the payload and digging force must be obtained during normal operation of the machine with at most a one second delay.

To be economically viable, a device to measure the payload and digging force should be applicable to any excavator without extensive modelling and fitting costs and should be constructed of robust, low cost sensors limited to measurement of angles between the components of the excavator and pressures in the hydraulic cylinders. One previous approach to the problem was to develop a detailed model of the dynamic behaviour of the machine's arm and bucket. In practice, however, it is not economical to fit such a detailed model to each machine. Further, there are a large number of unknown effects from unmodelled dynamics, such as friction and hydraulic elasticity, and from errors due to sensor limitations and background noise in pressure and angle data.

The ideal device would not require technical fitting or modelling of each machine, but would be a straightforward mechanical installation, with a single device suitable for any machine. The idea that a *black box* be installed and *trained* to each machine leads naturally to the use of parametric or non-parametric regression techniques to estimate the functional form of the dependence of the payload on the configuration of the machine. A non-parametric model assumes only that the payload can be expressed as a function of the independent variables, and that this function can be approximated by a weighted sum of a set of basis functions. Linear regression is used with a set of training data to estimate the weights. In contrast, a parametric model assumes that the payload can be expressed as a known function of the independent variables and several parameters. The training data is then used to estimate the parameters.

[1]University of Bath
[2]University of British Columbia
[3]RSI Technologies
[4]Simon Fraser University
[5]University of Victoria
[6]University of Calgary

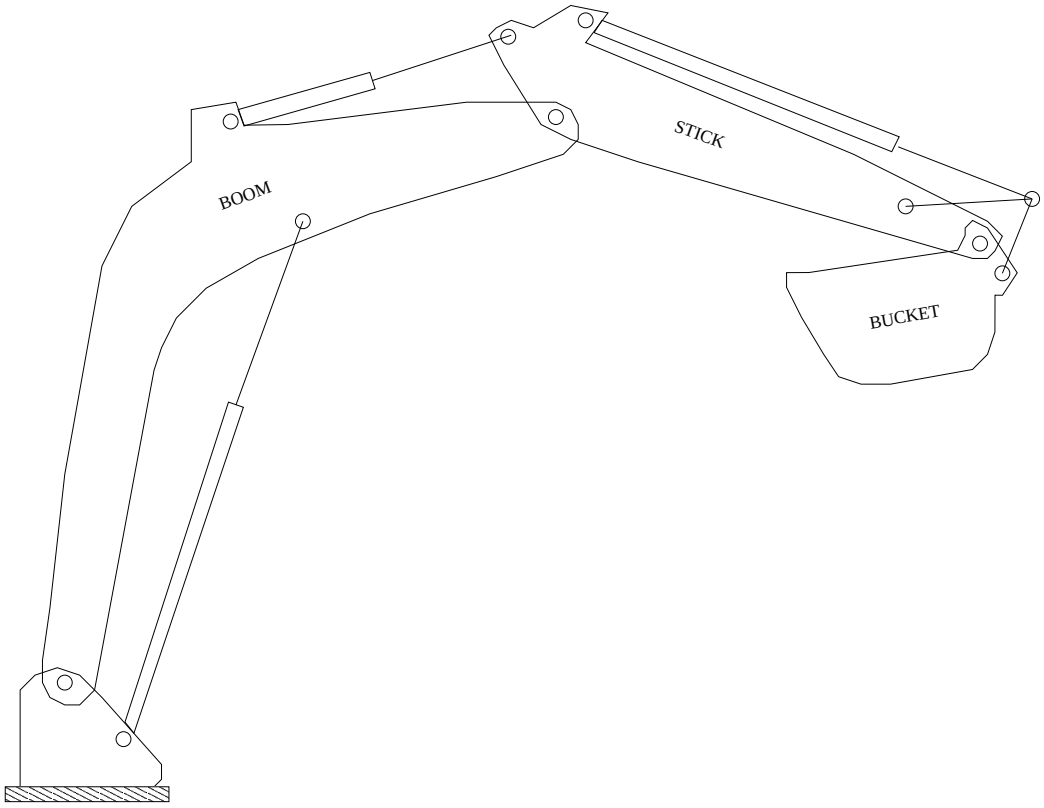


Figure 6.1: A typical excavator arm.

In this report we compare the use of parametric and non-parametric regression models and their suitability to the problem of determining the mass in the bucket. The problem of determining the force on the teeth of the bucket can be modelled in a similar fashion. However, it is not clear that there is a single force acting on the teeth of the bucket during digging and not a more complex contact between the bucket and the ground. For this reason, this report focuses on the problem of determining the payload mass. The problem of measuring the digging force on the bucket teeth is not dealt with in any detail.

The investigation of the non-parametric regression model indicates that the black box approach is impractical due to the large amount of training data needed to estimate the payload function. It would seem that both extremes, the detailed model on one hand, and the black box on the other, require too much time and effort to measure parameters to be of practical use. The parametric model requires far less data for training than the non-parametric regression model, and even though a detailed model of each type of machine must be developed, an exact measurement of the model parameters is not required. Hence, the parametric model appears to be the best approach.

6.2 Modelling approach

Figure 6.1 shows a typical excavator consisting of a boom, stick and bucket connected by pin joints and hydraulic cylinders. Using techniques from dynamics and robotics, the mass in the bucket can be specified as a function of the pressures in the hydraulic cylinders, the angles between each component and their derivatives (see Section 6.3). Hence, if all the parameters in this model could be determined, then determining the mass in the bucket would be a simple matter of evaluating the function given a time history of the angles and pressures. Unfortunately, this approach is impractical for several reasons. Obviously, even with a detailed model there are unmodelled features, such as friction in the pins and cylinders, and the hydraulic elasticity and inertia of the cylinders. There is also a large amount of noise in the data, particularly in the measurement of angular accelerations. However, the

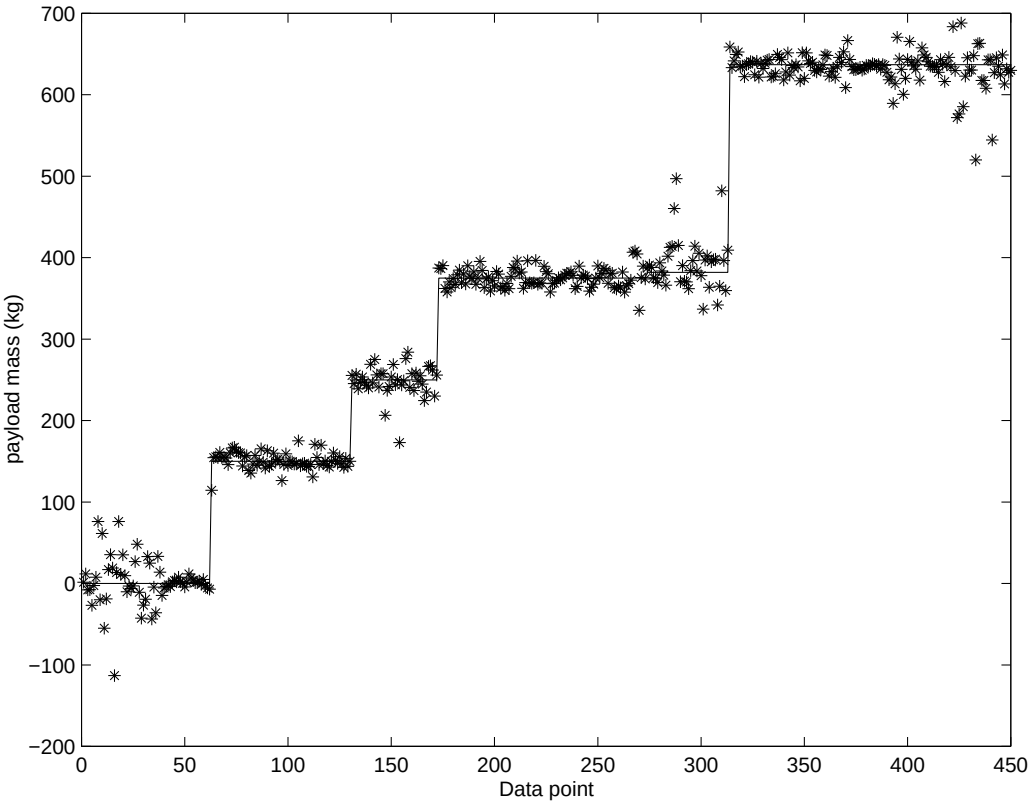


Figure 6.2: Approximation of the payload using radial basis functions. Evaluation of the function at the training points.

most significant drawback using a detailed model is the impracticality of measuring all the parameters for each machine. To be economically viable, a device to measure the payload should not require detailed measurements for each machine.

Another approach, diametrically opposed to the detailed model is to use non-parametric regression to estimate the payload function. Routines implementing these techniques are widely available. In fact, a simple search of the Internet revealed a set of public domain Matlab routines[7] for the approximation of functions using radial basis functions [2]. Briefly, this approach assumes that the function in question can be approximated by a weighted sum of radial basis functions and uses linear regression to determine the weights. Some sophistication is added to determine the most suitable set of basis functions. To test this method, data made available from RSI Technologies was used to construct an approximation to the payload function. The data was obtained by placing known masses in the bucket of an excavator and measuring the pressures and angles as the configuration of the arm was varied. To reduce the complexity of the problem, the testing was done only on a subset of this data for which the configuration was static. In this case, the payload function depends on the three angles describing the configuration of the arm and the pressure in one cylinder, but not on the angular velocities and accelerations. Figure 6.2 shows the plot of the approximated function for the data points used in the regression. To test the ability of the approximation to interpolate, data for a mass of 275kg was reserved and the model was trained on the remaining data. The resulting payload function was then evaluated for the 275kg data. The results are shown graphically in Figure 6.3. The method is clearly unable to interpolate between masses in the training data. Since the device is expected to be used for a wide range of payload masses, it is impractical to train the device to the level of detail required. We conclude, therefore, that the black box approach of non-parametric regression is not an economically feasible option.

