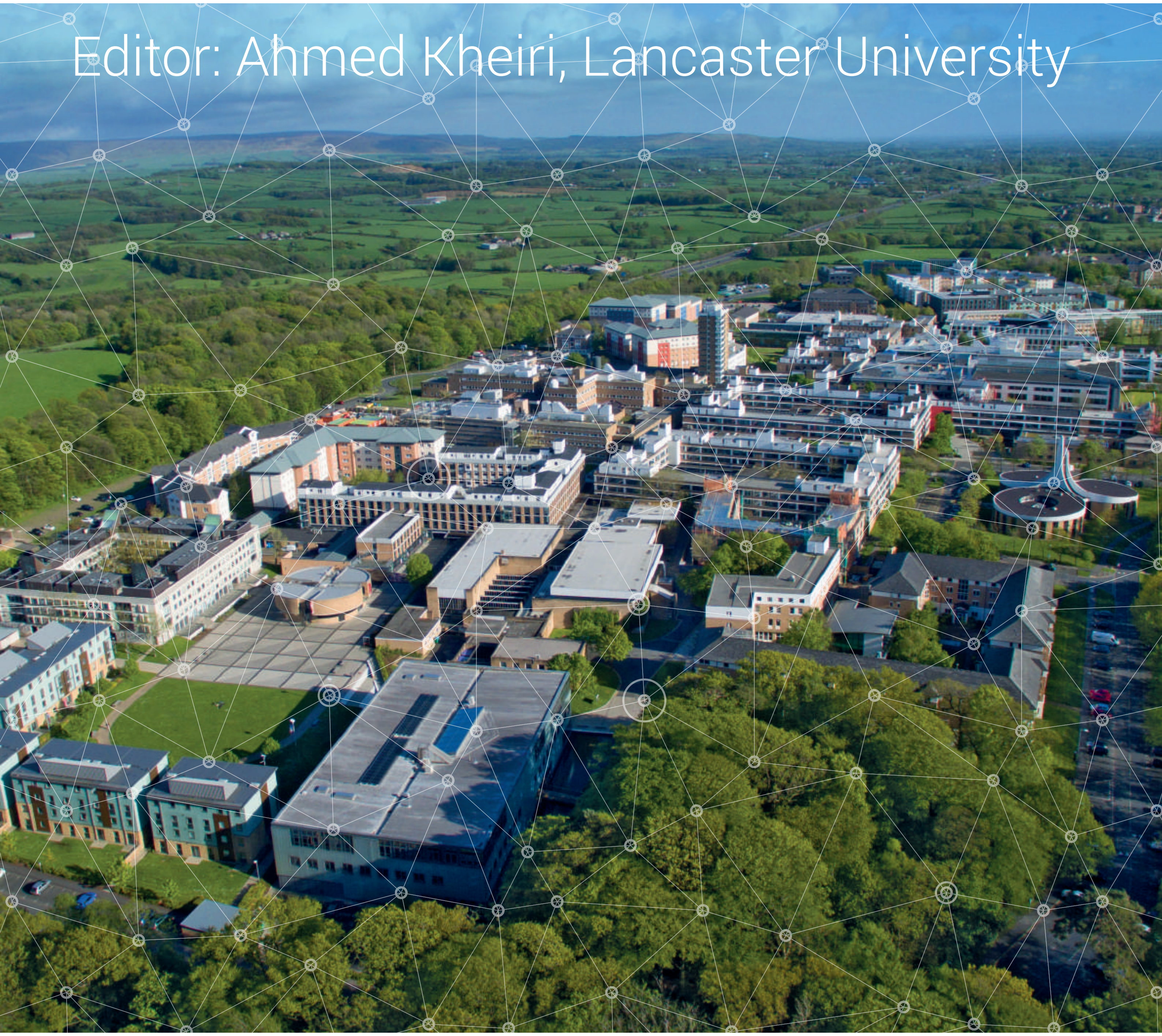
THE OPERATIONAL RESEARCH SOCIETY

OR60: The OR Society Annual Conference

11-13 September 2018 Lancaster University

Keynote Papers

Editor: Ahmed Kheiri, Lancaster University



Lancaster University

OR60 Annual Conference - Keynote Papers and Extended Abstracts
Lancaster University. Lancaster. 11-13 September 2018

Foreword

On behalf of the organising committee, I would like to welcome you to the OR Society's Annual Conference - OR60. This year we are hosted by Lancaster University, set in the county of Lancashire.

Lancaster University was established by Royal Charter in 1964, one of seven new universities created in the 1960s. Operational Research was one of the founding departments, the first in the country. The first professor was Pat Rivett, President of the OR Society at the time of his appointment. So, it is particularly significant that the conference returns to Lancaster to celebrate the Diamond anniversary.

In addition to three prominent plenary speakers, the conference is made up of presentations covering an impressive 38 streams. Many of these streams have keynote speakers who will present the latest developments in the field. These keynote papers surely include something that will appeal to all participants. They range from optimisation and simulation to forecasting. You can read about refugee resettlement, stochastic modelling of aircraft queues, energy systems and OR in defence and security.

I would like to take this opportunity to convey our special thanks to all the keynote speakers and other contributors who provided their work to be included in this book. Reading through each of these papers provides a great opportunity to find out about the latest developments across the field of OR. I would like to thank stream organisers for their extremely valuable contribution and hard work, which reflects their commitment and dedication to the profession. I would also like to thank all those who submitted their abstract for presentation in the streams, these contributions are central to making a successful conference.

In addition to the streams of papers, represented by these keynotes, there are many other aspects of the conference. 'Making an Impact' is packed with sessions aimed at helping everybody to become more successful in making an impact in their work. These sessions are designed to meet the needs and interests of practitioners as well as academics and are now a regular feature of our national conference but this year, based on previous feedback, will include more interactive sessions on key topics, all designed to make you personally more effective and your work to have a greater impact. The President's Medal presentations showcase the best applications of OR and Analytics practice.

An important aspect of the conferences is networking and being able to chat informally with others involved in OR. As always, the social program will offer lots of opportunities to meet up with old friends and to make some new ones.

OR60 then offers something for absolutely everyone.

I hope that you have a highly productive and thoroughly enjoyable conference.

With best regards,

Ahmed Kheiri, Lancaster University
Keynote Handbook Editor
September 2018

OR60 Annual Conference - Keynote Papers and Extended Abstracts
Lancaster University. Lancaster. 11-13 September 2018

Silver Screen: Basis for Decision - Selected Presentations by Professor Pat Rivett

OR60 offers an excellent opportunity to celebrate a champion of OR. On the USB stick is a selection of presentations by Professor Pat Rivett filmed in 1964. He was a President of the OR Society and took up the first chair of Operational Research at a British University that year at Lancaster. He also introduced OR, the science of complexity, to Britain in 1964, presenting “Basis For Decision - the Application of Operational Research”, three 45-minute episodes on the new television channel BBC2.

The three episodes were broadcast on Tuesday nights from 14 to 28 July 1964; “Cutting the Queue”, “Playing it Through” and “The Human Factor”. The first, second and part of the third programmes comprise case studies. These show a spectrum of approaches in use on a range of operations. The third programme includes a debate about the limitations of OR between leading practitioners and Denzil Freeth MP with John Barber, then Finance Director of the Ford Motor Company. The paper about OR perennials, starting on page 39 of this handbook, describes the episodes in more detail.

It is striking how many of Professor Rivett’s comments remain relevant in 2018. The OR60 USB stick includes a selection of films of his presentations, illustrated by scenes from the case studies and comments from other leading lights of OR in 1964. These include Sir Charles Goodeve, with Professors Brian Houlden, Reg Revans, Stafford Beer, Ronnie Shephard and Keith Tocher.

Recovering the three episodes of “Basis for Decision” to share at OR60 was a long process that began in 2010. I would like to acknowledge Richard Fitzgerald’s interest and tenacity which were vital to this, complemented by the generous support of the OR Society, the OR Society Archive at the Modern Records Centre, University of Warwick, the library of Monash University, the Kaleidoscope Group, Lancaster University, Danny Hardaker Video and the BBC.

Professor Rivett’s closing remark is a worthy perennial aspiration for OR:
“So that managers at all levels have a firmer basis for decisions on industry and government.”

With best wishes for OR60

Ian Mitchell FORS

Contents

OR60 Annual Conference - Keynote Papers and Extended Abstracts
Lancaster University. Lancaster. 11 - 13 September 2018

Air Traffic (Flow) Management: Issues, Challenges, and Research Opportunities

Guglielmo Lulli

Lancaster University, Department of Management Science, Lancaster, United Kingdom
g.lulli@lancaster.ac.uk

Abstract

Air Traffic Management is the overall collection of the air traffic system's management processes that support the ultimate goal of safe, efficient, and expeditious aircraft movement. Because the growth of demand has not been supported by a corresponding development of the air traffic system, the need for a more efficient use of capacity is imperative. In response to this need, prominent initiatives have been launched to modernise the Air Traffic Management system. In this paper, I provide an overview of some of the issues, challenges and research opportunities that are relevant to the analytics community in order to deploy a more flexible, resilient and scalable air transport system than today's. In particular, I will focus on Air Traffic Flow Management and issues related to *i*) wider participation of the airspace users in the decision process; *ii*) a higher degree of interaction among the different "ATM function" thus providing a seamless solution to air traffic management; and *iii*) a paradigm shift from the airways system to free airspace.

Keywords: Air Traffic Management; Trajectory Based Operations, Preferences and Priorities, Collaborative Decision Making, Short Term Initiatives, Equity and Efficiency.

1. Introduction

Aviation is one of the fastest growing sectors of the global economy. In 2017, the total number of passenger worldwide reached 4.1 billion with an increase of 7.3% in comparison to the previous year (IATA, 2018). Because the growth of demand has not been supported by a corresponding development of airports and related systems, when weather conditions are less than ideal, both the European and the United States air transportation systems suffer with a systemic level of congestion thus generating enormous costs for airlines and air passengers. A FAA-Nextor study estimated the 2017 annual costs of delays (direct cost to airlines and passengers, lost demand, and indirect costs) for US air carriers to be $26.6 billion. This picture is expected to be even gloomier in the near future, as air traffic demand forecasts for the next 20 years predict annual growth rates close to 6% (IATA, 2018). Therefore, the need to modernise the air transportation system is evident. Prominent initiatives have been launched in Europe, the US, and in Japan to develop a future air transport system that will be more flexible, resilient and scalable than today's.

More specifically, the European Union - together with Eurocontrol – established a public-private partnership to manage the Single European Sky Air Traffic Management Research (SESAR) Programme. The scope of SESAR is to modernise the European Air Traffic

Management (ATM) system by delivering innovation in a number of areas (SESAR, 2015), among others:

- High Performing Airport Operations, which also looks at the enhancement of runway throughput, integrated surface management, and total airport management.
- Advanced Air Traffic Services, and in particular the enhancement of arrivals and departures management, separation management, and trajectory and performance based free routing.
- Optimised ATM Network Services, which also includes research activities in the areas of advanced airspace management, advanced dynamic capacity balancing, and optimised airspace user operations.

These areas and the listed (sub)problems are – perhaps – the most relevant to the analytics community at large. All these problems are highly interconnected, and their solutions should be integrated into the overall solution that aims to deploy a harmonized air traffic system. The underpinning element to achieve a more effective ATM system is the deployment of a higher level of automation for all the stakeholders' decision-making processes. This calls for the development of new mathematical models, algorithms and in general decision support systems.

In the remaining part of this paper, I focus on "Optimised ATM Network Services" and in particular on Air Traffic Flow Management (ATFM), which is the set of planning procedures aiming to detect and resolve demand-capacity imbalances. A key element of future ATFM operating concepts is Trajectory Based Operations (TBO), which is expected to provide the flexibility needed by airspace users to optimise their operations while simultaneously ensuring the predictability needed at the ATM network level for maximum overall performance (SESAR, 2015). It is important to highlight, that any development including TBO should be envisioned in a Collaborative Decision Making (CDM) setting. The driving philosophy behind CDM is that a greater involvement of all the ATM stakeholders in the decision process as well as improved data exchange and communication between aviation transportation organisations will lead to better decision making in ATFM.

In this scenario, the implementation of Trajectory Based Operations raises a number of new issues that provide the scientific community with new territory to explore. Three of these issues relate to *i*) wider participation of the airspace users in the decision process; *ii*) a higher degree of interaction among the different "ATM function" thus providing a seamless solution to air traffic management; and *iii*) a paradigm shift from the airways system to free airspace.

In what follows, I describe in more detail some of the associated challenges and research opportunities. In the described context, prescriptive mathematical models for Trajectory Based Operations should be "flight-centric", meaning that they specify the trajectory of each single flight. There are several mathematical models on this subject in the literature, e.g., Sherali et al. (2002), Bertsimas et al. (2011) and Balakrishnan and Chandran (2014) to mention a few. Although these models are not able to provide the right answers to the new research gaps they provide the groundwork for this new development activity.

2. Priorities and Preferences within TBO

To grant airspace users the flexibility needed to optimise their operations, the opportunity to express airspace users' preferences and priorities is envisioned in the new concept of operations. The scope of priorities is to provide airlines with flexibility in re-arranging their flights to protect more valuable flights from delays at the expense of a larger displacement (delays) of other flights. Regarding preferences, although the ATM stakeholders did not reach a common consensus on the concept and definition, preferences should be conceived as a mechanism to absorb delays at the tactical level of ATM (OptiFrame, 2016). The ATFM mathematical model should incorporate airspace users' priorities and preferences to define best in class modifications to preferred 4D trajectories based on their internal business models; thus reconciling the specific interests of airspace users with the ATM system efficiency. Research on mathematical modelling in this specific domain of ATFM is still very limited. So far, very few approaches have been proposed to model preferences and priorities in ATFM, and to be exact only Pilon et al. (2016) and Djeumou Fomeni et al. (2017), to the best of the author's knowledge.

3. TBO in a free-route environment (Odoni, 2018)

Trajectory-based operations face very different challenges depending on whether they are conducted in a network-structured environment or in a free-route environment. In the former case, the principal challenges involve the allocation, coordination and optimization of flows along the arcs and O/D paths of the network, as well as the merging and dispersing of flows at the network's nodes. By contrast, in the case of free routing, the emphasis is on anticipating and resolving potential conflicts that may occur (at least theoretically) at any point in a volume of airspace, while minimising interference with the user-preferred free routes. In both cases, the ultimate objective is to ensure the safety of operations, while maximising flow throughputs between origins and destinations.

Future operating environments will increasingly be hybrids of network routing and free routing. Thus, it is essential to identify the analytical and modelling issues that must be addressed in estimating performance (both safety-related and flow-efficiency-related) in such a hybrid environment.

4. Equity versus efficiency in ATFM

The greater participation of stakeholders in the ATFM decision process requires a coordination process, based on clear CDM decision criteria, to orchestrate the interaction between all stakeholders and ensuring stability of the ATM System. In current operations, "rationing by schedule" (RBS) is the cornerstone principle for Ground Delay Programs (GDPs) and other CDM initiatives (Chiang et al, 2001). Under RBS, resources are assigned to flights in accordance with a first-scheduled, first-served (FSFS) priority discipline. It has been agreed by all CDM stakeholders that RBS provides a basis for GDP planning that is fair to all parties. It was this agreement that spurred the successful development of CDM during the mid-1990s. Indeed, FSFS which not only is “fair” in the intuitive sense, but can also be shown to have mathematical properties consistent with the notion of equity in GDPs delay allocation (Vossen and Ball, 2006). However, when considering the overall air traffic system, which can be represented as a network of sectors and airports, each with a limited capacity, the conflict between equity and efficiency is more fundamental (Lulli and Odoni, 2007). Adherence to simple principles, like FCFS, may result in highly inefficient solutions with respect to cost. This implies that it may prove far more challenging to reach stakeholder consensus on what would constitute “fair and equitable” priority rules during periods of congestion. In this setting, one of the challenges is the design of mechanisms for capacity allocation of sectors and airports that are consistent with the CDM philosophy. This area of research is quite unexplored and it calls for multidisciplinary research activities. Indeed, this topic of research is not only of interest for the analytics community, but it also includes key aspects specific to the economics disciplines. Approaches might vary from system-wide optimisation to market-based mechanisms.

5. Short Term ATFCM Measures (STAM)

The European ATM Master Plan anticipates a closer and more efficient coordination between ATC and the Network Management function (ATFM) to prevent and/or resolve situations where traffic demand exceeds available ATC and airport capacity. Short Term ATFM Measures (STAM) are conceived as the means to manage short-term and tactical capacity and should be implemented using collaborative decision-making to manage flow before flights enter a sector. The scope of STAM initiatives is to resolve demand-capacity imbalances without deteriorating the performances of the overall ATM system and, at the same time, meeting the requirements of airspace users as set by their agreed Reference Business Trajectory.

From the mathematical modelling perspective, the design of STAM initiatives, which are of “local nature”, calls for mathematical models with a higher degree of details. This affects the dimensions of the mathematical models and the computational effort needed to solve them. Thus, it may be necessary to develop either decomposition methods – which decompose the original problem into subproblems, as for instance in Balakrishnan and Chandran (2014), a model architecture similar in spirit to Dell’Olmo and Lulli (2003) or a combination of both. In addition to prescriptive models, the design of efficient STAM initiatives and hotspot resolution measures may require predictive models. Accurate trajectory predictions and “sensitive” hotspot detections are key components of the problem, and they require the development of new methods that fall in the recent area of research of Data Analytics applied to ATM. Consistent with other areas of application, data analytics is increasingly applied in the aviation domain and the literature is quickly expanding in this specific topic.

Acknowledgements

The author is grateful to James Bleach from the obex team (http://theobexproject.co.uk/) for his editorial support.

References

H. Balakrishnan and B. Chandran (2014). Optimal Large-Scale Air Traffic Flow Management, available at http://web.mit.edu/hamsa/www/publications.html .

D. Bertsimas, G. Lulli and A. Odoni (2011). An integer optimization approach to large-scale air traffic flow management. Operations Research, 59(1): 211--227.

Chang, K., K. Howard, R. Oiesen, L. Shisler, M. Tanino and M. C. Wambsganss (2001). Enhancements to the FAA Ground-Delay Program under Collaborative Decision Making. Interfaces 3: 57--76.

P. Dell'Olmo and G. Lulli (2003). A new hierarchical architecture for Air Traffic Management: Optimisation of airway's capacity in a free flight scenario. European J. of Operational Research 144 (1): 179--193.

F. Djeumou Fomeni, G. Lulli and K. Zografos (2017). An optimization model for assigning 4D-trajectories to flights under the TBO concept. Proceedings of the 12th USA/Europe ATM R&D Seminar: Seattle, United States.

IATA (2018). Fact Sheet. Industry Statistics, available at http://www.iata.org/pressroom/facts_figures/fact_sheets/Documents/fact-sheet-industry-facts.pdf

G. Lulli and A.R. Odoni (2007). The European Air Traffic Flow Management Problem, Transportation Science 41 (4): 1--13.

N. Pilon, S. Ruiz, A. Bujor, A. Cook and L. Castelli (2016) Improved flexibility and equity for airspace users during demand-capacity imbalance. D. Schaefer (Editor) Proceedings of the SESAR Innovation Days (2016) EUROCONTROL. ISSN 0770-1268

A.R. Odoni (2018). Personal communication.

SESAR (2015). European ATM Master Plan. Executive view (Ed. 2015), available at https://www.atmmasterplan.eu/

H. D. Sherali, J. C. Smith and A. A. Trani (2002). An airspace planning model for selecting flight-plans under workload, safety, and equity considerations. Transportation Science, 36(4): 378--397.

T.W.M. Vossen and M.O. Ball (2006). Optimization and Mediated Bartering Models for Ground Delay Programs, Naval Research Logistics, 53: 75--90.

OR60 Annual Conference - Keynote Papers and Extended Abstracts
Lancaster University. Lancaster. 11-13 September 2018

On Tractable Cases in Combinatorial Optimization

Vladimir Deineko

University of Warwick
Warwick Business School, Coventry, CV4 7AL
V.Deineko@warwick.ac.uk

This presentation will consist of two parts. The first part is an introduction to the exciting world of polynomially solvable cases of NP-hard problems (this is how we interpret here the term ”tractable cases”). Three well-known NP-hard combinatorial optimization (CO) problems will be discussed: the quadratic assignment problem, the travelling salesman problem and the bipartite travelling salesman problem. Some conditions under which these problems can be solved in polynomial time will be described. The second part of the presentation will illustrate a usage of CO algorithms in real life applications. We will reflect on our experience when easy solvable problems become ”non-tractable” in real life applications.

Given two $n \times n$ square matrices $A = (a_{ij})$ and $B = (b_{ij})$, in the quadratic assignment problem, QAP(A,B), the objective is to find a permutation π that minimizes the function

$$Z_\pi(A,B) \; := \; \sum_{i=1}^{n}\sum_{j=1}^{n} a_{\pi(i)\pi(j)} b_{ij}.$$

For an impressive list of the QAP applications we refer a reader to the book of Eranda Çela (*The Quadratic Assignment Problem: Theory and Algorithms*, 1998, Kluwer Academic Publishers, Dordrecht, The Netherlands).

In our presentation we describe specially structured matrices A and B when the QAP(A,B) can be solved in polynomial time. We will use a visualisation of these matrices as in the two illustrative examples in Figure 1.

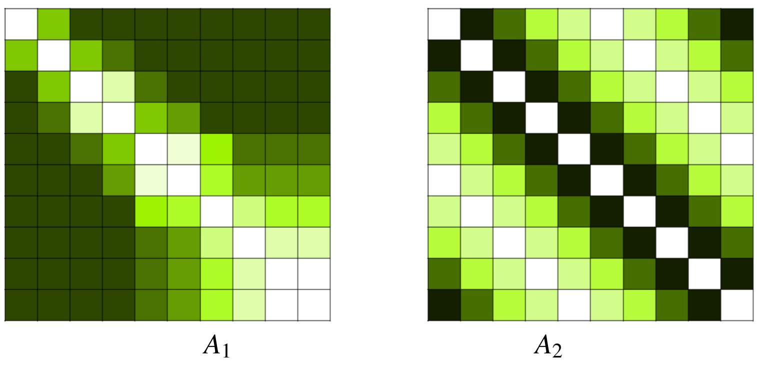


Figure 1: Visualisation of specially structured matrices for a QAP from Çela, Deineko & Woeginger (2018), *New special cases of the Quadratic Assignment Problem with diagonally structured coefficient matrices*, EJOR, 267, 3, 818-834.

Table 1: Classification of four-point conditions; further details can be found in Deineko, Klinz, Tiskin, Woeginger (2014) *Four-point conditions for the TSP: the complete classification.* Discrete Optimization 14, 147–159.

	$\mathscr{A} \le \mathscr{B}$	$\mathscr{A} \ge \mathscr{B}$	$\mathscr{A} \le \mathscr{C}$	$\mathscr{A} \ge \mathscr{C}$	$\mathscr{B} \le \mathscr{C}$	$\mathscr{B} \ge \mathscr{C}$
$\mathscr{A} \le \mathscr{B}$	$O(n^2)$	$O(n)$	$O(n^2)$	$O(1)$	$O(1)$	$O(1)$
$\mathscr{A} \ge \mathscr{B}$		NP-hard	$O(1)$	NP-hard	$O(n^2)$	$O(1)$
$\mathscr{A} \le \mathscr{C}$			$O(n^2)$	$O(1)$	$O(1)$	$O(1)$
$\mathscr{A} \ge \mathscr{C}$				NP-hard	$O(1)$	$O(n)$
$\mathscr{B} \le \mathscr{C}$					NP-hard	$O(n^2)$
$\mathscr{B} \ge \mathscr{C}$						$O(n^4)$

The objective in the travelling salesman problem (TSP) is to find a shortest tour through the set of n cities (points), i.e. a *cyclic* permutation σ that minimises the function $T_\sigma(C) := \sum_{i=1}^{n} c_{i\sigma(i)}$. The TSP can be formulated as the QAP(C,P), where C is a given matrix of the distances between the cities and P is a binary matrix which represents a cyclic permutation. The TSP is probably the best known problem in CO (we refer a reader to the classic book of Lawler, Lenstra, Rinnooy Kan, and Shmoys, (1985) *The Traveling Salesman Problem: A Guided Tour of Combinatorial Optimization*, Wiley, Chichester). No wonder that many researchers tried to identified polynomially solvable cases of the TSP. One of the best studied families of these cases can be characterised by the so-called *four-point* conditions. We use here a description of the four-point conditions from Deineko, Klinz, Woeginger and Tiskin, *Four-point conditions for the TSP: the complete classification*. (Discrete Optimization, 2014, 14, 147–159): Consider points i, j, k, l with $1 \le i < j < k < l \le n$. The symmetric distance matrix contains six different entries for these four points, which correspond to the six edges connecting the points. There are three pairings of these points into two disjoint edges $\{(i,j),(k,l)\}$, $\{(i,k),(j,l)\}$, $\{(i,l),(j,k)\}$, and we denote the lengths of these pairs as follows:

$$\mathscr{A} = c_{ij} + c_{kl}, \quad \mathscr{B} = c_{ik} + c_{jl}, \quad \mathscr{C} = c_{il} + c_{jk}.$$

Now a *four-point* condition simply specifies one or two inequalities which rank the values $\mathscr{A}, \mathscr{B}, \mathscr{C}$; these inequalities have to be satisfied for all possible choices of i, j, k, l with $1 \le i < j < k < l \le n$. Table 1 provides a summary of all possible cases defined by pairs of the four-point conditions.

As an example, we mention here the well known *Demidenko* matrix which is defined as a matrix that satisfies the condition $\mathscr{A} \le \mathscr{B}$. As stated in the table, the TSP with a Demidenko distance matrix can be solved in (n^2) time. An optimal TSP tour can be found in this case among the so-called *pyramidal* tours. Figure 2 illustrates a set of points with the distance matrix being a Demidenko matrix.

Given two points s and t from the set $\{1,2,\ldots,n\}$, in a variant of the TSP known as the Path-TSP one looks for a shortest tour that visits all points. The tour has to start at s and end at t. Surprisingly, the family of known polynomially solvable cases of the Path-TSP is not yet well populated. We will describe new polynomially solvable cases of the Path-TSP. As can be seen on the example of the shortest TSP path from node 5 to node 6 for the set of points shown in Figure 2 (which is $\langle 5,6,4,2,1,3,8,9,11,12,10,7 \rangle$), the structure of the path is not necessary pyramidal.

Vladimir Deineko

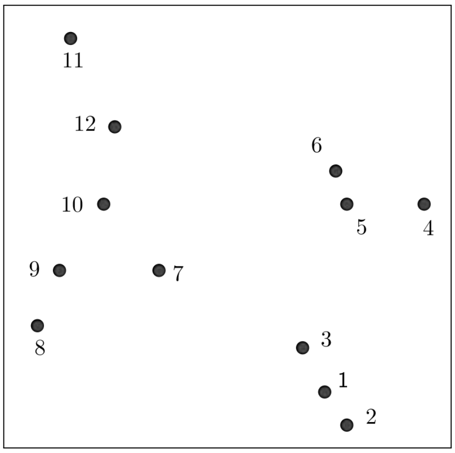


Figure 2: Euclidean Demidenko set of points; $\langle 1,2,4,5,6,11,12,10,9,8,7,3,1\rangle$ is the optimal tour.

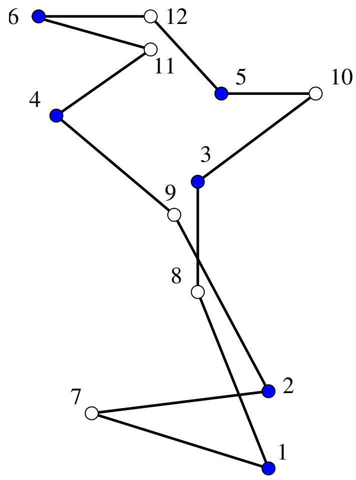


Figure 3: An instance of a solvable case of the BTSP from Deineko, Woeginger (2014) *Another look at the shoelace TSP : the case of very old shoes*, Fun with Algorithms, Volume 8496, 125-126.

In the *bipartite travelling salesman problem* (BTSP) the set of $n = 2k$ cities is partitioned into two subsets of equal size: the set K_1 of blue cities and the set K_2 of white cities, with $|K_1| = |K_2| = k$. The objective is to find the shortest tour which alternates between the blue and the white cities. In our presentation, we describe new polynomially solvable cases of the BTSP.

We conclude the presentation with reflection on our experience of dealing with CO problems in real life settings.

AUTHOR BIOGRAPHY

VLADIMIR DEINEKO is an associate professor (Reader) in Operational Research in the Warwick Business School. His research interests are within the area of algorithmic aspects of the problem solving process, with the main focus on the analysis of efficiently solvable cases of hard optimisation problems, in particular the travelling salesman problem and the quadratic assignment problem; he is interested in design and implementation of exact and approximate algorithms for CO problems from real life applications. www.wbs.ac.uk/about/person/Vladimir-Deineko

OR60 Annual Conference – Keynote Papers and Extended Abstracts
Lancaster University, Lancaster. 11–13 September 2018

OR in Defence and Security – Past and Present

Paul R. Syms [a]

[a] Defence Science and Technology Laboratory, Portsdown West, UK
prsyms@dstl.gov.uk

Abstract

Whether it has been branded 'operational research' or not, there is a very long history, starting with Archimedes and King Hieron II of Syracuse, of military leaders inviting independent and objective advice on important matters of defence and security. This paper discusses the criteria defining military operational research, then traces its history from ancient times to the present. Some episodes, including the Prussian *Kriegsspliels* and the advice on convoying merchant shipping in two world wars, are well documented. Others are less well known, but deserve to be included in the story for their impact or their innovation. The years since World War Two have seen a waxing and waning of different OR methods and applications, with wargaming in the ascendant today, while historical analysis (at least in the USA) has fallen out of favour. So, what of the future? That is for the rest of the Defence and Security Stream.

Keywords: Decision-making; History of OR; *Kriegsspiel*; Lanchester models; Wargaming

1. Introduction

As long as there have been warfare and abstract thought, humans will have sat around the camp fire, discussing how best to conduct military operations. There is archaeological evidence from China, long before Sun Tzu wrote his first treatise on warfare (around 500 BCE), that people made playing pieces, presumably representing the forces on each side of a battle, either as a kind of wargame, or to rehearse a future plan. They were, in the broadest sense, tools for thinking about the conduct of warfare.

