Assessing the impact of the bathtub curve failure rate on fleet performance using designed simulation experiments

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

 ^a Joint and Operations Analysis Division, Defence Science and Technology, Fishermans Bend, Victoria, Australia
^b Operations Research Department, Naval Postgraduate School, Monterey, California, USA Email: david.marlow@dst.defence.gov.au

Abstract: This work seeks to determine the impact of the 'bathtub curve' failure rate on the performance of a fleet of military aircraft. Previous work in this area using a simulation modelling approach (e.g., Mattila and Virtanen, 2014; Marlow et al., 2015) has assumed a constant failure rate for unscheduled maintenance throughout a fleet life. We seek to ascertain whether a fleet is able to absorb the potential impacts of a representative bathtub curve through other actions such as implementing better fleet management policies, and at what point the effect of the bathtub overwhelms any mitigation procedures.

The bathtub curve is a concept used in reliability engineering to describe the varying failure rates of a system over time. It consists of three main periods:

- the 'infant mortality' or 'burn in' period, where the failure rate is high as initial faults are identified, before decreasing over time, resembling the left-hand side of a bathtub (viewed in cross-section);
- the 'random failures' period in the 'middle' of the bathtub, where failures occur at a constant rate; and
- the 'wear-out' period, where the failure rate increases as the system ages, resembling the right-hand side of a bathtub.

We incorporate the bathtub curve effect into our aircraft fleet simulation model used in previous work for a naval helicopter fleet (Marlow et al., 2015) but applicable to any fleet. The model represents the movement of aircraft between states for each day of a 30 year fleet life. Aircraft may be embarked on ships or ashore; they may be serviceable (able to fly) or in various types of scheduled or unscheduled maintenance. Aircraft change states once particular criteria are met – e.g., entering maintenance after achieving a certain number of flying hours. The primary measures of effectiveness (MoEs) for the model are to achieve the required number of embarked aircraft each day, and the required annual embarked and ashore flying hours. In order to evaluate the impact of the bathtub curve on fleet performance, we include additional MoEs that calculate the percentage of years that the fleet does not meet its annual embarked and ashore requirements. The model also includes a range of pre-determined policies to assist the fleet in meeting its MoEs. There are six policy categories: flying allocation, maintenance allocation, crew rotation between maintenance areas, tail rotation between multiple squadrons, policies to balance the total flying hours across the fleet, and squadron sharing.

The two new input variables to the model represent the 'height' of the bathtub (corresponding with an increase in the unscheduled maintenance failure rate of 2^{height}) and the 'duration' in years of each end of the symmetrical bathtub. To represent the bathtub curve effect, we use a 'thinning' technique (Lewis and Shedler, 1979) to create time-dependent rates over the fleet life. We generate gamma-distributed potential failure events at a rate corresponding to the maximum bathtub height, and accept or reject the generated events with a probability determined by the height of the bathtub at the time of the event. The probability decreases linearly from 1 to $1/2^{\text{height}}$ during the infant mortality period; takes the value $1/2^{\text{height}}$ in the random failures period; and increases linearly from $1/2^{\text{height}}$ to 1 during the wear-out period. In order to test various bathtub heights and durations, we use a designed simulation experiment approach (Sanchez and Wan, 2015) over 512 design points, with the height varying from 0 to 3 and duration varying from 0 to 10 years.

We perform statistical analysis of the results on the two new MoEs. For the percentage of years that the embarked annual flying requirements are not met, flying at a higher tempo (up to 7 hours per day) has the largest impact on the results, more so than the bathtub variables. For the percentage of years that the ashore annual flying requirements are not met, various combinations of existing policies, particularly sharing squadron resources, can be implemented to reduce the impact of the bathtub curve. However, if the bathtub effects last for longer than around 5 years at each end, and the height of the bathtub exceeds approximately 1.3 (i.e., slightly more the double the failure rate), negative impacts on the fleet performance eventuate.

