

Testing aircraft fleet management policies using designed simulation experiments

D. O. Marlow^{a,b}, S. M. Sanchez^b and P. J. Sanchez^b

^a Joint and Operations Analysis Division, Defence Science and Technology Group, Fishermans Bend, Victoria, Australia

^b Operations Research Department, Naval Postgraduate School, Monterey, California, USA

Email: david.marlow@dsto.defence.gov.au

Abstract: The Australian Defence Force is in the process of receiving several new capabilities following recent acquisition decisions. In an era of declining government – and thus military – spending, there is an increasing imperative to maximize performance within existing Defence budgets. Therefore there is a desire – indeed obligation – to seek to manage these new aircraft fleets in the most efficient way possible. In practice, this translates into areas such as better maintenance processes and increased operational availability.

In this work we consider a new fleet of 24 MH-60R Seahawk naval combat helicopters being delivered to the Royal Australian Navy. The fleet is required to maintain eight aircraft embarked on ships at sea every day, and to achieve a certain number of ashore flying hours (across two squadrons) and embarked flying hours each year, over a 30 year life. Good fleet management would suggest that not only should these three primary requirements be met, but also the flying hours across the fleet should be somewhat balanced. In that way, no individual aircraft is too far ahead or below the rest of the fleet regarding total flying hours. Aircraft too far ahead would ultimately have to be retired early, reducing the fleet size. Similarly, each aircraft should have a reasonable split of embarked and ashore hours, as aircraft wear more rapidly when embarked due to increased exposure to corrosive salt-laden environments and the higher deployed operational tempo.

A fleet simulation model has been developed to test the capability of the fleet to meet these requirements. The model represents individual aircraft moving between various states, such as between ashore and embarked, and between serviceable (able to fly) and different types of maintenance, from regular short inspections to depot-level maintenance. Unscheduled maintenance is included as a random effect, with both the time between failures and the repair time represented as probability distributions. The model also incorporates various fleet management policies, to see which are most effective in meeting the three main requirements as well as the desire to provide a reasonably balanced fleet. Six policy types are tested: daily flying allocation; daily maintenance allocation; maintenance crew rotations within squadrons; tail rotation between squadrons; balancing the flying hours across the fleet; and sharing squadron resources.

There are a large number of inputs in the model, and while many are fixed, others are more ambiguous, such as the range of flying hours that an aircraft and the fleet may fly each day. When combined with the uncertainty in the values of the unscheduled maintenance distributions, as well as the range of fleet management policies described above, there is a large parameter space that can be explored. Given this, a simulation experimental design approach has been applied to this problem. This allows a thorough exploration of the parameter space in fewer runs than would otherwise be required. We apply this design to eleven continuous variables, two discrete variables, and the six fleet management ‘policies’ listed above. Each of these policies may have between two and five options. Ultimately, 1040 design points are used, with 50 replications at each point – by contrast, a full factorial design would require around 1E33 points.

The results show the strong influence of varying unscheduled maintenance on the ability of the fleet to meet the ongoing requirements, particularly if the frequency is greater than anticipated. It also demonstrates that particular fleet management policies, such as those that share squadron resources and rotate maintenance crews within squadrons, can be used to counteract an increase in unscheduled maintenance. For the requirements to balance the embarked, ashore and total flying hours across the fleet, tail rotation policies that enable more frequent rotations are most effective overall, especially for balancing embarked hours.

Overall, the experimental design technique provides insights to decision makers. In this fleet management example, it identifies the variables and policies of greatest influence. It thus assists decision makers in determining the best policies, and targeting the areas that may negatively impact on meeting requirements.

Keywords: *Aircraft fleet management, simulation, experimental design*

1. INTRODUCTION

This century, the Australian Government has acquired, or decided to acquire, several different new military capabilities for the Australian Defence Force (ADF). These include 72 F-35 Joint Strike Fighters (due to enter service in 2020), 24 F/A-18 Super Hornets (announced in 2006), and 24 MH-60R Seahawk naval combat helicopters (chosen in 2011). Each of these capabilities is expected to operate for 20-30 years.

