Proceedings of the 2018 Winter Simulation Conference
M. Rabe, A. A. Juan, N. Mustafee, A. Skoogh, S. Jain, and B. Johansson, eds.

OPTIMIZING RECRUITMENT TO ACHIEVE OPERATIONAL CAPABILITY CONDITIONAL ON APPETITE FOR RISK

Brendan Hill
David Kirszenblat

Department of Mathematics and Statistics
The University of Melbourne
Grattan Street
Parkville VIC, 3010, AUSTRALIA

Bill Moran

Department of Electrical and Electronic Engineering
The University of Melbourne
Grattan Street
Parkville VIC, 3010, AUSTRALIA

Ana Novak

Defence Science and Technology Group
506 Lorimer St
Fishermans Bend VIC, 3207, AUSTRALIA

ABSTRACT

This work is motivated by the need for the Australian Defence Force to produce the right number of trained aircrew in the right place at the right time. This necessitates the development of optimal recruitment strategies while sustaining squadron capability within some risk tolerance. The challenge is that Defence Aircrew training environments typically have highly variable failure rates and relatively small numbers of students. We investigate three receding horizon strategies, each of which use inflated notional targets with some deterministic assumptions to mitigate risk. The first strategy back-fills expected demand given fixed targets; the second strategy dynamically chooses targets using Monte Carlo simulations; and the third strategy incorporates Integer Linear Programming for partial solutions. We show that the first two strategies scale well and maintain steady states, and that the second strategy successfully incorporates the risk tolerance, resulting in an efficient and highly scalable strategy for the recruitment problem.

1 INTRODUCTION

Training manpower strategy typically refers to recruitment of new staff/students and scheduling of their training to deliver expected standards of qualification over a given time period. In the area of training for highly specialized skills, where the cost of training is high and available infrastructure and human resources are limited, the problem can become complex.

Optimal recruitment and manpower supply strategies have been studied for a number of decades using various approaches. Of particular interest is the degree to which Linear Programming models and Markov Decision Processes are scalable to real Military and Commercial problems.

Our focus is on the efficient generation of trained Australian Defence Force (ADF) pilots, observers and aircrewman. This problem has a number of unique features, one of which is the requirement that some graduated students become instructors rather than exclusively supplying manpower for squadrons, generating a feedback loop in the scheduling process. While instructors external to the ADF are sometimes available, they are costly, and their exclusive use has other drawbacks. Furthermore, recruitment into all courses is limited but controllable, student numbers are relatively small, and failure rates are significant and highly variable. These features, taken together, limit the usefulness of existing manpower planning and training scheduling approaches in this domain.

Linear Programming (LP) and Integer Linear Programming (ILP) have been common deterministic approaches to recruitment/manpower supply optimization. Here the search is for optimal values of a linear supply model, constrained by linear inequalities. In the ILP formulation, variables are additionally constrained to be integers. For example, Akinyele (Akinyele 2007) used ILP for the determination of effective manpower size incorporating global constraints such as production capacity/demand rate and allowable time-of-operation into the model to take into account differences between different countries' development times. Azimi (Azimi et al. 2013) used LP to allocate staff within a setting of an Iranian beverage company. Regarding the computational complexity of linear programming, it has been observed in practice that the the number of steps required by the simplex method to solve a problem with m equality constraints in n nonnegative variables is typically a small multiple of m, say $3m$ (Cottle 2003). Notably, Borgwardt, Smale and Haimovich have established that the simplex method is on average a linear-time method in a certain natural probabilistic model (Schrijver 2011). LP has unbounded complexity when dealing with real numbers (Traub and Woźniakowski 1982). Such results clearly indicate that the suitability of LP is largely determined by the data types, how the problem is formulated and, in particular, the number of constraints. Most importantly the LP approach solves the optimization problem in a single stage. That is, all recruitment, training continuum manpower supply (as in students per course per year) and instructor constraints collectively delineate the feasible region, which is a convex polyhedron, and the solution algorithm (typically the simplex method or a variant of Karmarkars algorithm, unless stochastic formulations are used for varying constraint parameters) performs the optimization in a single process. In this paper, we overcome the single step process by creating a simulation environment which interacts with the ILP optimizer.