One early form of wargame, presumably with its origins further east, was formalised as chess in Persia, and from there spread into western Europe. Chess was in some ways very sophisticated, in that it differentiates the opposing forces into playing pieces with different characteristics, potentially representing heavy and light infantry, cavalry, elephants, and command units. But chess also assumes that both sides are identical, except in the skills that the opposing players bring to the board; in reality, very few military encounters are between equal sides.

2. Defining OR, and its early application

Lane (2010) made a strong case that Archimedes of Syracuse, who in 213–212 BCE advised King Hieron II on how to defend the city from siege and amphibious assault during the Second Punic War, was in fact an early operational research (OR) practitioner. Archimedes was a civilian thinker who was recognised by the king as a wise advisor, but one who stood outside the military hierarchy. He was a polymath, being skilled in mathematics, physics, engineering, and other disciplines, and crucially, he was able to take a holistic view of a problem – in this case, defending a city. These days we might have called him a systems engineer. He designed short- and long-ranged mechanical artillery systems, and a novel ship-sinking crane to prevent an enemy from scaling the walls from the sea. He took care to inspect the defences himself, siting his weapons to afford interlocking fields of fire, denying the enemy a covered approach. He also conducted experiments, and fed back the results into the weapon designs and deployment. In doing so, Lane (2010) argued that Archimedes' work satisfied all the essential features of OR:

- he took an interdisciplinary approach, using 'whole system' thinking;
- he inspected defences himself, gathering input data;
- he used processes that approximated to modelling;
- he used experimentation to improve the system;
- it was an excellent example of civilian-military collaboration in a crisis;
- he reported directly to the senior decision-maker (DM).

2.1. Wargaming, professional and amateur

The story resumes in the seventeenth century, when Prussian Army developed the first *Kriegsspiels* or wargames, to train their officer corps, and to rehearse plans for battles and operations. These were played on maps of real terrain, using wooden blocks to represent the opposing forces, and combat was adjudicated by experienced umpires, with the assistance of tables of casualties for different circumstances. These wargames were gradually adopted by other nations, as their armies became professionalised, with the United States Military Academy at Fort Leavenworth adding the use of dice to simulate the uncertainties that were not accounted for in their attrition tables. Prussia used these *Kriegsspiels* to plan their successful campaigns against Austria in 1866, against France in 1870–71, and their opening operations in the two world wars.

The idea of wargaming caught on in civilian circles, as industrially produced lead soldiers, and cannons firing match-sticks, became an essential toy in the late nineteenth century. The novelist H.G. Wells famously published a book of formalised wargames rules in 1913, illustrated with pictures of games played with G.K. Chesterton and Jerome K. Jerome, on a grand scale in a gentleman's games room.

The end of the nineteenth century also saw the first quantitative enquiry into the conditions that were necessary for victory in battle, beyond the military assertions that had been made regarding the 'three-to-one' rule of combat (see Kress & Talmor, 1999). Thus in 1897, Otto Berndt compiled statistics on recent historical battles and sieges, and his work was extended by Thomas Harbottle (in 1904), and Gaston Bodart (in 1908). Together they laid the foundations of historical battle analysis.

3. First World War: the first steps in modern OR

While it was not yet formally called operational research, there were several developments immediately preceding and during the First World War (WW1) that share some, if not all, of the features of OR. Germany had continued to use its *Kriegsspiels* to plan its campaigns, both in the west and in the east. Britain applied analytical thinking, but not necessarily mathematical modelling, to the problem of combatting the German submarine threat, resulting in the decision to adopt convoys. But the lasting OR legacy of WW1 was in the formulation of what became known as the Lanchester models of combat.

3.1. Lanchester combat models

In 1902, US Navy Lieutenant J.V. Chase formulated an idea in textual form, that if the rate at which a fleet of ships could disable the opposing fleet was proportional to the firepower generated by those ships, then the side with more ships would reinforce its initial numerical advantage over time, so that all else being equal, numbers were more valuable than individual firepower (see Hausken & Moxnes, 2000). Chase's idea was implemented numerically by another USN officer, Commander Bradley Fiske, who presented the evolution of a sea battle using Chase's assumptions, but presenting the results as a time-stepped table of remaining strengths on each side.

But it was a British polymath, the automotive and aeronautical engineer Frederick W. Lanchester, who in 1914 first formulated two pairs of differential equations that described Chase and Fiske's ideas in analytical form, and which form the basis of the whole branch of Lanchester models of warfare. They have since been extended include non-combat losses, and a family of equations that can be fitted to the entire range of combat intensities, from all-out warfare through to stabilisation operations (e.g. see Taylor, 1980; Bracken, 1995). Incidentally, the same differential equations were arrived at independently by Mikhail Ospiov in Russia, in 1915.

3.2. OR by Canada and France?

Finan and Hurley (1997) argued that Canada was the first to apply OR to a front-line problem in WW1, that of using microphones and state-of-the-art recording equipment to locate enemy artillery batteries that lay out of line-of-sight. While cross-disciplinary methods were undoubtedly brought to bear on an operational problem, and senior decision-makers were advised as to the best solution, in the present author's opinion it fails the criteria set out by Lane (2010), that OR should contain a human element. The problem was fundamentally one of acoustic engineering, and while they provided a sound solution, it was to an entirely technical problem.

A less well-known case is that of *Général* Alexandre Percin, who in 1921 sought to understand why so many French soldiers had been killed and injured by their own artillery. Percin assembled 298 accounts of friendly fire incidents by interview and personal correspondence, and in his book he attempted to draw out the common causal threads, in order to improve French artillery training and doctrine, so that friendly fire could be minimised in the future.

While his quantitative analysis was limited, his method of addressing a highly complex socio-technical problem by sampling operational events and looking for common features still underpins military historical analysis today, and still provides the starting point for studies of friendly fire (e.g. see Shrader, 1982; Syms, 2012).

3.3. Did OR make a difference in WW1?

Depending on what is counted as OR, it had limited impact; there is no evidence that DMs sought their advice in the way that they would later in the century. Battleship proponents used Lanchester's equations to argue for greater numbers, and the German Imperial staff had (as always) planned their initial operations using *Kriegsspiels*. But OR had yet to be unified under a single method or title.

4. Between world wars: the focus is on air power

In 1921, the Italian General Giulio Douhet attempted to demonstrate that a modest fleet of the new long-range bombers would be able to destroy a city and kill or demoralise its inhabitants in a few raids, forcing a country to capitulate before its armies had a chance to engage each other. After the costly attritional battles of WW1, this new technology found favour with the politicians in Italy, and later Germany, persuading them to invest heavily in strategic bombing as a decisive weapon, and during WW2, Britain and the USA followed Douhet's ideas enthusiastically, and on an even grander scale.

The flaws in Douhet's calculations were first, that given the technology of the time, not to mention the weather and the enemy, meant that it was by no means easy to hit an average-sized city from the air. But the greater error was in his calculated lethality: given the average population density of cities at the time, Douhet's calculations imply that bombers would kill around 50 people per tonne of bombs dropped, whereas in practice in WW2 the figure was around 0.9 per tonne, an over-estimate by a factor of 55. And even when Britain increased the tonnage of bombs dropped per raid from 500 (a heavy raid on London in 1940) to 3000 (a heavy raid on Germany in 1943) it failed to secure a quick victory. One of the greatest discrepancies was that Douhet had assumed that the civilian population would remain in place, whereas in fact they took shelter (something he should have known from the first WW1 bombings; see Corum, 1998).

Arguably, we still live with Douhet's legacy, with air forces offering politicians the prospect of a quick and relatively cheap victory, without committing 'boots on the ground'. Most recently air forces have promised to deliver decision but failed to do so in Kosovo in 1999, and in Lebanon in 2006 (see Cordesman, 2006; Matthews, 2008). In each case, their effects have been roughly two orders of magnitude less than was calculated before the start of the campaign.

4.1. United Kingdom air defence, and OR is finally called OR

Not surprisingly, given the emphasis placed on air power by Italy and then Germany, British thoughts turned to air defence. The Air Ministry, recognising that a major challenge of the new air defence system was integrating the new Chain Home radar system with the fighter squadrons that were to intercept the incoming bomber formations, called upon the physicist Professor Patrick Blackett, and other leading academics, to offer an outside view on how the sensor-effector system should be commanded and controlled. Collectively they used the techniques of mathematical modelling to reduce the problem to its essentials, and offered advice to the senior DMs (see Kirby & Capey, 1997a; Kirby, 2003). And in 1937, it had finally been called ‘operational research’, the term being coined by either R. Watson Watt, or A.P. Rowe (which one remains unclear; see Williams, 1967).

4.2. New statistics, and the birth of civil OR

Meanwhile, Ronald Fisher developed analytical statistics for the purposes of improving agriculture, and William Gosset (writing under the name of ‘Student’) developed more for the purposes of improving beer; but that lies outside the scope of the present paper.

5. OR in the Second World War: from simplicity to complexity

Many believe that OR started in WW2, but as I have demonstrated, it was more that it came of age. But it still took several years to percolate from addressing high-level operational questions in the relatively simpler air and naval environments, into the chaos and complexity of ground warfare.

5.1. OR in the naval domain

While WW2 naval operations were anything but simple, they did have the advantage from the point of view of analysis of being conducted by discrete elements of force, and on a more or less homogenous plane. Early OR studies focussed on how best to organise and protect merchant convoys, and searching for the submarines that threatened them. The most lasting legacy of WW2 naval OR was the development of search theory by Bernard Koopman and his team, which started the entire science of search, patrolling, and the evaluation of sensors, which is still used today (see Koopman, 1946 & 1956; Wagner *et al*., 1999).

One of the greatest (but least well known) OR successes was in the development of detailed convoy escort tactics by the Western Approaches Tactical Unit in Liverpool (Williams, 1979; Strong, 2017). Under the guiding genius of Captain Gilbert Roberts, the school developed and honed anti-submarine warfare (ASW) tactics, minimising the impact of German submarines on Allied convoys. At one point they anticipated the deployment of the acoustic homing torpedo against escorts, and had a successful counter-tactic wargamed and disseminated to the escort fleet within days of its first use.

5.2. OR on air operations

OR continued to be applied to air operations, and particularly to the business of strategic bombing. Studies addressed the volume of bombardment required to achieve a given military effect, and the more complex issue of how best to disrupt the enemy’s military, industrial, and transport infrastructure. Unfortunately in this case, the early misapplication of OR persuaded national DMs that aerial bombardment would be effective, whereas their later misgivings were not heeded for political reasons, and the bombing campaign was continued (Kirby & Capey, 1997b).

5.3. OR on ground operations

Ground operations were a later arrival at the growing OR party, largely on account of their greater complexity. Forces were distributed over the ground, rather than in discrete packets, the terrain was anything but flat and uniform, and the interactions were less clear cut. Professor Solly Zuckerman summed it up when he said, "… the situations on field [i.e. land] warfare were infinitely more varied than in either sea or air war." Hence land OR started by gathering data on casualties, the effects of bombardment, tank losses, and other low-level battlefield phenomena, trying to make sense of a confused picture, and to draw lessons that might improve its conduct. Not surprisingly, it was noted that professional biologists, such as Zuckermann, were some of the most effective at land OR, eager to dive into the chaos and complexity that might have daunted the mathematicians and physicists.

By the closing years of WW2, an OR section had accompanied the Allied invasion of Europe, and another was sent to support the Far East campaign. OR had started to penetrate all domains and all levels of military decision-making, and was an integral part of operations.

5.4. Did OR make a difference in WW2?

The new OR in the hands of the Western Allies was a powerful tool, and arguably saved them from defeat on one or two occasions. At the very least it had hastened the end of the war, saving thousands if not millions of lives on both sides.

While warfare had changed immeasurably since Archimedes, and even since the first Prussian *Kriegsspiels*, the essential nature of this early OR remained the same. It was pioneered by cross-disciplinary civilians, offering independent advice to top-level military decision-makers in times of crisis. A holistic, systems view was taken, including both the weapons and their human users. It included an abstract representation of real-world systems, using an approach that was now called 'modelling', but for the first time it incorporated the statistical techniques that were developed between the two world wars.

6. OR in the Cold War: from Carmonette to COEIAs

Following its success in WW2, the methods of OR were applied enthusiastically across government, both in managing the changed realities of a nuclear-armed world, and in improving the efficiency of the newly nationalised heavy industries in the UK (see Kirby, 2003). The story was repeated in the USA, where OR was embraced by private industry, particularly car manufacturing.

6.1. The USA adds computers: we can, and so we will?

Digital computers were a particular gift to land force OR, where the sheer number of entities presented computational challenges to traditional analytical techniques. The first computerised battle model was called Carmonette, in 1953. It was a stochastic simulation, first with tens, and then hundreds of entities on each side. As computers became more capable, so the models became larger and more complicated, but not necessarily better. In some of the worst cases, theatre-level battles were modelled down to company level, with thousands of entities moving down tram-lined routes, betrayed by inflexibility and poor data (discussed by Wilson, 1968). It was a now-familiar story of the effort being put into the computing and the coding, at the expense of thinking about how battles proceed, and of gathering validated input data. Computers were also used to support interactive wargaming, as discussed by Shephard (1963), but the interfaces proved painfully slow, making them expensive to run.

Computer models were built of every operational activity, from one-on-one engagements to wars of global scope; the present author is aware of over 400 computer models in the land domain alone, and suspects that many more were missed from the official catalogues. However, their focus of application shifted from operational problems to procurement, driven initially by the US Defense Secretary, Robert McNamara, an OR champion who had come from the Ford Motor Company. He mandated that every new defence equipment be proven to cost-effective, and to do this required detailed computer models that were able to compare each option in a range of scenarios. Some of these models ran to a million lines of code, and took weeks to prepare mountains of input data. To a degree, analysts had become servants of the models, rather than independent thinkers. The age where military OR was dominated by the combined operational effectiveness and investment appraisal (COEIA) lasted from the 1960s into the 2000s.

6.2. The decline and revival of front-line OR

As the decision-makers who had benefitted from OR in WW2 retired, so the perceived need for independent advice at the front line declined. By the end of the Cold War, military OR had, in the UK at least, been side-lined. It was largely confined to equipment procurement, and aside from helping plan the long-range air strike on Port Stanley, took almost no part in the Falklands operations in 1982. So when the Coalition was assembling a force to liberate Kuwait following the Iraqi invasion, the OR capability that had accompanied Allied forces in WW2 was no longer there, and a front line capability was rapidly re-created from experienced procurement analysts and field trials scientists. We quickly re-learned why we had developed OR in the first place, and today, that capability is still valued, having proven itself in the Balkans (see Neighbour *et al*., 2002), in Iraq, and in Afghanistan, as well as addressing the problems created by terrorism in our homelands.

7. OR in the Cold War: a renaissance in wargaming?

While the USA and UK governments were applying computers to professional OR, a game designer called Charles Roberts had seen a RAND wargame where they had superimposed a hexagonal grid on a map, and had devised a combat resolution table (CRT) using dice to determine the outcome of each combat. His ground-breaking hex-and-counter wargame, called *Afrika Korps*, was published by Avalon-Hill (AH) in 1964.

7.1. The heyday of hobby wargaming

Afrika Korps was followed by *PanzerBlitz*, one of the first tactical wargames, which was designed by Jim Dunnigan. Dunnigan went on to found Simulations Publications Incorporated (SPI), and together with AH and other companies, inspired a generation of pre-computer young people to study military history through the lens of manual simulations: *Soldiers, Sixth Fleet, Squad Leader, Air War*, and many more. Without knowing it, they had trained the next generation of OR scientists.

7.2. Governments tap the hobby wargaming goldmine

In 1976, the US Department of Defense (DoD) commissioned SPI to produce a tactical training game called *Firefight*, which they hoped would encourage serving officers to play through the kinds of tactical battles that they could well have faced, had the Soviet Union invaded Western Europe. By the 1980s, the US DoD had tapped into this new resource, and Jim Dunnigan, David Isby, and others were all employed in government OR in some capacity, as game designers, analysts, or modellers. In the UK a fresh generation of hobby wargamers, including the present author, inspired by the wave of American boardgames, joined the ranks of Cold War OR scientists.

After a relative decline in the 1990s and 2000s, wargaming has enjoyed something of a renaissance in the 2010s, as its value in generating insight into complex military problems is appreciated. The US *Connections* wargaming conference crossed the Atlantic in 2013, re-enthusing the UK's professional wargaming scene, while the Ministry of Defence (MoD) has commissioned a series of wargaming studies to examine how best to respond to new defence and security threats, following withdrawal from Afghanistan. And at this conference, Colonel Ivor Gardiner reminded us of the value of manual wargaming as an effective officer training tool, incidentally using AH's *Squad Leader* (see Gardiner, 2016).

8. The ups and downs of historical analysis

The collection of battle data had started at the turn of the twentieth century, but quantitative historical analysis (HA) had to wait until the advent of the digital computer. Pioneers including Robert Helmbold and Trevor Dupuy collected and analysed large databases of battles, greatly improving our understanding of modern conflict, and demonstrating that Lanchester's models, and military 'rules of thumb' such as the 3:1 superiority needed to attack, should be treated with great caution (e.g. see Helmbold, 1961; Dupuy, 1988).

HA has proven a particularly effective tool for exploring the effects of human factors that cannot be modelled in computer simulations (see, for example, Rowland, 1987). But it fell out of favour with the US DoD in the 2000s, and also with the MoD, as customers questioned its relevance to the modern world. However, it has recently enjoyed a revival in the UK, as government started to demand that business cases be supported by a wider range of methods, not relying on (for example) computer simulations (see Macpherson, 2015). It also serves as a corporate memory for the armed forces, enabling them to draw on enduring lessons that were learned before the present generation of senior DMs served.

9. An emerging pattern?

If there is an emerging pattern, it is this. A city state, a nation, or one of its component forces finds itself in a perilous military situation; defeat looms. Its leaders, having exhausted the solutions offered by existing military doctrine, recognise that a fresh look at the problem by someone outside the military command structure might be beneficial. Academics and other independent thinkers either volunteer their services, because they can see another potential way of solving a problem, or they are brought in as an act of desperation. They are cross-disciplinary, and use the scientific method to solve problems, often using no more than 'back-of-envelope' models. OR is hailed as the saviour of the day.

Once the crisis is over, DMs are convinced of the value of OR, and institutionalize it in their management processes. But once these DMs retire, new ones with no experience of the *value* of OR, but who are all too aware of its cost, believe that they can make all the important decisions without independent advice. At best, OR becomes 'handle-turning' to support procurement, far removed from the front line; at worst the capability, or parts of it, is disbanded as a cost-saving measure. OR becomes marginalized, is ignored, or used on minor decisions, until ... another crisis!

And so our duty as OR practitioners is to remind those who might need our advice in a crisis of our illustrious history, and to keep our tools sharp for when they are needed. For if one thing is assured, the coming century will present several crises, both civil and military, and defence analysts will be called upon to advise.

Those interested in more detail on the history of OR are recommended the books by Wilson (1968), Kirby (2003), and Shrader (2006).

References

Berndt O (1897). *Die Zahl im Kriege: Statistische Daten aus der neueren Krigsgeschichte in graphischer Dartellung*. (Numbers in war: statistical data from modern military history in graphical form). Freytag & Berndt, Vienna: 169 pp.

Bodart G (1908). *Militär-historisches Kriegs-Lexikon (1618–1905)*. (Military-Historical Battle Dictionary 1618–1905). C W Stern, Vienna and Leipzig: 234 pp.

Bracken J (1995). Lanchester Models of the Ardennes Campaign. Naval Research Logistics 42(5): 559–577.

Cordesman A H (2006). Preliminary "Lessons" of the Israeli-Hezbollah War. Center for Strategic and International Studies, Washington, DC, draft paper.

Corum J S (1998). Inflated by Air: Common Perceptions of Civilian Casualties from Bombing. Research Report, Maxwell Air Force Base, Alabama.

Douhet G (1921). Il dominio dell'aria. Translated by D. Ferrari as, The Command of the Air. US Office of Air Force History, 1983: 394 pp.

Dupuy T N (1987). Understanding War: History and Theory of Combat. K S Giniger, New York; Leo Cooper, London: 312 pp.

Finan J S & Hurley W J (1997). McNaughton and Canadian operational research at Vimy. Journal of the Operational Research Society 48(1): 10–14.

Gardiner I (2016). Wargaming: An Overlooked Educational Tool. British Army Review 165: 65–73.

Harbottle T B (1904). Dictionary of Battles, from the Earliest Date to the Present Time. Dutton, New York: 303 pp.

Hausken K & Moxnes J F (2000). The Micro-Foundations of the Lanchester War Equations. Military Operations Research 5(3): 79–99.

Helmbold R L (1961). Historical Data and Lanchester's Theory of Combat. CORG Staff Paper CORG-SP-123.

Kirby M W & Capey R (1997a). The air defence of Great Britain, 1920–1940: an operational research perspective. J Operational Research Society 48(6): 555–568.

Kirby M W & Capey R (1997b). The area bombing of Germany in World War II: an operational research perspective. J Operational Research Society 48(7): 661–677.

Kirby M W (2003). Operational Research in War and Peace. Imperial College Press, London, and The Operational Research Society, Birmingham: 444 pp.

Koopman B O (1946). Search and Screening. OEG Report No. 56, Summary Reports Group, Columbia University Division of War Research; Center for Naval Analyses.

Koopman B O (1956). The theory of search II. Target detection. Operations Research 4(5): 503–531.

Kress M & Talmor I (1999). A new look at the 3:1 rule of combat through Markov Stochastic Lanchester models. J Operational Research Society 50(7): 733–744.

Lanchester F W (1914). Aircraft in Warfare: The Dawn of the Fourth Arm – No. V, the principle of concentration. Engineering 98: 422–423.

Lane D C (2010). High Leverage Interventions: Three Cases of Defensive Action and Their Lessons for OR/MS Today. Operations Research 58(6): 1535–1547.

Macpherson N (2015). The Aqua Book: guidance on producing quality analysis for government. HM Treasury PU1740.

Matthews M M (2008). We Were Caught Unprepared: The 2006 Hezbollah-Israel War. Occasional Paper 26. US Army Combined Arms Center, Fort Leavenworth, Kansas.

Neighbour M R, Bailey P, Hawthorn M, *et al*. (2002). Providing Operational Analysis to a Peace Support Operation: The Kosovo Experience. J Operational Research Society 53(5): 523–543.

Osipov M (1915). The Influence of Numerical Strength of Engaged Forces on their Casualties. Translated 1995 by R L Helmbold & A S Rehn. Naval Research Logistics 46(3): 435–490.

Percin A (1921). Le Massacre de Notre Infanterie 1914–1918. Michel Albin, Paris: 301 pp.

Rowland D (1987). The Use of Historical Data in the Assessment of Combat Degradation. J Operational Research Society 38(2): 149–162.

Shephard R W (1963). War Gaming as a Technique in the Study of Operational Research Problems. Operational Research Quarterly 14(2): 119–130.

Shrader C R (1982). Amicicide: The Problem of Friendly Fire in Modern War. US Command & General Staff College, Fort Leavenworth, Kansas, Combat Studies Institute Research Survey No. 1.

Shrader C R (2006). History of Operations Research in the Unites States Army. Office of the Deputy Under Secretary of the Army for OR, Washington, DC: 211 pp.

Strong P.E. (2017). Wargaming the Atlantic War: Captain Gilbert Roberts and the Wrens of the Western Approaches Tactical Unit. MORS Wargaming Special Meeting, Working Group 2, Validity and Utility of Wargaming.

Syms P R (2012). An Historical Analysis of Fratricide. In: Kirke C.M.StG. (ed). Fratricide in Battle: (Un)Friendly Fire. Continuum International, London: 266 pp.

Taylor J G (1980). Force-on-Force Attrition Modelling. (2 volumes) MAS/ORSA, Arlington, Virginia: 1426 pp.

Wagner D H, Mylander W C, & Sanders T J (1999). Naval Operations Analysis. Naval Institute Press, Annapolis Maryland: 421 pp.

Wells H G (1913). Little Wars. Frank Palmer Ltd., London; facsimile reprinted by Arms & Armour, London, 1970: 111 pp.

Williams E C (1967). The Origin of the Term "Operational Research" and the Early Development of the Military Work. Operational Research Quarterly 19(2): 111–113.

Williams M (1979). Captain Gilbert Roberts R.N. and the Anti-U-Boat School. Cassell & Co., London: 224 pp.

Wilson A (1968). The Bomb and The Computer. Barry & Rockliff; re-published as, War Games. Penguin Books Ltd., Harmondsworth, 1970: 202 pp.

P.R. Syms

OR60 Annual Conference - Keynote Papers and Extended Abstracts
Lancaster University. Lancaster. 11-13 September 2018

Refugee Resettlement via Machine Learning and Integer Optimization

Andrew C. Trapp

Worcester Polytechnic Institute
Worcester, MA
atrapp@wpi.edu

Alexander Teytelboym

University of Oxford
Oxford, UK
alexander.teytelboym@economics.ox.ac.uk

Narges Ahani

Worcester Polytechnic Institute
Worcester, MA
nahani@wpi.edu

Tommy Andersson

Lund University
Lund, Sweden
tommy.andersson@nek.lu.se

ABSTRACT

Around 100,000 refugees are resettled to dozens of countries from conflict zones every year. While there is growing evidence that the initial placement of refugee families profoundly affects their lifetime outcomes, there have been few attempts to optimize resettlement destinations within host countries. We describe how machine learning and integer optimization can be used to empower resettlement agencies to drastically improve refugee employment outcomes. We describe possible future work on multi-objective optimization, the dynamics of allocation, and the inclusion of refugee preferences.

Keywords:

refugees, humanitarian operations research, machine learning, integer optimization, multiple multidimensional knapsack problem, matching

1 INTRODUCTION

In 2017, there were 18.5 million refugees—the highest number ever recorded—under the mandate of the Office of the United Nations High Commissioner for Refugees (UNHCR 2018). Of those, the UNHCR considers 1.2 million to be in need of *resettlement*—permanent relocation from their asylum country to a third country (UNHCR 2017). In 2016, the number of resettlement submissions reached 165,000 (a twenty-year high) and 125,800 people departed for resettlement (UNHCR 2017). Clearly, resettlement places offered by host countries are incredibly scarce. Currently, most refugees departing for resettlement are Syrians who are seeking asylum in Jordan and Lebanon, but there are also thousands of resettled refugees from the Democratic Republic of the Congo, Iraq, Somalia, and Myanmar. Refugees in need of resettlement are particularly vulnerable: a quarter are survivors of torture and a third face persecution in their country of origin (UNHCR 2017, Annex 3).

Dozens of countries resettle refugees.[1] The United States (US) is by far the world's largest destination of resettled refugees with 78,340 admitted in 2016.[2] The US State Department delegates the resettlement process to nine resettlement agencies, known as Voluntary Agencies (VolAgs).[3] The VolAgs are responsible for developing their own networks of "affiliates"—communities that welcome refugees and help them integrate into a new life in the US. Affiliates (which we refer to as *localities*) offer resettlement capacity voluntarily.

There is ample empirical evidence that the initial placement of refugees determines their lifetime outcomes (Åslund and Rooth 2007, Åslund and Fredriksson 2009, Åslund, Östh, and Zenou 2010, Åslund, Edin, Fredriksson, and Grönqvist 2011, Damm 2014, Feywerda and Gest 2016). Therefore, ensuring the optimality of the initial match between the refugee family and the community is an important challenge from social, economic, and humanitarian perspectives. However, there is currently little evidence that resettlement capacity offered by localities is being used to maximize either the welfare of refugees or of the host population.

In this paper, we describe how machine learning and integer optimization can improve the decision-making of the resettlement agencies and outcomes of refugees. Like Bansak, Ferwerda, Hainmueller, Dillon, Hangartner, Lawrence, and Weinstein (2018), we aim to maximize the employment outcomes (at three months after arrival) of resettled refugees. However, we incorporate an important additional complexity in resettlement. Localities face a number of constraints on the services they can offer to refugees, such as whether they can provide single-family support, whether they have appropriate translators, and whether they can adequately house and help large families (with six or more members). Some of these constraints are hard zero-one constraints while others might be quantity constraints (for example, the number of spots in employment training).[4] As refugees typically arrive in large families with children, they often require myriad services. Therefore, localities are often constrained in the kinds of refugee families they can reasonably host. Hence, optimizing refugee outcomes, while meeting refugees' needs and respecting various hard constraints of the localities, turns out to be a complex operations research task. Here, we describe the basic framework of our empirical and optimization models. The design of the software, the preliminary results of the optimization, and the experiences of the practitioners are laid out in detail in Trapp, Teytelboym, Ahani, and Andersson (2018). We conclude with an outlook for future work in refugee resettlement optimization within the growing field of humanitarian operations research (Holguín-Veras, Pérez, Jaller, Van Wassenhove, and Aros-Vera 2013, Pedraza-Martinez and Van Wassenhove 2016).

2 MATCHING MODEL

The model of *matching with multidimensional constraints* is based on Delacrétaz, Kominers, and Teytelboym (2016).

There is a finite set of *refugee families* $\mathscr{F}$. A family $F \in \mathscr{F}$ is itself a finite set of size $|F|$. That is, a family F has $|F|$ members with a typical member denoted $f \in F$. There is a finite set of *localities* L. Families are inseparable so all family members must be placed in the same locality. Refugee families require *multiple units* of *different services* (for example, employment training slots, language support, single-parent support etc.) from a set S. We denote by ν the matrix of family *service needs*, with typical element $\nu_s^F \in \mathbb{Z}_{\geq 0}$ denoting the total number of units of service

[1] For refugee allocation mechanisms *across* countries, see Moraga and Rapoport (2014) and Jones and Teytelboym (2017a).

[2] In terms of per capita refugee resettlement, US is behind Canada, Norway, and Australia.

[3] These are: Church World Service (CWS), Ethiopian Community Development Council (ECDC), Episcopal Migration Ministries (EMM), Hebrew Immigrant Aid Society (HIAS), International Rescue Committee (IRC), Lutheran Immigration and Refugee Services (LIRS), US Committee for Refugees and Immigrants (USCRI), US Conference of Catholic Bishops (USCCB), and World Relief Corporation (WR).

[4] While most constraints are infrastructure constraints, there might also be social or political constraints.