In this era of high government debt and reduced spending (including military) around the world, the drive for finding greater efficiencies within existing Defence budgets is high. Several years ago the Australian Government released the Strategic Reform Program (Commonwealth of Australia, 2009) with the intent to fund future Defence capability purchases through internal departmental savings of \$20 billion over 10 years. Target areas include “logistics” and “smart sustainment,” with some of the aims of “improving...operational availability” and “improved maintenance processes.” Thus the need, indeed imperative, to maintain these newly-acquired capabilities in the most efficient way becomes paramount.

Most military aircraft fleet management work in the literature seeks to examine short-term planning and scheduling problems, using techniques such as mixed-integer linear programming (e.g., Hahn & Alexander, 2008; Kozanidis *et al.*, 2012) or simulation-optimisation (e.g., Mattila & Virtanen, 2014). The aims of this work are: 1) to test various fleet management policies over an entire fleet life (in this case, for the incoming MH-60R fleet) to ascertain which are the most robust to a range of circumstances; and 2) to determine which model input values have the most influence on the results. To accomplish this, we use a simulation model of the fleet (described in Section 2), coupled with simulation experimental design techniques (Section 3).

2. AIRCRAFT FLEET SIMULATION MODEL

2.1. Overview

The aircraft fleet simulation model used here extends an earlier fleet sizing model (Marlow & Novak, 2013). It represents a fleet of aircraft moving between various states over the fleet life. Aircraft may be serviceable (able to fly), in various types of maintenance, or other states, and move from one state to another when certain criteria are met. Aircraft may be embarked on ships, ashore at home base, or deployed for short periods for training. The model also includes changeover periods (when aircraft move between being embarked and ashore) and stand-down/holiday periods for the workforce.

The model includes various levels of scheduled maintenance: regular inspections, phased maintenance, and deep or depot-level maintenance. Regular inspections may be weekly and take 1-2 days; phased maintenance will occur after a certain number of flying hours and may take days or weeks; and deep maintenance will take place every few years and last weeks or months. The model also includes unscheduled maintenance, represented as probability distributions for both time between failures and time to repair. There are three different types of maintenance facility:

- Flight line maintenance lines. These handle regular inspections, and repairs on serviceable aircraft that have subsequently failed due to unscheduled maintenance;
- Phased maintenance lines. Along with aircraft in phased maintenance, unscheduled maintenance may need to be performed on these aircraft. Minor modifications may also be applied here.
- Deep maintenance lines. These facilities are operated by contracted rather than military personnel, and aircraft undergoing deep maintenance or major modifications will be serviced there.

An extension in this work (from that of Marlow & Novak, 2013) is that two squadrons are considered. One is an ‘operational’ squadron, responsible for embarking aircraft and preparing crews for such deployments. The other is a ‘training’ squadron, where crews are sufficiently trained in order to ultimately perform duties in the operational squadron. Each squadron has their own aircraft, along with maintenance facilities and personnel. The operational squadron has a small ashore component for both aircraft and maintainers, and otherwise is responsible for achieving all embarked requirements. The training squadron will solely achieve ashore hours, with any hours flown while undergoing training exercises counted towards this requirement.

2.2. Measures of Effectiveness

The three primary Measures of Effectiveness (MoEs) for this work are: 1) percentage of time that the minimum embarked aircraft requirement (eight aircraft each day) is not met; 2) annual embarked hours achieved; and 3) annual ashore hours achieved. These measures are the same as those for the earlier fleet sizing work (Marlow & Novak, 2013). With the inclusion of multiple squadrons, each squadron must also meet its own requirements.

A further consideration is not just whether the fleet flying hours are being met, but how they are being met. Ideally, the fleet would have a range (minimum to maximum) of flying hours that facilitates a phased withdrawal from service, in the same way that a fleet has a phased introduction to service. Having individual aircraft with total flying hours too far ahead of the rest of the fleet may ultimately force that aircraft to retire early, thus impacting the fleet size. Each aircraft should also have a balance of embarked and ashore hours, in order to equalise the exposure to embarked conditions: e.g., corrosive salt-laden environments and a higher operational tempo. Consequently, further MoEs are added in this work, these being the mean, standard deviation and range of embarked, ashore and total hours for the fleet.