In previous work, the authors successfully applied ILP approaches to the solution of related ADF timetabling problems. To simplify these problems, the variability associated with course pass rates was ignored, the pass rates being replaced by their respective means, i.e., expected values. This rendered the problems as deterministic and amenable to ILP methods. In particular, we applied, on the one hand, a novel combination of a standard ILP approach in conjunction with a fast technique for solving inexact cover problems (Nguyen et al. 2018) and, on the other hand, an advanced ILP method, namely column generation (Kirszenblat et al. 2017). It is worth noting that in the case where the pass rates are replaced by their respective means, the solution is referred to as the *expected value solution* (Birge 2011). One drawback of the expected value solution is that it may no longer be optimal, let alone feasible, when the real-world variability of course pass rates is taken into account. For example, in the context of optimizing recruitment, an unusually low course pass rate could result in a failure to meet target capabilities when following the expected value solution, whereas an unusually high course pass rate could result in a failure to respect class capacity constraints in subsequent courses when following the expected value solution.

A primary concern of this paper is the inclusion of risk in the problem formulation, in particular, the risk of failing to meet capability targets. As such, it is necessary to incorporate the random elements (pass rates and squadron attrition rates). Consequently, deterministic approaches such as ILP methods have been found to be inadequate when used alone. In order to mitigate the risk of failing to meet capability targets, it is clearly necessary to increase the number of students in the training continuum, which is at odds with the objective of minimizing the expected number of students in the system. To estimate risk, we have considered the introduction of probabilistic or chance constraints (Charnes and Cooper 1963) and Monte Carlo simulations. We have also considered local search and simulated annealing as ingredients of a stochastic optimization algorithm (Dimitris Bertsimas 1997).

Markov Models, on the other hand, model recruitment and manpower training as a stochastic dynamic system. One of the earlier developments of this approach is that of Bartholomew (Bartholomew 1971), where basic recruitment and progression-through-training are defined as prior probabilities of being accepted into a program and progression defined in terms of Markov transition matrices. This basic model has been applied to a wide range of domains, using Dynamic Programming (DP), and extended to model required training continuum numbers (see, for example, Mehlmann (Mehlmann 1980)).

For example, Udom (Udom 2013) has developed a Markov Decision Process (MDP) for optimal control of a Multi-level Hierarchical Manpower System (MHMS) dealing with department transfers and related movement of staff through organizations. The complexity of DP varies with the number of states and action and the search horizon (Chow and Tsitsiklis 1989). In contrast to LP, DP becomes exponential in complexity when states are continuous or even fractional. That is, they are best when dealing with in discrete, often categorical states. However, even with discrete states MDPs are not always scalable. In our previous work, we provided an elegant MDP solution to a simplified version of the aircrew recruitment problem. In concurrent work (Pike et al. 2018) we consider a model using Delta matrices.

In this paper we a) describe our modeling of the training continuum; b) consider three solutions to a full scale, real aviation training optimization problem in the Australian Defence Force; and c) provide results from experimentation with a demonstration scenario and discuss the relative merits and scalability of each model. Here, the objective is to minimize the number of recruits and the overall number of students in the system, subject to attaining squadron capability below a specified level of risk.

2 AIRCREW TRAINING PIPELINES

The Australian Defence Force provides training for various aviation-related personnel including aircrew, engineers and technicians. Training commences through one of the recruitment agencies such as the Australian Defence Force Academy (ADFA), the Royal Naval College (RNC), and others including direct entry. After recruitment, in the case of pilots, applicants undergo initial flight screening to assess their aptitude for the job. Successful applicants are then officially enrolled in the pilot training stream. In a temporal sense, the bulk of aircrew training consists of a linear sequence of major courses that constitute the backbone of the training continuum. In addition, students have to complete a number of mandatory short courses that are limited in prerequisite structure. Both short and long courses are run repeatedly throughout a term with a fixed number of repetitions.