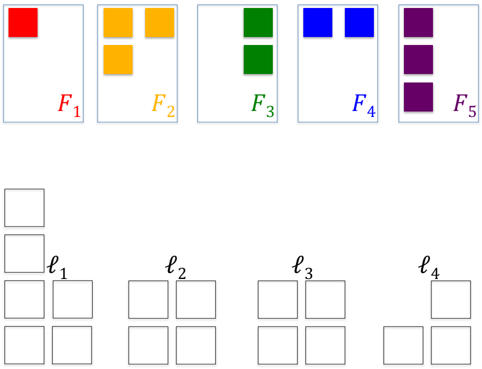


Figure 1: An instance of OOP.

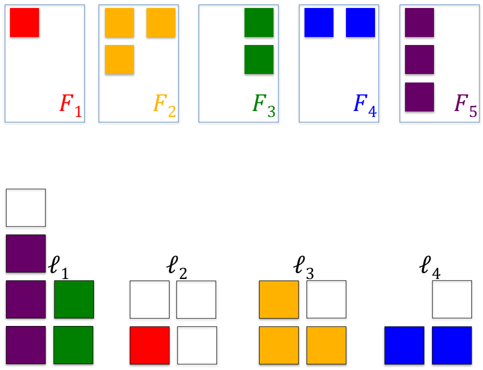


Figure 2: A feasible solution to the instance.

s required by family F. A refugee family F can only be assigned to locality $\ell \in L$ if ℓ can provide services to meet all the needs of F. We denote by κ the matrix of locality *service capacities*, with typical element $\kappa_s^\ell \in \mathbb{Z}_{\geq 0}$ denoting the number of units of service s that locality ℓ can provide.[5]

We summarize the estimated quality of each refugee-locality match with a single number called the *quality score*. Let $q : \mathscr{F} \times L \to \mathbb{R}_{\geq 0}$ denote the quality score function. Setting $q = 1$ (or any other constant) would maximize the total number of families that are resettled, while setting $q(F,\ell) = |F|$ would maximize the total number of refugees that are resettled. We will be interested in the case where q represents the employment outcome of family F in locality ℓ and can be estimated from data using observable locality and family characteristics.

Let us now state the problem of maximizing the overall observed quality of the match via integer optimization. We introduce a binary variable $x(F,\ell)$ which is equal to 1 if family F has been matched to locality ℓ and 0 otherwise. In order to maximize the outcome within the feasibility constraints, the social planner solves the following *outcome optimization problem* (OOP):

$$\max \sum_{F \in \mathscr{F}} \sum_{\ell \in L} q(F,\ell) x(F,\ell) \quad \text{subject to:} \quad \begin{aligned} &\sum_{F \in \mathscr{F}} \nu_s^F x(F,\ell) \leq \kappa_s^\ell && \forall \ell, s \\ &\sum_{\ell \in L} x(F,\ell) \leq 1 && \forall F \\ &x(F,\ell) \in \{0,1\} && \forall F, \ell \end{aligned}$$

The OOP is an example of a *0–1 multiple multidimensional knapsack problem* (Song, Zhang, and Fang 2008)—a combination of the 0–1 multiple knapsack problem (Martello and Toth 1980) and the 0–1 multidimensional knapsack problem (Fréville 2004). Clearly, the multiple multidimensional knapsack problem is NP-hard. Even so, instances of OOP involving tens of localities and hundreds of families run in just seconds on a modern laptop, even with open-source software..

Figure 1 illustrates an instance of the outcome optimization problem with five families, four localities, and two services (left and right). Here, family F_2 requires two units of the left service and one unit of the right service whereas locality ℓ_1 provides four units of the left service and two units of the right service. Figure 2 presents a feasible outcome—which satisfies all the constraints—for this instance (a feasible outcome always exists).

[5]The theoretical set-up is unchanged if we assume that $\nu_s^F \kappa_s^\ell \in \mathbb{R}_{\geq 0}$. As an example, we might wish to capture that a refugee may require $\frac{1}{7}$ of a dialysis machine because she needs to use it once a week.

3 PREDICTIVE MODEL

Our predictive model is based on recent work by Bansak, Ferwerda, Hainmueller, Dillon, Hangartner, Lawrence, and Weinstein (2018). Our empirical data come from a resettlement agency over the course of several years. The basic model is a (regularized) logistic regression. The dependent variable is a binary outcome $q(f,\ell)$ of whether an adult refugee f has managed to secure employment within three months of arrival to the US in locality ℓ. The independent variables include: gender, English-speaking ability, age, presence of a severe medical condition, whether the refugee required urgent resettlement, family size, education, as well as nationality dummies and time trends. We run a different logistic regression for each locality, dropping the localities with few observations. Feywerda and Gest (2016) argue that, since the assignment of refugees to locality is random conditional on observables, the model estimates should deliver the causal effect of initial placement in a locality on employment. Once the model has been estimated, we can use it to predict $\hat{q}(f,\ell)$ out-of-sample. This gives us an estimate of the probability of employment for every adult refugee in every locality. We then set $q(F,\ell) = \sum_{f \in F} \hat{q}(f,\ell)$ so the quality score for a family-locality pair is the sum of the expected employment probabilities of the adult members of the family. Our predictive model can explain heterogeneity in observed employment outcomes for different refugees within localities.

4 EXTENSIONS

There are at least three ways in which our approach can be extended. First, while localities agree on annual quotas, refugees arrive stochastically over the course of a year. Therefore, it is important to schedule the arrival of refugees. Andersson, Ehlers, and Martinello (2018) tackle this problem in the Swedish context. Second, we could include multiple objectives such as longer-term employment, income, health and education outcomes. Unfortunately, at the time of writing, no data on these objectives for resettled refugees arriving in the US appears to be systematically available. Third, we could include preferences of refugees and priorities of localities explicitly (Delacrétaz, Kominers, and Teytelboym 2016, Jones and Teytelboym 2017b, Aziz, Chen, Gaspers, and Sun 2017). Preferences could be collected during the refugee pre-arrival orientation using a questionnaire that elicits how refugees might trade off features of areas (for example, crime, climate, population density, amenities, quality of schools). However, including preferences while optimizing for a particular observable outcome can be a challenging problem (Andersson and Ehlers 2016, Biró and Gudmundsson 2018).

5 CONCLUSION

Refugee resettlement is a complex humanitarian problem which requires insights from a number of disciplines, including operations research, economics, statistics, political science, and sociology. Many new systems and approaches are urgently needed to improve the livelihoods of the resettled refugees and the communities into which they integrate. In this paper, we show how combining tools from machine learning and integer optimization can potentially improve refugees' outcomes within a particular hosting country. In forthcoming work, we report on our initial attempt to implement the model described in this paper in practice (Trapp, Teytelboym, Ahani, and Andersson 2018).

ACKNOWLEDGMENTS

We are incredibly grateful to a number of dedicated practitioners for helping us implement this research. Emily Aiken provided invaluable software development assistance. Scott Duke Kominers provided valuable comments on the first draft. Andersson gratefully acknowledges the financial support of the Jan Wallander and Tom Hedelius Foundation (Research Grant P2016-0126:1) and the

Ragnar Söderberg Foundation (E8/13). Teytelboym gratefully acknowledges the Skoll Centre for Social Entrepreneurship Research Accelerator Grant ("Implementing the Local Refugee Match"). This work was supported by the Economic and Social Research Council grant number ES/R007470/1. Trapp and Ahani are grateful for the support provided by the Foisie Business School at Worcester Polytechnic Institute.

REFERENCES

Andersson, T., and L. Ehlers. 2016. "Assigning Refugees to Landlords in Sweden: Efficient Stable Maximum Matchings". Technical report.

Andersson, T., L. Ehlers, and A. Martinello. 2018. "Dynamic Refugee Matching". Technical report.

Åslund, O., P.-A. Edin, P. Fredriksson, and H. Grönqvist. 2011. "Peers, neighborhoods, and immigrant student achievement: Evidence from a placement policy". *American Economic Journal: Applied Economics* 3 (2): 67–95.

Åslund, O., and P. Fredriksson. 2009. "Peer effects in welfare dependence quasi-experimental evidence". *Journal of Human Resources* 44 (3): 798–825.

Åslund, O., J. Östh, and Y. Zenou. 2010. "How important is access to jobs? Old question, improved answer". *Journal of Economic Geography* 10 (3): 389–422.

Åslund, O., and D.-O. Rooth. 2007. "Do when and where matter? Initial labour market conditions and immigrant earnings". *The Economic Journal* 117 (518): 422–448.

Aziz, H., J. Chen, S. Gaspers, and Z. Sun. 2017. "Stability and Pareto optimality in refugee allocation matchings". Technical report, Mimeo.

Bansak, K., J. Ferwerda, J. Hainmueller, A. Dillon, D. Hangartner, D. Lawrence, and J. Weinstein. 2018. "Improving refugee integration through data-driven algorithmic assignment". *Science* 359 (6373): 325–329.

Biró, P., and J. Gudmundsson. 2018, March. "Complexity of Finding Pareto-Efficient Allocations of Highest Welfare". Technical report.

Damm, A. P. 2014. "Neighborhood quality and labor market outcomes: Evidence from quasi-random neighborhood assignment of immigrants". *Journal of Urban Economics* 79:139–166.

Delacrétaz, D., S. D. Kominers, and A. Teytelboym. 2016. "Refugee resettlement". Technical report, MImeo.

Feywerda, J., and J. Gest. 2016. "Location, Location: Refugee Resettlement and Integration Outcomes in the United States". Mimeo.

Fréville, A. 2004. "The multidimensional 0–1 knapsack problem: An overview". *European Journal of Operational Research* 155 (1): 1–21.

Holguín-Veras, J., N. Pérez, M. Jaller, L. N. Van Wassenhove, and F. Aros-Vera. 2013. "On the appropriate objective function for post-disaster humanitarian logistics models". *Journal of Operations Management* 31 (5): 262–280.

Jones, W., and A. Teytelboym. 2017a. "The International Refugee Match: A System that Respects Refugees' Preferences and the Priorities of States". *Refugee Survey Quarterly* 36 (2): 84–109.

Jones, W., and A. Teytelboym. 2017b, forthcoming. "The Local Refugee Match: Aligning refugees' preferences with the capacities and priorities of localities". *Journal of Refugee Studies*.

Martello, S., and P. Toth. 1980. "Solution of the zero-one multiple knapsack problem". *European Journal of Operational Research* 4 (4): 276–283.

Moraga, J. F.-H., and H. Rapoport. 2014. "Tradable immigration quotas". *Journal of Public Economics* 115:94–108.

Pedraza-Martinez, A. J., and L. N. Van Wassenhove. 2016. "Empirically grounded research in humanitarian operations management: The way forward". *Journal of Operations Management* 45:1–10.

Song, Y., C. Zhang, and Y. Fang. 2008. "Multiple multidimensional knapsack problem and its applications in cognitive radio networks". In *MILCOM 2008-2008 IEEE Military Communications Conference*, 1–7. IEEE.

Trapp, A. C., A. Teytelboym, N. Ahani, and T. Andersson. 2018. "Resettlement Agency Uses Software to Optimize Placement of Refugees". Technical report, in preparation.

UNHCR 2017, June. "Global Resettlement Needs 2018". Technical report, United Nations High Commissioner for Refugees.

UNHCR 2018, March. "Mid-Year Trends, June 2017". Technical report, United Nations High Commissioner for Refugees.

AUTHOR BIOGRAPHIES

ANDREW C. TRAPP is Associate Professor of Operations and Industrial Engineering at Worcester Polytechnic Institute (WPI), and holds courtesy professorships in Mathematical Sciences and Data Science. His research focuses on using analytical techniques, in particular mathematical optimization, to identify optimal decisions to problems arising from a diverse cross-section of sectors such as humanitarian operations, sustainability, healthcare, and data mining. Andy develops new theory, models, and computational solution approaches to tackle such problems. http://users.wpi.edu/∼atrapp/

ALEXANDER TEYTELBOYM is Associate Professor of Economics at the University of Oxford. He is also a Tutorial Fellow at St. Catherine's College and Deputy Director of the Economics of Sustainability Programme at the Institute for New Economic Thinking at the Oxford Martin School, University of Oxford. http://www.t8el.com

NARGES AHANI is a doctoral student at Worcester Polytechnic Institute in Data Science program. She received a Bachelor's degree in Electrical Engineering at Shahroud University of Technology (2001—2005). She has ten years of experience in software development, application architecture, and hardware programming and design. She is completing an MS in Operational Analytics and Management at Foisie School of Business at WPI (2017—2018).

TOMMY ANDERSSON is Professor in Microeconomics at the Department of Economics, Lund University, and an External Fellow at the Center for Mechanism and Institution Design at University of York. He is a Board Member of the Society for Economic Design (2014—present), an elected Board Member of the Society for Social Choice and Welfare (2016—2022), and a member of the Scandiatransplant Steering Commitee for the Scandiatransplant Paired Kidney Exchange Program. https://sites.google.com/site/tommyanderssonlunduniversity/home

OR60 Annual Conference - Keynote Papers and Extended Abstracts
Lancaster University. Lancaster. 11 - 13 September 2018

Performance Modelling, Impact Assessment and Decision Analysis of Multi-vector Decentralised Energy Systems

Jian-Bo Yang, Ting Wu and Dong-Ling Xu

Alliance Manchester Business School
The University of Manchester, Manchester M13 9SS, UK
{jian-bo.yang, ting.wu-2, l.xu}@manchester.ac.uk

Fossil fuels like coal, oil and gas mainly utilized by large-centralized power plants are limited and inadequately distributed in the world. The consumption of conventional energy sources has a high impact on global warming in the form of CO2 emission. There has been growing recognition that renewable energy is essential for the world economy and the current and future generations' welfare.

Renewable energy contributes to attaining the general goal of energy security and environmental protection in a balanced way. For energy sustainability, requirements for future energy include long-term supply, stable prices, steady technology improvement and continuous installation and maintenance. Sustainable energy development should take into account not only energy saving, but also energy efficiency and flexible combination of different types of energy. The trend of future energy should therefore be the development of more renewable energy, while transferring a centralized energy system to a clean and decentralized energy system.

Decentralized energy (DE) is also called distributed energy. It is usually produced close to where it is consumed, in contrast to centralised energy which is produced at large plants and transmitted through the national grid. DE is regarded to be central to the world's future energy and economic strategies, and DE solutions are likely to have an increasingly important role to play in the national energy landscape of many countries. Currently, Europe is making a transition from a centralized and largely fossil-fuel and nuclear-based system delivering electricity to passive consumers toward a more decentralised energy system, which mostly uses renewable energy sources, such as small hydro, solar, wind, biomass, biogas, geothermal power, etc.

The main problem with renewable energy sources in a DE system is its high dependency on environmental conditions like wind speed and solar irradiance. Single renewable energy source in particular wind and solar does not provide continuous power supply because of its uncertainty and intermittent nature. This makes it necessary to integrate different renewable energy sources, including wind, solar, hydro, biogas and storage unit, to form a hybrid system for more reliable and environmentally friendly energy supply. Given the novelty and relatively short development history of DE systems, their performances and potential impact on world economy have not yet been studied systematically, and there are also challenges and barriers to renewable energy generation, distribution and consumption, which involve technical, economic, cultural and financial aspects. There is an urgent need to systemically model, analyse and assess the cost-effectiveness and the societal and environmental impact of various DE solutions which are based on different types of renewable energy. This requires the systematic and consistent handling of multiple factors of both a quantitative and qualitative nature under uncertainty, which in essence is a multiple criteria decision analysis (MCDA) problem but needs to make use of both numerical data and expert knowledge.

Among many MCDA methods, the evidential reasoning (ER) approach is a generic evidence-based MCDA approach and uses a belief structure or extended probability distribution to represent the assessment of an alternative on each attribute as a piece of evidence, whether it is qualitative or quantitative . The aggregation of multiple criteria in the ER approach is by combining the extended probability distributions. The weights and reliabilities of assessments collected from multiple sources can be taken into account. In this way, the ER approach can deal with various types of uncertainty in MCDA, form a solid basis for sensitivity analysis and provide a panoramic view for informative decision analysis. In this presentation, the ER approach will be discussed in the context of analysing the performance and impact of DE systems.

Modelling performance and analysing impact of hybrid DE systems using the ER approach includes the following main tasks: 1) investigating important issues about DE systems and their contributions to green economy, which are of common and widespread interests to many countries; 2) developing a hierarchical framework to evaluate the performances of individual renewable energy systems and also hybrid DE systems by taking account of technical, economic, environmental and social aspects; 3) developing MCDA models based on the ER approach to support the performance assessment and impact analysis; 4) collecting data from different sources, conducting empirical studies and validating the framework of modelling and assessing the performance, cost-effectiveness, societal and environmental impact of hybrid DE systems.

In this presentation, we will show the preliminary investigation into a case study about a big micro-grid project in Inner Mongolia, located at the northwest region of China. Recently, individual DE systems have been constructed for each area of one county and five cities in the above region, which are then connected to a big energy network. Every individual DE system is a hybrid system that includes various types of (multi-vector) energy including solar, wind and storage unit. The entire energy network covering the five areas is a complex hybrid system that includes energy generation, energy transmission and trading. The project was initially conducted mainly using experts' knowledge, rather than systematic performance modelling and impact analysis.

In this presentation, we will discuss the preliminary findings from the systematic performance modelling and impact analysis of DE systems including the above case study. We believe that many stakeholders can directly benefit from these research findings, including policy makers, energy suppliers and consumers, energy network owners, and DE investors and stakeholders in local communities, who have direct interests in the generation, transition and consumption of renewable energy.

OR60 Annual Conference - Keynote Papers and Extended Abstracts
Lancaster University. Lancaster. 11 - 13 September 2018

Strategic Predictors, Calibration Tests and Scoring Functions for Probabilistic Forecasts

James W. Taylor

Saïd Business School, University of Oxford, UK
james.taylor@sbs.ox.ac.uk

Abstract

Probabilistic forecasts are needed to support decision making in many applications. The forecast accuracy of different predictors can be compared using scoring functions, and insight can be provided by calibration tests. These tests evaluate the consistency of the prediction with the observations, and are used by some as the sole approach to evaluation. It is important that calibration tests cannot be gamed by predictors that, although clearly inadequate, have been strategically designed to pass the test. Drawing on previous results for quantile prediction, we show how strategic predictors exist for previously proposed calibration tests for quantile-bounded intervals, expectiles and event probability forecasts, and also for the widely-used test for distributional forecasts based on the probability integral transform. To address this, we introduce and extend existing regression-based calibration tests, which cannot be gamed. We make further contributions related to expectiles. Motivated by the attractive properties of expectiles, we introduce the idea of expectile-bounded interval forecasts. We provide interpretation for these intervals, and present a scoring function and calibration test. We also introduce, for distributional forecasts, a new scoring function based on expectiles. We use financial and environmental data to illustrate several of our new proposals.

Keywords: Prediction; evaluation; quantiles; expectiles; intervals; distributions

Evaluation of Probabilistic Forecasts

Probabilistic forecasts are needed to support decision making in many different contexts. In some applications, a prediction of the complete probability distribution is required, while in others, it is a point-valued summary of the distribution that is of interest. For example, quantile predictions may be needed as measures of tail risk, as bounds on prediction intervals, and as the optimal point forecasts when an asymmetric loss function is appropriate. Forecasts of expectiles are also of interest, and are receiving increasing attention. Expectiles can be viewed as generalisations of the mean, just as quantiles can be viewed as generalisations of the median. Another probabilistic measure of interest is an event probability, such as an exceedance probability.

The aim of distributional forecasting is to maximize sharpness subject to calibration (Gneiting et al. 2007). Sharpness relates to the concentration of the predictive distributions, while calibration concerns their statistical consistency with the observations. If a predictive distribution is calibrated, randomly sampled values from it will be indistinguishable from the observations (Gneiting and Katzfuss 2014). A scoring function summarises both calibration and sharpness, and is described as proper if minimized by the correct forecast. The use of a proper score encourages honest reporting by a forecaster (Gneiting and Raftery 2007). However, a score can be used only to rank predictors, while a calibration test can provide insight leading to improved accuracy.

Forecasts of quantiles and expectiles should also be evaluated in terms of proper scores and calibration tests (Nolde and Fissler 2017). A quantile predictor is conditionally calibrated if exceedances beyond the quantile forecasts occur independently, and with probability equal to the probability level of the quantile. A binary variable, indicating exceedance, should have zero autocorrelation, and a mean equal to the probability level, and this has been the focus of calibration tests. However, Engle and Manganelli (2004) present a quantile predictor that, although inadequate, is able to pass such a test. Their predictor can be viewed as a dishonest forecaster that has strategically manipulated the forecasts in order to pass the test. Just as propriety is necessary for a score to ensure honest reporting, calibration tests should not permit strategic behaviour. Engle and Manganelli (2004) provide a regression-based test that cannot be gamed by their strategic quantile predictor.

In this paper, we contribute to the literature on calibration testing by first showing that strategic predictors are a concern for existing tests of interval forecasts, expectile forecasts and binary event probability predictions. In each case, we show how including appropriate regressors in regression-based tests can guard against strategic predictors. We also present a strategic distributional predictor that, although clearly inadequate, is able to pass the widely-used calibration test of Berkowitz (2001) for distributional forecasts. This test focuses on the probability integral transform (PIT), which is the value of the distributional forecast at the observation. The PIT should be an independent random variable, uniformly distributed between 0 and 1. We present an augmented version of the Berkowitz (2001) test, which cannot be gamed.

Although strategic predictors will be exposed as inadequate by a proper score, there are several reasons why it remains a concern that a calibration test can be gamed. First, evaluating just calibration has been common for quantile and distributional forecasts. This is perhaps due to tradition, the intuitive simplicity and informative nature of calibration tests, or the perceived computational inconvenience of some scores. Second, comparing predictors may involve a trade-off between the results of a calibration test and a score, and so a seemingly calibrated strategic predictor may be viewed as dominating a competitor that has a better score but fails a calibration test. Third, a forecaster that has the behaviour of a strategic predictor to some extent, or for some of the time, may be competitive in terms of both a calibration test and score.

We also make further contributions to the literature on expectiles. Expectiles have a number of appealing theoretical and practical properties. Although tail quantiles have been dominant in communicating forecast uncertainty, they convey no information regarding the size of quantile exceedances. By contrast, the value of an expectile is dictated by the whole distribution. Unlike expectiles, quantiles are typically not uniquely defined, and this can be particularly apparent with discrete data. Motivated by the attractive properties of expectiles, we propose expectile-bounded intervals. We interpret such intervals, and provide a scoring function and calibration test. In considering a scoring function for more than one expectile, it is natural to extend the score for the full expectile function, which leads us to a new expectile-based score for distributional forecasts, similar to the popular continuous ranked probability score.

We use stock index and environmental data to illustrate the new strategic predictors, calibration tests and scoring functions that we propose in the paper. We limit our focus to one step-ahead prediction, but all results can be easily adapted for longer lead times.

References

Berkowitz J (2001). Testing density forecasts, with applications to risk management. Journal of Business and Economic Statistics 19(4): 465-474.

Engle R F and Manganelli S (2004). CAViaR: Conditional autoregressive value at risk by regression quantiles. Journal of Business and Economic Statistics 22(4): 367-381.

Gneiting T F, Balabdaoui A E and A E Raftery (2007). Probabilistic forecasts, calibration and sharpness. Journal of the Royal Statistical Society: Series B (Statistical Methodology) 69(2): 243-268.

Gneiting T and Katzfuss M (2014). Probabilistic forecasting. Annual Review of Statistics and its Application 1(1): 125-151.

Gneiting T and Raftery A E (2007). Strictly proper scoring rules, prediction, and estimation. Journal of the American Statistical Association 102(477): 359-378.

Nolde N and Ziegel J F (2017). Elicitiability and Backtesting: Perspectives for Banking Regulation. Annals of Applied Statistics 11(4): 1833-1874.

OR60 Annual Conference - Keynote Papers and Extended Abstracts
Lancaster University. Lancaster. 11 - 13 September 2018

Exact or Metaheuristic Methods or a Bit of Both – The Rise of Matheuristics

Jonathan M. Thompson

School of Mathematics, Cardiff University
Cardiff CF244AG, Wales, UK
thompsonjm1@cardiff.ac.uk

Abstract

When faced with a combinatorial optimisation problem, one needs to decide whether to employ an exact or a non-exact solution method. The usual rationale concerns the complexity and size of the problem. If it is possible to produce an exact solution in the required amount of time, then in most situations an exact method should be used. Otherwise a (meta) heuristic method is normally employed. However this often means ignoring the exact model completely rather than using this information to inform the metaheuristic approach. Matheuristics covers a broad range of hybrid exact/heuristic approaches but this paper will focus in the main on how exact models can be combined with metaheuristics using examples including scheduling and routing problems.

Keywords: Metaheuristics, linear programming, matheuristics

1. Introduction

Research into the solution of combinatorial optimisation problems has traditionally focused on either exact or approximate solution methods. The decision as to whether an exact or a heuristic approach should be used depends on many factors including the complexity of the problem, the size of the problem, the solution time available etc. It appears sensible to use an exact approach if possible; however there are many potential difficulties. Firstly although one problem of a given size may solve quickly on an exact solver, another instance of the problem of the same size may take considerably longer and it is not always clear for the reason for this difference. Secondly, a slight change in problem definition may have a large impact on the solution time required. Thirdly memory consumption can be high and lead to early termination of the algorithm (Dumitrescu and Stützle, 2003). However in accepting an approximate method, one is accepting a decrease in solution quality that is difficult to gauge precisely. Typically many researchers have a preferred approach and therefore, research into these two avenues has tended to develop in parallel rather than there being any cross-fertilisation of ideas. In recent years, matheuristics have developed which provide a platform for the two approaches to combine leading to better outcomes in many cases. In this extended abstract we will summarise ways in which exact and metaheuristic methods may be combined using examples from literature and from the author's experience. Section 2 will outline different classes of matheuristics before Sections 3 and 4 give specific examples from nurse scheduling and routing.

2. Different Structures of Matheuristics

Matheuristics consist of the hybridisation of exact and heuristic methods. Exact methods that find the optimal solution include branch and bound, dynamic programming and integer programming. Metaheuristic methods that find good but not necessarily optimal solutions include tabu search, genetic algorithms and ant colony optimisation. The ways in which the two methods are usually combined include the use of the mathematical model of the problem at hand to shape features of the algorithm. Puchinger and Raidl (2004) consider two categories of matheuristic:

(a) Collaborative combinations where the algorithms exchange information but are separate. Examples include Clements et al. (1997) who use a heuristic method to generate feasible solutions which form the columns within an exact column generation method. Klau et al (2004) use a memetic algorithm to find a solution to the prize collecting Steiner tree problem and then use an exact branch and cut method to improve it. Feltl and Raidl (2004) find the exact solution to the LP relaxation of the Generalised Assignment Problem which is then used as the basis of the starting population for a Genetic Algorithm.

(b) Integrative combinations where one method is embedded into the other. One example is where large neighbourhoods are evaluated exactly. Dumitrescu and Stützle (2003) point out that it is better to use larger neighbourhoods as these lead to better solutions in a single step than smaller neighbourhoods. However the corresponding disadvantage is that more time is required to search larger neighbourhoods. They consider two approaches. One is to model the searching of a neighbourhood as an optimisation problem and solve it exactly. The second is to fix part of the solution with the remaining part forming an optimisation problem. Burke et al. (2001) present a variable neighbourhood search method for the Asymmetric TSP and search neighbourhoods exactly in order to ensure that promising regions of the solution space are searched thoroughly. Marino et al. (1999) use an exact method as a crossover operator to solve the graph colouring problem. Others use heuristic methods to find high quality bounds within branch and bound approaches.

We now consider two matheuristics in more detail. The first employs a decomposition approach while the second divides the problem into several phases that can be solved exactly. In a meta-heuristic approach to the middle phase, the exact solution to a subsidiary problem is used to guide the search method to the feasible region.

3. Example One - Vehicle Routing with Time Windows

The Vehicle Routing Problem with Time Windows (VRPTW) consists of finding the set of shortest routes that visit a set of customers within a given time period. More precisely the problem consists of a set of customers and their locations, a depot and its location and a set of vehicles. Each vehicle has a given capacity and each customer has a stated demand. The service of each customer must begin within a certain time period.

All routes start and end at the depot. The aim is to construct a set of routes such that each customer is serviced once, all time-window and capacity constraints are satisfied and the overall distance travelled is minimised. Some exact solution methods have been proposed including Column Generation (Desrosiers et al, 1984) and Dynamic Programming (Kolen et al, 1987) but can only be used to solve very small problems. A number of meta-heuristic methods have been employed for this problem including Tabu Search (Solomon, 1987), Simulated Annealing (Chiang and Russell, 1996) and Genetic Algorithms (Thangiah et al., 1991).

One approximate method that has been applied to the Vehicle Routing Problem (without time-windows) is a decomposition approach called the Petal Algorithm. Here, the problem is divided into geographical sectors and each sector is then solved as a single route, typically as a travelling salesman problem. Where the sectors are sufficiently small, exact solutions can be found. These solutions are then glued together to create an overall solution. The algorithm produces good results because the main feature behind the structure of a solution to the VRP is the geographical solution of the customers. A typical solutions contains almost no crossing of routes and the structure is clear. However once time-windows are introduced, then the structure depends less on the geographical positions of the customers and more on their availability times leading to a more random looking structure as some vehicles travel between customers that are not geographically close.

Here we consider a decomposition approach where the problem is divided into sub-problems that are simpler and easier to solve. When the solutions to the sub-problems are combined, a feasible solution to the original problem should be produced. A possible decomposition approach for the VRPTW is to divide the customers into sectors according to their time windows. On this basis, we divide the n customers into m groups of approximately equal size $s \sim n/m$. The s customers with the earliest time windows are then considered and a solution produced for just these customers. The objective is to minimise the distance travelled though the distance to return to the depot is excluded. The solution can be produced exactly if s is sufficiently small. The next s customers are then considered and added to the problem. The problem is then to add these new customers to the existing routes or new routes such that the additional distance travelled is minimised. In these subsequent time periods vehicles are no longer at the depot but are positioned at the starting point of the last customer they visited in the previous time period. The process continues from one time period to the next until all customers have been included. A simple local search can be used to determine whether any simple changes can improve the solutions, in particular where the routes transition between time periods. Clearly this will give a suboptimal solution but experiments showed that it produced stronger results than using a stand-alone Ant Colony Optimisation algorithm using the same amount of computational effort. This is known as a rolling horizon approach and has been applied to several problems including aircraft scheduling (Samà et al, 2013) and operation room scheduling (Addis et al, 2016).