2.3. Fleet management policies

In order to try to meet the various MoEs, various policies are tested to see which policies or combinations of policies are most effective. The six policy types are described below.

Flying allocation

Flying allocation policies are used to determine which serviceable aircraft fly each day, and how many hours they fly. Two policies are chosen, as in Marlow & Novak (2013):

- Maximise serviceability ('serv') flies each aircraft at the ratio of: hours until its next phased service, to days until the next phased maintenance line becomes free. It starts with the aircraft with the fewest hours until its next phased service, and so on for the aircraft with the second-fewest hours. However, this policy does not ensure that the daily requirement (based on the annual requirement) will be met.
- Greedy policy ('phG') directly seeks to achieve the required daily flying hour requirement, by choosing the aircraft with the least number of hours until their next phased service and flying it to the maximum daily rate, and then the aircraft with the second-least hours, and so on, until the requirement is met.

Maintenance allocation

Maintenance allocation policies determine which aircraft in various types of maintenance have manpower applied to them. Three options are chosen:

- Greedy flight line maintenance ('fltG') seeks to allocate manpower to aircraft in flight line maintenance, starting with the least number of maintenance manhours (MMH) remaining until they become serviceable (up to the maximum per aircraft per day), then the aircraft with the second least hours, etc.
- Greedy flight line and phased maintenance policy ('G') is as above, although the method is also applied to aircraft separately in both flight line maintenance and phased maintenance.
- Greedy squadron policy ('sqnG') is as above, except in this case all of the squadron aircraft in flight line and phased maintenance are pooled and sorted, and the greedy method applied.

Maintenance crew rotation

Maintenance crew rotations allow manpower to be transferred between the flight line maintenance and phased maintenance areas under particular circumstances. Here a 'norot' policy represents no rotations, whereas the 'rotP' policy allows rotations, depending on the number of empty maintenance lines at each area. For example, if one area has all lines empty, all of that manpower may be utilized in the other area.

Tail rotation

Tail rotation policies allow an individual aircraft to move between squadrons. The aim is to attempt to balance the embarked hours across the fleet. The model considers three options: no rotation ('none'); rotation following deep maintenance ('DM'); or rotation after scheduled (phased or deep) maintenance ('sch'). This also requires some variation in the maximum number of available aircraft in each squadron.

If, following a maintenance type listed above, an aircraft is in the operational squadron and has more embarked hours than the fleet average, it can be rotated into the training squadron. Similarly, an aircraft in the training squadron that has less embarked hours than the rest of the fleet may be rotated into the operational squadron. Rotations can be forced or banned if the allowed squadron size is or will be breached.

Balancing fleet flying hours

Policies to try to balance the total flying hours across the fleet are implemented when allocating daily ashore flying and maintenance resources. The balance flying hours policies directly consider the total number of flying hours per aircraft. For flying allocation policies, a serviceable aircraft that has less total flying hours than another will be prioritized. In the 'serv' and 'phG' flying allocations, an aircraft with less total flying

hours, but more hours until the next phased service, will be prioritized over an aircraft with more total flying hours, but less hours until its next phased service. The intention is to hold back aircraft with too many flying hours, and push forward aircraft with too few. Similarly, with the maintenance allocation policies, aircraft with less total flying hours, but more manhours until they become serviceable, will be prioritized over those aircraft with more total flying hours, but less manhours required until they become serviceable.

The policies provided in this work are either no attempt (‘none’), or based on either trying to keep total flying hours within a certain range, e.g., 500, 1000 or 1500 hours (‘f500’, ‘f1000’, ‘f1500’), or within one standard deviation of the mean (‘f1stdev’). The policies are applied when aircraft are outside of these ranges.