[7] http://www.anc.ed.ac.uk/∼mjo/rbf.html

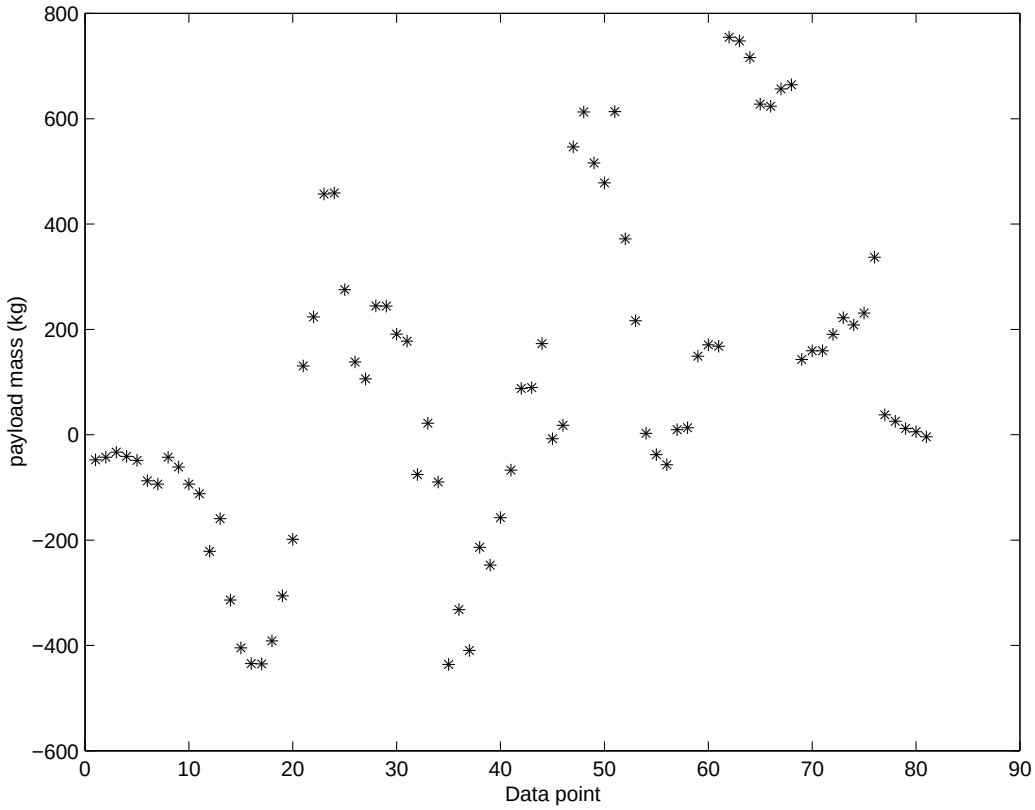


Figure 6.3: Approximation of the payload using radial basis functions. Evaluation of the function for masses not used in training.

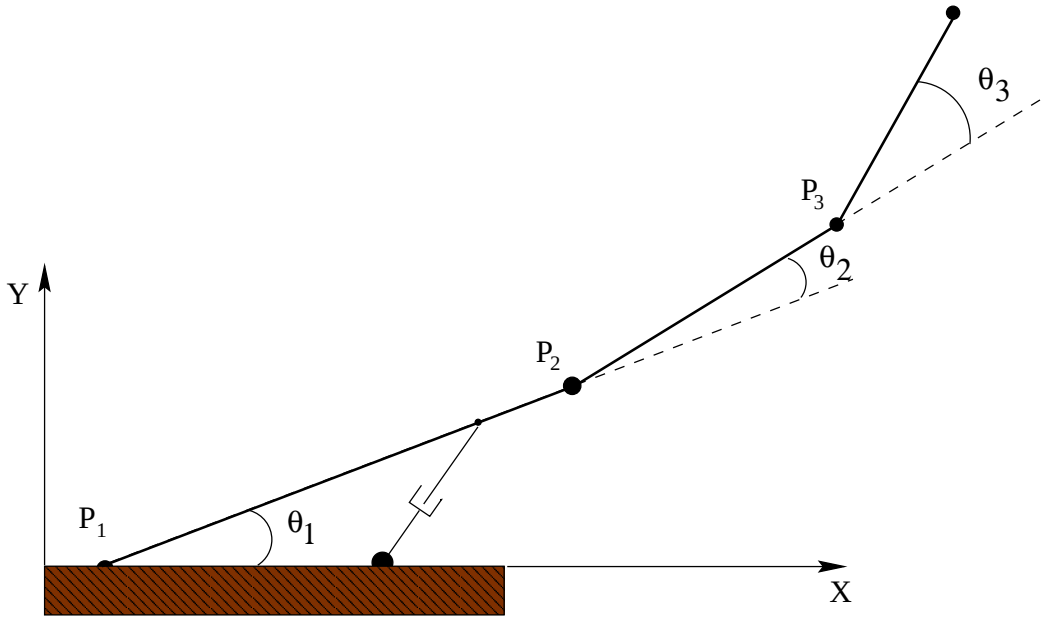


Figure 6.4: Coordinate system for the general model.

The poor ability of the non-parametric approach derives from the local nature of the approximation. In order to interpolate new masses not in the training data, some knowledge of the global form of the function is required. In the following section we develop a detailed model for an excavator arm. However, rather than measure parameters directly from the machine, we propose to use regression analysis of data for known weights to estimate the parameters in the model. In general, there is a non-linear dependence of the payload mass on the independent variables and the unknown parameters. To avoid the complications of using non-linear regression algorithms, the ideas are illustrated in Section 6.4 using a simple one-armed digger. For this model digger, the payload is a linear function of two parameters, allowing linear regression to be used for their estimation.

6.3 Model Excavator

The excavator shown in Figure 6.1 is generalized to the three link arm shown in Figure 6.4. A dynamic model of this generic excavator is developed using the Lagrangian approach. Each component of the digger is replaced by a single link. The links are connected by pins which allow each link a single degree of freedom and constrain all the links to lie in the same plane. The links are numbered from 1 through 3, with θ_1 being the angle of the first link with the horizontal plane and θ_i, $i > 1$, being the relative angle between links $i-1$ and i. The angle θ_0 is reserved for the rotation of the base around the Y axis.

Lagrange's equations of motion [1] specify that the energies satisfy the following system of equations:

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\theta}_i}\right) - \frac{\partial L}{\partial \theta_i} = Q_i, \qquad i = 0, \ldots, n, \tag{6.1}$$

where Q_i is the rate of work done on the system by the non-conservative forces with increases in the angle θ_i. The digger arm is subject to conservative forces, due to gravity, and non-conservative forces through the hydraulic cylinders between each link and the torque applied to the rotating base.

A coordinate system is attached to the fixed base of the first link. The X axis is aligned with the digger arm, the Y axis vertical to the base and the Z axis perpendicular to the arm in the horizontal plane. A translated coordinate system for link i is attached to the pin connecting link $i-1$ to link i. P_i is the position vector of the i^{th} link in the fixed coordinate system.

Each link is a rigid body undergoing rotation and translation. Link i rotates with angular velocity $\omega_i = (0, \dot{\theta}_0, \sum_{j=1}^{i} \dot{\theta}_j)$ and it's base (the point P_i) translates with velocity v_i, where

$$\begin{aligned} v_1 &= 0, \\ v_i &= v_{i-1} + (P_i - P_{i-1}) \times \omega_{i-1}. \end{aligned}$$

The kinetic and potential energy of each link depends on the mass m_i, centre of mass r_i and inertial tensor I_i. For convenience, the centre of mass and inertial tensors are taken relative to position P_i for each link. Obviously, since the link is rotating rigidly in this coordinate system, r_i and I_i depend in a known fashion on the link angles θ_1 through θ_i. The kinetic and potential energy of each link are given by the expressions

$$T_i = \frac{1}{2} m_i (v_i \cdot v_i) + m_i v_i \cdot (\omega_i \times r_i) + \frac{1}{2} \omega_i I_i \omega_i^t, \tag{6.2}$$

$$V_i = m_i (P_i + r_i) \cdot (0, g, 0), \tag{6.3}$$

where g is the acceleration due to gravity. That is, T_i and V_i are the kinetic and potential energies of link i and $L = \sum_{i=1}^{3} T_i - V_i$ is the total energy use in Equation 6.1.

Work is done by the non-conservative forces as the lengths of the hydraulic cylinders change. Let $h_i(\theta_i)$ be the length of the cylinder connecting link i with the previous link, or the base in the case of the first link. Since each cylinder connects neighbouring links, the work corresponding to changes in θ_i depends only on the function $h_i(\theta_i)$ and the forces on the i^{th} cylinder as follows

$$Q_i = F_i h_i'(\theta_i). \tag{6.4}$$

The force F_i is proportional to the pressure difference across the cylinder minus any frictional forces. It is hoped that these frictional forces are small enough to be neglected. However, in the event that they are not, they may be assumed to be in proportion to the velocity of the cylinder $h'(\theta_i)\dot{\theta}_i$.

The excavator bucket is represented as the third link. Hence the payload mass is m_3. Since the potential energy of the final link, P_i, depends on each of the angles, the payload mass appears in equation in the system given by Equation 6.1. That is, any one of these equations can be used to derive a function for the payload in terms of the angles, cylinder pressures and parameters of the system. Note, however, that each equation depends only on a single pressure. Hence, for the non-parametric approach, the payload function can be assumed to depend on the three angles, their time rates of change, and any single pressure.

For the parametric approach, we are interested in how the payload function depends on the unknown parameters. These unknown parameters are the masses, m_i, and the moments $m_i r_i$ and I_i of the links. Since the energies and generalized forces are linear in these parameters, the payload mass is a rational function of the parameters, linear in both the numerator and the denominator. Hence, the parametric approach entails a non-linear regression analysis.