It should be noted that there are similarities between the rolling horizon method described above and the dynamic VRPTW, where some customers only become known once lorries have embarked on their journeys and the problem is to incorporate these new customers into existing routes such that all constraints are still satisfied and the increase in distance travelled is minimised. A static VRPTW can be modelled as a dynamic problem leading to the problem being solved in a series of smaller problems rather than one large problem. Holborn et al. (2012) demonstrated the value of this approach when solving vehicle routing problems with pickups, deliveries and time windows.

Archetti and Grazia Speranza (2014) provide a summary of various matheuristics that have been applied to routing problems. These include various decomposition approaches including Cluster First Route Second methods where customers are clustered using an MILP method and each cluster is the routed using a metaheuristic approach. This method was extended by Federgruen and Zipkin (1984) to solve an Inventory Route problem where they solve an inventory allocation approach exactly, and then construct vehicle routes using a classical heuristic. The objective function to the Inventory Route problem is then modified to account for delayed services, and the two phases are repeated.

4. Example – Nurse Rostering Problem

Variants of the nurse rostering problem have been studied extensively in literature. The majority have employed metaheuristic approaches such as genetic algorithms (Aickelin and Dowsland, 2004), tabu search (Dowsland, 1988) and hyper-heuristics (Burke et al, 2003). The problem considered here involves allocating a set of nurses to a set of shifts such that there are sufficient nurses of the right grade on each shift, each nurse is assigned to the correct number of shifts and schedules are fair. The wards in question were open for 24 hours a day, with each being divided into two days shifts, denoted earlies and lates, and a night shift. Issues of fairness include satisfying nurse requests, ensuring each nurse has sufficient rest between shifts and giving nurses an approximately equal number of night shifts. Solving the problem exactly too difficult with 24 hours of run time still being insufficient to solve some problem instances using modern IP software. Therefore the problem was decomposed into a succession of phases that could be solved exactly.

Firstly a knapsack model was used to ensure there were sufficient nurses to cover all shifts. This was based on the fact that night shifts were longer than day shifts, meaning each nurse *i* worked either d_i day shifts or n_i night shifts . The knapsack model was used to maximise the number of night shifts subject to there being sufficient nurses on day shifts. If there were insufficient nurses, then temporary nurses should be employed. This problem was solved exactly using IP software and took no longer than 2 seconds on a wide variety of problem instances. The second phase then allocates each nurse to either day or nights shifts on particular days of the week. This was solved using a Greedy Randomised Adaptive Search Procedure (GRASP) with the solution from the Knapsack solution being used to guide the construction phase of the GRASP. The construction phase considers each nurse in turn and allocates them to a particular shift pattern, which may consists of either day or night shifts. This allocation is checked with the Knapsack solution to check that it is compatible. If it is, then the construction process continues. If not, then the Knapsack model is reconsidered to see if another solution can be found that is compatible with this allocation. If this is not possible, then the allocation is rejected and a new shift pattern is chosen for this nurse. In this way, feasibility is guaranteed and the construction algorithm can focus on identifying high quality solutions. These constructions are further improved with a simple local search algorithm and the search is diversified by using different feasible results from the knapsack model to produce different allocations of nurses to day and night shifts. The final phase allocates day nurses to either early or late shifts using an exact network model. This ensures that request are satisfied, nurses work pleasing shift patterns and the correct number of nurses with the correct skills are allocated to each shift. Experiments on a number of data sets from real life hospitals showed the advantage of using information from the exact knapsack model to guide the GRASP method, and that dividing the problem into three phases lead to better results than attempting to solve the entire problem at once.

5. Conclusions

For many years, researchers have tended to adopt metaheuristic solution methods for many combinatorial problems due to their complexity and size. However much valuable information can still be gleaned from the underlying model and formulation. Matheuristics are an exciting means for hybridising exact and heuristic solution methods. There are many examples in literature where matheuristics have produced solutions of higher quality than pure metaheuristic methods. There are a number of methods for hybridising exact and heuristic methods and more research is necessary to help practitioners understand how to best implement a matheuristic method, given the range of options available. As computing power continues to grow, the attraction of matheuristics is likely to increase also and may lead to significant improvements in the solution of many problems.

References

Addis B, Carello G, Grosso A and Tànfani E (2016). Operating room scheduling and rescheduling: a rolling horizon approach. Flexible Services and Manufacturing Journal 28, 206-232.

Aickelin U and Dowsland K A (2004). An indirect genetic algorithm for a nurse scheduling problem. Computers & Operations Research 31(5), 761-778.

Archetti C and Grazia Speranza M (2014). A survey on matheuristics for routing problems. EURO Journal of Computational Optimisation 2: 223-246

Burke E K, Cowling P I and Keuthen R (2001). Effective local and guided variable neighbourhood search methods for the asymmetric travelling salesman problem. . In E. Boers et al. (eds), Application of Evolutionary Computing: Evo workshops 2001, LNCS, 2037, 203-212.

Burke E K, Kendall G and Soubeiga E (2003). A tabu-search hyperheuristic for timetabling and rostering. Journal of Heuristics 9, 451-470.

Chiang W and Russell R (1996). Simulated annealing metaheuristics for the vehicle routing problem with time windows. Annals of Operations Research 63: 3-27.

Clements P, Crawford J, Joslin D, Nemhauser G, Puttlitz M and Savelsburgh M (1997). Heuristic Optimisation: A hybrid AI/OR Approach. In Proceedings of the Workshop on Industrial Constraint-Directed Scheduling

Desrosiers J, Sounis F and Desrosiers M (1984) Routing with time windows by column generation. Networks 14(4): 545-565.

Dowsland K A (1998). Nurse scheduling with tabu search and strategic oscillation. European Journal of Operational Research 106, 393-407.

Dumitrescu I and Stützle T (2003). Combinations of local search and exact algorithms. In G R Raidl et al (eds) Applications of Evolutionary Computation, LNCS, 2611, 211-223.

Federgruen A and Zipkin P (1984). A combined vehicle routing and inventory allocation problem. Operations Research 32: 1019-1037

Feltl H and Raidl G R (2004). An improved hybrid genetic algorithm for the generalized assignment problem. In H M Haddadd et al. (eds) Proceedings of the 2003 ACM Symposium on Applied Computing, ACM Press, 990-995.

Holborn P, Lewis R and Thompson J (2012). Combining heuristic and exact methods to solve the vehicle routing problem with pickups, deliveries and time windows. In Evolutionary Computation in Combinatorial Optimisation (Lecture Notes in Computer Science 7245, Springer-Verlag, Berlin, 63-74.

Klau G, Ljubic I, Moser A, Mutzel, P, Neuner P, Pferschy U, Raidl G and Weiskircher R (2004). Combining a memetic algorithm with integer programming to solve the prize-collecting Steiner tree problem. In K. Deb et al. (eds) Genetic and Evolutionary Computation GECCO, LNCS, 3102, 1304-1315.

Kolen A, Rinnooy Kan A and Trienkens H (1987). Vehicle routing problems with time windows. Operations Research 35: 266-273.

Marino A, Prugel-Bennett A and Glass C A (1999). Improving graph colouring with linear programming and genetic algorithms. Proceedings of Eurogen 1999, 113-118.

Puchinger J and Raidl G R (2004). An evolutionary algorithm for column generation in integer programming: an effective approach for 2D bin packing. In X Yao et al (eds) Parallel Problem Solving from Nature PPSN VIII, LNCS, 3242: 642-651.

Samà M, D'Ariano A and Pacciarelli D (2013). Rolling horizon approach for aircraft scheduling in the terminal control area of busy airports. Procedia – Social and Behavioural Sciences 80, 531-552.

Solomon M (1999). Algorithms for the vehicle routing and scheduling problems with time window constraints. Annals of Operational Research 86: 585-607.

Thangiah S R, Nygard P L and Juell P L (1991). Gideon: A genetic algorithm system for vehicle routing with time windows. In Proceedings of the Seventh IEEE Conference on Artificial Intelligence Applications, Miami, Florida, 322-328.

OR60 Annual Conference - Keynote Papers and Extended Abstracts
Lancaster University. Lancaster. 11 - 13 September 2018

A Principled Basis for Decision: Enduring Operational Research (OR) subjects and principles for Quality Assured Analytics

Ian M. Mitchell [a]

[a] Department for Business Energy and Industrial Strategy , London, UK
Ian.Mitchell@beis.gov.uk

Abstract

In a world of change, some things don't. This presentation proposes that efficiency, effectiveness, human interactions and behaviour remain perennial foci for much OR analytics and that quality analytics depend on perennial principles of assurance. Professor Pat Rivett's three subjects for his 1964 BBC documentary about Operational Research: Cutting the Queue, Playing It Through and The Human Factor, are examples of analysis of efficiency, effectiveness and behavioural OR. These remain major subjects for OR although the means to analyse them have evolved. Confidence in analysis is the other perennial aspect. The presentation describes the Department for Business Energy and Industrial Strategy (BEIS) approach to model assurance. Analytical tools for Big Data and Artificial Intelligence provide wider domains for OR. In that wider domain clarity of purpose and assurance of quality will remain essential enablers for a principled basis for decision.

Keywords: Quality Assurance; efficiency; effectiveness; interaction; behaviour; Basis for Decision ; Pat Rivett

1. Introduction

2018 sees OR60, the Diamond Anniversary of the annual OR conferences in Britain. The occasion is an opportunity to celebrate the longevity of the discipline and the success of the pioneers who embedded OR as a basis for decision into British industry, academia and government.

Professor Pat Rivett was President of the OR Society in 1962-63 and took up the first chair of Operational Research at a British University in 1964. He introduced OR, the science of complexity, to Britain in 1964, presenting “Basis For Decision”, three 45-minute episodes on the new television channel BBC2.

Watching these it is striking how many of his comments remain relevant in 2018. Some subjects of OR also remain perennial as does the essential need for confidence in any analysis to adopt it as the basis for decision.

This paper reviews the content of the “Basis For Decision” programmes and describes the Modelling Integrity approach in use by the Department for Business Energy and Industrial Strategy (BEIS). It discusses perennial principles, concluding that OR remains a pioneering socio-technical process.

2. “Basis For Decision” – the Television Programme

2.1 Recovering History

In 2010 the author joined the Operational Research Unit (ORU) at the Department for Business Innovation and Skills. Maurice Kirby’s book prompted conversation with Richard Fitzgerald of the ORU about a British Broadcasting Corporation (BBC) programme on Operational Research. Richard had obtained BBC records of what had been shown at 8pm on Tuesdays starting on 14 July 1964, the Programme As Broadcast (PAsB), and confirmed that the BBC no longer had copies of “Basis For Decision.” The three episodes were “Cutting the Queue”, “Playing it Through” and “The Human Factor”.

On browsing the OR Society Archive listing the author noted that there were two untitled 16mm film cans each containing 45 minutes of footage. Inquiry to the archivist revealed that these were the first two episodes of “Basis For Decision”. The archive had no means to play them so the BBC was offered the films to restore part of their corporate history and to make these available within the OR Society on modern technology.

In 2012 the first two episodes arrived on DVD from the BBC. They were impressive in scope and inspired the hunt for the missing third episode. The International Labour Organisation (ILO) held copies until the 1980s; “Basis for Decision” had achieved international impact and it was this impact that led to the recovery of the third episode. In 2013 we learned that Australia’s Sydney and Monash Universities both had copies of “The Human Factor.” Monash University was willing to send their copy and so in April 2016 the third episode at last became available in a viewable format in the UK.

The first, second and part of the third programmes comprise case studies. These show a spectrum of approaches in use on a range of operations. The last part of the third programme is a debate about the uses and limitations of OR. The emphasis on the social aspects of OR increases through the series. The importance of value judgements becomes clear in the third part as do the challenges from political pressures.

The next sub-sections describe each of the episodes.

2.2 Episode 1 Cutting the Queue

The first programme opens showing men in suits around a table, some smoking, talking about a pressing problem. Their problem is the number of ships idling in Swansea Bay waiting to unload for the Steel Company of Wales. From the studio Rivett introduces the significance of the problem; the idle ships drive the cost of living: Transport costs add to the cost of steel, so this affects the wider UK economy because so many things contain steel or are made on machines themselves made from steel. He mentions the origin of OR as the scientific study of military operations, which spread to scientific observation and analysis to solve complex management problems, and that “Complexity is my business.”

I.M. Mitchell

He suggests that complexity is only skin deep, and that there are eight types of structure, a skeleton of cause and effect. The Port Talbot problem can be addressed by using the queuing structure of In Service Out. Figure 1 is the example figure for the template for this abstract; it is an example of a perennial OR subject and approach, describing the same structure that Rivett used to model the problem in Port Talbot in 1964.

Figure 1 Input/output diagram

Figure 2 shows the diagram from “Basis For Decision”, representing the unloading of a ship.

Figure 2 In Service Output diagram from “Basis For Decision”

Rivett explains that OR uses an In Service Out structure to model the ore unloading problem. The same structure can be applied to similar situations such as busses, trains, post offices and factories; anywhere where losing time is the key factor. Port Talbot is a tidal port, so the size of ships interacts with tides, and the layout of the port to determine the amount of iron ore delivered. There are two interested organisations, a shipping company and steel company. Reduction of time in queues is the problem for both. The cost of the waiting ships has a trade off with the additional costs from unloading the ships faster so finding a fair share of the costs is essential to the success of both companies.

Rivett then introduces John Murdoch from the Cranfield College of Aeronautics who demonstrates a mechanical model of a factory store, with figurines moving from a queue to a counter then away. Matching average rates directly between input and output results in lengthening queues of the figurines at the factory store. A graph of the optimal solution for

I.M. Mitchell

the system design (Figure 3) shows that a compromise between acceptable queue length “not too long” and acceptable idle time “not too high” for the storeman has to be found.

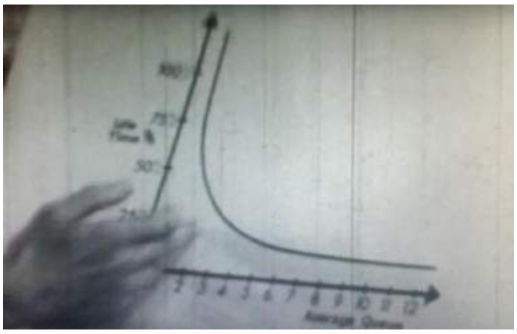


Figure 3 Diagram of average queue length (x) to average idle storeman time (y)

In passing Rivett mentions the use of statistics or the mathematics of probability and that all post offices are going over to a system with each desk offering all services in order to reduce queues. Ray Jackson describes his model of a Doctor’s Surgery to compare waiting times for patients and doctors resulting from an appointment-based system using 4-minute, 5 minute and 6-minute intervals compared with the former first come first served approach, that could lead to waiting times of over 90 minutes. A doctor concludes that he would never return to this previous system. That such appointment systems are now near universal is a reminder of how an innovation becomes the new normal.

Observing that “Whenever we have a problem in OR we build a model of it” Rivett demonstrates the use of a model of a system representing the essential features of the problem, demonstrating a physical analogue model (Figure 4) to optimise the location of a factory acting as a distribution centre for the UK.

Figure 4 A Physical Model for Factory Location

A metal ring is positioned by the combined pulls of weights representing the number of lorryloads going to each of several destinations, 10 to Newcastle, 40 to Manchester, 20 to Bristol and 40 to London. Where the ring ends up shows the most efficient location for the distribution centre; Coventry would be the cheapest location. This physical linear program can show how much additional force representing additional travel costs is needed to move the ring to somewhere more attractive, such as Stratford on Avon.

Mathematical models enable experiments to be done quickly at low cost. Reconciliation of conflicting aims through analysis can reduce conflict between different people with different objectives. David Owen and Roger Grison of the British Iron and Steel Research Association (BISRA) discuss Port Talbot, with footage from Port Talbot and Newport showing how ships can be unloaded using different types of crane. Port costs and shipping costs offset one another and graphs of the time to unload show a shared optimal point is a compromise at 50% idle time for the berths.

A Mr Thomas describes a model used for experiments, of the whole port of Port Talbot, considering spring and neep tides and how these affect delay times for ships then their unloading at berths. Historic records showed the patterns of how ships arrived providing input data. The Swansea Bay model required one-hour computer time to produce six months of schedule giving shipping delays. Experiments included effects of dredging on arrival rates to assess whether the consequent increase in despatch bonuses offset costs of dredging. Mr Cartwright of management, in suit and hard hat on the dock side, endorses the analysis: “OR has greatly assisted in reducing the area of uncertainty surrounding many of these decisions.”

Rivett describes stock control as a problem with levels of stock being 30% higher than necessary, recommending use of the mathematics of variability.

Ian Davidson interviews National Coal Board Chairman Lord Robens of Woldingham who identifies OR as one of the “greatest tools of management” saving hundreds of thousands of pounds per year with application to any complex managerial problem, including stocks. From his place on the National Economic Development Council (NEDC) Lord Robens’s view is that if every firm in this country used OR the result would be a 4% increase in Gross National Product (GNP) “quite easily”. Rivett concludes the episode defining the purpose of OR to: “provide the people controlling these operations with a firmer “Basis For Decision”.”

2.3 Episode 2 Playing It Through

Rivett uses The Great Train Robbery to illustrate the technique of Critical Path Scheduling (CPS) or Network Planning as shown in Figure 5.

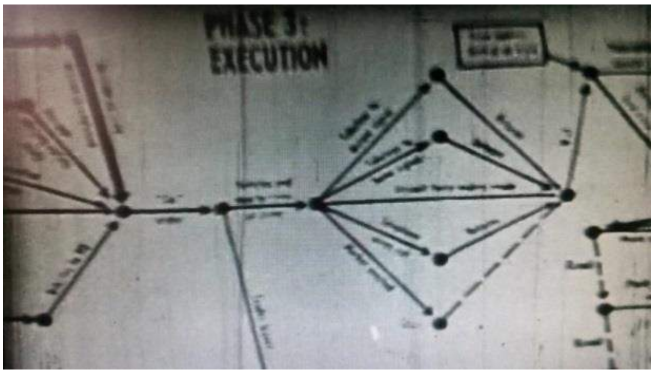


Figure 5 A Network Model for Armed Robbery

He moves to construction of a factory, then preparation of a Sunday Lunch, translating critical path network to time chart before introducing the use of CPS for construction of buildings, then merchant and passenger ships.

A Mr Morrison talks over ship blueprints and charts of "ladders" grouping task lists. Computer based scheduling enables confidence in offering delivery of a ship for a firm contract date. These task diagrams assisted in communication between firms. Lithgow's work on the Oriana installed an extra crane and made use of pre-cut sections to speed up the build. Networks were also used for Kincaid's work on engine building for the Oriana. This illustrated that by reviewing the situation as it developed, Network Planning suggested ways to change the operation in order to preserve the schedule for the benefit of both organisations.

Rivett summarises shipbuilding as a large scale but pretty simple scheduling structure. He contrasts this with the complex food production where variability in the rates of flow causes problems beyond the reach of calculation. This introduces the need for experiments. He uses the example of a pedestrian trying to cross Picadilly Circus to introduce experiments with physical or mathematical models; simulations that can develop understanding of the likely range of outcomes to inform planning and implementation.

Dr KD Tocher introduces the United Steel sequence of processes as a complete "system" of steel production with the Rolling Mill as the limiting step. He explains the challenges arising from the variable nature of different processes and the need for decisions hours in advance of implementation. Experiments in simulation explored how to avoid "a vicious spiral" of delay that could lead to disastrous loss of production. Using a computer simulation needed a whole new language to build a computer model of the operations. To see whether the computer model was valid real managers provided a realistic plan to test it. The simulations are also shown as a way to exercise new managers testing their decisions.

The last section describes wargaming at West Byfleet in the Army Operational Research Establishment (AORE). Rivett introduces Ronald Shephard, later Professor Ronnie Shephard, the founder of the International Symposia for Military Operational Research (ISMOR) as shown in Figure 6.

I.M. Mitchell

Figure 6 Ronald Shephard beside the West Byfleet Wargame Control map

This section illustrates the pioneering use of games for analysis for situations, driven by value judgements from humans, for example trading ground for casualties or time. Information available is limited to the players based on what their forces are modelled as being able to see. The outcome of many actions in battle are subject to the laws of chance. The computer here provides a random number generator to resolve chance interactions tabulated in a classified book of rules. Shephard notes that "The validity of the book of rules is of the utmost importance" depending on field trials, historical analysis and military judgement.

Rivett summarises that problems of complexity may be beyond calculation but experimental simulations can provide guidelines for decision.

2.4 Episode 3 The Human Factor

Episode 3 opens with Ian Davidson interviewing Lord Robens on Pit Closures and OR's contribution to manpower problems. Rivett introduces the idea of 4 balls to juggle with in any management problem: illustrating these as 4Ms - Machines –Materials- Money and Men.

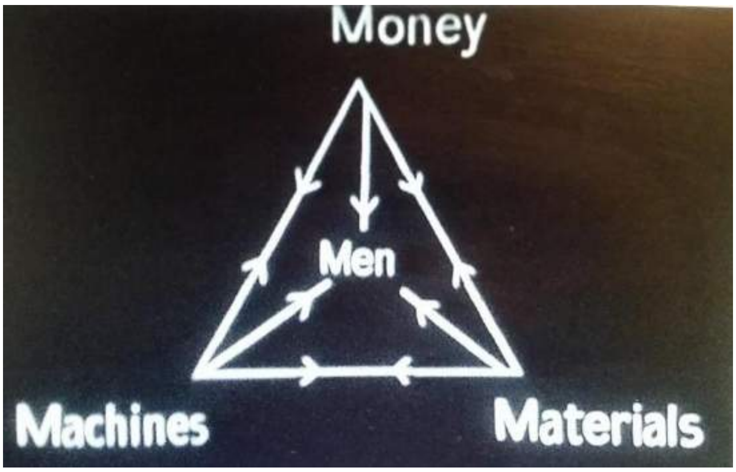


Figure 7 The 4Ms

He notes that the first two episodes considered machines, materials and money as the dominant factors. The fourth M stands for men, by which Rivett means everyone on the payroll; those running the organisation and working in it. Where the human factor is the dominating factor in a problem OR must tackle it directly so there are overlaps with social science. He emphasises OR as a "practical action science" solving problems rather than building an academic understanding of human nature.

Dr BT Houlden Director of OR at the National Coal Board describes a study of 60 factors covering thousands of miners and ex-miners of which distance to work from home and length of service were the most important factors to estimate numbers of miners willing to move between pits. This informed recruiting strategy but not the wider social impact. Lord Robens explains the operational impact of these individual decisions on the pit closures under consideration. Rivett compares the benefit of using purely external factors to inform mine planning moving to Professor RW Revans of Manchester's study of internal factors on the morale on a hospital's effectiveness and reduce waiting lists.

Figure 8 Observer and observed in Hospital

The study used Activity Sampling for the first time in this country. Observations were made at random intervals selected with a random number chart to achieve a random sample. The metrics were student nursing wastage rates and duration of patient stays by type of illness. The latter varied between 7.5 to 12.1 days in a direct correlation to the wastage rate. The next stage looked for a factor driving these.

The programme shows female observers talking to Ward Sisters and student nurses in interviews. These informed the design of structured questionnaires with agree or disagree scales to measure attitudes to superiors and subordinates. A study of 15 Hospitals found strong correlation between the quality of Sisters' attitudes (morale) and the metrics of effectiveness (patient stays) and efficiency (nurse wastage). In a low morale hospital communication suffers as orders go down the hierarchy but little valuable information returns so the organisations cannot adapt to change because they lack feedback and become extremely anxious places to work.

Rivett suggests future uses of OR may move to marketing, the effect on sales of advertising and discounts. This is in keeping with OR's introduction of measurement to areas not previously measured. He suggests that Government is the biggest potential area for OR and introduces three speakers on local government; central government; and governments of undeveloped countries.

Ray Ward describes Local Government OR on efficiency of routine and clerical operations such as stock control and school transport looking 10 years ahead planning by estimating children populations and the extent of built up areas to generate these.

Stafford Beer notes "a strange void when it comes to government", a lack of scientific appraisals for decisions on policy making in civil government. He emphasises that OR is at its most worthwhile addressing large and complex systems too involved to fit into the head. Despite accepting OR models of flows of traffic past a roundabout, he notes a steadfast refusal to discuss the national transport problem as a whole, "Government needs to make models of most aspects of the economy"; OR is not just any type of "coherent thinking".

Beer mentions his work for Puerto Rico and notes use of OR for US Defence administration, city planning, reorganisation of the Patent Office, Post Office operations and public health. In the United States funding drives studies at the national level. In France economic recovery was supported by the Commissariat aux Plans. This has urban and regional development plans for Rennes, Nantes, Grenoble, La Rochelle and Paris. These cover hospitals, schools, transportation, fuel, power and agriculture. They informed the Common Market negotiations. Russia may be slow with OR within the individual firm but is using models from cybernetics to study the redesign the structure of the state.

Sir Charles F Goodeve notes governments deciding to act without precise knowledge are like "Driving in a Fog". His experience in India, found prime objectives were focused on national needs and so were more obvious than in the UK where questions turn on the division of affluence. He describes steel forecasts where studies found that the new factor of positive feedback that would come from initial demand for steel would lead to a landslide in demand so raising the forecast need for steel by a factor of 10 to 1 million ingot tons per year.

Rivett introduces a debate on the limits to OR. Figure 9 shows Ian Davidson in the centre chairing, from left, Denzil Freeth MP, Parliamentary Secretary for Science, and John Barber, Financial Director of the Ford Motor Company leading on the limits to OR, with Professor GA Barnard, the President of the OR Society and Stafford Beer to argue against these.

I.M. Mitchell

Figure 9 Debating the limitations of OR

Freeth proposes three limitations to the applicability of OR; unknowns, value judgements and unique decision points.

Where there are many unknowns or a key factor must be unknown, the government has to take a hunch decision. For example, for the British and French Governments to decide whether to invest in Concorde, a key factor was whether it would pay which depended on competition, yet there was no way of knowing if Russia would enter this field. Barnard replies that certain factors may be unknowable or unmeasurable but creating a model makes explicit the relationships involved, exposing many otherwise hidden factors. Statistics, defined as decision making in the face of uncertainty, can also be helpful.

Freeth argues that decisions that are value judgements are not amenable to being put into a computer. The selection of route of the M4 motorway over the Berkshire Downs turned on amenity and aesthetic judgements to put against economic effects. Beer replies that economic benefit and disturbance to amenity can be measured; 90% of what people regard as important can be measured although people do not believe that this can be done. Freeth emphasises the aesthetics; in theory one can produce a situation on paper to upset the least number of people for the best effect for each route. In life it is often the people who are rarely interested who are the ones who need to be satisfied.

Freeth's second point was that decisions in government or industry are often unique and taken with great speed, such as whether to go to war or not, for the firm to react to a takeover bid, and for the individual on whether to marry or not. Barnard replies that Dynamic Programming in OR considers how to trade delay in making a decision versus the quality of decision made. During the Suez Crisis oil companies used computer models of movements to reroute tankers rapidly moving oil via alternative routes.

Barber notes that the most common complaint is the lack of data. Companies that naturally produce data like chemical and oil also have high capital expenditure for expensive assets. These produce data by their nature so are where OR has had its greatest effect; people heavy organisations are more difficult to help because the facts are not there, lending support for the establishment of a British Business School. Barnard notes that Management tend to run businesses to make products but businesses should generate information about their products to improve these.

Rivett sums up OR as an applied science that deals with two forms of problem: intellectual ones about models and practical ones about working with people in their organisations. Measuring things that have not been measured before is where OR can make the greatest contribution; this is its pioneering aspect. Rivett concludes with the hope that:

"as our research continues that our help to management will expand not only at the executive level but at the policy level so that Managers at all levels have a firmer Basis for Decisions on industry and government."

The final frame is shown in Figure 10 "Basis for Decision" overlaid on the Houses of Parliament.

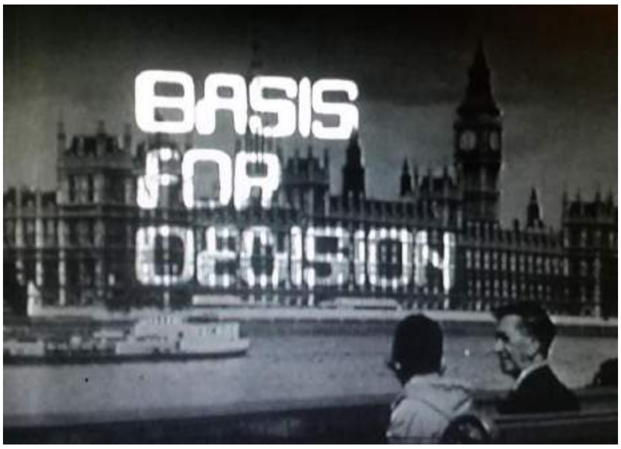


Figure 10 Closing image from "The Human Factor"

3. Quality Assurance

There are many forms of model. Rivett demonstrated a variety of physical analogues in "Basis For Decision" as well as mathematical models. The forms of model change with time, location, culture and technology; there were no spreadsheets in 1964 although some of the models described could have been implemented in these. Assurance of the reliability of any model output used to inform a decision is critical to its value. Kirby cites Ronald Shephard on the question of confidence:

"A decision reached by OR methods is not necessarily different from the decision that would be reached by other methods; a right decision is a right decision, irrespective of whether it is obtained by sticking a pin into a list of all the alternatives, or by a piece of OR Work. The basic difference is the degree of confidence that can be placed in the correctness of the result."

Quality Assurance of models is a perennial enabler for OR. Basis for Decision mentions the importance of validity in each episode, particularly for the simulations used to play through experiments. As George Box remarked: "Remember that all models are wrong; the practical question is how wrong they have to be to not be useful."

At OR40, also held at Lancaster in 1998, Jim Shalliker presented "Spreadsheet Models: the Good, the Bad and the Ugly." Some terminology is different from 2018 but the fundamentals are similar especially regarding documentation, structure and clarity, verification and validation. The current BEIS resources for assurance of models addresses the requirements of the Aqua Book, the Treasury's guidance on producing quality analysis for government. The majority of models in BEIS are implemented in Excel spreadsheets so the system addresses these directly.

The BEIS Model Report template includes Scope, Specification, Model Map, and Technical Guidance. As a handover document from one analyst to a successor preserves corporate memory. The Quality Assurance (QA) Log available from Gov.uk provides a way to assess the quality of the model using a Multi Criteria Analysis:

Documentation – Formally identify and agree on model purpose and deliverables to enable a robust and coherent design, and enhance maintainability/adaptability/ease of handover to new analysts.