The ADF aircrew training pipelines are susceptible to a great deal of variability, from changes to student intake, through to changing policies and aircraft types. Each change triggers a chain of events that unfold over terms and years into the future. Taking into account these interdependencies and how to manage them is critical for planning. The difficulty in managing the aviation training continuum, particularly for pilots, is further exacerbated by the high and extremely variable course failure rates, making it difficult to construct reliable predictions of numbers of students feeding into operational squadrons. Given that adequate resourcing of training schools is crucial to achieving capability, even small fluctuations in supply can be extremely costly, which makes obtaining an optimal recruitment strategy even more important.

We model the ADF training continuum as a directed pseudo-graph $G = (V, A)$ where the vertices are made up of intake pools (P), courses (C) and squadrons (Q), and the set of arcs (A) defines the possible progressions of each of the recruit types (R) through the program (see Figure 1).

Here, the intake pools are assumed to have an unlimited number of recruits available at any time. Each course repetition constitutes a session ($s \in S$), where session start/end times (σ_s, ε_s), maximum capacities (M_s), and further capacity limits per recruit type ($M_{s,r}$) are predefined known quantities. The number of enrolled students who pass is sampled from a Beta-Binomial distribution with mean μ_i and variance fitted to historical data. If recruits pass the course, they enter a waiting buffer, otherwise they leave the system.

Each squadron has a target manpower (O_v), where dropping below this target is considered a failure to meet capability. To reduce the probability of this event, the target may be increased (by β_v) to form a notional target. Squadrons are subject to random annual attrition of proportion ξ_v on average.

We define risk as the probability of a squadron going below the capability target in a given year; and the risk tolerance ρ^* as the acceptable risk of this event in each year over the horizon ($\mathbf{Y} = 10$ years / $\mathbf{T} = 120$ months). Therefore, the objective is to determine the minimum number of recruits, along with optimal movements of recruits throughout the continuum, such that the risk tolerance is attained by all squadrons. A solution is found for the full 10 years but presented as a 12 month schedule of movements for each arc in the graph ($X^t_{a,r}$). An updated solution is created each year, forming a receding horizon.

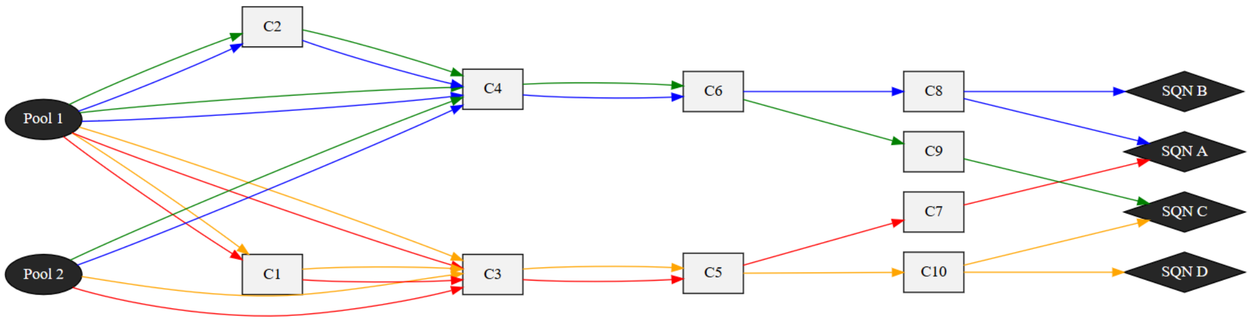

Figure 1: Example of a training continuum.

A real world concern is to support individual morale of recruits in training by preventing situations in which they spend large amounts of time waiting between courses. Since our model represents people in aggregate, this concern does not feature explicitly in our model; instead it is addressed indirectly via the minimization of recruits in the system. We also exclude consideration of individuals' eligibility to leave the system, and their individual likelihood of leaving based on rank, demography and other factors; instead we ensure that known attrition rates are attained on average over the horizon. This excludes the possibility of using a higher fidelity attrition model, but facilitates modeling numbers in aggregate - a key simplification.