6.4 PIMS Digger

To compare the parametric and non-parametric approaches, consider the simple digger shown schematically in Figure 6.5. The base is fixed, and the configuration changes only with the angle θ. As there is only a single link of length l which does not rotate, the subscripts are dropped. The centre of mass of the link is assumed to lie on the link at a distance x from the base. However, in addition to the mass of the link, a mass M representing the payload is attached to the end of the link. Thus, $v_1 = 0$, $\omega_1 = \dot{\theta}$ and $r_1 = (x\cos\theta, x\sin\theta)$, and Equations (6.2), (6.3) and (6.4) become

$$T = \frac{1}{2}\dot{\theta}^2 I + \frac{1}{2} M l^2 \dot{\theta}^2 \tag{6.5}$$

$$V = (Ml + mx)\sin\theta \tag{6.6}$$

$$Q = h'(\theta) p A \tag{6.7}$$

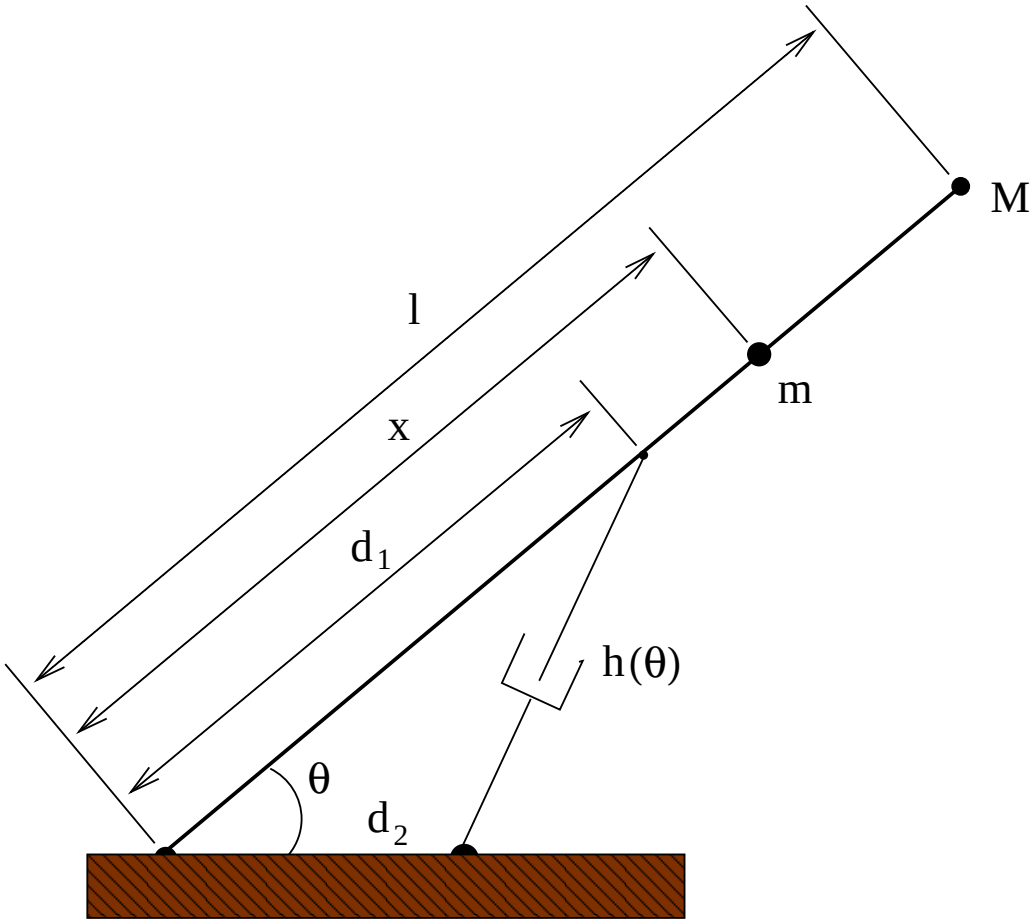


Figure 6.5: The PiMS Digger: a simple, single degree of freedom digger.

Substituting the required derivatives into Equation 6.1, leads to the relation

$$\left(I + Ml^2\right)\ddot{\theta} + g\left(Ml + mx\right)\cos\theta = h'(\theta)pA,$$

which can be solved for the payload mass. This is expressed as

$$M(p,\theta) = \frac{h'(\theta)pA - gmx\cos\theta - I\ddot{\theta}}{l^2\ddot{\theta} + gl\cos\theta}. \tag{6.8}$$

The length h is given by the expression

$$h(\theta) = \sqrt{d_1^2 + d_2^2 - 2d_1d_2\cos\theta},$$

and so

$$h'(\theta) = d_1d_2\sin(\theta)/h(\theta).$$

The lengths d_1 and d_2, which describe the configuration of the piston, are assumed to be known to any desired accuracy. Whereas, c, m I, and l can only be estimated. This is representative of the real case where the centre of mass of the payload cannot be determined in advance simply because the distribution of the load in the bucket is not known. Further, the centre of mass and the inertial tensors of each link cannot be determined for each machine due to economic constraints.

For the dynamic equation, the mass is a non-linear function of the three parameters. To avoid the complications of non-linear regression consider the static case. In this case, the mass can be represented as a function of the pressure, angle and the two unknown parameters, c_1 and c_2 as and Equation 6.8 becomes

$$M(p,\theta) = c_1p\tan\theta/h(\theta) - c_2. \tag{6.9}$$

where $c_1 = d_1d_2A/gl$ and $c_2 = mx/l$. Note, that since many of the unknown parameters cannot be determined to any desired accuracy, non-dimensionalization is not practical. For simplicity, the units kilograms, meters and seconds will be used, as these are relevant to the scale of the mining excavators. Figure 6.6 shows the contours of constant mass for the simple model as a function of the single angle θ and the pressure, p in the single cylinder using the parameter values $c_1 = 10$ m^{2}s^2 , $c_2 = 2$ kg, $d_1 = 2$ m, $d_2 = 1$ m.

In this simple model we have the option of using radial basis function to approximate this function, or using linear regression with *training* data for M, θ and p to estimate the parameters.

To illustrate the effectiveness of the parametric and non-parametric approaches to the problem, two data sets were generated for testing. The first data set was generated by choosing 1000 random points in the p-θ plane, computing the mass, $M(p,\theta)$ from Equation 6.9 and adding a small amount of noise to all three variables. This data is shown in Figure 6.7a. The 'plus' signs represent data used for training the model and the circles represent data used to test the resulting payload function. A second set of data, shown in Figure 6.7b, was generated using only three values of the mass M (20 kg, 30 kg and 40 kg). Equation 6.9 was used to determine the pressure p as a function of the mass M and angle θ. This function was then used to determine the pressures corresponding to the three masses and a random sample of angles. Once again, noise was added to the resulting points to represent real errors in measurement. The data points corresponding to the largest and smallest masses were used for training, and the data for the middle mass was used for verification.

The function $M(p,\theta)$ obtained using approximation by a radial basis functions trained on these two data sets is shown in Figure 6.8. Comparing the two contour plots to the exact contour plot of Figure 6.6 clearly shows the need to train the approximation on a large, uniformly distributed sampling of data.

Since the model is linear in the two parameters, it is straightforward to perform a linear regression using Equation 6.9 to estimate both c_1 and c_2 for each data set. Figure 6.9 shows the contour plots of the approximate functions using the two data sets. Although the errors are much larger for the second set, the parametric method produces far better results than the non-parametric method.

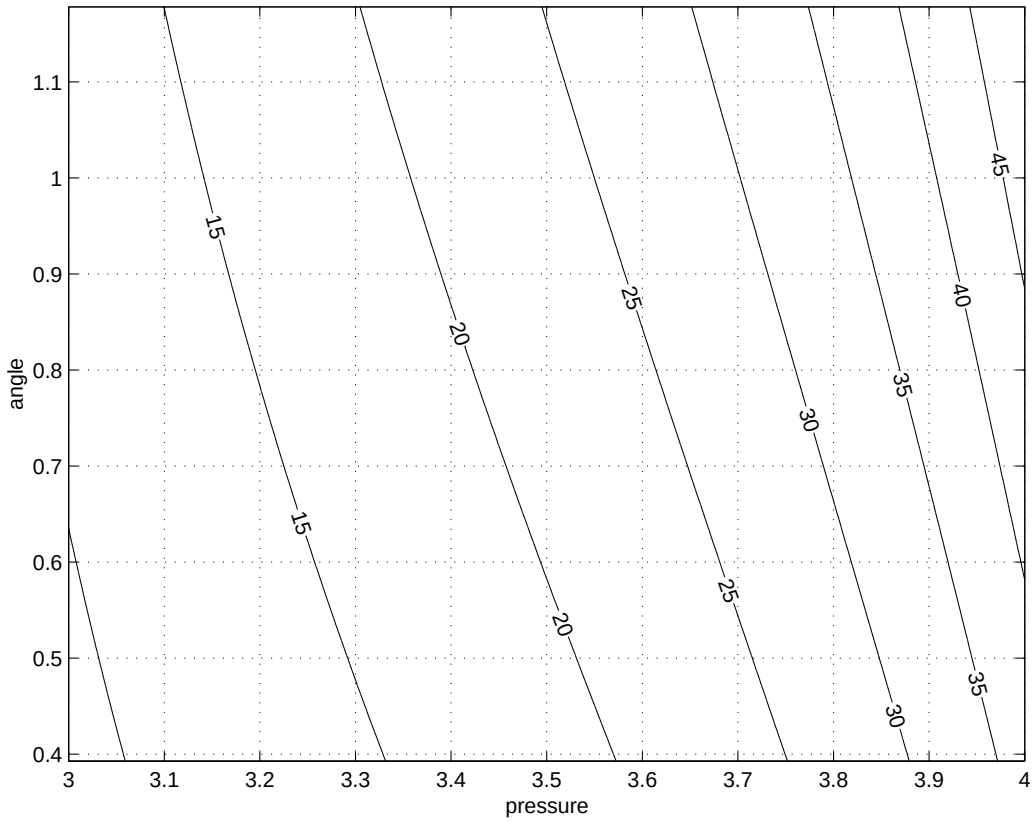
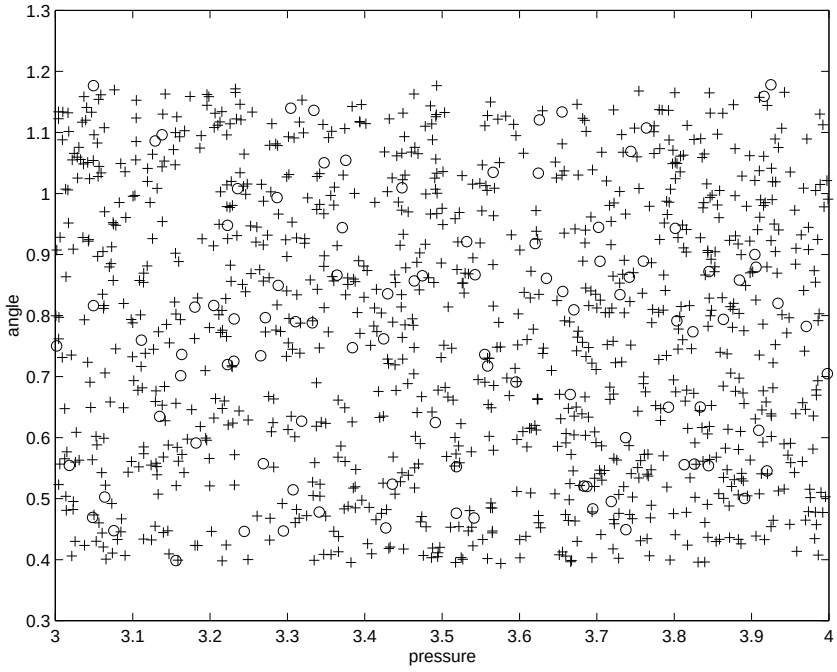
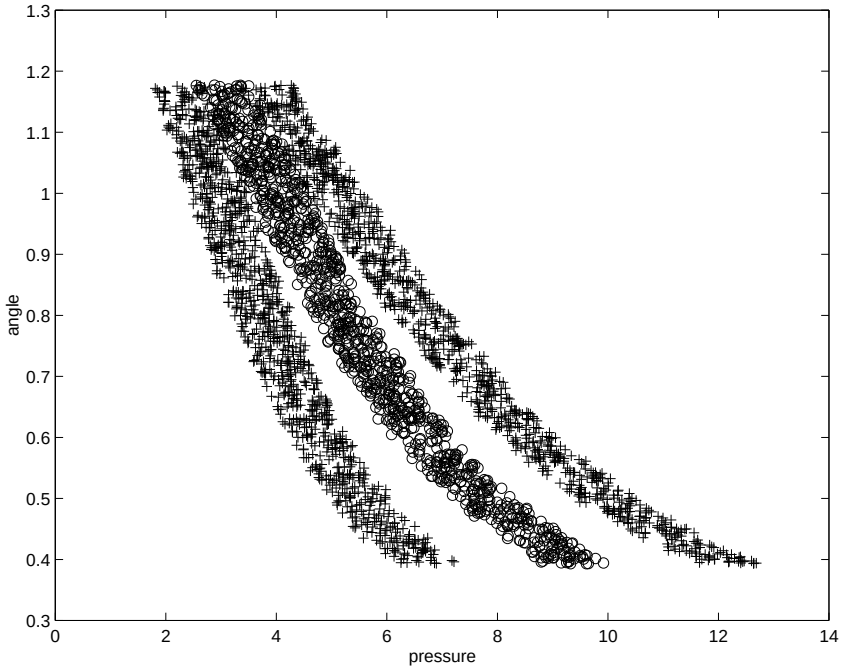