Keywords: Bathtub curve, aircraft fleet simulation model, unscheduled maintenance, failure rate

U.S. Department of Defense Distribution Statement: Approved for public release; distribution is unlimited. The views expressed in this document are those of the authors and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

1. INTRODUCTION

In an era of increasing budget pressures for Western governments, there is increasing pressure on their military budgets, driving the demand for greater efficiencies within the constraints of those budgets. The Australian Defence Force (ADF) is facing these pressures while also undertaking a major capability improvement program through the acquisition of F-35 Joint Strike Fighters, MH-60R Seahawk naval combat helicopters, P-8 Poseidon maritime patrol aircraft, and other platforms. As these aircraft progressively enter service, the Australian Government will expect the ADF to maximize the utility of these capabilities.

Modelling and simulation techniques have the potential to provide substantial assistance in meeting these aims, e.g., via simulation-optimisation methods (such as Mattila and Virtanen, 2014). This paper builds on previous work that described a simulation model used to represent a fleet of aircraft operating over an entire fleet life: firstly to determine the minimum fleet size to meet various criteria (Marlow and Novak, 2013) and then extended to test the effectiveness and influence of various pre-determined fleet management policies (Marlow et al., 2015) using designed simulation experiments.

However, the work described above assumes that the unscheduled maintenance parameters are constant over the fleet life. In this work we introduce to our model the effect of a 'bathtub curve'; a concept used extensively in reliability engineering to describe the varying failure rates of a system over time.

1.1. The bathtub curve

The bathtub curve consists of three main periods:

- The 'infant mortality' or 'burn-in' period, where failures initially occur at a high rate as early faults are identified, before settling down over time;
- The 'random failure' period, where failures occur at a generally constant rate; and
- The 'wear out' period, where failure rates increase as the age of the component(s) increases.

Figure 1 provides an illustration of a representative bathtub curve. The h(t) term represents the hazard function or the failure rate, which equates to:



Figure 1. Illustration of the bathtub curve (adapted from Klutke et al., 2003)

$$h(t) = \frac{f(t)}{1 - F(t)}$$

where f(t) is the probability density function (pdf) and F(t) is the cumulative distribution function (cdf) for the relevant probability distribution(s) of the bathtub.

1.2. Literature review

The reliability literature contains many papers about the bathtub curve. Some of these question the merits of even using a bathtub curve, e.g., to represent the infant mortality period (e.g., Klutke et al., 2003). In this work we do not seek to enter the debate nor make any judgement on the validity of a bathtub representation – rather, we assume that such an effect occurs and seek to generate a representative bathtub model.

Numerous other papers seek to generate appropriate probability distributions for the bathtub curve. Due to its complexity, often more than one distribution is required to represent all bathtub periods: one example uses a combination of an exponential distribution with a decreasing failure rate for the infant mortality period, an exponential distribution with a constant failure rate for the random failure period, and a normal distribution for the wear-out period (Briand et al., 2007). Others (such as Xie et al., 2002) exploit the inherent flexibility of the Weibull distribution as the basis for generating a single distribution to represent the entire bathtub. However, these papers only consider the bathtub effect for a population of a single component or part, which fail once in their life and are not repaired. We are interested in investigating the bathtub effect for an entire fleet of aircraft, with multiple separate components that can and are being repaired, but which still mimic the bathtub effect in the failure rates.

1.3. Aims of this work

The primary aim of this work is to determine the impact of the bathtub curve on the performance of an aircraft fleet over its life. Assuming that a bathtub curve-type effect exists for an aircraft fleet, we want to know if and when the bathtub curve will affect the operations of the fleet – and if so, by how much. These impacts will of course depend on the 'height' of the bathtub at the beginning and end of the fleet life (i.e., the start of the infant mortality period and end of the wear-out failure period), and for how many years these periods will last. Since bathtub height and duration are unknowns for a given fleet at the start of its life, we use a designed simulation experiment approach to provide us with insights as to what values of bathtub height and duration will have a notable impact on fleet performance.

2. AIRCRAFT FLEET SIMULATION MODEL

2.1. Overview

The aircraft fleet simulation model used for this analysis has been described extensively in previous work (e.g., Marlow and Novak, 2013). In that work and here we consider a fleet of naval combat helicopters, although any fleet may be represented. It is designed as a whole-of-life simulation model that can test the impact of particular policies on fleet operations and can be used for sensitivity analysis on other variables.

The model represents a fleet of aircraft that can move between various states over its life. Aircraft can be serviceable (i.e., able to fly), unserviceable (in scheduled or unscheduled maintenance), or unavailable (in depot maintenance). Aircraft may be embarked on ships where they operate at a higher tempo (if modelling fixed-wing aircraft, this may be considered analogous to a deployment), or ashore for training. The model can represent aircraft in multiple squadrons, and the movement of aircraft between those squadrons.