Structure and Clarity – Designing models as modules with the input, processing, output structure shown in Figure 1 and discussed at length in "Cutting the Queue".

Verification – Auditing of formulas and code - are the sums being done right?

Validation - Are those the right sums to be doing for the question at hand?

Data and Assumptions – Where did these come from? Who from? How certain are they? How credible are they? As Shephard observed of the wargame's book of rules the data and assumptions are of the "utmost importance." Freeth's remark about potential Russian competition to Concorde is striking in that the Tupolev-144 did exist yet was to have no impact on Concorde's viability. In contrast the American approach to air transport is not mentioned.

4. Perennials of OR

The operational problems of efficiency, effectiveness and human behaviour, especially individual preferences, occur in many OR studies. They overlap with descriptive, predictive and prescriptive forms of analytics. The use of models is universal and perennial.

Paul Harrison's diagram at Figure 11 illustrates the process of OR translating a real problem from its initial context into a mathematical representation of the problem, a model.

I.M. Mitchell

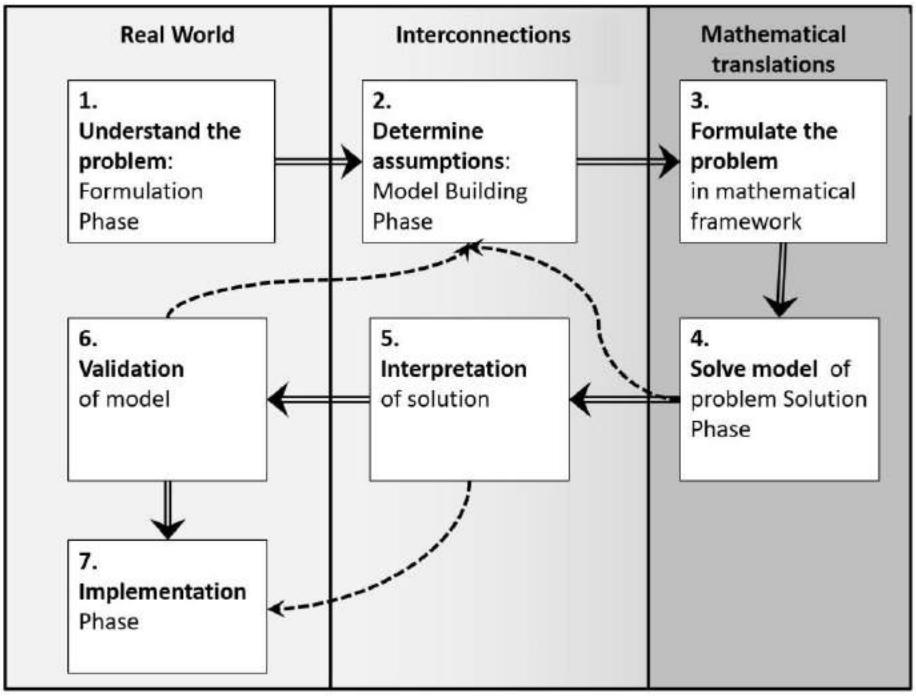


Figure 11 Harrison's diagram of OR Methodology: a perennial process

The mathematical representation is solved and then translated back into the original context where action is implemented. During solution and validation dotted lines show possible feedback loops to model building. There is also the destructive temptation to jump from interpreting the solution to implementation, avoiding validation.

The division between the real world and the world of mathematical models brings out OR as a support to operations in the real world; the model is not an end in itself. The decisions supported depend on the decision maker's concerns, some of which are perennial.

Efficiency- changing the ratio of outputs to inputs; These take many forms, often financial but also in terms of time or the Loss Exchange Ratio of battle models. "Basis For Decision"'s models identified overall benefits in the operation of docks, post offices and surgeries through cutting the queue. The OR Society IMPACT magazine of Spring 2018 has two cover articles on cutting the queues, one at Dover the other on decontamination.

Effectiveness – exploring the achievement of an organisation's aims, models act as experiments, some in virtual environments with human decision making. Tactics, techniques, procedures and technologies can be tested by playing it through. The June 2018 PHALANX magazine cover story is a Cyber Wargaming workshop and a feature on strategic wargaming.

The Human Factor and its connections to the measurements of harder items are perennials with much pioneering work on Problem Structuring Methods as summarised by Rosen head. Freeth's point about value judgements on route selection is particularly interesting. The 2017 consultation by Highways England on the route of the A303 past Stonehenge used a multi criteria structure and explored trade-offs through ranges of weights applied to Freeth's criteria of economic benefit, aesthetic and amenity.

Rivett's 4Ms; Materials, Machines, Money and Men, reflect the culture of 1964 but the generic areas remain of interest. Translating the 4th M, Men, to People would require Ps for the others; perhaps Physics, Power, Pounds and People. This scope seems compatible with Raworth's calls in "Doughnut Economics" for as broad an approach to models to that suggested by Stafford Beer's comments.

OR is a socio-technical process. The debate on the utility of OR shows the vital importance of the sponsor of OR. The decision maker is a gate keeper, deciding whether OR is used or not. Stirrup observed lost opportunities in the lack of traction within the new Royal Air Force in 1918 experienced by Lord Tiverton for his studies on strategic bombing, a prototype of OR 20 years ahead of its time. As the resident OR adviser to a directorate of equipment capability in the Ministry of Defence the author observed his client move from a view of OR as a "Necessary Evil" to survive scrutiny to a useful basis for decision on balance of investment.

Rivett's closing remark is a worthy perennial aspiration for OR: "So that managers at all levels have a firmer basis for decisions on industry and government."

Acknowledgements

Richard Fitzgerald's interest and tenacity were vital to the recovery of the three episodes of "Basis For Decision" complemented by the generous support of the OR Society, the OR Society Archive at the Modern Records Centre, University of Warwick, the library of Monash University, the Kaleidoscope Group, Lancaster University and the BBC.

References

Box GEP and Draper NR (1987) Empirical Model-Building and Response surfaces John Wiley and Sons New York p 74 cited by https://en.wikiquote.org/wiki/George_E._P._Box#Empirical_Model-Building_and_Response_Surfaces_(1987)

British Broadcasting Corporation (1964) "Basis For Decision" - The use of Operational Research

GOV.UK (2018) BEIS Modelling Quality Assurance tools and guidance Tools and guidance to help people with Quality Assurance of analytical models. https://www.gov.uk/government/collections/quality-assurance-tools-and-guidance-in-decc

Harrison (1983). Operational Research. Kluwer Academic Publishers: London. Page 37

Her Majesty's Treasury (2015) The Aqua Book Guidance on producing quality analysis for government

Highways England (2017) A303 Stonehenge Amesbury to Berwick Down Technical Appraisal Report Volume 1 HE551506 page 84 5.2.40

Kirby, M W (2003) Operational Research in war and peace The British Experience from the 1930s to 1970 Imperial College Press: London and the Operational Research Society: Birmingham

Military Operations Research Society (2018) PHALANX The Magazine of National Security Analysis Volume 51 No 2

OR Society (2018) IMPACT Spring 2018

Raworth K J (2017) Doughnut Economics Seven ways to think like a 21st-Century economist Random House London

Rosenhead J (1989) Rational Analysis for a Problematic World - Problem Structuring Methods for Complexity, Uncertainty and Conflict

Shalliker J (1998) Spreadsheet Models: The Good. The Bad and the Ugly. OR40 Lancaster handout

Shephard R W(1961) "Some aspects of Measurement in Operational Research. Modern records Centre University of Warwick. Cited by Kirby at page 25)

Stirrup J, (2005) "Inside and Outside the Box: Individual Creativity versus Military Orthodoxy" Journal of the Royal United Services Institute for defence studies Volume 150 No 6 page 46

OR60 Annual Conference - Keynote Papers and Extended Abstracts
Lancaster University. Lancaster. 11 - 13 September 2018

Writing about Simulation

Christine S.M. Currie

University of Southampton, UK
christine.currie@soton.ac.uk

Abstract

The focus of the article is on the second half of a simulation project, after the model has been built. While less glamorous than model-building, the analysis of results and the descriptions of what have been done are still vital for a successful simulation project. We concentrate on three main themes: reporting simulation models; accounting for output variability and decision-making via simulation. Discrete Event Simulation, Agent Based Modelling and System Dynamics are all discussed and so we hope that the recommendations will be paradigm-agnostic.

Keywords: Simulation, Output Analysis, Decision Analysis

1. Introduction

We concentrate here on how simulation models and their results can best be communicated to ensure that they have maximum benefit. Whilst perhaps less fun than actually building the model, this part of a simulation project is where the real value can be created and, as such, is vital to a successful modelling exercise. The discussion is intended to be paradigm-agnostic in the sense that the majority of it will hold for all three of discrete event simulation (DES), agent based modelling (ABM) and system dynamics (SD). The split between these three disciplines is shrinking with the major international simulation conferences such as Winter Simulation Conference and the UK Simulation Workshop welcoming research from all three areas, as well as the Journal of Simulation. As a result, it is important to develop methodologies that work well across all three methodologies.

Typically, the first stage of a simulation project involves scoping and building the model. This is not something that is covered here, although it could be argued that many successful simulation projects have most impact in these early stages, during the facilitated workshops and conceptual model building, and it is certainly true that a great deal of communication about the simulation model takes place in this phase of most projects. In a practical (rather than academic) setting, this may be where many of the important stakeholders gain the most insight into how a simulation model works and how best to interpret its results. Robinson (2008) gives an overview of conceptual modelling for simulation and how this can be useful in a complex simulation project. While his focus has been on DES models, the general principles hold true for ABM and SD.

Clear descriptions of simulation models allow for a better understanding of the compromises made during the modelling process and consequently a more intelligent interpretation of the results. Of equal importance, providing all of the necessary details enables a model to be reused and adapted for subsequent modelling projects and for results to be checked by a third party, where necessary. Model reuse saves time and money in the commercial world and is a valuable piece of knowledge management, but it is also important in academia to introduce transparency into the research process and enable other researchers to reproduce the results. Open science is defined by the FOSTER project as "… the practice of science in such a way that others can collaborate and contribute, where research data, lab notes and other research processes are freely available, under terms that enable reuse, redistribution and reproduction of the research and its underlying data and methods." (www.fosteropenscience.eu). It has become a hot topic in recent years and ensuring the reproducibility of academic simulation studies is crucial for encouraging high quality research in the area. Checklists are a good method for ensuring that all of the necessary details have been included in a model description and the STRESS (Strengthening the REporting of Simulation Studies) guidelines have recently been introduced to do this for DES, ABM and SD (Monks et al. 2018). We go over the basics of STRESS in the next section.

Accurate reporting of the simulation model and of the simulation results ensures that the reader, who may also be the decision-maker, can appreciate where approximations were made in the modelling and how these might have affected the final recommendations. It helps with building trust between the modeller and the model-user.

Honesty is important in all reporting, but particularly when discussing the inherent variability in results from a simulation study. We discuss methods for quantifying the true uncertainty in a set of simulation results taking into account the uncertainty coming from the simulation process and the uncertainty coming from parameter estimation. Given the nature of the paper, we do not go into details of any particular methods but instead discuss relevant research in the area.

Insightful results help with providing good support for decision-makers as, after all, this is generally the main purpose of a simulation modelling study. Making simulation results insightful requires a thorough understanding of the problem but can also involve a set of well-chosen simulation optimisation routines. While in the past these have been accompanied by a long list of rather stringent assumptions that must be met, more recent general purpose simulation optimisation routines are designed to work on a wide range of problems and we introduce these briefly.

In the remainder of the paper, we progress through the reporting of a simulation model in Section 2, where we discuss the STRESS guidelines for reporting on simulation models, and then go on to discuss how to analyse model output and use it for informing decision processes in Section 3. In Section 4 we discuss how simulation results should be used to make decisions, including an overview of simulation optimisation. Finally we conclude in Section 5 with a suave and sophisticated acronym to describe the key features of an effective analysis of a simulation project.

2. Reporting Simulation Models

Accurately reporting what a simulation model does and how it works is important for fully understanding and trusting its results, as well as allowing it to be reused or reproduced. Reproducibility has become a buzzword in science in recent years and this is also being seen in Operational Research (OR) and Management Science (MS). From an academic perspective, applied research that cannot be reproduced is not particularly useful. Similarly, a practitioner needs to provide a thorough description of a simulation project as a form of knowledge management.

The STRESS checklists have been designed to ensure that all of the important aspects of a model have been included in either the academic paper or technical report (Monks et al. 2018). They come in three flavours – DES, SD and ABM – with each of them based on six principles for good reporting:

1. State the purpose of the study and the model's intended use
2. Provide enough detail to reproduce the results of the base run of the model and any simulation experiments conducted as part of the study.
3. Ensure that descriptions of the model are software and hardware independent.
4. Include data for verification and parameter values. Where proprietary or ethical issues prevent the inclusion of data, 'hypothetical' test data should be included for verification purposes.
5. Document all software and where necessary hardware-specific implementation.
6. Provide additional visualisation of model logic or algorithms using a recognised diagramming approach.

The checklists can be downloaded from the publishers' website using the doi given in the reference list and we hope will become ubiquitous when describing applications of simulation in academic publications. The main reason for developing them was a frustration at the paucity of detail in some applied modelling papers, as discussed in several recent surveys within the OR/MS literature (Boylan et al., 2015; Dalle, 2012; Grimm et al., 2006; Kendall et al., 2016; Kurkowski et al., 2005; Rahmandad & Sterman, 2012). Nonetheless, we can see that there is also a need within a practical setting for a well-established template or checklist for ensuring that adequate information is stored about important simulation models.

3. Accounting for Output Variability

The key point in this section, which applies to DES, ABM and SD, is to be honest about the variability in the output of simulation models. This includes the natural variability of the output of a stochastic simulation model, such as DES or ABM output, but also the *input uncertainty*, where this is defined as our uncertainty about what the model parameters should be. For example, when building a model of a factory assembly line, there may be data available on the breakdown rates for each machine and this can be used either to fit a parametric statistical model or as a direct input to the simulation model (so called EDF input). In both cases, the fact that we have finite data with which to work will mean that the distributions used in the model are not a perfect representation of the real system. This makes our inputs uncertain and this uncertainty should carry through to the model outputs. Quantifying the effect of input uncertainty on model outputs is not straightforward, particularly for stochastic simulation models.

Being honest about uncertainty in this context means reporting confidence intervals on results, rather than just mean values and ensuring that sufficient runs of the simulation model have been made to justify any recommendations. Russell Cheng will often quote an early tutorial of Averill Law at the Winter Simulation Conference in which he advocated running the (discrete-event) simulation model not once, but three times. Since then, the textbooks have become a little more demanding on the number of runs but in essence, the argument that no decision can be made on just one run still holds. For those looking for a basic introduction to output analysis, the simulation textbooks by Law (2007) and by Robinson (2004) provide some excellent "how-to" descriptions on setting the length of the warm up period, determining the appropriate number of replications and for optimising the run time of simulation models using variance reduction techniques such as common random numbers. Getting the output analysis right will mean that the results being reported are an accurate reflection of the simulation model. There is no point spending a significant amount of time building a superb model of a process if the analysis of that model misrepresents the outputs.

Input uncertainty requires a little more thought because with anything other than the very simplest models, it is impossible to ascertain analytically how varying one or more input parameters will affect the output. Song et al. (2014) provide an excellent overview of input uncertainty for stochastic simulation models and give details of a practical method for diagnosing its presence. The algorithm they propose, a variation of which has been employed within the Simio package, can be used to determine the importance of input uncertainty to the overall variability in an output. Reporting this helps both to provide a full picture of the uncertainty in the mean output and to identify where more data are needed to improve the reliability of the simulation model results.

4. Using Simulation to Make Decisions

The established simulation analysis methods work well for solving relatively clear-cut optimisation problems and the recent book by Fu (2015) is a good guide to the range of methods available.

Optimisation via simulation has traditionally focused on optimizing a single output measure and Nelson and Hong (2009) provide a useful classification for simulation optimization problems, dividing them into three main groups:

1. **Ranking and selection problems** in which there are a small number of solutions; sufficiently small that it is possible to test all of them in the simulation model and select

the solution with the best result, e.g. deciding between different set-ups for a factory or hospital to maximise throughput.

2. **Continuous optimisation problems** for which the vector of decision variables, x, is continuous, e.g., choosing the optimal price to charge for airline tickets in a network.
3. **Discrete optimisation problems** where the vector of decision variables, x, is discrete and integer ordered, e.g. choosing the optimal number of call centre staff on duty to minimise costs subject to a constraint on response times.

The methods used to solve the three different classes of optimisation problem vary significantly but all take into account the stochastic nature of the output.

We focus here on ranking and selection methods as these tend to be more common when making decisions using simulation. Ranking and selection methods aim to choose one of several systems, based on a single system output. Branke et al. (2007) give a good overview of the popular methods and compare their performance. As the review shows, methods are split between those that optimise the allocation of computational budget to different options to ensure the most statistically significant result can be obtained (Chick and Inoue, 2001) (the OCBA approach) and methods that allocate sufficient replications to different options to ensure that the optimal solution can be found with a relatively high probability (Kim and Nelson, 2006) (the indifference zone approach). This work is important and deserves to be better known in practice as it can significantly improve the quality of the recommendations coming out of simulation projects.

There is a substantial body of work on the psychology of decision-making suggesting that emotion can play an important part (e.g. see Bechara and Damasio , 2005) in how the decision is made. Simulation, when presented well, has the potential for shaping emotions and providing a more holistic picture of the uncertainty associated with each option. Difficult decisions involve a messy trade-off between a number of factors, and the visual tools that simulation offers could help significantly with multi-criteria problems. While some work has been done in identifying Pareto sets of solutions (Lee et al. 2010) or using chance constraints to account for secondary performance indicators (e.g. see Hong et al. 2015), the results can be hard to interpret. As Belton and Stewart (2002) discuss, a simulation model can allow decision makers to take account of multiple performance measures intuitively. Such methods, and more formal ideas from MCDA implemented with simulation, have the potential for high impact in policy decisions and should form the basis of future work.

5. Conclusions

All good papers come up with an acronym and I am undecided as to whether the choice of word says more about the author or the subject matter. Here we suggest CHAI (as in the lightly spiced tea) to denote the important aspects of talking about a simulation project: Clear, Honest, Accurate, Insightful. While not as imaginative as MR POTATOHEAD (Parker et al. 2008) or as apposite as the STRESS test (Monks et al.

2018), it hopefully brings some sophistication to simulation reporting, through better descriptions of modelling projects and the uncertainties inherent in the results; a good understanding of the statistics through accurate output analysis; and more insights coming from an appropriate application of simulation optimisation techniques.

References

Bechara A and Damasio A R (2005). Emotion, decision making and the orbitofrontal cortex. Cerebral Cortex 10, 295-307.

Belton V and Stewart T (2002). Multiple Criteria Decision Analysis: an Integrated Approach. Kluwer Academic Publishers.

Boylan, J E, Goodwin, P, Mohammadipour, M, & Syntetos, A A (2015). Reproducibility in forecasting research. International Journal of Forecasting, 31(1): 79-90.

Branke, J, Chick S E, and Schmidt C. (2007). Selecting a selection procedure. Management Science, 53:1916–1932.

Chick, S E, and Inoue K (2001). New Two-Stage and Sequential Procedures for Selecting the Best Simulated System. Operations Research 49:732–743.

Dalle, O (2012). On reproducibility and traceability of simulations. In Proceedings of the 2012 Winter Simulation Conference, Berlin, Germany, 2763-2774.

Fu, M C (2015). Handbook of Simulation Optimization. Springer-Verlag, New York.

Grimm, V, Berger, U, Bastiansen, F, Eliassen, S, Ginot, V, Giske, J, De Angelis, D L (2006). A standard protocol for describing individual-based and agent-based models. Ecological Modelling, 198(1–2): 115-126.

Hong, L J, Luo J, and Nelson B L (2015). Chance constrained selection of the best. INFORMS Journal on Computing 27:317–334.

Kendall, G, Bai, R, Błazewicz, J, De Causmaecker, P, Gendreau, M, John, R, Yee, A (2016). Good Laboratory Practice for optimization research. Journal of the Operational Research Society, 67(4): 676-689.

Kim, S-H, and Nelson B L (2006). Selecting the best system. In Elsevier Handbooks in Operations Research and Management Science: Simulation, edited by S. G. Henderson and B. Nelson. Waltham, MA: Elsevier.

Kurkowski, S, Camp, T, & Colagrosso, M (2005). MANET simulation studies: the incredibles. SIGMOBILE Mob. Comput. Commun. Rev., 9(4): 50-61.

Law, A M (2007). Simulation Modeling and Analysis (Fourth Edition). McGraw-Hill.

Lee, L H, Chew E P, Teng S and Goldsman D (2010). Finding the Non-Dominated Pareto Set for Multi-Objective Simulation Models. IIE Transactions 42:656–674.

Monks, T, Currie C S M, Onggo B S, Robinson S, Kunc M and Taylor S J E (2018). Strengthening the reporting of empirical simulation studies: Introducing the STRESS guidelines. Journal of Simulation advance online publication 14 February, https://doi.org/10.1080/17477778.2018.1442155

Nelson, B L and Hong L J (2009). A brief introduction to optimization via simulation. In Proceedings of the 2009 Winter Simulation Conference, USA, 75-85.

Parker, D, Brown D, Polhill, J G, Manson S M and Deadman P J (2008). Illustrating a new conceptual design pattern for agent-based models of land use via five case studies - the MR POTATOHEAD framework. In A. Lopex Paredes & H. I (Eds.), Agent-based modelling in natural resource management (pp. 23-51). Valladolid: INSISOC.

Rahmandad, H, & Sterman, J D (2012). Reporting guidelines for simulation-based research in social sciences. System Dynamics Review, 28(4): 396-411.

Robinson, S. (2004). Simulation: The Practice of Model Development and Use. John Wiley and Sons Ltd.

Robinson, S (2008). Conceptual modelling for simulation Part II: a framework for conceptual modelling. Journal of the Operational Research Society 59(3): 291-304.

Song, E, Nelson B L and Pegden C D (2014). Advanced tutorial: input uncertainty quantification. In Proceedings of the 2014 Winter Simulation Conference, USA, 162-176.

OR60 Annual Conference - Keynote Papers and Extended Abstracts
Lancaster University. Lancaster. 11 - 13 September 2018

Stochastic Modelling of Aircraft Queues: A Review

Rob Shone, Kevin Glazebrook, Konstantinos Zografos

Lancaster University, Department of Management Science, Lancaster, United Kingdom
r.shone@lancaster.ac.uk, k.glazebrook@lancaster.ac.uk, k.zografos@lancaster.ac.uk

Abstract

In this paper we consider the modelling and optimal control of queues of aircraft waiting to use the runway(s) at airports, and present a review of the related literature. We discuss the formulation of aircraft queues as nonstationary queueing systems and examine the common assumptions made in the literature regarding the random distributions for inter-arrival and service times. These depend on various operational factors, including the expected level of precision in meeting pre-scheduled operation times and the inherent uncertainty in airport capacity due to weather and wind variations. We also discuss strategic and tactical methods for managing congestion at airports, including the use of slot controls, ground holding programs, runway configuration changes and aircraft sequencing policies.

Keywords: Aviation; Queueing systems; Stochastic modelling

1. Introduction

Many of the busiest airports around the world experience very high levels of traffic congestion for lengthy periods of time during their daily operations. This is due to a rapid growth in demand for air transport services, combined with physical and political constraints which usually prevent the expansion of airport infrastructure in the short-term. Congestion increases the likelihood of flights being delayed, and these delays may propagate throughout an airport network, with serious financial consequences for airlines, passengers and other stakeholders (Pyrgiotis *et al*, 2013). As airport slot coordinators and traffic controllers strive to improve the efficiency of their operations, there is considerable scope for new and innovative mathematical modelling techniques to offer valuable insight.

The capacity of the runway system represents the main bottleneck of operations at a busy airport (de Neufville and Odoni, 2013). When demand exceeds capacity, queues of aircraft form either in the sky (in the case of arriving aircraft) or on the ground (in the case of departures). The purpose of this paper is to present a concise review of the methods used by researchers to model aircraft queues since research in this area began in earnest about 60 years ago. Aviation in general is currently a very active research area, and our review will touch upon some of the wider topics that are closely related to aircraft queue modelling, including demand management strategies and the potential of strategic and tactical interventions to improve the utilisation of scarce resources at airports. Thus, we intend to discuss aircraft queues not only from a mathematical modelling perspective, but also in the context of the optimisation problems frequently posed in the literature.

Of course, queueing theory itself is also a vast topic and there is no common agreement on which of the classical models (if any) are most appropriate in the context of air traffic. Classical queueing theory texts such as Kleinrock (1975) tend to focus on models which are mathematically tractable, such as those with Markovian distributions for customer inter-arrival times and/or service times. Generally, closed-form "steady-state" expressions for expected queue lengths, waiting times and other performance measures are available only in cases where the parameters of these distributions are stationary and customer arrival rates do not exceed service rates. However, demand for runway use at a typical airport varies considerably during the day according to the schedule of operations, and runway throughput rates may be affected by weather conditions, sequencing rules and other factors. Moreover, demand rates may exceed capacity limits for extended periods of time at busy hub airports (Barnhart *et al*, 2003). We therefore need to consider *time-dependent* queues, for which steady-state results are of limited practical use (Schwarz *et al*, 2016).

In general, two of the most important characteristics of a queueing system are the customer arrival and customer service processes. We therefore organise this paper in such a way that these processes are discussed in Sections 2 and 3 respectively. Other, more application-specific aspects of modelling air traffic queues, including the effects of weather conditions and approaches for modelling airport networks, are discussed in Section 4. In Section 5 we provide a summary and discuss possible directions for future research.

2. Modelling demand for runway usage at airports

This section is divided into two parts. The first part focuses on the modelling assumptions often made regarding demand processes at airports, and the second part discusses related optimisation problems which frequently attract attention in the literature.

2.1. Modelling assumptions

Throughout this section we are concerned with the processes by which aircraft join queues waiting to use the runway(s) at airports. In the case of departing aircraft, these queues are located on the ground, usually at the threshold of the departure runway(s). Arriving aircraft, on the other hand, must wait in airborne "holding stacks" which are usually located near the terminal airspace, although in some cases they may also be "held" at other stages of their journeys by air traffic controllers (to control the flow of traffic into a congested air sector, for example). In many cases, a plane which lands at an airport will take off again (not necessarily from the same runway) within a couple of hours. This implies that the demand processes for arrivals and departures are not independent of each other, but in fact it is quite common in existing mathematical models for arrivals and departures to be treated as independent queues with time-varying demand rates which are configured according to the schedule of operations. The assumption of independence is undoubtedly an oversimplification, but it may not be particularly harmful if one considers a large airport with separate runways being used for arrivals and departures (this system is referred to as "segregated operations" and is used at London Heathrow, for example).

Nonhomogenous Poisson processes (i.e. those with time-varying demand rates) were first used by Galliher and Wheeler (1958) to model the arrivals of landing aircraft at an airport. They used a discrete-time approach to compute probability distributions for queue lengths and waiting times. Subsequently, the Poisson assumption became very popular. Koopman (1972) considered the case of arrivals and departures sharing a single runway and modelled the Poisson arrival rates for both operation types as not only time-dependent but also state-dependent, with the two-dimensional state consisting of the queue lengths for arrivals and departures. This model allows for the possibility of "controlled" demand rates, whereby the demand placed on the system is reduced during peak congestion hours.

Hengsbach and Odoni (1975) extended Koopman's approach to the case of multiple-runway airports, and claimed that the nonhomogeneous Poisson model was consistent with observed data from several major airports. Subsequently, Dunlay and Horonjeff (1976) and Willemain *et al* (2004) used case studies to provide further evidence in support of the Poisson assumption. In the last few decades, nonhomogeneous Poisson processes have been widely adopted for queues of arrivals and departures at single airports (Kivestu, 1976; Bookbinder, 1986; Jung and Lee, 1989; Daniel, 1995; Hebert and Dietz, 1997; Fan, 2003; Mukherjee *et al*, 2005; Lovell *et al*, 2007; Stolletz, 2008; Jacquillat and Odoni, 2015a; Jacquillat *et al*, 2017) and also at networks of airports (Malone, 1995; Long *et al*, 1999; Long and Hasan, 2009; Pyrgiotis *et al*, 2013; Pyrgiotis and Odoni, 2016).

In case studies which rely on the nonhomogeneous Poisson model, the question arises as to how the demand rate functions for arrivals and departures – which we will denote here by $\lambda_a(t)$ and $\lambda_d(t)$ respectively – should be estimated. The schedule of operations for a single day at an airport can be used to aggregate the numbers of arrivals and departures expected to take place in contiguous time intervals of fixed length – for example, 15 minutes or one hour. The approach of Hengsbach and Odoni (1976) was to model $\lambda_a(t)$ and $\lambda_d(t)$ as piecewise linear functions, obtained by aggregating scheduled operations over each hour and then connecting the half-hour points using line segments, as shown in Figure 1. Various alternative data-driven methods can be devised. Jacquillat *et al* (2017) modelled $\lambda_a(t)$ and $\lambda_d(t)$ as piecewise constant over 15-minute intervals, while Bookbinder (1986) used hourly data but relied on a three-point moving average method to remove "jump discontinuities" in the demand rates which would otherwise occur at the end of each hour.

Of course, airlines operate flights according to pre-defined schedules, so it is reasonable to question whether the Poisson assumption (which implies memoryless inter-arrival times) actually makes sense in this context. Various arguments can be put forward to make the case that, in practice, inter-arrival (and inter-departure) times are 'sufficiently random' for the Poisson model to be valid. For example, Pyrgiotis (2011) argues that large deviations from scheduled operations can occur as a result of flight cancellations, delays at "upstream" airports, gate delays for departures, variability of flight times due to weather and winds, etc. These deviations have the effect of "randomising" actual queue entry times.