Figure 6.6: Contours of the exact payload function for the PiMS Digger

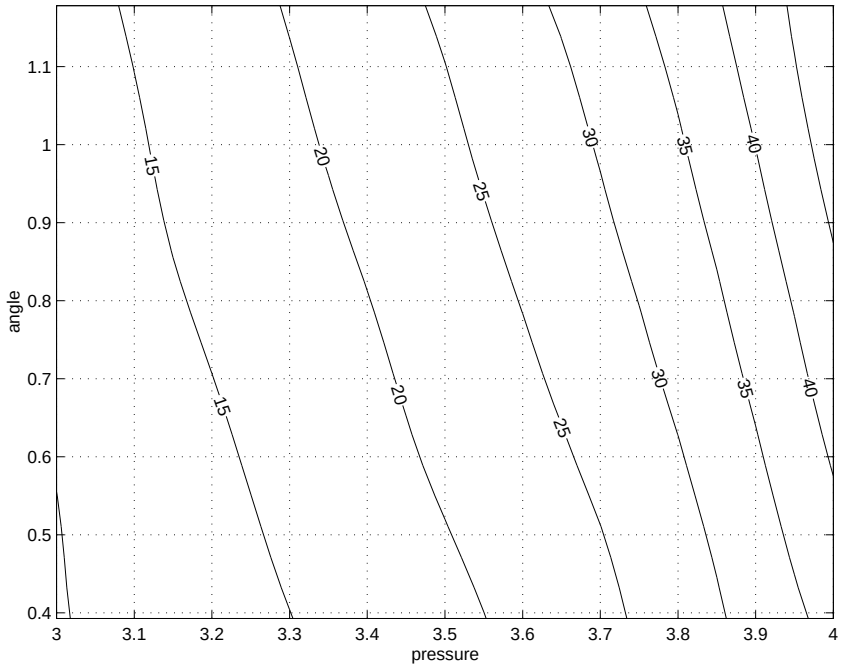


(a) Data set A

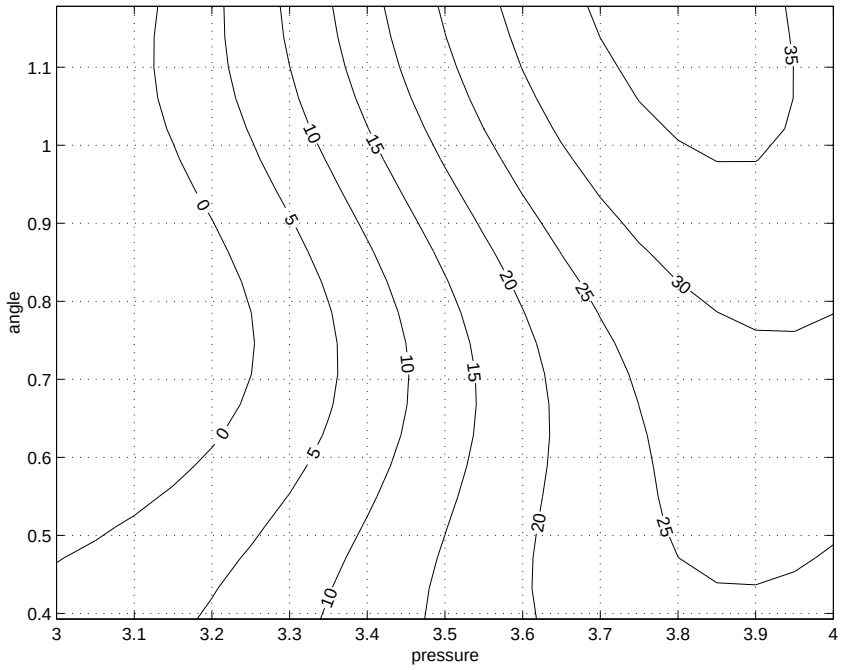


(b) Data set B

Figure 6.7: Data sets for training and testing of the models. The training data is marked by cross-hairs and the testing data by circles.

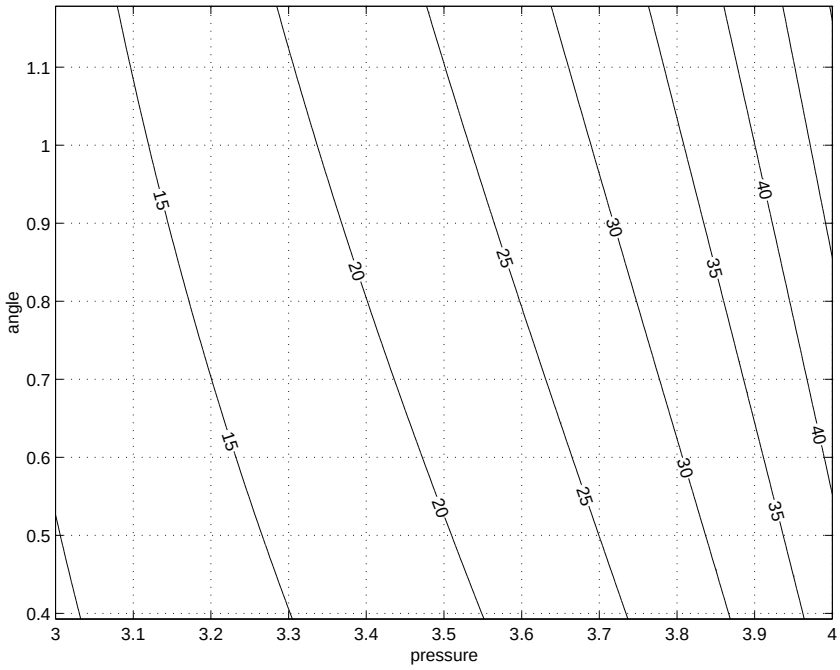


(a) Data set A

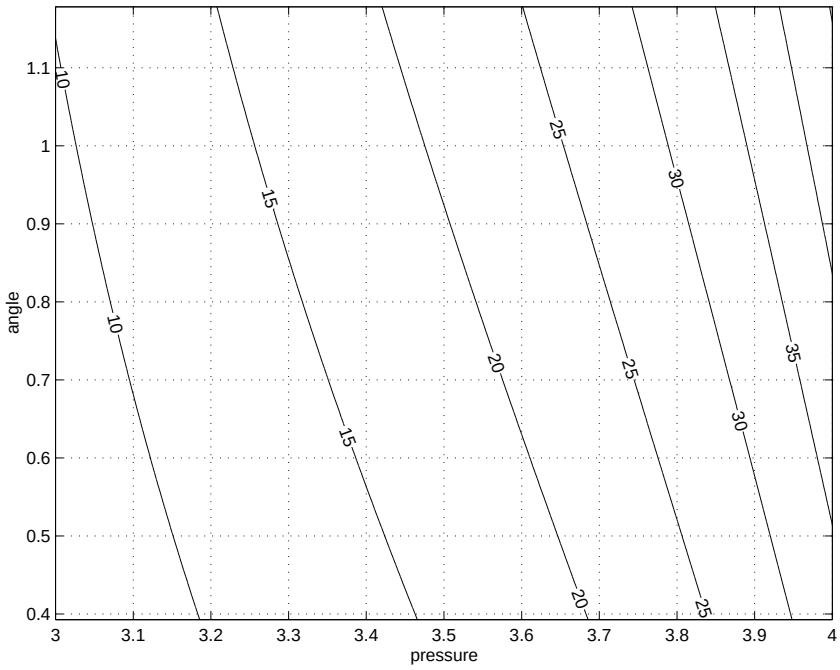


(b) Data set B

Figure 6.8: Contours of the non-parametric approximations to the payload function for the PiMS Digger. Top, data set A, bottom, data set B



(a) Data set A



(b) Data set B

Figure 6.9: Parametric approximations to the payload function for the PiMS Digger. Top, data set A, bottom, data set B.