Scheduled maintenance includes three types. Regular inspections are usually frequent (e.g., daily, weekly etc.) and occur after consuming a certain number of flying hours or elapsed time, and are short in duration (e.g., hours or days). Phased maintenance is required after a higher number of flying hours and may take days or weeks. Deep or depot maintenance occurs every few years and may take weeks or months to complete. Unscheduled maintenance occurs randomly in both time between failures and duration. Maintenance is performed at separate facilities: flight line maintenance lines undertake regular inspections and unscheduled maintenance discovered on previously serviceable aircraft, while phased and depot maintenance lines handle their respective services. Unscheduled maintenance may occur at any of these facilities if additional defects are discovered during scheduled maintenance.

The measures of effectiveness (MoEs) for this model are:

- 1. Percentage of time that the minimum embarked aircraft requirement is not met;
- 2. Annual embarked hours achieved;
- 3. Annual ashore hours achieved;
- 4. Mean and range (difference between min and max) of embarked, ashore and total hours for the fleet;
- 5. Percentage of years that the desired annual embarked and ashore hours are not met.

The first three MoEs are the primary MoEs and measure the fleet's ability to meet overall and ongoing requirements. The fourth set of MoEs is concerned with measuring the balance of flying hours across the fleet and across each squadron. To align with the aims of this work, we add a fifth set of MoEs to ascertain the impact of the bathtub at various stages of the fleet life. We allow a slack of 1% from the desired value each year as a judgement call that meeting 99% of the hours would be acceptable to a fleet commander.

2.2. Fleet management policies

The model can also been used to test various fleet management policies. The policies tested are described in more detail elsewhere (Marlow et al., 2015) and can be summarized as follows into six categories:

- Flying allocation used to determine which aircraft fly each day and how much. These include policies that seek to maximise serviceability ('serv') without explicitly seeking to meet the daily flying hours required (which are based on the annual requirement), or greedily choose aircraft to fly ('phG') which may cause aircraft to queue for phased maintenance later on.
- Maintenance allocation works in a similar way to flying allocation. All policies greedily apply maintenance man hours to aircraft with the least amount of maintenance remaining, whether they are in flight line maintenance ('fltG'), separately to aircraft in flight line and phased maintenance ('G'), or pooling aircraft in both flight line and phased maintenance ('sqnG').

- **Maintenance crew rotation** allows manpower to be transferred between the flight line maintenance and phased maintenance areas under particular circumstances, and thus assumes a multi-skilled workforce. Put simply, the 'norot' policy represents no rotations; the 'rotP' policy allows rotations.
- Tail rotation policies provide conditions for aircraft to move between squadrons, with the aim to balance the squadron hours across the fleet. The three options are no rotations ('none'), following depot maintenance ('DM'), and following scheduled maintenance ('sch') which includes both phased and depot maintenance. Generally an aircraft exiting maintenance with less hours than average in one squadron will transfer to another squadron, whereas those with more hours than average will not.
- **Balancing fleet flying hours** similarly to the tail rotation policies, these seek to balance the total flying hours across the fleet. They monitor the range of flying hours between the aircraft with the least and most hours in the fleet, and attempt to keep it below a specified number ('f500', 'f1000', 'f1500'), one standard deviation from the fleet mean ('f1stev'), or ignore it ('none'). These policies are manifested in the flying and maintenance allocation policies, with the policy to balance total hours prioritising decisions on which aircraft to fly or maintain.
- **Squadron sharing** these seek to improve resource use across the fleet if squadrons are co-located. For example, if squadron A is unable to meet flying or maintenance requirements on a particular day while squadron B can, then squadron B can share any spare resources with squadron A (aircraft or maintenance personnel) to assist squadron A. The squadron sharing option is on ('need') or off ('none').

Part of the intent of this work is to ascertain whether there is sufficient scope in these policies, or other model variables (e.g., flying rates), to offset or reduce the impact of the bathtub curve on fleet performance.

3. MODELLING APPROACH

3.1. Modelling the bathtub curve

We assume a symmetrical bathtub, so the specified height and duration apply to both the infant mortality and wear-out stages of the bathtub. We also only apply the bathtub effect to unscheduled maintenance.