Sharing squadron resources

The aim of squadron sharing policies is to improve resource use across the fleet. Resources can be swapped as required (‘need’) if one squadron has met its requirements and has spare resources; and the other squadron has not met its requirements and has used all of its resources. Both maintainers and aircraft may be swapped. Spare maintainers in one squadron can provide extra manpower to the other squadron, reducing maintenance times. Similarly, spare aircraft in one squadron can assist the other in meeting its daily flying requirements.

3. DESIGNED SIMULATION EXPERIMENTS

Clearly, the number and variety of inputs to this model have the potential to yield complex behaviors and a very large number of combinations need to be considered. We do so with “data farming” – a term that has been used in the Defence community for over a decade to capture the notion of purposeful data generation from simulation models. It is different from data mining in much the same way that farming differs from mining. Miners seek to separate out valuable nuggets from their less valuable surroundings; data miners seek valuable information from large volumes of data. By contrast, farmers modify the growing conditions to promote the greatest yield obtainable from their resources; data farmers use designed experimentation to maximize the information obtainable from a limited set of simulation experiments. An additional benefit of data farming is that designed experiments are performed prospectively. Unlike observational data used in data mining, which can only note correlations, designed experiments can legitimately identify root cause-and-effect relationships between the input factor settings and the resulting outcomes. Using modern large-scale designs developed within the last decade, we can now explore massive input spaces and their relationship to complex simulation response surfaces. Data farming has enabled quantum leaps in the breadth, depth, and timeliness of the insights that can be obtained from simulation models.

To illustrate the basic concepts of designed experiments, consider the notional results of a “capture the flag” simulation shown in **Figure 1**. Green squares, yellow triangles, and red balls represent good, medium, and bad outcomes, respectively. The two designs in **Figure 1** have serious drawbacks, but are sadly representative of the way many simulation studies are conducted. In the first case, the factors have been varied in such a way (called confounding) that it is impossible to separate the effects of speed and stealth. In the second, when the factors are varied separately from each other, some may conclude (incorrectly) that neither speed nor stealth is important. However, putting them all together as in Figure 1(c) shows an interaction—there *is* a possibility of winning if both speed and stealth are varied simultaneously.

The designs in Figure 2 all allow interactions to be studied, but have different strengths and weaknesses. The 2^2 factorial (1(c)) has only $2^2=4$ design points, but does not reveal the behavior in the corners of the factor space. The 11^2 factorial (1(d)) provides a wealth of information, but at the cost of $11^2=121$ design points. The Latin hypercube (LH) design (1(e)) provides an intermediate level of

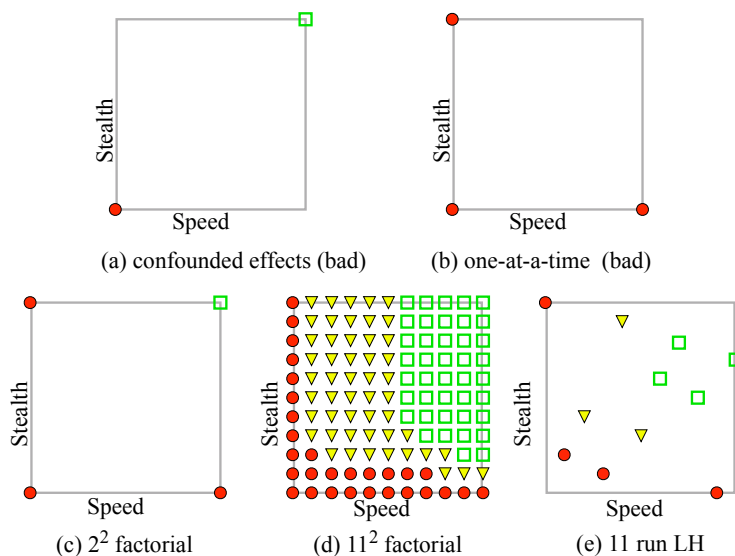


Figure 1. An illustration of bad (a-b) and good (c-e) experimental design using the “capture the flag” example

information with only 11 design points. The benefits of efficient designs, including LH designs, are even more dramatic when investigating large numbers of factors, while a brute-force approach quickly falls victim to the curse of dimensionality. Further details and references on designing large-scale simulation experiments are available (Sanchez *et al.*, 2012; Sanchez & Wan, 2012).