6.5 Conclusions

Our testing of both parametric and non-parametric methods of reconstructing the mass function from data indicate that the non-parametric *black box* approach to the problem is not likely to be economically feasible. Although in theory the approach is simple to implement and can be applied to a variety of machines with no additional modelling, in practice, the number of data points necessary for the training of each box is too large. On the other hand, it is clear that the parametric approach can be used to estimate the mass function without knowledge of the exact geometry of the machine, so that a single, parametric, 'grey' box can be applied to a wide variety of machines with the same basic geometry.

In this report we considered only a simple linear model to illustrate the advantages of parametric regression over non-parametric regression for this particular application. For the full machine, it is likely that non-linear regression techniques must be applied. It remains to be shown that these methods will converge for the data available for training. Since the normal operation of the excavator is not to be interfered with, the actual device would necessarily use the full dynamic model rather than the simpler static model explored here. Interestingly, there is less uncertainty in the dynamic motion of the digger arm than in the static situation. In the kinetic case, the frictional force in a cylinder is a function of the piston velocity. In contrast, the magnitude of the static frictional force will depend on the history of motion. That is, even in theory, when frictional forces are included, the mass in the bucket can not be determined from pressure and angle data alone. A continuum of configurations are possible for the same mass.

Bibliography

[1] Jerry H. Ginsberg. *Advanced Engineering Dynamics.* Harper & Row, New York, 1988

[2] Mark Or. *Introduction to Radial Basis Function Networks.* Technical report, Institute of Adaptive and Neural Computation, Division of Informatics, Edinburgh University, 1996

Chapter 7

Classification of Chemical Compound Pharmacophore Structures

Lynn Batten[1], C. Sean Bohun[2], Kell Cheng[3], Tom Doman[4], John Drew[5], Rod Edwards[2], Susan Kutay[3], Devon McCrea[2], Wendy Myrvold[2], Frank Ruskey[2], Joe Sawada[2], Pauline van den Driessche[2], Jana Vander Kloet[3], Kathryn L.B. Wood[2].

Report prepared by Andrej Bona[3] and Claude Laflamme[3].

7.1 Introduction

There is a staggeringly large number of chemical compounds that can be made; even when one considers only those hypothetical molecules that are "drug-like", there are still probably on the order of 10^{100} unique molecular structures. To date, only on the order of 10^7 compounds have been made and characterized. There is consequently no hope of producing the totality of known chemical compounds at any time in the future.

However, many of the compounds that have been synthesized, as well as those that could be imagined in the size 10^{100} set, are structurally close to each other and would be expected to interact with biological molecules such as enzymes (according to Fischer's "lock & key" concept) in a similar fashion. One promising notion of structural similarity is the concept of a "pharmacophore", where compounds matching the same pharmacophore often interact with biological molecules (enzymes, receptors, etc) in a similar way. For example, the pharmacophore fingerprint can provide valuable information in the search for novel active compounds, in the measure of diversity of a screening library, and of structure-activity relationships between these compounds.

We are also particularly interested in classifying low molecular weight chemical compounds for drug discovery projects that can inhibit the action of an enzyme by lodging themselves within its active site. Since the geometry of these compounds must match the configuration of the receiving protein, the pharmacophore information could be used to improve the success rate.

A pharmacophore is a structural abstraction of the interactions between various functional group types in a compound. They are described by a spatial representation of these groups as centres (or vertices) of geometrical polyhedra, together with pairwise distances between centres. The corresponding groups under consideration are typically *hydrogen bond acceptor* (HACC), *hydrogen bond donor* (HDON), *negative* (NEG) and *positive* (POS) charges, and *hydrophobic* (HYD) groups. Other groups can easily be incorporated into the analysis. The centre-to-centre distances have been estimated at between 2Å to 15Å,[7] but again, other distance ranges can be easily

[1]University of Manitoba
[2]University of Victoria
[3]University of Calgary
[4]Searle Corporation
[5]College of William and Mary
[6]University of Calgary
[7]1Å = 10^{-10}m

adapted. Typically, the edge lengths are divided into distance intervals, where every pharmacophore with edge lengths in the same specified intervals yields similar chemical properties.

We provide an analysis that facilitates counting 3 and 4 centre pharmacophores, including a mathematical model for distance interval ratios, triangle and other inequality requirements for feasible triangles and tetrahedra, and symmetries. We also include some special cases as an indication of the sheer number of relevant pharmacophores even under very specific limiting circumstances.

Beside spatial symmetries and distance similarities for each edge of the polyhedra, there does not appear to be any other relevant structural similarity feature between two pharmacophores that can be used to reduce the classification of a typical compound.

7.2 Symmetries of the tetrahedron

To get very rough, but quite general estimates on the number of possible 3-centre and 4-centre pharmacophores, a well known mathematical counting principle called Burnside's lemma can be used. This formula counts the exact number of non-isomorphic triangles and tetrahedra with k possible edge types and c possible types of vertices (centres), taking into account equivalence classes of symmetries given by a specified group G acting on a triangle or tetrahedron, as given by:

$$T(k,c) = \frac{1}{|G|}\sum_{g\in G}\psi(g),$$

where $\psi(g)$ is the number of configurations fixed by g.

In the case of the 3-centre pharmacophore (represented by a triangle $\triangle$), symmetries correctly include mirror images. For the 4-centre case (represented by a tetrahedron $\boxtimes$), mirror images (a reflection through any one of the faces) are distinct in the chemical sense, and must be excluded from the computation.

The resulting estimate on the number of 3 and 4 centre pharmacophores is given by the following table:

symmetry type:	description	num ($\triangle$)	$\psi(g)$	num ($\boxtimes$)	$\psi(g)$
id	identity	1	k^3c^3	1	k^6c^4
(ab)(cd)	antipodal edges			3	k^4c^2
(abc)	triangular face	2	k^1c^1	8	k^2c^2
(ab)	flip	3	k^2c^2		
Total	$\|G\|$	6		12	

Table 7.1: Pharmacophores with 3 and 4 centres.

Note that for the tetrahedron each of the resulting permutations is even (the product of an even number of transpositions) and there are 12 of them. Thus the group is the alternating group, A_4. This fact is used later in the program as it computes canonical representatives of each tetrahedron.

For the triangle, the Burnside computation gives us

$$T(k,c)_{\triangle} = \frac{1}{6}(k^3c^3 + 3k^2c^2 + 2kc).$$

For the tetrahedron we get

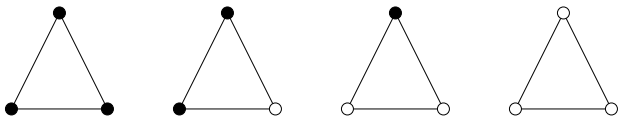
$$T(k,c)_{\boxtimes} = \frac{1}{12}(k^6c^4 + 3k^4c^2 + 8k^2c^2).$$

Here are listed different values of $T_{\triangle}$ and $T_{\boxtimes}$ for various values of k and c:

k	c	$T_{\triangle}(k,c)$	$T_{\boxtimes}(k,c)$
1	2	4	5
1	5	35	75
1	7	84	245
2	1	4	12
31	1	5456	74190161
31	5	632710	46229938075
31	7	1726669	177586039365

Table 7.2: Values of $T_{\triangle}$ and $T_{\boxtimes}$.

The actual configurations for $T(1,2)$ are shown below.



7.2.1 More than 4 centres

To create an $n+1$-centre pharmacophore from an n-centre pharmacophore requires taking 3 additional edge distances and a particular face of the previously constructed polyhedron; this is sufficient to determine each new centre's location in space. The dominating term obtained from Burnside's lemma is therefore at most

$$\frac{1}{|G|}k^{3(n-2)}c^n,$$

where again n is the number of the centres, G is the group of automorphisms when these centres are placed in the "most symmetric way" on a sphere.[8] As an example, the dominating term in the number of 8-centre pharmacophores ($n=8$, giving $|G|=24$), with 30 possible edge types ($k=30$), and 5 centre types ($c=5$), is:

$$\frac{1}{24}30^{18}5^8 \simeq 6.3\times 10^{30}.$$

7.2.2 Which distances give triangles and tetrahedra?

A more realistic count must consider the distances between centres, because an arbitrary set of distances is not necesarily achievable by triangles or tetrahedra. In the triangle case, a necessary and sufficient condition for numbers $a \le b \le c$ to be realized as side lengths is that $c \le a+b$ (triangle inequality). For the tetrahedra, this condition holding on each triangular face is clearly necessary, but it is not sufficient. There are some necessary and sufficient conditions for distances to be achievable by a tetrahedron, but the following approach is more suitable for our purposes. Assume that two triangular faces have been determined — the triangles with sides a,b,c and b,d,e shown in Figure 7.1. Possible values that f can have must be determined.

The figure shows the two extreme cases for the length of f; in both cases the tetrahedron lies flat in the plane. Thinking of b as a hinge, all the possible tetrahedra with all distances except f fixed will be obtained. Noting that the angles are

$$\gamma = \cos^{-1}\left(\frac{a^2+b^2-c^2}{2ab}\right)$$

[8] For chemical reasons we can assume that all centres lie on the convex hull of those centres (vertices).