To generate representative bathtub curves, we use a 'thinning' technique (Lewis and Shedler, 1979) applied to our probability distribution for unscheduled maintenance time between failures. The thinning procedure works by taking an appropriate distribution (i.e., one that represents independent events) and accepting or rejecting numbers drawn randomly from that distribution with some probability. By changing that probability over time, we change the rate at which those numbers are accepted, and thus can represent our bathtub.

Due to the thinning requirements to have a distribution representing independent events, we need to modify our extant probability distribution representing unscheduled maintenance time between failures from a lognormal distribution (we are not beholden to using a lognormal distribution as we use notional data to represent unscheduled maintenance). We choose a gamma distribution as the closest appropriate distribution given its right-tailed form, and use a maximum likelihood estimation technique to generate parameters for the closest gamma distribution to our original lognormal distribution.

The challenge of using this technique to generate a bathtub curve effect is to ensure that the mean time between failure (MTBF) during the random failures stage is the same for all instances, and then ensuring that the bathtub reaches the desired height at the beginning and end of the fleet life. Therefore, the thinning procedure is implemented as follows:

- Using the MTBF for the gamma distribution, and the bathtub curve height, determine the new MTBF for our gamma distribution that we wish to thin. The height is represented such that the change to the mean time between failure (MTBF) for the distribution is divided by 2^{height}. For example, an original MTBF of 32 and a bathtub curve height of 1 will lead to an MTBF of 16; a height of 2 generates an MTBF of 8, etc.
- Generate a new gamma distribution using that MTBF, with the same ratio of the gamma parameters a and b (given that a*b = MTBF for the gamma distribution). This will be the distribution to thin.
- At the beginning and end of the fleet life, the probability of acceptance of the numbers drawn from our new gamma distribution will be 1. In the middle random failures section of the bathtub, the probability of acceptance is $1/2^{\text{height}}$. The probability of acceptance then decreases linearly from 1 at the start of the fleet life to $1/2^{\text{height}}$ after the specified duration in years, and increases linearly from $1/2^{\text{height}}$ back to 1 once the duration in years from the end of the fleet life is reached. In this way, the MTBF in the middle of the bathtub will be the same as the original gamma distribution with no bathtub effect.

3.2. Designed simulation experiments

Data farming using designed simulation experiments (Sanchez and Wan, 2015) is an increasingly popular technique for generating purposeful data from large-scale simulation models, particular in the Defence modelling and simulation community. It allows the exploration of a large number of input variables in a simulation model – something that was previously insurmountable given the computational curse of dimensionality. It also discourages the use of 'one-at-a-time' sensitivity analysis, where the model owner uses their best guess to choose which variable(s) they believe to be the most influential on the results, and then varies each variable, in turn, while holding all others constant. This technique often feels forced upon the model owner due to the expected enormous number of cases to run, the model run time, and the time available to do a study. Conversely, data farmers use designed experimentation to maximize the information obtainable from a limited set of simulation experiments – so rather than *choosing up front* those variables are important. Good references on the theoretical foundations for and practical applications of designing large-scale simulation experiments are readily available (Sanchez et al., 2012; Sanchez and Wan, 2015).

4. EXPERIMENT SETUP AND RESULTS

4.1. Input data and design setup

For consistency, we use the same problem set as in our previous work (Marlow et al., 2015): i.e., a fleet of 24 helicopters in two squadrons, where one squadron is based ashore and used for training, while the other 'operational' squadron handles the 8 embarked aircraft plus a small ashore component.

Table 1 shows the variables used in our design as well as their ranges of values. Overall we have 21 input variables: 13 continuous, 2 discrete, and 6 categorical variables representing the 6 policies described in Section 2.2; further descriptions of these variables may be found in our previous work (Marlow et al., 2015)... We have also added the 2 new continuous variables encapsulating the bathtub curve effect to this model.

We vary the bathtub curve height from 0 and 3 (i.e., up to 1/8*MTBF of the original distribution) and the bathtub curve duration from 0 to 10 years. The duration applies *at each end* of the bathtub, so for a duration of 10 years, with our 30 year fleet life, the random failures period would only be 10 years.