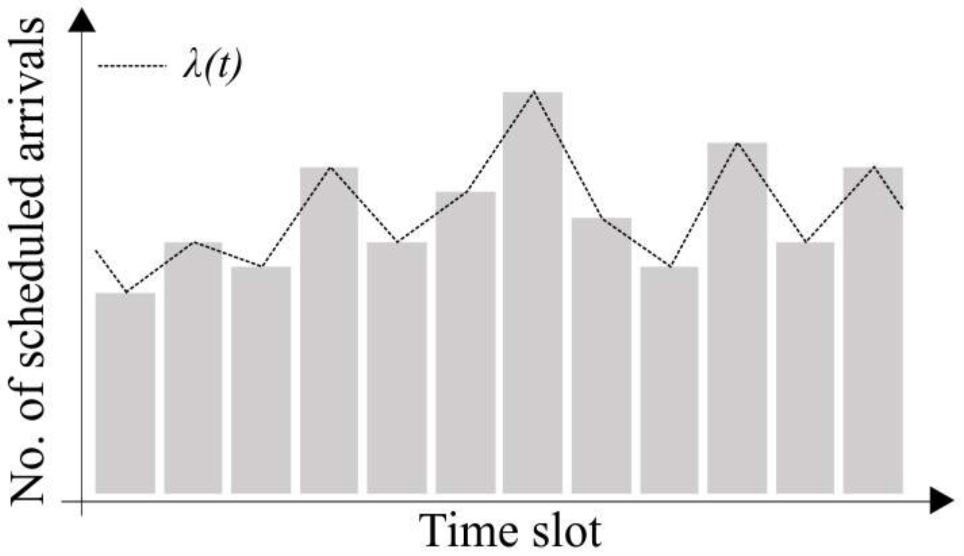


Figure 1: A piecewise linear, continuous function $\lambda(t)$, obtained by interpolating between half-hour points on a bar chart showing hourly demand. The function $\lambda(t)$ can be used as the demand function for a nonhomogeneous Poisson process.

Nevertheless, it is no surprise that various authors have challenged the Poisson assumption. In recent years, several authors have cited the development of the Next Generation Air Transportation System (NextGen) in the US as a possible reason for abandoning the Poisson model in the future. The NextGen system, which is expected to be fully in place by 2025, will allow four-dimensional *trajectory-based operations* (TBO). This should allow arrivals and departures to meet their scheduled operating times with greater precision (Joint Planning and Development Office, 2010). There is a similar ongoing project in Europe, known as Single European Sky ATM Research or SESAR (European Commission, 2014). In the light of these developments, there is considerable interest in modelling demand processes which have less variability than Poisson processes. Nikoleris and Hansen (2012) argued that the Poisson model cannot capture the effects of different levels of trajectory-based precision, because the variance in inter-arrival times is simply determined by the rate parameter. In a related piece of work, Hansen *et al* (2009) considered deterministic and exponentially-distributed inter-arrival times (both with time-varying rates) as two opposite extremes for the level of precision in meeting pre-scheduled operation times, and used case studies to show that the deterministic case could yield delay savings of up to 35%.

One type of demand process which has gained significant attention in recent years is the "pre-scheduled random arrivals" (PSRA) process. In PSRA queueing systems, customers have pre-scheduled arrival times but their actual arrival times vary according to random earliness/lateness distributions; for example, deviations from scheduled times may be normally or exponentially distributed. PSRA queues have been studied since the late 1950s (Winsten, 1959; Mercer, 1960), but their application to aircraft queues is a relatively recent development. An advantage of using the PSRA model is that variances of arrival and departure times can be controlled by choosing appropriate parameters for the earliness/lateness distributions, and this may be useful for modelling the more precise operation times expected under the NextGen and

SESAR systems. One disadvantage, however, is that PSRA queues are more difficult to study analytically, and indeed they are quite different from many of the classical models usually studied in queueing theory since inter-arrival times are neither independent nor identically distributed.

Guadagni *et al* (2011) made explicit comparisons between Poisson and PSRA demand processes and pointed out that PSRA queues exhibit negative autocorrelation, in the sense that time intervals which experience fewer arrivals than expected are likely to be followed by time periods with more arrivals than expected. Jouini and Benjaafar (2011) also made some progress in proving analytical properties of PSRA systems with heterogeneous customers and possible cancellations, although their model assumes that earliness/lateness distributions are bounded in such a way that customers are guaranteed to arrive in order of their scheduled times, which may not be suitable in an airport context. Caccavale *et al* (2014) used a PSRA model to study inbound traffic at Heathrow Airport, and argued that Poisson processes are a poor model for arrivals at a busy airport since, in practice, the arrivals stream is successively rearranged according to air traffic control (ATC) rules. Gwiggner and Nagaoka (2014) compared a PSRA model with a Poisson model using a case study based on Japanese air traffic, and found that the two models exhibited similar behaviour in systems with moderate congestion, but deviated from each other during high congestion. Lancia and Lulli (2017) studied the arrivals process at eight major European airports and found that a PSRA model with nonparametric, data-driven delay distributions provided a better fit for the observed data than a Poisson model.

Although time-dependent Poisson, deterministic and PSRA processes are by far the most popular choices for modelling aircraft queues found in the literature, a handful of other approaches have also been proposed. Krishnamoorthy *et al* (2009) considered “Markovian arrival processes” (MAPs), which generalise Poisson processes and can be studied using matrix analytic methods. Some authors have used observed data to fit nonparametric distributions for arrival and/or departure delays (Tu *et al*, 2008; Kim and Hansen, 2013). Finally, although our discussion throughout this section has focused on the use of time-dependent distributions, a number of authors have considered stationary demand processes (e.g. homogeneous Poisson processes) and attempted to gain insight by modelling aircraft as customers of different job classes (Rue and Rosenshine, 1985; Horonjeff and McKelvey, 1994; Bolender and Slater, 2000; Bauerle *et al*, 2007; Grunewald, 2016).

2.2. Optimisation Problems

Demand-related optimisation problems at airports are based on managing patterns of demand in such a way that the worst effects of congestion are mitigated, while at the same time the level of service provided (in terms of flight availability, punctuality, etc.) remains acceptable to passengers and other airspace users. Demand management strategies can be implemented at the strategic level, as part of an airport’s schedule design (which usually takes place several months in advance of operations) or at the tactical level, by making adjustments to aircraft flight plans in real time in order to prevent particular airports or airspace sectors from becoming heavily congested at certain times of day.

The busiest airports outside the US fall into the category of slot-controlled (level 3) airports, which means that airlines intending to use these airports for take-offs or landings must submit requests for time slots (typically 15 minutes long) during which they have permission to use the runways and other airport infrastructure. Although the US does not implement slot controls in the same manner, a small number of its airports are subject to the 'high density rule', which imposes hourly capacity limits (Madas and Zografos, 2006). Since slot allocation is usually carried out with a broad set of objectives in mind (including the need to design schedules which satisfy airlines' requirements as equitably as possible), the resulting schedules do not always insure effectively against the danger of severe operational (queueing) delays occurring in practice. For example, if too many flights are allocated to a small set of consecutive time slots, the consequences for airport congestion levels may be catastrophic. Thus, there is a need for demand management strategies to ensure that congestion mitigation is included as part of the slot allocation procedure.

Various authors (Barnhart et al, 2012; Swaroop et al, 2012; Zografos et al, 2012) have commented on the inherent trade-off that exists between schedule displacement and operational delays, as illustrated by Figure 2. At slot-controlled airports, certain time slots tend to be more sought-after by airlines than others. As a result, flight schedules which conform closely to airline requests are likely to result in large 'peaks' in demand at certain times of day. These schedules incur only a small amount of schedule displacement, since the requests from airlines are largely satisfied; however, severe operational delays are likely to be caused by the peaks in demand. Conversely, operational delays can be reduced by smoothing (or 'flattening out') the schedule to avoid such peaks, but this generally involves displacing flights to a greater extent from the times requested by airlines.

A useful survey of demand management strategies that have been implemented around the world is provided by Fan and Odoni (2002). These strategies can generally be divided into two categories: administrative and market-based. Administrative strategies involve setting 'caps' on the numbers of runway operations that can take place at an airport in a single time period, or a number of consecutive time periods. These 'caps' may apply to arrivals, departures or both, and are usually referred to in the aviation community as declared capacities (Zografos et al, 2017). The relevant optimisation problems involve deciding how these caps should be set optimally in order to ensure a satisfactory trade-off is achieved between congestion levels (which are usually modelled stochastically) and airlines' operational needs (Swaroop et al, 2012; Churchill et al, 2012; Corolli, 2013). On the other hand, market-based strategies are based on using economic measures such as congestion pricing and slot auctions to relieve congestion during peak periods (Andreatta and Odoni, 2003; Fan, 2003; Pels and Verhoef, 2004; Mukherjee et al, 2005; Ball et al, 2006; Andreatta and Lulli, 2009; Pellegrini et al, 2012). A number of authors have directly compared administrative and market-based strategies using analyses and/or case studies (Brueckner, 2009; Basso and Zhang, 2010; Czerny, 2010; Gillen et al, 2016).

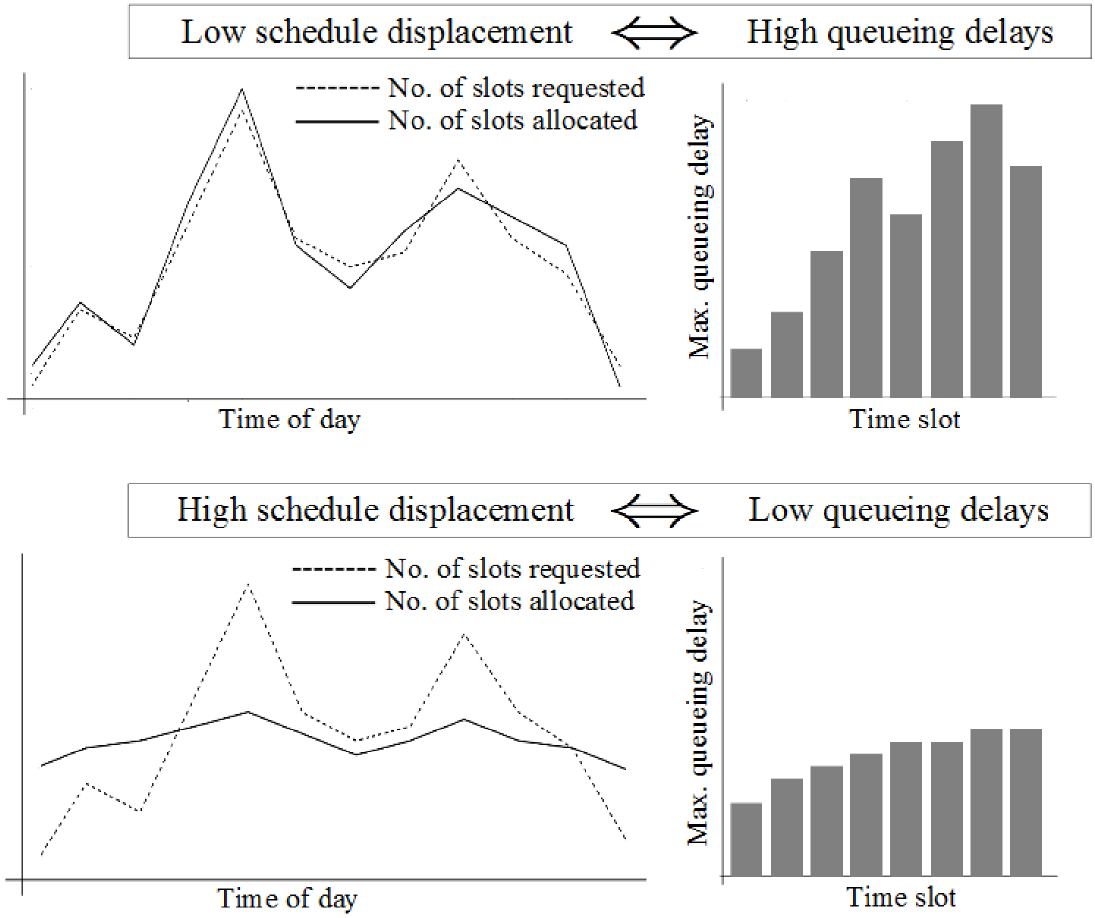


Figure 2: The trade-off between schedule displacement and operational delays.

As mentioned earlier, demand management can also be done at a tactical level. Ground-holding programs can be used to delay departing aircraft in order to ensure that they do not arrive at their destination airports during periods of high congestion. This not only relieves congestion at busy airports, but also has the benefit of preventing aircraft from wasting too much fuel by being forced to wait in airborne holding stacks. In the broader realm of air traffic flow management (ATFM), aircraft can be directed by air traffic controllers to delay their arrivals at congested airports or air sectors by adjusting their speeds or routes. Optimisation problems related to ground-holding programs and ATFM operations have been well-studied in the literature. These problems usually do not involve stochastic queue modelling, but they do commonly take account of the uncertainty caused by weather conditions, enroute congestion etc. by incorporating probabilistic “capacity profiles” for destination airports and incorporating the probabilities for different capacity scenarios within optimisation models such as integer linear programs (ILPs). Some notable references include Terrab and Odoni (1992), Richetta and Odoni (1993), Vranas et al (1994), Richetta (1995), Bertsimas and Stock (1998), Hall (1999), Ball et al (2003), Inniss and Ball (2004), Kotnyek and Richetta (2006), Lulli and Odoni (2007), Mukherjee and Hansen (2007), Liu et al (2008), Balakrishnan and Chandran (2014).

3. Modelling capacity and runway throughput at airports

This section is organised in a similar way to Section 2. The first part focuses on modelling assumptions related to capacity and runway throughput rates at airports, and the second part discusses some relevant optimisation problems.

3.1. Modelling assumptions

An airport's *capacity* can be defined as the expected number of runway movements (either arrivals or departures) that can be operated per unit time under conditions of continuous demand (de Neufville and Odoni, 2013). It is very important to estimate an airport's capacity accurately, since long queues of aircraft waiting to use the runway(s) may form as a result of imbalances between demand and capacity, and therefore capacity modelling must be used to inform the demand management strategies discussed in Section 2.2. However, an airport's capacity is not simply a fixed quantity, but instead is time-varying and depends on a number of factors. Adverse weather conditions might increase the separation requirements between consecutive arriving aircraft, while strong winds may prevent certain runways from being used. Additionally, runway movements might be restricted at certain times of day due to noise considerations. For example, at Heathrow Airport, the period between 11:30pm and 6:00am is known as the "Night Quota Period", with traffic restrictions imposed by the Department for Transport (Heathrow Airport, 2018).

Blumstein (1959) produced a seminal paper in which he explained how to calculate the landing capacity of a single runway (i.e. when it is used for arrivals only) based on aircraft speeds and separation requirements. Hockaday and Kanafani (1974) generalised Blumstein's work by deriving expressions for the capacity of a single runway under three different modes of operation: arrivals only, departures only and mixed operations. Newell (1979) showed how to extend these analyses to airports with multiple runways under various different configurations. A key principle which emerged from these early contributions was the importance of taking into account different possible fleet mixes and sequencing strategies. When one runway movement is followed by another, the movements in question are subject to a minimum time separation which depends not only on the types of movements involved (arrivals or departures), but also on the types of aircraft. To be more specific, aircraft can be categorised into different 'weight classes'. Heavy aircraft generate a lot of wake turbulence, which can be dangerous to lighter aircraft following too closely behind (Newell, 1979). Therefore, in airport capacity calculations, one must take into account the relative expected frequencies of different 'weight pairs' (e.g. heavy-light, heavy-heavy, etc.) and use these to calculate average time separations between movements.

Gilbo (1993) developed the idea of the *runway capacity curve* (referred to by subsequent authors as a "capacity envelope"), as shown in Figure 3. This curve represents the departure capacity of an airport as a convex, nonincreasing function of the arrival capacity. The shape of the curve depends on various time-varying factors, including weather conditions, the runway configuration in use and the aircraft fleet mix. However, the essential principle is that each point on the capacity envelope represents a feasible pair of capacity values for arrivals and departures during the time period for which the envelope applies. Various authors have provided detailed descriptions of how airport capacity envelopes can be constructed using both empirical and analytical methods (Lee *et al*, 1997; Stamatopoulos *et al*, 2004; Simaiakis, 2013) and these capacity envelopes have been incorporated into various types of optimisation problems, which we discuss further in Section 3.2.

The capacity of an airport is naturally related to the concept of a *service rate* in queueing theory, since it specifies how many runway movements (which we can think of as 'services' of aircraft) can be achieved in a given time interval. Several early studies modelled the queueing dynamics at airports using nonstationary deterministic models, with the arrival and service rates defined according to flight schedules and capacity estimates respectively (Kivestu, 1976; Hubbard, 1978; Newell, 1979). However, at the same time, interest was developing in modelling aircraft service times stochastically. Koopman (1972) proposed that the queueing dynamics of an airport with a nonhomogeneous Poisson process for arrivals and k runways (modelled as independent servers) could be bounded by the characteristics of the $M(t)/D(t)/k$ and $M(t)/M(t)/k$ queueing systems. The former system – in which the service process is nonstationary and deterministic – can be regarded as a "best-case" scenario, since queueing delays are shorter in the case of predictable service times. The latter system – with exponentially-distributed service times – is a "worst-case" scenario, in which highly variable service times cause average queueing delays to increase. Koopman used numerical solution of the Chapman-Kolmogorov equations (assuming a finite queue capacity) to estimate queue length probability distributions.

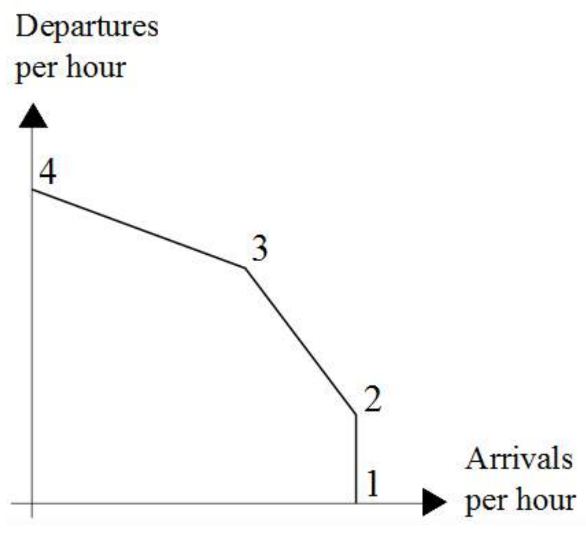


Figure 3: A piecewise linear capacity envelope for a particular time interval, adapted from Stamatopoulos *et al* (2004). Each point on the envelope represents a feasible pair of capacity values. Points 1 and 4 represent "all arrivals" and "all departures" policies respectively. Point 2 represents a sequencing strategy where departures are freely inserted during large inter-arrival gaps, and Point 3 is a "mixed operations" point.

Kivestu (1976) proposed an $M(t)/E_k(t)/s$ queueing model for aircraft queues, in which the service time distribution is Erlang with k exponentially-distributed service phases. This approach is closely related to that of Koopman (1972), since the cases $k = 1$ and $k = \infty$ represent exponential and deterministic service times respectively. However, Kivestu also introduced a fast, practical numerical approximation method for the time-dependent queue length probabilities in an $M(t)/E_k(t)/1$ queue, which became known as the DELAYS algorithm. Subsequently, DELAYS – as well as the $M(t)/E_k(t)/1$ model for aircraft queues itself – has become very popular, and has been used for estimating queueing delays in a variety of settings (Abundo, 1990; Malone, 1995; Fan and Odoni, 2002; Stamatopoulos *et al*, 2004; Mukherjee *et al*, 2005; Lovell *et al*, 2007; Churchill *et al*, 2008; Hansen *et al*, 2009; Pyrgiotis and Odoni, 2016). An advantage of using DELAYS is that it estimates the full probability distribution for the queue length at any given time. This is useful because, in practice, airports are often interested in tail-based performance measures such as the expected number of queueing delays that will exceed a given threshold.

The $M(t)/E_k(t)/1$ model for aircraft queues can be regarded as somewhat macroscopic, since it does not explicitly take into account fleet mixes and separation requirements between different aircraft types. Instead, it assumes that such considerations are implicitly accounted for via the use of an Erlang distribution for service times, whose variance can be controlled by adjusting the parameter k (with larger values implying less variance). Models which explicitly consider runway occupancy times for different classes of aircraft have been proposed by a number of authors. Hockaday and Kanafani (1974) and Stamatopoulos *et al* (2004) modelled these using normal distributions, while Jeddi *et al* (2006) suggested beta distributions and Nikoleris and Hansen (2015) used Gumbel random variables. In models where different aircraft types are considered explicitly, there is certainly some justification for using service time distributions with very small variances – or even deterministic service times – since the time lapse between two consecutive aircraft entering the runway depends on separation guidelines which are enforced by air traffic controllers according to the aircraft weight classes. Indeed, a number of authors have used stochastic queueing formulations in which service times are deterministic, such as $M(t)/D(t)/1$ (Galliher and Wheeler, 1958; Daniel, 1995; Daniel and Pahwa, 2000) or $PSRA/D/1$ (Caccavale *et al*, 2014; Gwiggner and Nagaoka, 2014). The use of trajectory-based operations in the future (as discussed in Section 2.1) may provide further justification for considering deterministic service times.

In traditional queueing system formulations, the description of the service process includes not only the service time distribution but also the number of servers, finite queue capacity (if applicable) and the queue discipline. We therefore address the relevant modelling assumptions here in an aviation context. The assumption of a single server (as in the $M(t)/E_k(t)/1$ model, for example) is surprisingly common in the literature, even when the airport being modelled has more than one runway. One possible explanation for this is that even when an airport has multiple runways, there is usually some inter-dependence between them, which implies that it is inappropriate to model them as independent servers (Jacquillat, 2012). For example, runways may intersect each other – or even if they do not, they may be too closely-spaced to allow independent operations, since the effects of wake turbulence may create “diagonal separation requirements" between aircraft on different runways (Stamatopoulos *et al*, 2004). Nevertheless, the single-server assumption – which effectively models the runway system as a 'black box' processing arrivals and departures – is undoubtedly a simplification which, arguably, has been over-used in the literature.

A finite queue capacity is usually not considered an essential component of an airport queueing model, since in practice it is rare for aircraft to be denied access to an airport due to over-congestion (this would be referred to as 'balking' in queueing theory). Nevertheless, it should be noted that certain numerical methods for estimating queue length probability distributions, including the DELAYS algorithm and numerical solution of the Chapman-Kolmogorov equations, must assume a finite queue capacity for computational purposes. In practice, the queue capacity used in these methods is chosen to be large enough to ensure that it has very little impact on estimates of performance measures.

Finally, models which assume independent queues for arrivals and departures use the first-come-first-served queue discipline (FCFS) almost universally, unless they are intended to examine the effects of different sequencing policies. The FCFS assumption is largely consistent with ATC procedures in practice (Pyrgiotis, 2011). Some authors, however, have considered priority queues in which arrivals are given priority for service over departures (Roth, 1979; Horonjeff and McKelvey, 1994; Grunewald, 2016).

3.2 Optimisation problems

Optimisation problems related to service rates in aircraft queues may involve the strategic or tactical control of runway configurations, the dynamic balancing of service rates between arrivals and departures, the sequencing of aircraft using the runway(s) or some combination of these. In this subsection we provide examples from the literature.

As mentioned in Section 3.1, the shape of the airport capacity envelope (see Figure 3) depends on a number of operational factors which may vary during a day of operations. One such factor is the *runway configuration*. Airports with multiple runways may control which ones are active at any given time, although sometimes this choice is constrained by wind conditions which make it unsafe for aircraft to take off or land in a particular direction (Jacquillat and Odoni, 2015a). Ramanujam and Balakrishnan (2015) used empirical data to analyse the runway configuration selection process at US airports and aimed to predict the configurations chosen under different wind, weather and demand conditions. As discussed earlier, any point on the capacity envelope associated with a particular runway configuration represents a pair of attainable *capacity values* for arrivals and departures. It is natural to interpret capacity values as *service rates* which can be incorporated within queueing models. Various authors have considered optimisation problems in which an airport capacity envelope (or sequence of envelopes) is given, and the objective is to choose a sequence of points (i.e. service rate pairs) on these envelopes which will optimise a performance measure related to queue lengths or flight operation times (Gilbo, 1993; Gilbo, 1997; Hall, 1999; Dell'Olmo and Lulli, 2003). Other authors have extended this approach by modelling the runway configuration as a decision variable, so that the decision-maker must jointly optimise runway configurations and arrival/departure service rates (Li and Clarke, 2010; Weld *et al*, 2010; Bertsimas *et al*, 2011). Jacquillat *et al* (2017) (see also Jacquillat and Odoni, 2015b) also considered a similar problem, but made an important contribution by including *stochastic* queueing dynamics (based on an $M(t)/E_k(t)/1$ formulation) in their model. Prior to this, deterministic queueing dynamics had generally been assumed for such problems, with solutions found using ILPs.

Aircraft sequencing (also known as runway scheduling) problems involve planning the order that arriving and/or departing aircraft will use the runway(s) in such a way that a certain performance measure is optimised. As discussed earlier, the time separations between consecutive runway movements depend on the types of aircraft involved, and significant amounts of time can be lost if smaller aircraft often have to follow heavier ones. Throughput rates will generally be maximised if groups of similar aircraft are allowed to take off or land consecutively, but other constraints and objectives must also be taken into account. For example, individual aircraft might have to take off or land within fixed time windows, and the objective(s) might include allowing aircraft to take off (land) as close as possible to pre-specified ‘preferred’ take-off (landing) times. “Constrained position shifting” (CPS), whereby an aircraft’s position in the sequence is allowed to deviate by only a certain maximum number of places from its position in a “first-come-first-served” sequence, is another common way of enforcing constraints (Dear, 1976). “Static” aircraft sequencing problems are those in which the sequence of runway movements is optimised only once, and does not change in response to any subsequent events (Psaraftis, 1978; Beasley *et al*, 2000; Artiouchine *et al*, 2008; Balakrishnan and Chandran, 2010). On the other hand, in the dynamic version of the problem, the sequence is re-optimised every time new aircraft enter the terminal control area and become available for sequencing (Dear, 1976; Beasley *et al*, 2004; Murca and Muller, 2015; Bennell *et al*, 2017). Both versions of the problem are usually formulated as deterministic optimisation problems and solution approaches may include dynamic programming, branch-and-bound methods and metaheuristics (Potts *et al*, 2009). There has also been some interest in formulating *stochastic* runway scheduling problems. In these problems, the random variables may include the arrival times of aircraft in terminal areas, pushback delays for departures and taxiway times. Two-stage stochastic optimisation has been employed as a solution method (Anagnostakis and Clarke, 2003; Solveling *et al*, 2011; Solak *et al*, 2018).

4. Other modelling considerations at airports

In this section we discuss certain other aspects of modelling airport operations which have been touched upon only briefly in the previous sections.

Firstly, we address the subject of weather. One of the most obvious reasons for using stochastic (as opposed to deterministic) models for airport operations is the fact that weather and wind conditions can never be anticipated with complete confidence. Poor weather conditions cause visibility problems which can increase the separation requirements between consecutive runway operations and runway occupancy times, thereby effectively reducing airport capacities. Gilbo (1993) described how empirical data could be used to construct separate capacity envelopes for different weather categories. Other authors (Simaiakis, 2013; Jacquillat and Odoni, 2015a) have noted that, in practice, a distinction exists between “Visual

Meteorological Conditions" (VMC) and "Instrumental Meteorological Conditions" (IMC), which indicate "good" and "poor" weather respectively. Based on this distinction, VMC and IMC envelopes can be constructed for each possible runway configuration, with the IMC envelopes being smaller than the VMC ones but similar in shape.

Of course, knowing how to estimate airport capacity envelopes under different weather conditions is one thing, but simulating random weather changes within decision problems is quite another. When designing stochastic models for weather evolution, it makes sense to consult historical data in order to estimate the relative frequencies for different weather states. Modelling the random transitions between weather states can be done in different ways. Jacquillat and Odoni (2015a) used a nonstationary two-state Markov chain, with the transition probabilities from state "VMC" to "IMC" and vice versa estimated using historical data (see Figure 4). Other authors have opted for a semi-Markov model, in which the time spent in a particular weather state has a non-exponential distribution (Abundo, 1990; Peterson *et al*, 1995). In the literature on ground-holding problems discussed in Section 2.2, it is common practice to represent an airport's capacity profile probabilistically by specifying probabilities for different weather scenarios (see, for example, Richetta and Odoni, 1993). Liu *et al* (2008) and Buxi and Hansen (2011) have discussed the use of clustering techniques for generating probabilistic capacity profiles. In problems where the decision-maker has the ability to switch between different runway configurations, it is also important to note that wind conditions may prevent certain configurations from being used. Jacquillat *et al* (2017) described the use of a Markov chain model for transitions between 13 different wind states in a case study based on JFK Airport in New York.

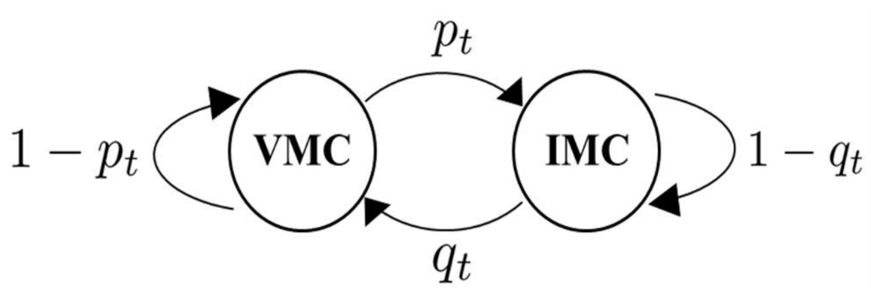


Figure 4: A nonstationary two-state Markov chain model for weather evolution. The parameters $p_t, q_t \in [0,1]$ depend on the discrete time interval *t*.

The previous sections of this paper have focused mainly on the modelling of aircraft queues at single airports. However, research has also been done into modelling *airport networks*. The relevant papers tend to focus on the propagation of delays around a network, referred to as the "ripple effect". Long *et al* (1999) (see also Long and Hasan, 2009) developed the "LMINET model", in which airports are modelled as a network of $M(t)/E_k(t)/1$ queues. Pyrgiotis *et al* (2013) developed the "Approximate Network Delays" (AND) model (first conceptualised in Malone (1995)), which iterates between a network queueing engine and a delay propagation algorithm for modelling network delays. The queueing engine is based on a network of

$M(t)/E_k(t)/1$ queues and relies upon the DELAYS algorithm discussed in Section 3.1, while the delay propagation algorithm explicitly considers individual aircraft itineraries. Baspinar *et al* (2016) used a similar model to investigate the effects of local disturbances (e.g. strike action or severe weather) at European airports. Czerny (2010) compared slot constraints with congestion pricing as alternative methods for managing demand in a network of airports, while Vaze and Barnhart (2012) used the AND model to test the effects of demand management strategies under different capacity scenarios. Campanelli *et al* (2016) discussed the use of agent-based simulations for modelling network delays. It should also be mentioned that all of the literature on ground holding problems (see Section 2.2) is network-related, since the decision to delay an aircraft's departure from one airport is made with the intention of improving congestion at another. However, several of these papers consider simplified "star-shaped" networks in which a single "hub" airport is assumed to be the only one prone to congestion, and ground holding decisions made at other airports are based entirely on managing congestion at the hub airport.