3 THREE MODELS

In this paper we propose three models for solving the recruitment problem:

1. The *Proportional Back-Filling* (PBF) algorithm rapidly back-propagates expected demand, based on mean pass rates and mean attrition rates, given session capacities and initial recruits in the system. No risk calculation is performed, instead a predefined inflation figure determines the notional targets.
2. The *Proportional Back-Filling — with Risk* (PBF-R) strategy utilizes the PBF algorithm, but replaces the hard coded target inflation figure with a dynamic value, determined by estimating risk via Monte Carlo simulation against the risk tolerance ρ^*.
3. The *Integer Linear Programming — with Risk* (ILP-R) algorithm employs PBF for the preliminary region of the graph, and solves the remainder using Integer Linear Programming. Notional targets are determined dynamically in the same manner as PBF-R.

3.1 Proportional Back-Filling (PBF)

The PBF algorithm back-propagates expected demand at each node, at each time-step. This is computed in a single pass across the training graph, from squadrons to intake pools, along a reverse topological sort.

Each squadron's demand for recruits is modeled on mean attrition rates, with a notional target inflated by $\eta = 10\%$. The squadron's current state (below/at/exceeding target) influences the initial demand in order to correct the difference from the notional target.

Demand is back-propagated along the courses, temporally offset according to session lengths, and in accordance with session capacities. The preference is to satisfy outgoing demand using recruits already in the waiting buffers, or currently in training. Otherwise, enrollments in future sessions are scheduled, and demand is pushed back to earlier nodes.

The demand eventually calculated for the intake pools represents the numbers of new recruits to introduce into the training continuum at each time-step.

The back-propagation is computed using incoming demand vectors ($\phi(v)$) and outgoing demand vectors ($\psi(v)$) for nodes, and demand vectors $\lambda(a)$ for arcs. Pseudo-code for the PBF algorithm is provided in Algorithm 1, and full details of the demand vector calculations are provided in Appendix A.

Input: P, C, Q, A, s_0 (the initial state)
Output: ϕ, ψ, λ
for $q \in Q$ **do**
  $\psi(q) \leftarrow calculateSquadronOutgoingDemand(\lambda')$
  $\phi(q) \leftarrow calculateSquadronIncomingDemand(\psi(q), s_0)$
  $\{\lambda(a_1), \dots, \lambda(a_n)\} \leftarrow calculateArcDemandApportionment(\phi(q), \pi)$, for $a_i \in in(q)$
end
for $c \in reverse(toplogicalOrder(C, A))$ **do**
  $\psi(c) \leftarrow \sum_{a \in out(c)} \lambda(a)$
  $\phi(c) \leftarrow calculateCourseIncomingDemand(c, \psi(c), s_0)$
  $\{\lambda(a_1), \dots, \lambda(a_n)\} \leftarrow calculateArcDemandApportionment(\phi(c), \pi)$, for $a_i \in in(c)$
end
for $p \in P$ **do**
  $\psi(p) \leftarrow \sum_{a \in out(p)} \lambda(a)$
end
return ϕ, ψ, λ

Algorithm 1: Pseudo-code of PBF algorithm.

3.2 Proportional Backfilling with risk (PBF-R)

The PBF-R model replaces the fixed inflationary value of PBF (η) with a dynamically determined notional target specific to each squadron. These notional targets are determined by estimating squadron risk via Monte Carlo simulation, then incrementing the notional targets for squadrons still at risk. This is repeated until all squadrons satisfy the risk tolerance, or a maximum number of iterations is reached. Full details of the risk estimation and notional target incrementing are provided in Appendix C.

During risk estimation, within each Monte Carlo run, the PBF algorithm is rerun each year. This leads to greater responsiveness to the realization of stochastic events, and a more accurate risk estimate.

Aside from the dynamically determined notional targets, PBF-R is identical to PBF. In situations where the notional targets have been precomputed for a given graph, it is sufficient to run the PBF algorithm with those notional targets fixed.

3.3 Integer Linear Programming with risk (ILP-R)

The training graph can be divided into two subgraphs: the region in which the subgraphs for the individual recruit types each form a tree structure (which includes the squadrons), and the remainder of the graph (which includes the intake pools).

We formulate an ILP model which solves the optimal movements across the tree part of the graph. The objective function minimizes the number of recruits in the system, subject to satisfying the session timings, capacities, the initial state the system, and squadron notional targets. Squadron attrition is modeled based on average yearly attrition. Since it is sometimes unavoidable that squadrons are below target, squadron target constraints are implemented as soft constraints, where dropping below target is heavily penalized. The full formulation of the ILP is provided in Appendix B.