4. FLEET SIMULATION MODEL DESIGN

As stated in Section 1, the aim of this experiment using the aircraft fleet simulation model is to test not only the most effective policies, but determine the most influential variables. We use notional input data and requirements, as our intent is to demonstrate the utility of the methodology.

4.1. Inputs

For this experiment, there are 19 input variables: 11 continuous, 2 discrete, and 6 categorical. The 6 categorical variables are the various policy options described in Section 2.3. The chosen continuous and discrete variables and their values are provided in **Table 1**. Most of these variables are directly controllable in practice, with the exception of unscheduled maintenance (but which has controllable aspects). The daily increase allowance represents the amount by which the daily flying rate can be increased over the *pro rata* rate, where the *pro rata* rate, if followed, would allow the ashore aircraft to meet the annual requirement. The minimum maintenance duration represents the value above which an aircraft would need to be removed from service for at least a day. All other variable names are self-evident.

Table 1. Quantitative factors for simulation model experiment (notional data)

Variable	Values
Maximum ashore tail daily flying rate (fly hrs)	3-10
Minimum ashore tail daily flying rate (fly hrs)	0.5-1
Maximum total daily flying rate (sqn 1) (fly hrs)	10-20
Maximum total daily flying rate (sqn 2) (fly hrs)	25-35
Daily increase allowance (%)	0-1*pro-rata
Maximum embarked daily flying rate (fly hrs)	3-10
Minimum embarked daily flying rate (fly hrs)	0.5-1
Unscheduled maintenance mean time between failures (days)	0.8-1.2*mean
Unscheduled maintenance mean duration (days)	0.8-1.2*mean
Maximum ashore daily maintenance (man hours/day)	50-100
Minimum maintenance duration (man hours)	5-20
Maximum variation in number per squadron	0-2 (discrete)
Maximum days idle for serviceable aircraft	2-10 (discrete)

A number of other model inputs are not varied. These include the three primary requirements: for the fleet to have eight aircraft embarked every day of the fleet life, and achieve 4000 embarked hours and 4000 ashore hours each year. Other inputs include maintenance manpower rates-of-effort, time between services, and time to repair. Lognormal distributions are used for both the time between failures and the duration of unscheduled maintenance.

4.2. Design

We use a custom nearly orthogonal-and-balanced design (Vieira Jr. *et al.*, 2013) with 1040 design points for the 19 factors, with 50 replications at each design point. Continuous factors are varied at 512 distinct levels. The near-orthogonality is beneficial for analysis; the design has good space-filling behavior for the continuous factors, and samples the levels of each discrete or categorical factor in roughly equal proportions. The total set of runs took 34 hours on a single Mac pro with a 12 core Xeon E5 processor, capable of running 24 threads. Had we attempted a full factorial experiment, the number of design points that would have been required for comparable numbers of levels is 6.19E33. Even if we limited the continuous-valued factors to 10 levels each, a full factorial would have 9.27E14 design points—and would keep our machine busy for over 3.5 million years. Clearly, efficient designs have made it possible to run and analyze experiments that would otherwise be unthinkable.

5. RESULTS

The results can be analysed using a range of statistical techniques such as regression and partition trees (see, e.g., Hand *et al.*, 2001). These are used to determine the most influential factors on the MoEs. Given the space constraints, a summary table only showing the factors that directly impact the results, either individually or via interactions, is shown in **Table 2**.

Table 2. Summary of results showing the influential factors on various MoEs based on contributions to R². Categories determined by inverse powers of 2; VS is used for a very strong influence (>0.25), S for strong (>0.125), M for moderate (>0.0625), W for weak (>0.03125), and VW for very weak.