Finally, some interesting papers have arisen from considering the differences in demand management and ATFM strategies at US and European airports. Odoni *et al* (2011) compared the demand-to-capacity relationships at Frankfurt International Airport (which is a slot-coordinated airport) and Newark Liberty Airport in New York. Frankfurt is subject to much stricter demand regulation than Newark, and consequently it performs better with respect to average flight delays, but the paper suggests that the economic benefits of increasing slot limits may outweigh the costs of increased delays. Swaroop *et al* (2012) investigated the slot controls in use at the four slot-controlled airports in the US and found that the costs of airport congestion were too high to justify the relatively relaxed slot constraints. Both of the aforementioned papers support the general view that slot controls in Europe are too strict, whilst in the US they tend to be too liberal. Campanelli *et al* (2016) investigated the differences in network delays between the US and European air traffic systems which are caused by different aircraft sequencing strategies.

5. Summary

Methods for modelling aircraft queues are continuously evolving. Nonstationary models based on classical queueing theory are still employed frequently. For example, the $M(t)/E_k(t)/1$ model continues to attract a lot of attention. With the ongoing development of systems based on trajectory-based operations (in particular, the NextGen system in the US and SESAR in Europe), we anticipate that models which allow the variances of inter-arrival times and/or service times to be controlled at a finer level – through the use of queue entry times based on pre-scheduled random arrivals (PSRA), for example – are likely to become more popular. In addition, we suggest that some of the simplifying assumptions that have been adopted almost universally over the last few decades – such as the single-server assumption for multiple-runway airports and the independence of queues for arrivals and departures – are likely to be relaxed as researchers increasingly aim to incorporate high-fidelity models of airport operations into their optimisation procedures.

This paper has touched upon some of the broader issues related to the stochastic modelling of aircraft queues, including demand management strategies and the tactical control of aircraft take-off and landing sequences. It is clear from our discussion that, in reality, the queueing dynamics at airports are influenced by a diverse range of factors, including the decisions made at different points in time by multiple stakeholders. From a strategic point of view, the decisions made regarding slot controls (or congestion pricing) at airports and the slot requests submitted by airlines are important for determining, several months in advance, the daily demand profiles at airports. However, the tactical decisions made by airports and air traffic controllers in 'real time' – which may be related to sequencing patterns, ground holding delays and runway configuration changes – are also critical for managing congestion. We conclude that the modelling and optimisation of queues and congestion levels at airports is a complex task which should be informed by field analyses and engagement with industry practitioners in order to maximise research impact.

References

Abundo S F (1990). An Approach for Estimating Delays at a Busy Airport. Master's thesis, Operations Research Center, Massachusetts Institute of Technology, Cambridge, MA.

Anagnostakis I and Clarke J-P (2003). Runway operations planning: a two-stage solution methodology. Proc. 36th annual Hawaii international conference on system sciences (HICSS'03), Big Island, HI.

Andreatta G and Odoni A R (2003). Analysis of market-based demand management strategies for airports and en route airspace. In: Ciriani T A, Fasano G, Gliozzi S and Tadei R (eds), Chapter 14, Operations Research in Space and Air, Kluwer, pp 257-278.

Andreatta G and Lulli G (2009). Equitable Demand Management Strategies for Different Classes of Customers. International Journal of Pure and Applied Mathematics, 57(1):1-22.

Artiouchine K, Baptiste P and Durr C (2008). Runway sequencing with holding patterns. European Journal of Operational Research, 189(3):1254-1266.

Balakrishnan H and Chandran B G (2010). Algorithms for Scheduling Runway Operations Under Constrained Position Shifting. Operations Research, 58(6):1650-1665.

Balakrishnan H and Chandran B G (2014). Optimal Large-Scale Air Traffic Flow Management. Accessed May 21, 2018. http://web.mit.edu/Hamsa/www/pubs/BalakrishnanChandran_ATFM.pdf.

Ball M O, Hoffman R, Odoni A R and Rifkin R (2003). A Stochastic Integer Program with Dual Network Structure and Its Application to the Ground-Holding Problem. Operations Research, 51(1):167-171.

Ball M O, Donohue G and Hoffman K (2006). Auctions for the safe, efficient and equitable allocation of airspace system resources. In: Cramton P, Shoham Y and Steinberg R (eds), Combinatorial Auctions, Cambridge: MIT Press, pp 507-538.

Barnhart C, Belobaba P and Odoni A R (2003). Applications of Operations Research in the Air Transport Industry. Transportation Science, 37(4):368-391.

Barnhart C, Fearing D, Odoni A R and Vaze V (2012). Demand and capacity management in air transportation. EURO Journal on Transportation and Logistics, 1(1-2):135-155.

Baspinar B, Kemal Ure N, Koyuncu E and Inalhan G (2016). Analysis of Delay Characteristics of European Air Traffic through a Data-Driven Airport-Centric Queuing Network Model. IFAC-PapersOnLine, 49(3):359-364.

Basso L J and Zhang A (2010). Pricing vs. slot policies when airport profits matter. Transportation Research Part B, 44:381-391.

Bauerle N, Engelhardt-Funke O and Kolonko M (2007). On the waiting time of arriving aircrafts and the capacity of airports with one or two runways. European Journal of Operational Research, 177(2):1180-1196.

Beasley J E, Krishnamoorthy M, Sharaiha Y M and Abramson D (2000). Scheduling Aircraft Landings – The Static Case. Transportation Science, 34(2):180-197.

Beasley J E, Krishnamoorthy M, Sharaiha Y M and Abramson D (2004). Displacement Problem and Dynamically Scheduling Aircraft Landings. Journal of the Operational Research Society, 55(1):54-64.

Bennell J A, Mesgarpour M and Potts C N (2017). Dynamic scheduling of aircraft landings. European Journal of Operational Research, 258(1):315-327.

Bertsimas D J and Stock Patterson S (1998). The Air Traffic Flow Management Problem with Enroute Capacities. Operations Research, 46(3):406-422.

Bertsimas D J, Frankovich M and Odoni A R (2011). Optimal Selection of Airport Runway Configurations. Operations Research, 59(6):1407-1419.

Blumstein A (1959). The Landing Capacity of a Runway. Operations Research, 7(6):752-763.

Bolender M A and Slater G L (2000). Evaluation of scheduling methods for multiple runways. Journal of Aircraft, 37(3):410-416.

Bookbinder J (1986). Multiple Queues Of Aircraft Under Time-Dependent Conditions. INFOR: Information Systems and Operations Research, 24(4):280-288.

Brueckner J K (2009). Price vs. quantity-based approaches to airport congestion management. Journal of Public Economics, 93(5-6):681-690.

Buxi G and Hansen M (2011). Generating Probabilistic Capacity Profiles from weather forecast: A design-of-experiment approach. Proc. 9th USA/Europe Air Traffic Management R&D Seminar, Berlin, Germany.

Caccavale M V, Iovanella A, Lancia C, Lulli G and Scoppola B (2014). A model of inbound air traffic: The application to Heathrow airport. Journal of Air Transport Management, 34:116-122.

Campanelli B, Fleurquin P, Arranz A, Etxebarria I, Ciruelos C, Eguiluz V M and Ramasco J J (2016). Comparing the modeling of delay propagation in the US and European air traffic networks. Journal of Air Transport Management, 56:12-18.

Churchill A M, Lovell D J, Mukherjee A and Ball M O (2012). Determining the Number of Airport Arrival Slots. Transportation Science, 47(4):526-541.

Churchill A M, Vlachou K and Lovell D J (2008). Filtering and aggregation schemes for delay model calibration. Proc. 3rd International Conference on Research in Air Transportation (ICRAT), Fairfax, VA.

Corolli L (2013). Deterministic and Stochastic Optimization for Heterogeneous Decision Phases in the Air Traffic Domain. PhD thesis, Universita degli Studi di Milano-Bicocca.

Czerny A I (2010). Airport congestion management under uncertainty. Transportation Research Part B, 44(3):371-380.

Daniel J I (1995). Congestion Pricing and Capacity of Large Hub Airports: A Bottleneck Model with Stochastic Queues. Econometrica, 63(2):327-370.

Daniel J I and Pahwa M (2000). Comparison of Three Empirical Models of Airport Congestion Pricing. Journal of Urban Economics, 47(1):1-38.

Dear R G (1976). The Dynamic Scheduling of Aircraft in the Near Terminal Area. Technical report R76-9, MIT Flight Transportation Laboratory.

Dell'Olmo P and Lulli G (2003). A dynamic programming approach for the airport capacity allocation problem. IMA Journal of Management Mathematics, 14(3):235-249.

Dunlay W J and Horonjeff R (1976). Stochastic properties of enroute air traffic – An empirical investigation. Journal of Aircraft, 13(5):376-381.

European Commission (2018). SESAR 2020: developing the next generation of European Air Traffic Management. Accessed May 21, 2018. https://ec.europa.eu/research/press/jti/factsheet_sesar-web.pdf .

Fan T P (2003). Market-based Airport Demand Management – Theory, Model and Applications. PhD thesis, Massachusetts Institute of Technology, Cambridge, MA.

Fan T P and Odoni A R (2002). A practical perspective on airport demand management. Air Traffic Control Quarterly, 10(3):285-306.

Galliher H P and Wheeler R C (1958). Nonstationary Queuing Probabilities for Landing Congestion of Aircraft. Operations Research, 6(2):264-275.

Gilbo E P (1993). Airport Capacity: Representation, Estimation, Optimization. IEEE Transactions on Control Systems Technology, 1(3): 144-154.

Gilbo E P (1997). Optimizing Airport Capacity Utilization in Air Traffic Flow Management Subject to Constraints at Arrival and Departure Fixes. IEEE Transactions on Control Systems Technology, 5(5):490-503.

Gillen D, Jacquillat A and Odoni A R (2016). Airport demand management: The operations research and economic perspectives and potential synergies. Transportation Research Part A, 94:495-513.

Guadagni G, Ndreca S and Scoppola B (2011). Queueing systems with pre-scheduled random arrivals. Mathematical Methods of Operations Research, 73(1):1-18.

Grunewald E (2016). Incentive-based Slot Allocation for Airports. Transportation Research Procedia, 14:3761-3770.

Gwiggner C and Nagaoka S (2014). Data and queueing analysis of a Japanese air-traffic flow. European Journal of Operational Research, 235(1):265-275.

Hall W (1999). Efficient capacity allocation in a collaborative air transportation system. PhD thesis, Massachusetts Institute of Technology, Cambridge, MA.

Hansen M, Nikoleris T, Lovell D J, Vlachou K and Odoni A R (2009). Use of Queuing Models to Estimate Delay Savings from 4D Trajectory Precision. Proc. 8th USA/Europe Air Traffic Management R&D Seminar, Napa, CA.

Heathrow Airport (2018). Night flights. Accessed May 21, 2018. https://www.heathrow.com/noise/heathrow-operations/night-flights.

Hebert J E and Dietz D C (1997). Modeling and Analysis of an Airport Departure Process. Journal of Aircraft, 34(1):43-47.

Hengsbach G and Odoni A R (1975). Time dependent estimates of delays and delay costs at major airports. Technical report, Dept. of Aeronautics and Astronautics, Massachusetts Institute of Technology, Cambridge, MA.

Hockaday S L M and Kanafani A K (1974). Developments in Airport Capacity Analysis. Transportation Research, 8(3):171-180.

Horonjeff R and McKelvey F X (1994). Planning and design of airports. McGraw-Hill, 4th edition, New York.

Hubbard H B (1978). Terminal Airspace/Airport Congestion Delays. Interfaces, 8(2):1-14.

Inniss T R and Ball M O (2004). Estimating One-Parameter Airport Arrival Capacity Distributions for Air Traffic Flow Management. Air Traffic Control Quarterly, 12(3):223-251.

Jacquillat A (2012). A queuing model of airport congestion and policy implications at JFK and EWR. Master's thesis, Massachusetts Institute of Technology, Cambridge, MA.

Jacquillat A and Odoni A R (2015a). An Integrated Scheduling and Operations Approach to Airport Congestion Mitigation. Operations Research, 63(6):1390-1410.

Jacquillat A and Odoni A R (2015b). Endogenous control of service rates in stochastic and dynamic queuing models of airport congestion. Transportation Research Part E, 73:133-151.

Jacquillat A, Odoni A R and Webster M D (2017). Dynamic Control of Runway Configurations and of Arrival and Departure Service Rates at JFK Airport Under Stochastic Queue Conditions. Transportation Science, 51(1):155-176.

Jeddi B G, Shortle J F and Sherry L (2006). Statistics of the Approach Process at Detroit Metropolitan Wayne County Airport. Proc. 2nd International Conference on Research in Air Transportation (ICRAT), Belgrade, Serbia and Montenegro.

Joint Planning and Development Office (2010). Concept of Operations for the Next Generation Air Transportation System, Version 3.0, Washington, DC.

Jouini O and Benjaafar S (2011). Queueing systems with appointment-driven arrivals, non-punctual customers, and no-shows. Working paper, Ecole Centrale Paris, Chatenay-Malabry, France.

Jung M and Lee E S (1989). Numerical Optimization of a Queueing System by Dynamic Programming. Journal of Mathematical Analysis and Applications, 141(1):84-93.

Kim A and Hansen M (2013). Deconstructing delay: A non-parametric approach to analyzing delay changes in single server queuing systems. Transportation Research Part B, 58:119-133.

Kivestu P (1976). Alternative Methods of Investigating the Time-Dependent M/G/k Queue. S.M. thesis, Dept. of Aeronautics and Astronautics, Massachusetts Institute of Technology, Cambridge, MA.

Kleinrock L (1975). Queueing Systems, Volume 1: Theory. John Wiley & Sons Inc., New York.

Koopman B O (1972). Air-Terminal Queues under Time-Dependent Conditions. Operations Research, 20(6):1089-1114.

Kotnyek B and Richetta O (2006). Equitable Models for the Stochastic Ground-Holding Problem under Collaborative Decision Making. Transportation Science, 40(2):133-146.

Krishnamoorthy A, Babu S and Narayanan V C (2009). The MAP/(PH/PH)/1 queue with self-generation of priorities and non-preemptive service. European Journal of Operational Research, 195(1):174-185.

Lancia C and Lulli G (2017). Data-driven modelling and validation of aircraft inbound-stream at some major European airports. Preprint available from arXiv database. Accessed May 21, 2018. https://arxiv.org/abs/1708.02486.

Lee D A, Kostiuk P F, Hemm R V, Wingrove W E and Shapiro G (1997). Estimating the effects of the terminal area productivity program. NS301R3, Logistics Management Institute, VA.

Li L and Clarke J-P (2010). A stochastic model of runway configuration planning. AIAA Guidance, Navigation and Control Conference, Toronto, Canada.

Liu P B, Hansen M and Mukherjee A (2008). Scenario-based air traffic flow management: From theory to practice. Transportation Research Part B, 42(7-8):685-702.

Long D, Lee D, Johnson J, Gaier E and Kostiuk P (1999). Modeling Air Traffic Management Technologies with a Queuing Network Model of the National Airspace System. Technical report NASA/CR-1999-208988, NASA Langley Research Center, Hampton, VA.

Long D and Hasan S (2009). Improved Prediction of Flight Delays Using the LMINET2 System-Wide Simulation Model. AIAA Aviation Technology, Integration, and Operations Conference (ATIO).

Lovell D J, Churchill A, Odoni A R, Mukherjee A and Ball M O (2007). Calibrating Aggregate Models of Flight Delays and Cancellation Probabilities at Individual Airports. Proc. 7th USA/Europe Air Traffic Management R&D Seminar, Barcelona, Spain.

Lulli, G and Odoni A R (2007). The European Air Traffic Flow Management Problem. Transportation Science, 41(4):431-443.

Madas M A and Zografos K G (2006). Airport slot allocation: From instruments to strategies. Journal of Air Transport Management, 12(2):53-62.

Malone K M (1995). Dynamic Queuing Systems: Behavior and Approximations for Individual Queues and for Networks. PhD thesis, Operations Research Center, Massachusetts Institute of Technology, Cambridge, MA.

Mercer A (1960). A Queueing Problem in which the Arrival Times of the Customers are Scheduled. Journal of the Royal Statistical Society, Series B (Methodological), 22(1):108-113.

Mukherjee A and Hansen M (2007). A Dynamic Stochastic Model for the Single Airport Ground Holding Problem. Transportation Science, 41(4): 444-456.

Mukherjee A, Lovell D J, Ball M O, Odoni A R and Zerbib G (2005). Modeling delays and cancellation probabilities to support strategic simulations. Proc. 6th USA/Europe Air Traffic Management R&D Seminar, Baltimore, Maryland.

Murca M C R and Muller C (2015). Control-based optimization approach for aircraft scheduling in a terminal area with alternative arrival routes. Transportation Research Part E, 73:96-113.

de Neufville R and Odoni A R (2013). Airport Systems: Planning, Design and Management. McGraw-Hill, 2nd edition, New York.

Newell G F (1979). Airport Capacity and Delays. Transportation Science, 13(3):201-241.

Nikoleris T and Hansen M (2012). Queueing Models for Trajectory-Based Aircraft Operations. Transportation Science, 46(4):501-511.

Nikoleris T and Hansen M (2015). Effect of Trajectory Prediction and Stochastic Runway Occupancy Times on Aircraft Delays. Transportation Science, 50(1):110-119.

Odoni A R, Morisset T, Drotleff W and Zock A (2011). Benchmarking Airport Airside Performance: FRA vs. EWR. Proc. 9th USA/Europe Air Traffic Management R&D Seminar, Berlin, Germany.

Pellegrini P, Castelli L and Pesenti R (2012). Secondary trading of airport slots as a combinatorial exchange. Transportation Research Part E, 48(5):1009-1022.

Pels E and Verhoef E T (2004). The economics of airport congestion pricing. Journal of Urban Economics, 55(2):257-277.

Peterson M D, Bertsimas D J and Odoni A R (1995). Models and Algorithms for Transient Queueing Congestion at Airports. Management Science, 41(8):1279-1295.

Potts C N, Mesgarpour M and Bennell J A (2009). A Review Of Airport Runway Optimization. University of Southampton, UK.

Psaraftis H N (1978). A Dynamic Programming Approach to the Aircraft Sequencing Problem. Technical report R78-4, MIT Flight Transportation Laboratory.

Pyrgiotis N (2011). A Stochastic and Dynamic Model of Delay Propagation Within an Airport Network for Policy Analysis. PhD thesis, Dept. of Aeronautics and Astronautics, Massachusetts Institute of Technology, Cambridge, MA.

Pyrgiotis N, Malone K M and Odoni A R (2013). Modelling delay propagation within an airport network. Transportation Research Part C, 27:60-75.

Pyrgiotis N and Odoni A R (2016). On the Impact of Scheduling Limits: A Case Study at Newark Liberty International Airport. Transportation Science, 50(1):150-165.

Ramanujam V and Balakrishnan H (2015). Data-Driven Modeling of the Airport Configuration Selection Process. IEEE Transactions on Human-Machine Systems, 45(4):490-499.

Richetta O (1995). Optimal Algorithms and a Remarkably Efficient Heuristic for the Ground-Holding Problem in Air Traffic Control. Operations Research, 43(5):758-770.

Richetta O and Odoni A R (1993). Solving Optimally the Static Ground-Holding Policy Problem in Air Traffic Control. Transportation Science, 27(3):228-238.

Roth E (1979). An Advanced Time-Dependent Queueing Model for Airport Delay Analysis. Technical report R-79-9, MIT Flight Transportation Laboratory.

Rue R C and Rosenshine M (1985). The Application of Semi-Markov Decision Processes to Queueing of Aircraft for Landing at an Airport. Transportation Science, 19(2):154-172.

Schwarz J A, Selinka G and Stolletz R (2016). Performance analysis of time-dependent queueing systems: Survey and classification. Omega, 63:170-189.

Simaiakis I (2013). Analysis, modeling and control of the airport departure process. PhD thesis, Massachusetts Institute of Technology, Cambridge, MA.

Solak S, Solveling G, Clarke J-P and Johnson E L (2018). Stochastic Runway Scheduling. Transportation Science, Articles in Advance, pp 1-24.

Solveling G, Solak S, Clarke J-P and Johnson E L (2011). Runway operations optimization in the presence of uncertainties. Journal of Guidance, Control and Dynamics, 34(5):1373-1382.

Stamatopoulos M A, Zografos K G and Odoni A R (2004). A decision support system for airport strategic planning. Transportation Research Part C, 12(2):91-117.

Stolletz R (2008). Non-stationary delay analysis of runway systems. OR Spectrum, 30(1):191-213.

Swaroop P, Zou B, Ball M O and Hansen M (2012). Do more US airports need slot controls? A welfare based approach to determine slot levels. Transportation Research Part B, 46(9):1239-1259.

Terrab M and Odoni A R (1992). Strategic Flow Management for Air Traffic Control. Operations Research, 41(1):138-152.

Tu Y, Ball M O and Jank W S (2008). Estimating Flight Departure Delay Distributions – A Statistical Approach With Long-Term Trend and Short-Term Pattern. Journal of the American Statistical Association, 103(481):112-125.

Vaze V and Barnhart C (2012). An assessment of the impact of demand management strategies for efficient allocation of airport capacity. International Journal of Revenue Management, 6(1-2):5-27.

Vranas P B, Bertsimas D J and Odoni A R (1994). The Multi-Airport Ground-Holding Problem in Air Traffic Control. Operations Research, 42(2):249-261.

Weld C, Duarte M and Kincaid R (2010). A runway configuration management model with marginally decreasing transition capacities. Advances in Operations Research, http://dx.doi.org/10.1155/2010/436765.

Willemain T R, Fan H and Ma H (2004). Statistical analysis of intervals between projected airport arrivals. DSES Technical Report 38-04-510, Rensselaer Polytechnic Institute, Troy, New York.

Winsten C B (1959). Geometric Distributions in the Theory of Queues. Journal of the Royal Statistical Society, Series B (Methodological), 21(1):1-35.

Zografos K G, Salouras Y and Madas M A (2012). Dealing with the efficient allocation of scarce resources at congested airports. Transportation Research Part C, 21(1):244-256.

Zografos K G, Madas M A and Androutsopoulos K N (2017). Increasing airport capacity utilisation through optimum slot scheduling: review of current developments and identification of future needs. Journal of Scheduling, 20(1):3-24.

OR60 Annual Conference - Keynote Papers and Extended Abstracts
Lancaster University. Lancaster. 11 - 13 September 2018

Data as an Artefact in Supply Chain Literature: A Conceptual Framing

Thanos Papadopoulos [a], Konstantina Spanaki [b]

[a] Kent Business School, University of Kent, Canterbury, UK
[b] School of Business and Economics, Loughborough University, Leicestershire, UK
A.Papadopoulos@kent.ac.uk , K.Spanaki@lboro.ac.uk

Abstract

During the last decade, studies in the context of data management within operations have mostly focused on 'Supply Chain Analytics (SCA)'. SCA is presented as a way to develop supply chain strategies and efficiently managing supply chain operations at tactical and operational levels with the use of data analytics. While the role of data within SCM literature has been highlighted, research should also focus on data as an artefact, setting data as a raw material/resource triggering processes, analogous to the product/service manufacturing processes. Our study reviews and discusses these two approaches, namely: a) the use of data for supply chains and b) the supply chains of data, while trying to highlight their conceptual differences. The implications and different approaches for these two concepts are analysed and further research directions are proposed.

Keywords: Big Data, Data artefact, Data resources, Data flows, Conceptual framework

1. Introduction

Nowadays, there is an ongoing discussion around Big Data amongst academics and practitioners because of the multiple opportunities and challenges this data evolution can bring for the individuals and organizations (Hazen et al., 2016; Sivarajah et al., 2017; Spanaki et al., 2017; Wamba et al., 2015; Wang et al., 2016). The initial conceptualisation of the 'Big Data' as unprecedented volumes of data, beyond the capacity of extant tools and analytics to manage (Manyika et al., 2011), has been superseded by more comprehensive views (Boyd and Crawford, 2012). Boyd and Crawford (2012, p.663) presented the concept of 'Big Data' as a cultural, technological, and scholarly phenomenon that rests on the interplay of technology, analysis, and mythology with the aura of truth, objectivity, and accuracy. However, many of the conceptualisations are still based on the V-framework of Laney (2001) and propose three, four (Jin et al., 2016; Lycett, 2013), five (Erevelles et al., 2016; Opresnik and Taisch, 2015) or even more Vs for the description of 'Big Data'(Wang et al., 2016). In addition to Volume, these conceptualizations include: Variety, Velocity, Veracity and Value. Due to its disruptive and transformatory potential (Baines et al., 2017; Spanaki et al., 2017; Yoo et al., 2012), the data revolution offers new challenges – such as developing new organizational capabilities (Davenport et al., 2012; McAfee et al., 2012) - and opportunities for organizations to innovate, create strategic advantage and generate new business value from the data (Gandomi and Haider, 2015). As an organizational resource, data has grown in importance and the effective management of data streams, originating from the activities inside and outside the organizations is imperative, data manufacturing and processing could provide new streams of value for organizations as well as the development of novel 'data products'(Karafili et al., 2018; Spanaki et al., 2017). This is because as data flows from inbound and outbound activities, can provide a competitive advantage for the organizations to create value through the use of these data streams and transform the data in 'valuable and innovative outcomes' (Spanaki et al., 2017). This is evident in areas such as manufacturing where disruption is triggered mainly by the increasing role of data and technology in the production processes. Cyber-physical systems, Internet of things (IoT), cloud computing, analytics as well as data exchange between actors and processes have transformed in addition to traditional organizational structures the manufacturing landscape.

We propose here a discussion and a conceptualization of the *data as an artefact*, while identifying the role of data within Supply Chain and Operations Management literature. Data as an *artefact* is identified in the literature with a dual focus: *data about the Supply Chain* and *data moving through the Supply Chain* - this double focus will be further explained and discussed in this study.

2. Background

Data as an artefact is described usually as a source of competitive advantage exploited through Analytics for knowledge creation (Agarwal and Dhar, 2014). This direction in the literature suggests that there will be a radical Supply Chain transformation through the use of data (Feki et al., 2016; Wang et al., 2016). The transformation of Supply Chains, operations and processes, aims at exploiting insights from the data to be used in decision making so as value can be realized (Sharma et al., 2014). The value creation can develop knowledge through the use and evaluation of data (Agarwal and Dhar, 2014), but also through cross-boundary industry disruptions, as new forms of business strategy (Bharadwaj et al., 2013). Additionally, knowledge can be created through information strategy transformation (Constantiou and Kallinikos, 2015; Gandomi and Haider, 2015), while reshaping fundamentally traditional business strategy models (Bharadwaj et al., 2013) or through business model innovation (Ferrando-Llopis et al., 2013). Transformation of business strategies (Digital Business Strategy) focuses on the scope, scale, speed of value creation (Bharadwaj et al., 2013) as well as innovative work patterns designed across boundaries of time, distance and function and propose a new collaborative regime–ecosystem approach. (Ferrando-Llopis et al., 2013).

2.1. Data about and across the Supply Chain

There is plenty of evidence that companies from multiple industries invest on analytics and data solutions, with the view to improve their performance and their business processes (Feki et al., 2016; Matthias et al., 2017; Papadopoulos et al., 2016). Waller and Fawcett (2013) proposed the use of data in supply chain management (SCM) to improve and expand the production of enterprises by using their analytical skills in viewing and optimizing their supply chains of their core business. Although data can be used along with the core business focus in different industries, we can observe that the recent data evolution expanded the business scope and disrupted the operating models providing opportunities to work solely on data as the main “raw material”, processing this data, creating new products and services and also reselling, exchanging data and transforming the organizational and operational landscape (Spanaki et al., 2017).

Reviews of the literature in the context of data management within logistics and SCM mostly focus on “Supply Chain Analytics (SCA)” (Wang et al., 2016) as a way in developing supply chain strategies and efficiently managing supply chain operations at tactical and operational levels (Wang et al., 2016) with the use of data analytics functions. This literature direction focuses on how the analytics can be applied to strategic decisions related to SCM (Wamba et al., 2015; Wang et al., 2016), how efficiency and effectiveness of supply chains can be improved through the use of data (Waller and Fawcett, 2013) as well as data oriented strategies used for value creation in product-service systems fields for tailored services/products (Song, 2017) and innovative servitization models (Baines et al., 2017; Opresnik and Taisch, 2015).

In the review of Cerchione and Esposito (2016) supply chain is a multi-objective system (economic, productive, strategic, environmental, social, etc.) crossed by a variety of flows (financial, material, information, technology, etc.). In such a complex network, the management of the processes of adoption, creation, storage, transfer, sharing and application of knowledge, information and data appears to be the necessary response to the new challenges posed by globalization and sustainability issues (Cerchione and Esposito, 2016; Esposito et al., 2015), and therefore, the potential of data and its role is highlighted in capturing business value (Wamba et al., 2015). Business value is realised in the sense of enhanced decision-making through valuable information extraction, providing new revenue streams for competitive advantage (Sivarajah et al., 2017). Data resources as revenue streams can support the continuously evolving manufacturing from concept development to methods and tools available for the production of goods for use or sale (Esmaeilian et al., 2016). Supply chain analytics (SCA) and specifically data can provide unique insights into, inter alia, market trends, customer buying patterns, and maintenance cycles, as well as into ways of lowering costs and enabling more targeted business decisions (Wang et al., 2016) and therefore transform traditional manufacturing supply chain patters to follow the future requirements (Esmaeilian et al., 2016). Analytics involves issues quite aside from data management, number crunching, technology use, systematic reasoning, and so forth (Holsapple et al., 2014), thereby including smart manufacturing practices based on IoT technology (Addo-Tenkorang and Helo, 2016), cloud manufacturing (Jede and Teuteberg, 2016) and collaborative supply chains (Spekman and Davis, 2016), as well as smart supply chains (Wu et al., 2016).