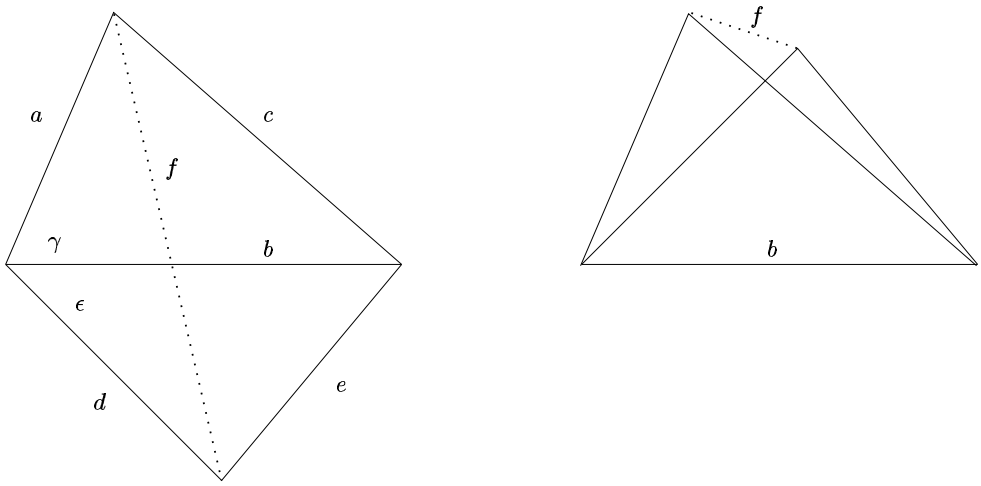


Figure 7.1: Making a tetrahedron. Hinge is b. Extreme values for f shown.

and

$$\epsilon = \cos^{-1}\left(\frac{d^2 + b^2 - e^2}{2db}\right),$$

the possible values of f are bounded as follows

$$a^2 + d^2 - 2ad \cdot \cos(\gamma - \epsilon) \leq f^2 \leq a^2 + d^2 - 2ad \cdot \cos(\gamma + \epsilon). \tag{7.1}$$

Adding the triangle inequality constraints and (7.1) greatly reduces the number of possible pharmacophores. The exact magnitude of the reduction will depend on the number and lengths of the intervals on the edge lengths.

7.3 Chemistry

To further reduce our estimate of the number of pharmacophores, some possible chemical restrictions are considered in this section. A pharmacophore is a representation of the generalized molecular features that are considered to be responsible for a desired biological activity.[9] To improve the library of potential molecules of pharmacological activity, Searle is interested in developing a method that would improve the testing efficiency of drug-like molecules.

As stated in the introduction, the following five general categories of molecular features available for interaction with an enzyme are considered: hydrophobic [HYD], hydrogen bond donating [HDON], hydrogen bond accepting [HACC] and ionic sites which consist of positively [POS] and negatively [NEG] charged functional groups. Commercially, six to seven groups have been used to describe the active sites; however, we have constrained the number of pharmacophoric groups to five by grouping the aromatic and hydrophobic categories and by not considering acidic and basic functional groups as separate from the other features.

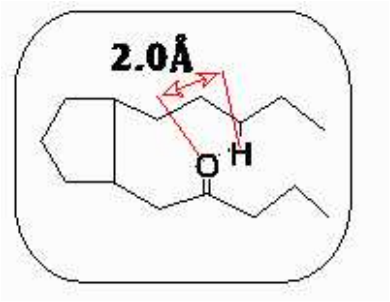
Lower bounds for the distances between the various possible combinations of these five types of sites need to be defined in terms of distances between the groups on a particular molecule based on bond lengths, bond angles and approximate degrees of rotation. When the active sites interact through space rather than the molecular structure of the compound, the sum of the van der Waals radii would be the limiting case in the calculation for the distance of closest approach. Here are listed some van der Waals radii for some important atoms: O: 1.4Å; N: 1.5Å; H: 1.2Å.

Pharmacophore space is defined by the relative distance between the centres of each site based on the binding opportunities of the enzyme active sites. The upper bound was defined by a reasonable distance (15Å) for the low molecular weight enzyme inhibitors. The basis for the distances are justified as follows:

[9]For more information please consult the web-site at `http://Intral.msi.com/weblab/medchem/mc_align.html`

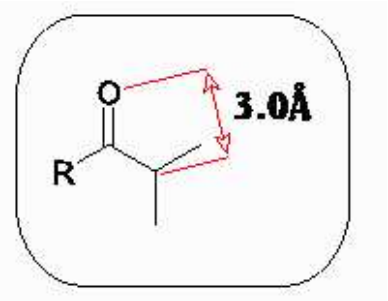
- HDON or HACC site with another HDON or HACC site:

  Given the possibility of this molecular structure, two sites of the HDON or HACC type could be located on the hydrocarbon tails, which due to the flexibility of the tails, could approach distances as small as 2Å. Even a change in distance of 0.15Å could diminish or remove the beneficial interaction with the enzyme receptor because of the sensitivity of these interactions.



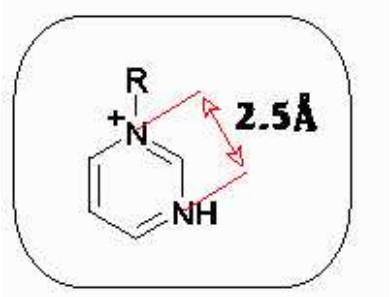
- An HDON - HACC site with a HYD site:

  The limiting distance between a hydrophobic centre and either an HDON or an HACC was determined by examination of a ketone/aldehyde site bonded to a hydrocarbon propyl chain. The centre of the HYD propyl group was based on the approximate centre of mass and the shortest distance between these sites was determined by the bond lengths and angles.



- An HDON - HACC site with a POS site:

  The closest approach between these two types of sites was found to be 2.5Å based on the bond angles and bond lengths of the ring illustrated. The figure shows that one nitrogen atom could behave as a positively charged site while the other could act as an HDON group.

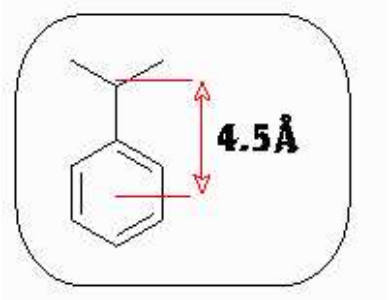


- An HDON - HACC site with a NEG site:

  The closest approach between these two types of sites can be determined by the sum of the van der Waals radii. The example used to determined the distance of closest approach (3Å) was a the negatively charged oxygens of a phosphate group and a common HDON-HACC sites, such as hydroxide or ketone functional groups.

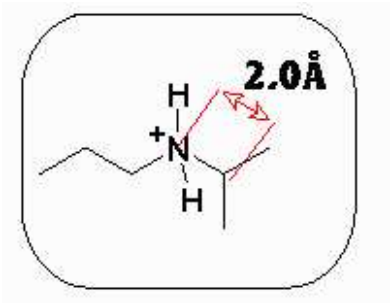
- An HYD side with another HYD site:

Given a sample molecule containing two hydrophobic groups, one a benzene ring and the other a propyl moiety, the approximate radius across the ring is 3Å and the approximate distance from the ring to the centre of propyl group is 1.5Å. The choice of a ring and a three carbon hydrocarbon chain as the model for a HYD-HYD interaction is justified since aromatic rings are common to drug molecules and the hydrocarbon chains with HYD functionality may contain as few as 3 carbons. The large margin for error of ± 1.0Å is acceptable since the hydrophobic interaction is diffuse and therefore not as directionally specific.



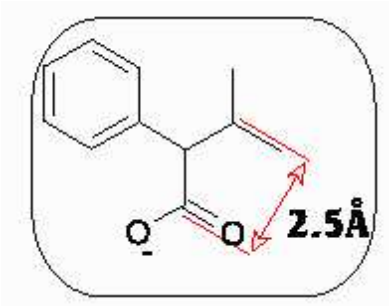
- An HYD site with a POS site:

  The shortest distance between a HYD and a POS site was determined under consideration of a POS amine group ($-NH_2^+-$) bonded to a propyl HYD group. The distance was calculated based on the centre of the propyl group (as mentioned previously) and the position the nitrogen atom.



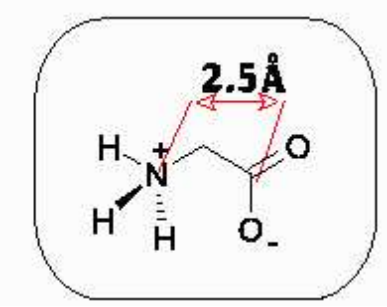
- An HYD site with a NEG site:

  The molecule illustrated here demonstrates a net negative charge at the COO^- group and the centre of the HYD group near the junction carbon of the propyl group. Thus, bond lengths (C-C) were used to approximate the shortest possible distance under these conditions.



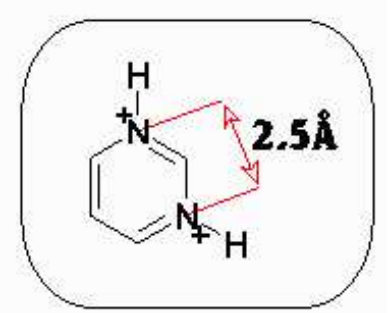
- A POS site with a NEG site:

  The shortest distance between these two centres was calculated based on the bond lengths of C-N and C-C and the angle of 109^o between these bonds.



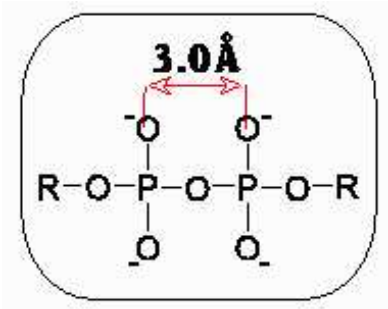
- A POS site with a POS site:

In the six member ring illustrated, both nitrogen atoms could accept a hydrogen atom, leaving them positively charged. This span was again calculated using C-N bond lengths and the angle of 120^o between these bonds.



- A NEG site with a NEG site:

  This common example of a functional group containing multiple negative charges has a larger minimal distance (3Å) between the various points of charge - the oxygen atoms - than the positive-positive charge distance (2.5Å). The sum of the van der Waals radii was utilized under these circumstances.



In the following table is given a summary of above results. Here is denoted the minimal distance between two active centres by d and minimal tolerance by δ.

	HDON/HACC	HYD	POS	NEG
HDON/HACC	0.15/2.0	0.25/3.0	0.25/2.5	0.25/3.0
HYD		1.0/4.5	0.25/2.0	0.25/2.5
POS			0.25/2.5	0.25/2.5
NEG				0.25/3.0

Table 7.3: Values of δ/d for various pairs of centres.