Variable	%t< min emb	Ann ash hrs	Ann emb hrs	Emb tail hr mean	Emb tail hr range	Ash tail hr mean	Ash tail hr range	Total tail hr mean	Total tail hr range
Maximum ashore tail daily flying rate		VW				VW			
Maximum embarked daily flying rate			S	S					
Unscheduled maintenance time btw fails	VS	VS	VS	VS		VS		VS	
Unscheduled maintenance duration	S	S	S	S		S		S	
Maximum variation in number/sqn					M		M		W
Maintenance crew rotation policies		M				M		W	
Tail rotation policies					VS		VS		VS
Balance flying hour policies							W		W
Squadron sharing policies	W	VW							

The most striking feature of the analysis is the impact of variable unscheduled maintenance on the results for the three primary MoEs and the mean results for the tail hours distributions. A reduction in time between failures (and thus an increase in frequency), coupled with an increase in duration leads to a large increase in unserviceability and thus a reduced ability to meet the annual ashore hours requirement, and in turn the embarked requirements. However, the sheer magnitude of its influence compared to other factors is notable. Unscheduled maintenance data for a new capability are difficult to predict, and these results suggest that an estimate that is incorrect by 10-20% may lead to a serious decrease in performance.

The influence of the various fleet management policies can be observed at various levels. For the less than percentage minimum embarked MoE, using the ‘need’ squadron sharing heuristic can help to reduce the impact of unscheduled maintenance, in some cases by a factor of 2. Similarly, for meeting the annual embarked hours requirement, the results suggest that increasing the maximum daily embarked flying rate can help to offset the reduced serviceability levels from increasing unscheduled maintenance. For meeting the annual ashore hours requirement, even with a higher frequency of unscheduled maintenance, using crew rotation and squadron sharing policies can allow this requirement to be achieved to within 1%. In each case the policies that encourage greater sharing or allocation of resources prove to be most effective. The ‘need’ squadron sharing policy and the ‘rotP’ crew rotation policy are preferred to no action.

For the range results, the tail rotation and balance flying hours policies have their greatest influence, as they are designed to address these issues specifically, whereas the other policies are designed to help meet the overall requirements. The results indicate that the tail rotation policies for embarked aircraft have the greatest influence. If no tail rotation policy is implemented, then only aircraft in the operational squadron will accrue embarked flying hours, leading to a fleet-wide imbalance. The effect of tail rotation policies therefore overwhelms that of any balance flying hour policies.

For balancing the embarked hours, the ‘sch’ tail rotation policy is preferred to the ‘DM’ and ‘none’ policies. Tail rotation after scheduled maintenance allows more opportunities for rotations to occur, and thus more flexibility in allowing the fleet to balance the embarked hours. This is shown starkly in **Figure 2**. While the maximum mean values are reached by all three policies, the desired lower range values are attained primarily through policies that permit greater rotation. Allowing a greater variation in the number of aircraft per squadron also facilitates more opportunities for embarked hours to be balanced.

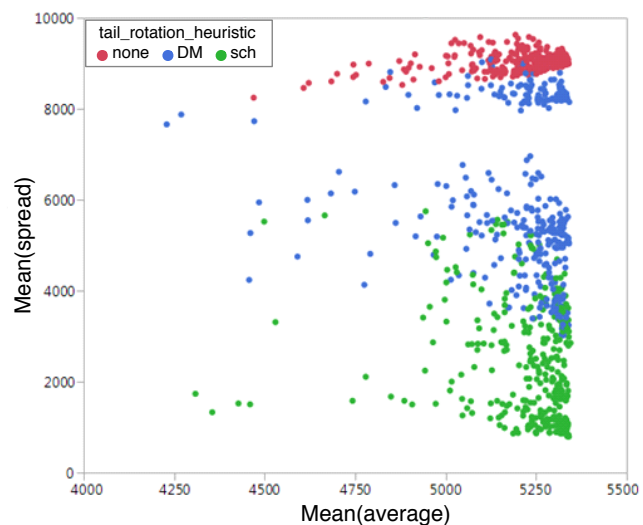


Figure 2. Impact of tail rotation policies on the range of embarked tail hours

Another interesting result is the fact that 8 of the 13 continuous and discrete variables in the experiment, as well as 2 of the 6 policies, have no significant impact overall. Along with the influential variables, a decision maker can combine these results to prioritise resources into the areas of most importance.