Our earlier work was specifically designed to test our various fleet management policies using a designed simulation experiment technique (Marlow et al., 2015), as well as to demonstrate the applicability of the simulation experimental design approach to such a problem. The results of this work provided two significant observations:

- The effect of unscheduled maintenance dominated the other variables and policies;
- Not having a tail rotation policy (i.e., the 'none' option) leads to an unacceptable balance of flying hours in each squadron.

Given these outcomes, and the fact that the bathtub curve directly impacts unscheduled maintenance,

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

Variable	Range of 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)	0.95-1.05*mean	
Unscheduled maintenance mean duration (days)	0.95-1.05*mean	
Maximum ashore daily maintenance (man hours/day)	50-100	
Minimum maintenance duration (man hours)	5-20	
Bathtub curve height (multiplier)	0-3	
Bathtub curve duration (years)	0-10	
Maximum variation in number per squadron	0-2 (discrete)	
Maximum days idle for serviceable aircraft	2-10 (discrete)	

we modify our variable ranges for unscheduled maintenance time between failures to between 0.95 and 1.05. We also remove the 'none' option from our list of tail rotation policies. All other variables and policy settings remain unchanged. Other inputs not varied include those involving the maintenance regime (e.g., manpower rates-of-effort, time between scheduled services, time to repair, and maintenance capacity).

The three primary requirements are: 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. Our experimental design is a custom

nearly orthogonal-and-balanced design (Vieira Jr. et al., 2013) with 512 design points, and we use 50 replications at each design point. The results generated from MATLAB 2014a took 60 hours to run using 8 parallel cores on an Intel[®] Xeon[®] with 2.1 GHz processing speed and 64 GB of RAM running Windows 7.

4.2. Results and analysis

Due to space constraints, we focus on the two new MoEs when presenting the results: i.e., the percentage of years that the desired annual embarked hours (represented as %yr<embhrs) and ashore hours (%yr<ashhrs) are not met. These MoEs explicitly incorporate the second and third primary MoEs. Statistical analysis is performed using the JMP[®] software package from SAS[®].

The values for %yr<embhrs are relatively high overall, with a median of 62.1 and a 10th percentile of 34.5. This is a function of how the model represents flying rates for embarked aircraft, with those aircraft split into three different types flying at different tempos (Marlow and Novak, 2013). When long unscheduled maintenance events occur towards the end of an embarkation or the end of a year, it is difficult for the remaining embarked aircraft to fly at the required rates to still meet the annual requirement, usually causing a shortfall of more than the 1% allowed.

Nevertheless, we can still undertake statistical analysis on the relative changes to this MoE. A partition tree splits the data into exclusive subsets based on variable combinations that yield the greatest separation among the subset averages. We find that the maximum embarked hours per day dominates other variables, responsible for 78% of the observed R^2 value (= 0.835 after 10 splits). The average %yr<embhrs is 34.3 when the maximum embarked hours per day exceeds 6.8, the bathtub duration is less than 5 years, and the bathtub height is less than 1.3. However, if the maximum embarked hours per day exceeds 5 and the bathtub duration is greater than 5 years, the outcome is degraded to an average of 48.3 when bathtub heights are less than 1.0 (i.e., an initial doubling of the failure rate), and degraded even further for higher bathtub heights.

For the ashore hours MoE (%yr<ashhrs), the values have a mode of 0 as desired, with a median of 27.6. Statistical analysis using a partition tree over 10 splits shows that the squadron sharing heuristic has the predominant effect overall, responsible for 30% of the observed R^2 value (= 0.637 after 10 splits). The maintenance crew rotation, maintenance allocation, and flying allocation policies also have significant effects, again more so than the bathtub curve variables. The best leaf (average %yr<ashhrs of 5.5) requires the squadron sharing and maintenance crew rotation policies to be implemented, along with a bathtub height of less than 1.44. The partition tree yields an average %yr<ashhrs of 22.7 for the leaf where the bathtub heights are greater than 1.44 and the bathtub durations are less than 6.2 years. Without squadron sharing, the best leaf has an average %yr<ashhrs of 10.9, and corresponds to the 'phG' flying allocation heuristic, the 'sqnG' maintenance allocation heuristic, and a bathtub duration of less than 5.6 years.