The cloud computing services available for various applications within SCM for example, may strengthen competitive advantage in a collaborative pattern of work (Jede and Teuteberg, 2016). Extended enterprise across organizational boundaries and collaborative supply chains depict the future state of supply chain management and shape the way in which organizations interact with their customers and suppliers (Spekman and Davis, 2016). As traditional supply chains are increasingly becoming collaborative (Melnyk et al., 2010; Spekman and Davis, 2016) and intelligent with more objects embedded with sensors and better communication, intelligent decision making and automation capabilities, smart supply chains present unprecedented opportunities for achieving cost reduction and enhancing efficiency improvement (Wu et al., 2016).

Data strategies and their impact in manufacturing and operational areas still remain a fragmented, early-stage domain of research in terms of theoretical grounding, methodological diversity and empirically oriented work (Frizzo-Barker et al., 2016). Data resources should be explored thoroughly in order to enable industrial managers and businesses executives to make pre-informed strategic operational and management decisions for increased return-on-investment (ROI) and it could also empower organizations with a value-adding stream of information to have a competitive edge over their competitors (Addo-Tenkorang and Helo, 2016). The current direction of SCM reflects that the research should analyse mostly the use of data in supply chains; nevertheless we posit that more value lies on the supply chains of data, setting data as a raw material/resource triggering a manufacturing process, analogous to the product/service manufacturing processes.

2.2. Data as a resource- Opportunities and Challenges

There are four distinct areas where data resources can be identified within operations and SCM:

1. Data utilized in Supply Chains for improving core supply chain processes (while increasing firm’s performance- combining internal and external data resources and providing tailored, customized and innovative products/services).

2. Data generated, captured, collected, and stored by smart supply chains (in this occasion, data resources provide value through the Operational and Business Model Innovation they create but also through the sale of the products/services as outputs of the smart supply chain).

3. Data generated and captured by smart objects (in this occasion, data resources provide value through the Operational and Business Model Innovation).

4. Data processed as a main supply chain, as a core business process of the firm (creating value through Operational and Business Model Innovation as well as by selling data/ information products/ services as the output of the supply chain).

In the first occasion, we have the traditional manufacturing supply chain, which is enhancing its performance with the outputs (information products/services) of data flows which are triggered in parallel. This kind of data flow is developed in parallel with the core supply chains and provides/ feeds the basic operations of the firm with data resources generated from internal transactions -but also in combination with external data. The outputs in these cases are tailored, customized and innovative products/services improving the firm's ROI, competitive posture and performance.

In the second occasion, we have data flows triggered by a smart object/product, which is a physical entity triggering data flows and ending up in providing information/data services as outputs. Here we have initiatives as Amazon Echo, Google NEST or Fitbit trackers, where the main product is a physical smart object but this product triggers data flows for more intelligent, smart, data-oriented information services. In this occasion, data resources provide value through the Operational and Business Model Innovation.

In the third occasion, we have a smart supply chain which generates, captures, collects and stores data from smart objects, IoT technology, RFID-enabled processes and therefore there the smart supply chain (which includes the data flows as subcategory along with physical, traditional manufacturing processes happening to the material flows in the smart supply chain). Data flows in this category are developed in parallel with the smart supply chain. The outputs are products or services intelligently developed, improving the performance of the supply chain and providing value through the sale of final products and services developed. The higher value-adding part of these cases is the Operational and Business Model Innovation happening by the use of data flows in the smart supply chain context.

In the fourth occasion, data flows are the core supply chain the firm operates (e.g. Experian develops financial and credit information products/services based on multiple data - or Uber which provides information services for travelling around cities, combining people, resources and customers through the use of data). These newly developed supply chains create value through Operational and Business Model Innovation as well as by selling data/ information products/ services as the output of the supply chain.

Data flows across the supply chain have some distinctive characteristics, compared to other supply chain approaches, therefore we can characterise these data flows as: 1) information-oriented, 2) outcome-driven, 3) interconnected, 4) intelligent, 5) tailored, 6) time-dependent, 7) integrated, 8) innovative.

Smart devices/ objects, analytics and data generation and collection techniques, along with collaborative and integrated patterns of activities and processes play a critical role in data flows as value is created through the use, reuse, combination, purchase and sale of data. In fact, data can create new value streams by developing business models, improving business processes, as well as by accelerating firm's performance, competitive advantage and innovative initiatives.

Table 1 Data as a resource – conceptualization of data flows

Concept	Description
Information-oriented	Information and knowledge is constantly generated by the flows of data, since the data raw material starts as an input and triggers the activities of the supply chain till final output – the information products/services.
Outcome-driven	Data flows can deliver varying degrees of outcomes — the traditional cost-related benefit plus responsiveness, security, sustainability, resilience and innovation based on key customers' needs.
Interconnected	The entire supply chain consisted of raw materials, organizational entities and relationships, analytics/tools, processes, systems, products and services are all connected forming flows of data.
Intelligent	Data flows support optimization, visualization of decision-making and analytic functions.
Tailored	There is a high degree of customization based on the demand of customers, and therefore the whole strategy can be formed around the customer's requirements (what will be the ideal combination of data, how and from where to be collected/ bought, with whom to collaborate/co-create value etc.)
Time-dependent	Data raw materials are constantly generated and collected, and therefore time-dependent (real time data but also archival data - depending on the combination of data resources, their use and the industry sector)
Integrated	Data flows are processed through integrated supply chains that involve collaboration across supply chain stages, decision making based on data, interconnected systems, and information sharing, using multiple (internal and external) data sources across organizational boundaries
Innovative	Innovation through ideas, practices, and business models; value is not solely created from an information product/ service but also through the disruption of the existing business and operational models.

3. Conceptualization and Way forward

The conceptualization of *data as an artefact* in the context of SCM warrants further investigation in fields related to the dimensions of this supply chain and what is the novelty around this concept. Following this direction and expanding the conceptualization, we propose here a few questions that can provide further insight about the aspects and the dimensions related to data artefact in supply chain context. The questions are based on aspects and attributes examined in previous studies of supply chain field (e.g. Cooper et al., 1997; Storey et al., 2006) and can provide a basis for comparative approaches/purposes in the research of data as an artefact when we study it in the context of supply chain, namely:

1. *What is the content and focus of the supply chain?*

The element of exchange of the supply chain

2. *What is the strategy used for this supply chain (Strategic focus)?*

The goals each supply chain has to fulfil – the motivation shaping this specific supply chain

3. *What is the role of knowledge and information for this supply chain?*

The role of information and knowledge, as well as the way each supply chain extracts and applies knowledge and manages information (e.g. to develop new products, to share information with peers etc.)

4. *What is the integration level of this supply chain?*

The relationships and collaborative pattern maintained by each supply chain internally and externally (e.g. collaboration with supply chain members, competition relationships, work pattern across boundaries etc.)

5. *What tools/ methods are used for this supply chain?*

The technological context applied for each supply chain (e.g. analytics techniques and methods, data extraction tools etc.)

6. *What manufacturing approach is used?*

The method the supply chain's manufacturing system follows (e.g. sustainability approaches for manufacturing processes, agility through the use of data etc.)

7. *What is the customization level used for this supply chain?*

The possibility of the output to be customized (e.g. tailored products to meet customer requirements, customization according to customer's data patterns etc.)

The previous questions along with their description are developed as an example research agenda for this topic. This research agenda could be explored further for identifying the role of data as an artefact in the supply chain context and can provide an initial framing for associated areas.

References

Addo-Tenkorang, R., Helo, P.T., 2016. Big data applications in operations/supply-chain management: A literature review. Comput. Ind. Eng. 101, 528–543. https://doi.org/10.1016/j.cie.2016.09.023

Agarwal, R., Dhar, V., 2014. Editorial—Big Data, Data Science, and Analytics: The Opportunity and Challenge for IS Research. Inf. Syst. Res. 25, 443–448.

Baines, T., Ziaee Bigdeli, A., Bustinza, O.F., Shi, V.G., Baldwin, J., Ridgway, K., 2017. Servitization: revisiting the state-of-the-art and research priorities. Int. J. Oper. Prod. Manag. 37, 256–278. https://doi.org/10.1108/IJOPM-06-2015-0312

Bharadwaj, A., Sawy, O.A. El, Pavlou, P.A., Venkatraman, N., 2013. Digital business strategy: toward a next generation of insights. MIS Q. 37, 471–482.

Cerchione, R., Esposito, E., 2016. A systematic review of supply chain knowledge management research: State of the art and research opportunities. Int. J. Prod. Econ. 182, 276–292. https://doi.org/10.1016/j.ijpe.2016.09.006

Constantiou, I.D., Kallinikos, J., 2015. New games, new rules: big data and the changing context of strategy. J. Inf. Technol. 30, 44–57.

Cooper, M.C., Lambert, D.M., Pagh, J.D., 1997. Supply chain management: more than a new name for logistics. Int. J. Logist. Manag. 8, 1–14.

danah boyd, Crawford, K., 2012. Critical Questions for Big Data. Information, Commun. Soc. 15, 662–679. https://doi.org/10.1080/1369118X.2012.678878

Davenport, T.H., BARTH, P., BEAN, R., 2012. How “Big Data” Is Different. (cover story). MIT Sloan Manag. Rev. 54, 43.

Erevelles, S., Fukawa, N., Swayne, L., 2016. Big Data consumer analytics and the transformation of marketing. J. Bus. Res. 69, 897–904. https://doi.org/http://dx.doi.org/10.1016/j.jbusres.2015.07.001

Esmaeilian, B., Behdad, S., Wang, B., 2016. The evolution and future of manufacturing: A review. J. Manuf. Syst. 39, 79–100. https://doi.org/10.1016/j.jmsy.2016.03.001

Esposito, C., Ficco, M., Palmieri, F., Castiglione, A., 2015. A knowledge-based platform for Big Data analytics based on publish/subscribe services and stream processing. Knowledge-Based Syst. 79, 3–17. https://doi.org/http://dx.doi.org/10.1016/j.knosys.2014.05.003

Fawcett, S.E., Waller, M.A., 2013. Considering supply chain management’s professional identity: The beautiful discipline (Or, “we don’t cure cancer, but we do make a big difference”). J. Bus. Logist. 34, 183–188. https://doi.org/10.1111/jbl.12018

Feki, M., Boughzala, I., Wamba, S.F., 2016. Big Data Analytics-Enabled Supply Chain Transformation: A Literature Review, in: Proceedings of the 2016 49th Hawaii International Conference on System Sciences (HICSS). IEEE Computer Society, pp. 1123–1132.

Ferrando-Llopis, R., Lopez-Berzosa, D., Mulligan, C., 2013. Advancing value creation and value capture in data-intensive contexts. IEEE, pp. 5–9.

Frizzo-Barker, J., Chow-White, P.A., Mozafari, M., Ha, D., 2016. An empirical study of the rise of big data in business scholarship. Int. J. Inf. Manage. 36, 403–413. https://doi.org/10.1016/j.ijinfomgt.2016.01.006

Gandomi, A., Haider, M., 2015. Beyond the hype: Big data concepts, methods, and analytics. Int. J. Inf. Manage. 35, 137–144. https://doi.org/http://dx.doi.org/10.1016/j.ijinfomgt.2014.10.007

Hazen, B.T., Skipper, J.B., Ezell, J.D., Boone, C.A., 2016. Big data and predictive analytics for supply chain sustainability: A theory-driven research agenda. Comput. Ind. Eng. 101, 592–598. https://doi.org/http://dx.doi.org/10.1016/j.cie.2016.06.030

Holsapple, C., Lee-Post, A., Pakath, R., 2014. A unified foundation for business analytics. Decis. Support Syst. 64, 130–141. https://doi.org/10.1016/j.dss.2014.05.013

Jede, A., Teuteberg, F., 2016. Towards cloud-based supply chain processes: Designing a reference model and elements of a research agenda. Int. J. Logist. Manag. 27, 438–462. https://doi.org/10.1108/IJLM-09-2014-0139

Jin, J., Liu, Y., Ji, P., Liu, H., 2016. Understanding big consumer opinion data for market-driven product design. Int. J. Prod. Res. 54, 3019–3041. https://doi.org/10.1080/00207543.2016.1154208

Karafili, E., Spanaki, K., Lupu, E.C., 2018. An argumentation reasoning approach for data processing. Comput. Ind. 94, 52–61. https://doi.org/10.1016/j.compind.2017.09.002

Laney, D., 2001. 3d data management: Controlling data volume, velocity and variety. Gartner.

Lycett, M., 2013. “Datafication”: making sense of (big) data in a complex world. Eur. J. Inf. Syst. 22, 381–386. https://doi.org/10.1057/ejis.2013.10

Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., Byers, A.H., 2011. Big data: The next frontier for innovation, competition. Tech. report, McKinsey Glob. Inst.

Matthias, O., Fouweather, I., Gregory, I., Vernon, A., 2017. Making sense of Big Data – can it transform operations management? Int Jrnl Op Prod Mnagemnt 37, 37–55. https://doi.org/10.1108/IJOPM-02-2015-0084

McAfee, A., Brynjolfsson, E., Davenport, T.H., Patil, D.J., Barton, D., 2012. Big Data: The management revolution. Harv. Bus. Rev. 90, 61–88.

Melnyk, S.A., Davis, E.W., Spekman, R.E., Sandor, J., 2010. Outcome-driven supply chains. MIT Sloan Manag. Rev. 51, 33.

Opresnik, D., Taisch, M., 2015. The value of big data in servitization. Int. J. Prod. Econ. 165, 174–184. https://doi.org/10.1016/j.ijpe.2014.12.036

Papadopoulos, T., Gunasekaran, A., Dubey, R., Altay, N., Childe, S.J., Fosso-Wamba, S., 2016. The role of Big Data in explaining disaster resilience in supply chains for sustainability. J. Clean. Prod. https://doi.org/10.1016/j.jclepro.2016.03.059

Sharma, R., Mithas, S., Kankanhalli, A., 2014. Transforming decision-making processes: a research agenda for understanding the impact of business analytics on organisations. Eur. J. Inf. Syst. 23, 433–441.

Sivarajah, U., Kamal, M.M., Irani, Z., Weerakkody, V., 2017. Critical analysis of Big Data challenges and analytical methods. J. Bus. Res. 70, 263–286. https://doi.org/10.1016/j.jbusres.2016.08.001

Song, W., 2017. Requirement management for product-service systems: Status review and future trends. Comput. Ind. 85, 11–22. https://doi.org/10.1016/j.compind.2016.11.005

Spanaki, K., Gürgüç, Z., Adams, R., Mulligan, C., 2017. Data supply chain (DSC): research synthesis and future directions . Int. J. Prod. Res. https://doi.org/10.1080/00207543.2017.1399222

Spekman, R., Davis, E.W., 2016. The extended enterprise: a decade later. Int. J. Phys. Distrib. Logist. Manag. 46, 43–61. https://doi.org/10.1108/IJPDLM-07-2015-0164

Storey, J., Emberson, C., Godsell, J., Harrison, A., 2006. Supply chain management: theory, practice and future challenges. Int. J. Oper. Prod. Manag. 26, 754–774.

Waller, M.A., Fawcett, S.E., 2013. Data science, predictive analytics, and big data: A revolution that will transform supply chain design and management. J. Bus. Logist. 34, 77–84. https://doi.org/10.1111/jbl.12010

Wamba, S.F., Akter, S., Edwards, A., Chopin, G., Gnanzou, D., 2015. How 'big data' can make big impact: Findings from a systematic review and a longitudinal case study. Int. J. Prod. Econ. 165, 234–246. https://doi.org/10.1016/j.ijpe.2014.12.031

Wang, G., Gunasekaran, A., Ngai, E.W.T., Papadopoulos, T., 2016. Big data analytics in logistics and supply chain management: Certain investigations for research and applications. Int. J. Prod. Econ. 176, 98–110. https://doi.org/10.1016/j.ijpe.2016.03.014

Wu, L., Yue, X., Jin, A., Yen, D.C., 2016. Smart supply chain management: A review and implications for future research. Int. J. Logist. Manag. 27, 395–417. https://doi.org/10.1108/IJLM-02-2014-0035

Yoo, Y., Jr, R.J.B., Lyytinen, K., Majchrzak, A., 2012. Organizing for innovation in the digitized world. Organ. Sci. 23, 1398–1408.

OR60 Annual Conference - Keynote Papers and Extended Abstracts
Lancaster University. Lancaster. 11-13 September 2018

Two Example Optimisation Problems from the World of Education

Rhyd Lewis

School of Mathematics, Cardiff University
Cardiff CF24 4AG, WALES
LewisR9@cf.ac.uk

ABSTRACT

This work considers two distinct combinatorial optimisation problems related to education, namely lecture timetabling and school bus scheduling, both of which are known to be NP-hard. Our research into these problems has centred around the design of various high-performance heuristics that are able to produce good quality solutions to these problems in short amounts of time. To do this, we propose that it is necessary to "get to the heart" of these problems by identifying their underlying sub-problems. This, in turn, helps to inform the design of algorithmic operators that are able to exploit these structures and help to produce the solutions we need. In this extended abstract these problems are briefly considered in turn.

1 Lecture Timetabling

The first problem we consider is lecture timetabling, which has crossovers with the classical operational research problems of graph colouring and bipartite matching. According to Lewis (2008), university timetabling "can be considered the task of assigning a number of events, such as lectures, exams, meetings, and so on, to a limited set of timeslots (and perhaps rooms) in accordance with a set of constraints". For this research we consider the "post enrolment" timetabling model, where the constraints of the problem are specified by student enrolment data—that is, each student gives a list of the subjects they would like to take and the final timetable must seek to accommodate these choices. Specifically, we use the model proposed in the second international timetabling competition (www.cs.qub.ac.uk/itc2007/), which can be briefly summarised as follows:

The problem involves a number of "hard" constraints whose satisfaction is mandatory, together with three "soft" constraints, whose satisfaction is desirable, but not essential. We are then required to assign the set of events to 45 timeslots (five days, with nine timeslots per-day) according to these constraints. The hard constraints for this problem are as follows. First, for each event there is a set of students who are enrolled to attend—thus events should be assigned to timeslots such that no student is required to attend more than one event in any one timeslot. Next, each event also requires a set of room features (e.g. a certain number of seats, specialist teaching equipment, etc.), which will only be provided by certain rooms; thus each event needs to be assigned to a suitable room that exhibits the room features that it requires. The double booking of rooms is also disallowed. Hard constraints are also imposed stating that some events cannot be taught in certain timeslots. Finally, precedence constraints—stating that some events need to be scheduled before or after others—are also stipulated. The soft constraints for this problem, meanwhile, are as follows: first, students should not be asked to attend an event in the last timeslot of each day (i.e. timeslots 9, 18, 27, 36, or 45); second, students should not have to attend events in three or more successive timeslots occurring in the same day; third, students should not be required to attend just one event in a day.

Arguably, most of the current best-performing algorithms for this problem have been based on neighbourhood search methods (Lewis and Thompson 2015). To design such a method, three ingredients are needed: a *solution space*, which defines the set of all possible candidate solutions; a *cost function*, which measures the quality of each member of the solution space; and a *neighbourhood operator*, which allows us to move from one solution to another in the solution space. A neighbourhood search algorithm operates by using the neighbourhood operator (or operators) to move around the solution space, seeking the solution with the smallest cost. Examples of such algorithms applied to this particular problem include simulated annealing (Cambazard, Hebrard, O'Sullivan, and Papadopoulos 2012), tabu search (Lewis and Thompson 2015), ant colony optimisation (Nothegger, Mayer, Chwatal, and Raidl 2012) and evolutionary algorithms (Jat and Yang 2011).

Many of the most successful algorithms for this problem have also operated using a two-staged approach. Here, the soft constraints are effectively ignored in the first stage, and the algorithm only seeks a feasible solution (i.e., one satisfying all of the hard constraints). Once this is achieved, the second stage can then be invoked, which will seek to make changes to a solution in order to reduce the number of soft constraint violations, while not re-violating any of the hard constraints. Note in particular that many potential changes to a feasible solution (e.g. moving an event from one timeslot to another) will indeed cause a hard constraint to be violated; hence, movements in the solution space can often be very restricted because only a small number of paths, if any, can be taken from a particular solution.

To look at this issue more closely, consider an algorithm whose solution space comprises all candidate solutions, feasible and infeasible, for a particular instance of this timetabling problem. Now consider the graph $G = (V, E)$ where each vertex $v \in V$ represents a different member of the solution space, and an edge between two vertices, denoted by $(u, v) \in E$, indicates the existence of a move under our neighbourhood operator that, when applied to solution u, will result in solution v. As noted, in Stage 2 of the two-staged approach infeasible solutions are not allowed, so our solution space now becomes the graph induced by V', where $V' \subseteq V$ is the set of all feasible solutions. This new graph will clearly contain far fewer edges than G; indeed, it may also be disconnected, meaning that if all of the "good" solutions to a particular problem instance happen to be in different components to that of our current solution, we will never be able to reach them. It is therefore in our interests to try to increase the connectivity of this solution space. In this research, this is achieved using two ideas that exploit the underlying structures of this problem. This allows a greater proportion of proposed neighbourhood moves to maintain solution feasibility and facilitates less restricted movements in the solution space. These are now outlined.

Maximum Matching: As noted above, an important constraint with this problem is ensuring that events are assigned to suitable rooms. Given a set of events assigned to a particular timeslot, the existence of a suitable room allocation can actually be determined in polynomial time by solving a maximum bipartite matching problem. If a neighbourhood move seeks to insert a new event into a particular timeslot, we can therefore use this property to determine whether the room allocations of this timeslot can be appropriately rearranged to allow this extra event to be accommodated.

Graph Colouring: In graph colouring, a Kempe chain is a connected sub-graph containing vertices of just two colours. It is also known that, given a proper colouring, the interchanging of vertex colours in a Kempe chain will result in a new proper colouring with different colour assignments (Lewis 2016). For this timetabling problem, the ideas behind Kempe chains can be used to achieve additional movements in the solution space that are not possible under more simple operators. More specifically, if the movement of an event from timeslot A to timeslot B were to cause a solution to become infeasible (e.g., by requiring a student to be in two places at once), Kempe chains might allow us to overcome this by identifying some events in timeslot B that, when also moved to timeslot A, will maintain solution feasibility.

Our research has clearly indicated that when the connectivity of the underlying solution space is increased through the application of these operators, the quality of the returned solutions in also enhanced. We have also looked at some other related techniques, such as using additional "dummy rooms" to help aid the search, though the results of these are much more mixed. Further details are reported by Lewis and Thompson (2015).

2 School Bus Scheduling

The second problem we consider in this paper concerns the design of school transport schedules (Lewis, Smith-Miles, and Phillips 2018). This involves compiling a list of eligible school children and then organising their transport to and from school; in particular, it requires the selection of a suitable set of pick-up points (bus stops), the assignment of students to these bus stops, and the design of bus routes that visit these stops while getting students to school on time. In doing this, a number of factors should be considered:

1. We should try to reduce economic costs by minimising the number of vehicles used;
2. Each student should be picked up from a bus stop within walking distance of their home address;
3. We cannot assign more students to a bus than there are available seats on the bus;
4. Journey times of students should not exceed a specified limit (e.g., one hour).

School bus routing problems belong to a wider family of vehicle routing problems (VRPs), which involve identifying routes for a fleet of vehicles that are to serve a set of customers. Traditionally, VRPs can be expressed using an edge-weighted directed graph $G = (V,E)$, where the vertex set $V = \{v_0, v_1, \ldots, v_n\}$ represents a single depot and n customers (v_0 and $v_1, \ldots, v_n$ respectively), and the weighting function $d(u,v)$ gives the travel distance (or time) between each pair of vertices $u,v \in V$. As we might expect for a such an important and pervasive problem, VRP formulations come in very many guises and can include many additional features such time-windows, limitations on routes lengths, ensuring suitable vehicles for each route, partitioning customers into pick-up and delivery locations, and the dynamic recalibration of routes subject to the arrival of new customer requests during the transportation period (Laporte 2009, Pillac, Gendreau, Guéret, and Medagila 2013).

One of the main differences between this school bus routing problem and more "traditional" VRPs is that, when planning school bus routes, it is not always clear which set of stopping points we should actually use. In general, a schoolchild may live within walking distance of a number of bus stops, and it makes sense to choose the stop that suitably balances the child's walking distance with the resultant time added to the bus's journey. It may also be advantageous to assign larger groups of students to the same bus stop in order to reduce the number of times that a bus needs to slow down, pick up children, and then rejoin the traffic stream. This problem therefore involves the additional complication of choosing a feasible subset of bus stops from the set of all available bus stops. A VRP along the more traditional lines mentioned above might then be considered using this subset.

The task of choosing a feasible subset of bus stops can be viewed as a set covering problem. Set covering involves taking a set of integers $U = \{1, 2, ..., n\}$ known as the "universe" together with a set S whose elements are subsets of the universe. The task is to then identify a subset $S' \subseteq S$ whose union equals the universe. With regards to the bus scheduling problem, S corresponds to the set whose elements are the addresses within walking distance of each bus stop. A feasible subset

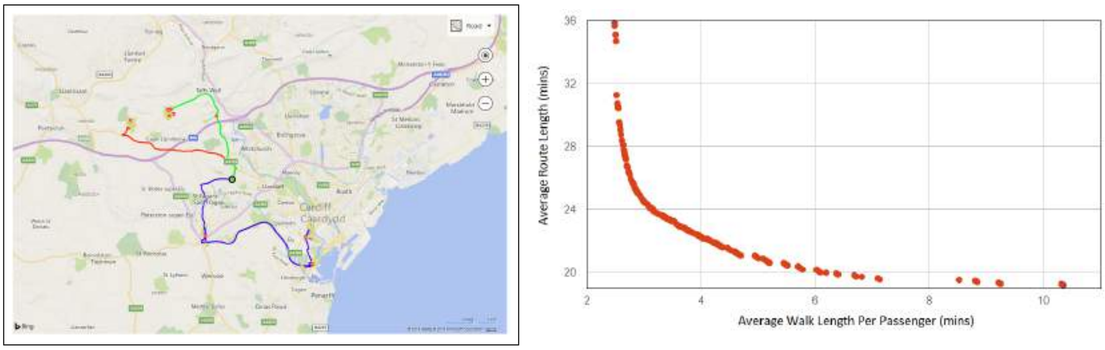


Figure 1: An example solution using three buses for a school in Cardiff (left). The chart on the right shows an approximate Pareto front for the same problem.

of bus stops (i.e., one that features at least one bus stop within walking distance of each address) then corresponds to a subset $S' \subseteq S$ that covers all addresses.

In our experience, the main consideration of school transport planners is to minimise the number of vehicles used, which has the potential to offer significant financial savings. Our strategy for this problem involves using a fixed number of routes k and searching through the solution space defined by all solutions obeying Constraints 2 and 3 above. We then apply specialised VRP-based operators to try and shorten journey times, therefore also hopefully satisfying Constraint 4. If this cannot be achieved at a certain computation limit, k is increased by one, and the algorithm is repeated (initially k equals a lower bound, calculated as the total number of students divided by the bus capacity, rounded up to the nearest integer). This process is also interlaced with an operator based on set covering heuristics, which makes alterations to the set of bus stops being used by a solution, while ensuring that Constraints 2 and 3 remain satisfied.

Our experiments with this algorithm have involved both artificial and real-world problem instances (Lewis, Smith-Miles, and Phillips 2018). In general, we find that larger numbers of vehicles are needed when the maximum journey times are low and/or the maximum walking distances are low. This is quite natural because stricter journey limits imply the need for additional routes in feasible solutions, while shorter walk limits mean that more bus stops need to be visited. For real-world problem instances we have also found that, once a feasible solution using as few routes as possible has been achieved, stakeholders are also interested in making further adjustments in order to make the service both cost efficient and convenient for the users. From the financial point-of-view, administrators are interested in keeping the lengths of each route as short as possible because this will attract lower quotes from the commercial bus companies that want to provide the service. On the other hand, users want to be assigned to bus stops close to their homes, keeping walks short and discouraging parents from driving their children to the bus stop. Clearly these objectives are in conflict, because a bus route that visits many stops (allowing shorter walks) will also tend to be longer and therefore more expensive to run. Decision makers are therefore interested in looking at a set of non-dominating solutions that, for a fixed number of vehicles k, shows how these objectives influence one another, allowing them to choose a solution seen to be an appropriate compromise of the two. This naturally points to the suitability of multiobjective-based optimisation methods. A number of Pareto front approximations for our real world problem instances, generated by our bespoke methods, can be found at www.rhydlewis.eu/sbrp. Examples from this resource are provided in Figure 1.

R. Lewis

REFERENCES

Cambazard, H., E. Hebrard, B. O'Sullivan, and A. Papadopoulos. 2012. "Local search and constraint programming for the post enrolment-based timetabling problem". *Annals of Operational Research* 194:111–135.

Jat, S., and S. Yang. 2011. "A hybrid genetic algorithm and tabu search approach for post enrolment course timetabling". *Journal of Scheduling* 14:617–637.

Laporte, G. 2009. "Fifty Years of Vehicle Routing". *Transportation Science* 43:408–416.

Lewis, R. 2008. "A Survey of Metaheuristic-based Techniques for University Timetabling Problems". *OR Spectrum* 30 (1): 167–190.

Lewis, R. 2016. *A Guide to Graph Colouring: Algorithms and Applications*. Springer International Publishing.

Lewis, R., K. Smith-Miles, and K. Phillips. 2018. "The School Bus Routing Problem: An Analysis and Algorithm". In *Combinatorial Algorithms*, Volume 10765 of *Lecture Notes in Computer Science*, 287–298. Springer-Verlag.

Lewis, R., and J. Thompson. 2015. "Analysing the Effects of Solution Space Connectivity with an Effective Metaheuristic for the Course Timetabling Problem". *European Journal of Operational Research* 240:637–648.

Nothegger, C., A. Mayer, A. Chwatal, and G. Raidl. 2012. "Solving the post enrolment course timetabling problem by ant colony optimization". *Annals of Operational Research* 194:325–339.

Pillac, V., M. Gendreau, C. Guéret, and A. Medagila. 2013. "A Review of Dynamic Vehicle Routing Problems". *European Journal of Operational Research* 225:1–11.

AUTHOR BIOGRAPHY

RHYD LEWIS is a senior lecturer in operational research at the School of Mathematics, Cardiff University, Wales. He holds a PhD in computing and operational research (Edinburgh Napier University, 2006) and is the author of the book *Graph Colouring: Algorithms and Applications* (Springer, 2015). Further publications are listed at his website www.RhydLewis.eu.

THE OPERATIONAL RESEARCH SOCIETY

ISBN: 978-0-903440-64-6