6. DISCUSSION AND CONCLUSIONS

This paper has demonstrated the utility of simulation experimental design techniques. Using a greatly reduced number of runs compared to a full factorial design, combined with various statistical methods, the most influential factors on the MoEs have been found. For this particular application to aircraft fleet management (noting the invented scenario), the tremendous influence of unscheduled maintenance has been revealed. From a practical viewpoint, this suggests that major efforts should be made to reduce its impact, through seeking to reduce maintenance frequency (e.g., through risk mitigation) and durations (e.g., by increasing manpower). It also demonstrates that particular fleet management policies can be applied that either provide superior solutions to some MoEs (e.g., the choice of policy to reduce the range of flying hours across the fleet) or assist in raising the overall annual flying hours (e.g., through maintenance crew rotations).

A limitation of this work is that it seeks to apply the same policy for the entirety of the fleet life. In reality, it may well be a case of ‘horses-for-courses,’ in that particular policies work best for particular situations at various times over the fleet life. Additional work is being undertaken in this area to find the best policies under those particular circumstances, using various simulation-optimisation approaches.

ACKNOWLEDGMENTS

This work was undertaken while the lead author was undertaking a fellowship at the Naval Postgraduate School in Monterey, California, USA. The author would like to thank the DST Group leadership team for agreeing to and funding this fellowship, and to the staff in the NPS Operations Research department – particular the former Chairman, Professor Rob Dell – for accommodating the author during the fellowship. The authors would also like to thank Steve Upton for his assistance in setting up the cluster runs.

REFERENCES

- Commonwealth of Australia (2009). THE STRATEGIC REFORM PROGRAM 2009. Delivering Force 2030. ISBN: 978-0-642-29707-5.
- Hahn, R. A. and Newman, A. M. (2008). Scheduling United States Coast Guard helicopter deployment and maintenance at Clearwater Air Station, Florida. *Computers & Operations Research*, **35**, 1829-1843.
- Hand, D. J., Mannila, H., and Smythe, P. (2001). *Principles of Data Mining*. Boston, Massachusetts: MIT Press.
- Kozanidis, G., Gavranis, A. and Kostarelou, E. (2012). Mixed Integer Least Squares Optimization for Flight and Maintenance Planning of Mission Aircraft. *Naval Research Logistics*, **59**, 212-229.
- Marlow, D. and Novak, A. (2013). Using discrete-event simulation to predict the size of a fleet of naval combat helicopters. In Piantadosi, J., Anderssen, R.S. and Boland J. (eds), *MODSIM2013, 20th International Congress on Modelling and Simulation*. Modelling and Simulation Society of Australia and New Zealand, December 2013, 2506–2512.
- Mattila, V. and Virtanen, K. (2014). Maintenance scheduling of a fleet of fighter aircraft through multi-objective simulation-optimization. *Simulation: Transactions of the Society of Modeling and Simulation International*, **90** (9), 1023-1040.
- Sanchez, S. M. and Wan, H. (2012). Work Smarter, Not Harder: A Tutorial on Designing and Conducting Simulation Experiments. In C. Laroque, J. Himmelspach, R. Pasupathy, O. Rose, and A. M. Uhrmacher (eds), *Proceedings of the 2012 Winter Simulation Conference*. Institute of Electrical and Electronic Engineers, New Jersey, 1929-1943.
- Sanchez, S. M., Lucas, T. W., Sanchez, P. J., Nannini, C. J., and Wan, H. (2012). Designs for large-scale simulation experiments, with applications to defense and homeland security. In Hinkelmann, K. (ed), *Design and Analysis of Experiments: Special Designs and Applications* (1st ed.), Volume 3, Chapter 12, 413–441. New York: John Wiley & Sons.
- Vieira Jr, H., Sanchez, S. M., Kienitz, K. H. K., and Belderrain, M. C. N. (2013). Efficient, nearly orthogonal-and-balanced, mixed designs: An effective way to conduct trade-off analyses via simulation. *Journal of Simulation*, **7** (4), 264–275.