7.4 Bin Sizes

Two molecules define the same pharmacophore if they have the same configuration of centres and the set of distances between active parts of the molecules are relatively close. These differences may become larger with increasing distance between the centres in the sense that centres that are far apart should allow more freedom in the actual position of the centres involved. Ideally, we would like to index the different classes of pharmacophores. This suggests that the distances should scale as

$$\frac{dy}{y} = b\,dx$$

where b depends on the nature of the centres involved, y is the total distance between any two centres, dy is the maximum difference in distance that defines the same pharmacophore and dx is the difference between adjacent types. Solving for y we get

$$y(x) = a\,e^{bx} \tag{7.2}$$

for some value of a. This is a two parameter model and a and b can be determined by assigning to the first pharmacophore, $x = 1$, the minimal distance between the centers, d. As another constraint, the smallest allowable change in y at this minimal distance denoted as $\delta = y(2) - y(1)$ can be specified. The values of d and δ are empirical quantities which depend upon the chemical interaction involved. These empirical quantities are listed in Table 7.3.

Given a maximum distance between any two centres, D, the corresponding upper bound on the number of bins, N, can be computed by solving $y(N) = D$. This gives

$$N = \left\lfloor \frac{\ln(D/d)}{\ln(1 + \delta/d)} \right\rfloor + 1 \tag{7.3}$$

where $\lfloor\ \rfloor$ denotes the floor of the value. The last two columns of Table 7.4 list this value of N for each pair in Table 7.3.

Since the distance function (7.2) is exponential, a significant reduction in the value of N can be expected with only a slight change in the value of δ. However, just what is considered to be a slight change is open to interpretation and should depend strongly on the nature of the centres. Table 7.4 is included for demonstration of sensitivity of N with respect to δ/d.

δ/d	N	δ/d	N
0.1/2.0	42	0.15/2.0	29
0.2/3.0	26	0.25/3.0	21
0.2/2.5	24	0.25/2.5	20
0.2/2.0	22	0.25/2.0	18
1.0/4.5	7	1.0/4.5	7

Table 7.4: Values assume $D = 15$Å for all pairs.

7.5 Implementation

We have developed a program to estimate the number of three and four centre pharmacophores under user specified conditions. These include not only restrictions on the number of possible group labels, but also relative distance intervals and ranges for each of the pairs of centres. The program eliminates all duplicate pharmacophores due to possible symmetries in two or three dimensions. It also verifies the required conditions for feasible constructions of triangles and tetrahedra, taking into account margins of errors indicated in each of the distance bins.

Here are listed some of our results.
The final number of pharmacophores with:

- 3 centres, 5 labels (including exactly two HYD type centres) with a table of distances as in Table 7.3 is 3,704
- 3 centres, 5 labels (including exactly two HYD type centres) with a table of distances corresponding to first column of Table 7.4 is 5,775
- 3 centres, as previous, but any 5 centres is 295,535
- 4 centres, 5 labels (including exactly two HYD type centres) with a table of distances as in Table 7.3 is 27,953,097
- 4 centres with all properties as above, but instead of a min. increment for distance bins between HDON/HACC - HDON/HACC $\delta = 0.15$ we put $\delta = 0.12$, we get 29,941,835
- 4 centres, 5 labels (including exactly two HYD type centres) with a table of distances corresponding to first column of Table 7.4 is 88,178,007.

7.6 Conclusion

In this report is given an estimate on the size of the "pharmacophoric space" with some restrictions on types of possible pharmacophores. These restrictions were derived both from geometrical arguments and chemical properties of usual drug-like compounds. These estimates give a sense of the feasibility in building a screening library covering all 3 and 4 centre pharmacophores (to date there are only on the order of 10^7 compounds which are made and characterized), which could play an important role in drug searches especially if one can isolate extra restrictions in specific contexts.

A counting argument for five centre or larger pharmacophores would be very similar, but would involve a more delicate symmetry analysis of the pharmacophore. Rough estimates suggested by the calculations of Section 2.1 indicate that a library incorporating all 5 centre pharmacophores, for example, is not possible with current technologies. In these cases, new practical ideas to reduce the pharmacophore space are absolutely necessary.

Bibliography

[1] J. Greene, S. Kahn, H. Savoj, P. Sprague, S. Teig. Chemical Function Queries for 3D Database Search, *J. Chem. Inf. Comput. Sci.*, 1994, *34*, 1297-1308.

[2] D. Clark, G. Jones, P. Willett, P. Kenny, R. Glen. Pharmacophoric Pattern Matching in Files of Three-Dimensional Chemical Structures: Comparison of Conformational-Searching Algorithms for Flexible Searching, *J. Chem. Inf. Comput. Sci.*, 1994, *34*, 197-206.

[3] S. Pickett, . Mason, I. McLay. Divesity Profiling and Design Using 3D Pharmacophores: Pharmacophore-Derived Queries (PQD), *J. Chem. Inf. Comput. Sci.*, 1996, *36*, 1214-1223.

[4] M. Ashton, M. Jaye, J. Mason. New perspectives in led generation II: Evaluating molecular diversity, *Drug Discovery Today*, 1996, *1*, No. 2, 71-78.

[5] M. McGregor, S. Muskal. Pharmacophore Fingerprinting. 1. Application to QSAR and Focused Library Design, *J. Chem. Inf. Comput. Sci.*, 1998.

Appendix A

Third Annual PIMS Industrial Problem Solving Workshop – List of Participants

Rita Aggarwala	University of Calgary	*rita@math.ucalgary.ca*
Eric Agyekum	University of Victoria	*agyekum@math.uvic.ca*
Janez Ales	Simon Fraser University	*janez@sfu.ca*
Seema Ali	Univ. of British Columbia	*seema@unixg.ubc.ca*
Andreea Amariei	University of Alberta	*andreeamama@yahoo.com*
Ben Amjoun	OCENS	*banjou@aol.com*
Hassan Aurag	University of Montreal	*aurag@dms.umontreal.ca*
Lynn Batten	University of Manitoba	*batten@cc.umanitoba.ca*
Jeremy Bell	University of Victoria	*jbell@math.uvic.ca*
Jesse D Bingham	University of Victoria	*jbingham@gulf.csc.uvic.ca*
C. Sean Bohun	University of Victoria	*bohun@math.uvic.ca*
Andrej Bona	University of Calgary	*bona@math.ucalgary.ca*
Chris Bose	University of Victoria	*cbose@math.uvic.ca*
James Brannan	Clemson University	*jrbrn@math.clemson.edu*
Chris Budd	University of Bath, UK	*cjb@maths.bath.ac.uk*
Dan Calistrate	University of Calgary	*calistra@math.ucalgary.ca*
Myriam Caprioglio	University of Montreal	*capriogl@crm.umontreal.ca*
Brenda Caughlin	Chemex Laboratories	*bcaughlin@chemex.com*
Jie Cheng	University of Alberta	*jcheng@cs.ualberta.ca*
Kell Cheng	University of Calgary	*khfcheng@math.ucalgary.ca*
Kelly Choo	University of Victoria	*chook@math.uvic.ca*
Nicola Costanzino	Simon Fraser University	*ncostanz@math.sfu.ca*
Ellis Cumberbatch	Claremont Graduate Univ.	*ellis.cumberbatch@cgu.edu*
Tamara Dakic	Simon Fraser University	*dakic@cs.sfu.ca*
Adriana Davidova	University of Victoria	*adriana@math.uvic.ca*
Simon Di	University of Victoria	*sdi@uvic.ca*
Tom Doman	Searle	*TOM.N.DOMAN@monsanto.com*
John Drew	College of William & UVic	*jhdrew@math.uvic.ca*
Rod Edwards	University of Victoria	*edwards@math.uvic.ca*
Natalie Foorde	University of Victoria	*nataf@math.uvic.ca*
Brenda Hawkins	University of Alberta	*hawkins@sif.math.ualberta.ca*
Huaxiong Huang	PIMS/UBC	*hhuang@york.ca*
Ella Huszti	University of Alberta	*huszti@math.ualberta.ca*
Veselin Jungic	Simon Fraser University	*vjungic@sfu.ca*
Mounia Kjiri	University of Montreal	*kjiri@dms.umontreal.ca*
Tamara Koziak	University of Alberta	*koziak@math.ualberta.ca*

Nathan Krislock	University of Regina	*krislock@math.uregina.ca*
Susan Kutay	University of Calgary	*smkutay@ucalgary.ca*
Claude Laflamme	University of Calgary	*laf@math.ucalgary.ca*
Vincent Lemaire	University of Montreal	*lemaire@dms.umontreal.ca*
Margaret Liang	Univ. of British Columbia	*mliang@physics.ubc.ca*
David Lokhorst	RSI Technologies	*lokhorst@compuserve.com*
Yves Lucet	CECM, Simon Fraser Univ.	*lucet@cecm.sfu.ca*
Scott MacLachlan	Univ. of British Columbia	*scottmac@interchange.ubc.ca*
Devon McCrea	University of Victoria	*devon@math.uvic.ca*
Jason McVean	Merak	*Jasonm@merak.com*
Fausto Milinazzo	University of Victoria	*fmilinaz@math.uvic.ca*
Marni Mishna	Simon Fraser University	*mmishna@math.sfu.ca*
Bryant Moodie	University of Alberta	*bryant.moodie@ualberta.ca*
Tim Myers	Cranfield University, UK	*t.g.myers@cranfield.ac.uk*
Wendy Myrvold	University of Victoria	*wendym@csr.csc.uvic.ca*
Emmanuel Ngembo	University of Montreal	*ngembo@dms.umontreal.ca*
Gordon O'Connell	University of Victoria	*goconne@gulf.csc.uvic.ca*
Dale Olesky	University of Victoria	*dolesky@csr.uvic.ca*
Todd Oliynyk	University of Alberta	*toliynyk@math.ualberta.ca*
Luz Palacios	University of Calgary	*palacios@msfs.math.ucalgary.ca*
Marc Paulhus	Univ. of Calgary/PIMS	*paulhusm@math.ucalgary.ca*
Douglas Pickering	Brandon University	*Pickering@BrandonU.ca*
Shelly Pinder	Univ. of British Columbia	*pinder@math.ubc.ca*
Miro Powojowski	University of Calgary	*powojo@math.ucalgary.ca*
Keith Promislow	Simon Fraser University	*kpromisl@sfu.ca*
Bill Reed	University of Victoria	*reed@math.uvic.ca*
Frank Ruskey	University of Victoria	*fruskey@csr.csc.uvic.ca*
Mateja Sajna	Simon Fraser University	*msajna@cs.sfu.ca*
Joe Sawada	University of Victoria	*jsawada@csr.csc.uvic.ca*
Donald Scott	Enbridge	*Don.Scott@cnpl.enbridge.com*
Hope Serate	University of Calgary	*hope@math.ucalgary.ca*
Nikhil Shah	Northern Financial Techn.	*pns25@aol.com*
Peilin Shi	University of Victoria	*shi@math.uvic.ca*
Doug Smith	Charles Howard and Assoc.	*doug.smith@chal.bc.ca*
Grace So	University of Toronto	*graces@utstat.toronto.edu*
John Stockie	Simon Fraser University	*jms@sfu.ca*
Cristina Stoica	University of Victoria	*cstoica@math.uvic.ca*
Holger Teismann	University of Victoria	*teismann@math.uvic.ca*
Satoshi Tomoda	University of Calgary	*tomoda@math.ucalgary.ca*
Min Tsao	University of Victoria	*tsao@math.uvic.ca*
Stan Tuller	University of Victoria	*stuller@office.geog.uvic.ca*
Pauline Van den Driessche	University of Victoria	*pvdd@math.uvic.ca*
Jana Vander Kloet	University of Calgary	*kjvander@ucalgary.ca*
James Watmough	University of Victoria	*watmough@math.uvic.ca*
Rex Westbrook	University of Calgary	*westbroo@ucalgary.ca*
Brian Wetton	Univ. of British Columbia	*wetton@math.ubc.ca*
Paul Wiebe	University of Alberta	*paully@v-wave.com*
JF Williams	Simon Fraser University	*jfwillia@math.sfu.ca*
Kathryn Wood	University of Victoria	*kathryn@math.uvic.ca*
Zili Wu	University of Victoria	*wzl@math.uvic.ca*
Yang Zhao	University of Victoria	*yzhao@math.uvic.ca*
Julie Zhou	University of Victoria	*jzhou@math.uvic.ca*