The impact of the squadron sharing policy is clearly seen in Figure 2. The contrast implementing between а squadron sharing policy and not doing so is starkly demonstrated in these results. Implementing a squadron sharing policy greatly reduces the number of years when the annual ashore flying hours are not met. It also clearly shows how as both the bathtub curve height (x-axis) and the duration (coloured dots) the %yr<ashhrs increase, MoE increases from around 10% on average (i.e., ~3 years of the fleet life) with no bathtub curve to 40% with the bathtub height of 3; if squadron sharing is not implemented, the average



Figure 2. Graphs showing the influence of the squadron sharing heuristic on the fleet's ability to meet its annual ashore flying hours requirements (%yr<ashhrs) against the bathtub's height and duration (black line and grey shading represent regression line mean and 95% confidence band)

starts at almost 40% with no bathtub curve.

5. DISCUSSION AND CONCLUSIONS

This work has considered the impact of generalized bathtub curve failure rates on the ability of an aircraft fleet to meet ongoing requirements over a fleet life. For the MoEs that capture the fleet's ability to meet annual embarked and ashore flying requirements, the analysis has revealed ways to mitigate the impact of the bathtub curve on fleet performance: primarily by flying more embarked hours, and by allowing squadrons to share aircraft and maintainers. However, it has also demonstrated that these mitigation procedures have limits to their capabilities. From the analysis, the critical values for the bathtub curve variables equate to a height of 1.3-1.4 and a duration of 5-6 years. Beyond these values, either separately or together, the negative impact on the MoEs becomes increasingly apparent.

The outcomes of this work may be used by fleet planners to consider what type of bathtub curve effect may be tolerable to a particular fleet. If the bathtub effect experienced is likely to be above the critical values identified here, additional policies or measures (such as increasing the number of maintenance personnel or the maintenance capacity) may be required to allow operations to continue to meet requirements during these periods. Within these critical values, the policies and other methods identified above should be sufficient to alleviate any bathtub effects.

ACKNOWLEDGMENTS

The authors would like to thank MAJ Manuel Loewer (German Army) for the conversation that inspired the idea for this work.

REFERENCES

- Briand, D., Campbell, J. E. and Huzurbazar, A. V. (2007). Updating a User Friendly Lifetime Failure Distribution. In *Reliability and Maintainability Symposium (RAMS) 2007*. IEEE, 311-316.
- Klutke, G.-A., Kiessler, P. C. and Wortman, M. A. (2003). A Critical Look at the Bathtub Curve. *IEEE Transactions on Reliability*, 52(1), 125-129.
- Lewis, P. A. and Shedler, G. S. (1979). Simulation of nonhomogeneous Poisson processes by thinning. *Naval Research Logistics*, 26(3), 403-413.
- Marlow, D. O. 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.
- Marlow, D. O., Sanchez, S. M. and Sanchez, P. J. (2015). Testing aircraft fleet management policies using designed simulation experiements. In Weber, T., McPhee, M.J. and Anderssen, R.S. (eds), MODSIM2015, 21st International Congress on Modelling and Simulation. Modelling and Simulation Society of Australia and New Zealand, December 2015, 917-923.
- Mattila, V. and Virtanen, K. (2014). Maintenance scheduling of a fleet of fighter aircraft through multiobjective simulation-optimization. *Simulation: Transactions of the Society of Modeling and Simulation International*, 90(9), 1023-1040.
- Sanchez, S. M. and Wan, H. (2015). Work smarter, not harder: A tutorial on designing and conducting simulation experiments. In L. Yilmaz, W. K. V. Chan, I. Moon, T. M. K. Roeder, C. Macal, and M. D. Rossetti (eds), *Proceedings of the 2015 Winter Simulation Conference*. Institute of Electrical and Electronic Engineers, New Jersey, 1795-1809.
- 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.), John Wiley & Sons (New York), Volume 3, Chapter 12, 413–441.
- Vieira Jr, H., Sanchez, S. M., Kienitz, K. H. K. and Belderrain, M. C. N. (2013). Efficient, nearly orthogonaland-balanced, mixed designs: An effective way to conduct trade-off analyses via simulation. *Journal of Simulation*, 7(4), 264–275.
- Xie, M., Tang, Y. and Goh, T. N. (2002). A modified Weibull extension with bathtub-shaped failure rate function. *Reliability Engineering and System Safety*, 76, 279-285.