Appendix B

The Organization of an Industrial Problem Solving Workshop

Introduction

History

The "Study Groups" with Industry started in Oxford in 1968, with the aim of creating a mutually beneficial link between researchers in industry and academic applied mathematicians. The formula of setting aside a week for intensive study of a few real world problems proved so successful that the Study Group became an annual fixture in many countries around the world. The success of such workshops can be attributed to a number of reasons, for example:

- they help foster contacts between academia and industry, sometimes leading to research contracts,
- they frequently lead to challenging new research areas, which have a direct bearing on physical problems,
- they are an excellent source of research topics for graduate students (as well as academics). Further, they allow companies to become acquainted with students and evaluate them for future employment.

Then there are the obvious reasons: the company wanting a problem solved and the pleasure of working with enthusiastic colleagues.

Set-up of the meetings

The usual format of the meetings is:

- Monday morning: problem presentations by industrial representatives,
- Monday afternoon to Thursday: intensive workshop sessions,
- Friday: summing up by academics.

Some meetings also have a regular summing up of the day's progress.

It is considered very useful to have at least one company representative present during the whole week, to answer any questions that may arise. Although, if this is not possible they should at least be present on Monday and Friday and be easily contactable throughout the week.

How To Contact Industry

Personal contacts are clearly the best way to get in touch with a company. These may come about from previous workshops (many companies become regulars), or in the case of PIMS, from the industrial working seminars. Many academics have their own contacts, particularly from graduate students that have moved into industry. As a group's experience with industrial modeling increases so does its reputation and number of contacts.

To foster new contacts a letter, fax or phone call to a company's research department may work. This should outline:

- how the meeting is set up,
- the benefits for the company and how much (or little) it will cost them,
- the previous item may make them suspicious and question the worth of the meeting, so explain the benefits to academics as highlighted in the history section,
- in a fax or letter you should include a list of problems tackled at previous meetings, to indicate what can be done (feel free to bias this list to the companies interests).

A selection of problems investigated at recent PIMS workshops is given in the Appendix.

Screening The Problems

Once problems start to flood, or trickle, in, they require filtering. Industrial researchers may not be familiar with recent academic work and so the problems they submit may be already solved, trivial, reasonable or insoluble (however, even insoluble problems may be of interest, *i.e.* a different approach may be suggested). Presuming the coordinator is not an expert on every subject, they should try to enlist expert help to determine which category a problem falls into. Familiarity with, or knowledge of, experts in different fields is essential for a coordinator. Screening may also be carried out by searching the literature, using a science citation index.

Even if a problem appears tractable to the coordinator, it must be suitable for the participants. However, experts can usually be enlisted (lured) if they know a problem of specific interest will be presented.

The number of problems chosen should be based on the the number of participants. Typically ten people per problem is a realistic number. Bear in mind that only a small proportion of participants will be knowledgeable in any given field. Many people attend to learn new topics or through curiosity.

Report/Proceedings

By Thursday it should have been decided who is to present the results the following day. This person will also be in charge of the writing up, whether they do it all themselves or coordinate efforts by group members is up to them, provided the report gets completed. Two months is the standard time for this task; the more time drags on, the less committed people are to dealing with it. The companies should be allowed to screen the report before it is finalized.

A member of the organization committee will be in charge of seeing that the deadlines are obeyed and also to put the final report together.

Organization

Most requirements for organizing a Problem-Solving Workshop are similar to those of a normal academic meeting. The following highlights a number of differences that should be kept in mind.

Location

Such a workshop must be held in the vicinity of a university, as both library and computer access are essential. For the problem presentation on the first day and the summing-up on the last day, a big room is required, with overhead projector facilities and a board. For the workshop's discussions, several small rooms are preferred with boards. All rooms should be close together, as some participants will want to flit between the different problem sessions.

Accommodation

Academics should be housed close together, allowing the discussions to keep going into the evening (and frequently in the bar). Note, industrial people will probably prefer hotel accommodation to university dorms; these should be close to the university so the representatives are not 'left out'.

Fees

Most meetings require companies to pay a fee, to help 'defray' costs for the academics (the fee from approximately 8 companies is unlikely to cover travel and accommodation costs for everybody). Typical charges range from $1600 to $4000. Government and other types of research grants can also be useful; emphasizing the benefits to be gained from interaction between universities and industry, such as increased competitiveness, can help in obtaining funds. Academics, and their universities, may also be willing to pay their own costs.

Facilities

The library should be accessible for all participants to consult journals and make xerox copies. If applicable, arrange for copy cards or access codes and ask library members attending the meeting to help with the borrowing of books.

Computer access is a must for e-mail junkies, but also to provide facilities such as telnet, maple, mathematica, matlab, fortran etc., plotting & print-outs.

Food and drink

Accommodation and meals should be provided to all participants. Coffee and tea breaks are invaluable for regular summing up and to get honest opinions and fresh ideas from people who are working on other problems. It's nice to point out some restaurants, pubs and bars in the vicinity of the workshop rooms, so that the discussions can continue in the evening.

References

Useful world-wide contacts on the study group meetings:

Canada: Dr. Nassif Ghoussoub *Director*, Pacific Institute for the Mathematical Sciences.
UK: Dr John Ockendon *OCIAM Research Director* University of Oxford.
UK: Dr Tim Myers *Industrial Post-doc* Cranfield University.
Australia: Dr Kerry Landman *MISG Director* Melbourne University.
US: Dr Ellis Cumberbatch *Professor* Claremont Graduate School.
US: Dr Avner Friedman *Director* IMA Minnesota, Minneapolis.
US: Donald Schwendeman *Associate Professor* Rensselaer Polytechnic Institute (RPI).

Web sites:

http://www.pims.math.ca
http://macserver.maths.mu.oz.au/misg/
http://www.maths.bath.ac.uk/ESGI97

http://www.maths.ox.ac.uk/ociam (see newsletter in particular)
http://www.math.rpi.edu
http://www.indmath.uni-linz.ac.at
http://www.mat.dtu.dk/ECMI
http://www.siam.org

A Selection of Problems Investigated During Recent Workshops at PIMS

IPSW-1; August, 1997 – University of British Columbia

- Inversion for Anisotropic-Velocity Parameter (Petro-Canada)
- Fingerprint Identification (Kinetic Sciences Inc.)
- Modeling Bronchial Epithelial Lesions (BC Cancer Labs)
- Optimally Cutting Logs (MacMillan Bloedel)
- Modeling Stress Intensity in a Thermoroll (MacMillan Bloedel)
- Measuring the Stress Intensity of a Composite Vessel (Powertech Labs)
- Evaluating Computational Methods in Fuel Cell Systems (Ballard Powersystems)
- Rapid Thermal Processing of Semiconductors (Vortek Industries Ltd.)
- Optimal Trading Strategies for Electricity Trading (Powerex Ltd.)

IPSW-2; June, 1998 – University of Calgary

- An Optimal Strategy for Maintaining Excess Capacity (Boeing Corp.)
- Inventory Optimization using a Renewal Model for Sales (Boeing Corp.)
- A Problem in Petroleum Reservoir Simulation (Computer Modelling Group Ltd.)
- On Seismic Imaging: Geodesics, Isocrons, and Fermat's Principle (The Geomechanics Project, University of Calgary)
- Trip Wire Detection for Land Mines (ITRES Research Ltd.)
- Torsion in Multistrand Cables (Powertech Labs Inc.)
- Automatic Detection of Egg Shell Cracks (VisionSmart Inc.)

PIMS Contact Information

email: pims@pims.math.ca
http://www.pims.math.ca

Director: **N. Ghoussoub**
email: director@pims.math.ca
phone: (604) 822-3922
fax: (604) 822-0883

Scientific Executive Officer: S. Rutherford
email: sandy@pims.math.ca

Business Development Officer: Lindsay Hughes
email: lphughes@pims.math.ca

Financial Officer: K. Soh
email: katrina@pims.math.ca

Industrial Facilitator in Calgary: M. Paulhus
email: paulhusm@math.calgary.ca

Education Coordinator: K. Hoechsmann
email: hoek@pims.math.ca

SFU-Site Director: P. Borwein
email: sfu@pims.math.ca

UBC-Site Director: D. Rolfsen
email: ubc@pims.math.ca

UVic-Site Director: F. Diacu
email: uvic@pims.math.ca

UC-Site Director: M. Lamoureux
email: uc@puns.math.ca

UA-Site Director: B. Moodie
email: ua@pims.math.ca