A solution to the ILP model yields expected demand at the midpoints of the graph onwards, forming a partial plan. From there, the PBF algorithm is used to back-propagate the demand to the initial intake pools, augmenting the ILP solution to form a complete plan.

ILP-R incorporates risk in the same manner as PBF-R, by estimating risk, and incrementing notional targets, as described in Appendix C. However, the computational cost of solving the ILP makes it infeasible to re-plan each year within each Monte Carlo run during risk estimation. Note, however, that ILP-R is still re-run each year (as per receding horizon model) when evaluating the strategy in the simulator.

4 RESULTS

4.1 Simulation engine

A purpose built simulation engine written in Java 1.8 forecasts the results of following each strategy.

One stochastic element is the random pass rates for each session of each course. This may naturally be modeled with a Binomial distribution with mean pass rate μ_i estimated from previous sessions. However, an analysis of historical data revealed that the actual pass rates exhibited greater variance than the Binomial distribution. So instead, we model pass rates with a Beta Binomial distribution where α, β are fitted to historical data using Maximum Likelihood Estimation and Method of Moments where necessary. This method still attains the mean historical pass rate, but with increased overall variance, since μ_i is sampled from a Beta(α, β) distribution each time.

The other stochastic element is squadron attrition. In the real world scenario, there are constraints on when individuals can leave based on individual deployments and minimum time spent in the system. Since the simulator models recruits in aggregate, we instead model attrition according to a Poisson distribution with parameter $\lambda = \xi_v B_v^{(t)}/12$ at time t, which attains the mean yearly attrition ξ_v observed in historical data. To counteract squadron attrition, newly trained recruits are sent to squadrons immediately. In the case where multiple squadrons draw from the same terminal waiting buffer, we calculate the target delta $\Delta_v = B_v^{(t)} - O_v$ for each squadron, and send each newly trained recruit to the squadron with the least Δ_v. This ensures that squadrons most significantly below target have priority for new recruits, and also results in a fair distribution when squadrons are above target.

4.2 Scenario Description

Due to the sensitivity of the real world data, we present a simplified demonstration scenario with 4 recruit types, 2 pools, 10 courses and 4 squadrons, shown in Figure 1. Average pass rates range between 50%-100%, and squadron average attrition rates range from 10% - 15% per year. The initial state is determined by filling each ongoing session to 50% capacity, and waiting buffers at 50% of the most recently completed session capacity. Squadrons are initialized at varying levels below, at, and above their respective targets.

Each of the strategies PBF, PBF-R, and ILP-R are run for 20 Monte Carlo runs with a risk tolerance of $\rho^* = 10\%$. The squadron sizes over time, relative to the target, are recorded, along with the average size across the Monte Carlo runs. This scenario is also used to explore the effect of modifying the risk tolerance in the range $1\% - 50\%$, when using the PBF-R algorithm.

4.3 Scenario Results and Discussion

The results for each strategy at 10% risk tolerance can be seen in Figure 2, showing the individual and average sizes for each squadron/type over the horizon, relative to the target number.

At the 10% risk tolerance, the PBF and PBF-R algorithms achieve a steady state for all squadrons. Subsequent experimentation showed that this steady state is maintained over longer horizons up to 30 years. The buffer created on average by PBF is indeed approximately 10% higher than the target, as per the η parameter used. It is unclear whether the positions of the steady states attained by PBF-R differ significantly from those of PBF, despite the notional targets being dynamically determined.

In contrast, the ILP-R results show significantly greater variability, and higher numbers on average, and it is unclear whether steady states are achieved within the horizon. We suggest two possible reasons:

1. The risk estimation method used by ILP-R does not re-plan within each Monte Carlo run, instead following a fixed plan over the 10 year horizon. When stochastic events are realized, it is harder to stay above target when constrained by a fixed plan, compared to regular re-planning. Hence, ILP-R will tend to produce higher risk estimates compared to PBF-R (which re-plans yearly, within each Monte Carlo run), and aim for higher notional targets, as appears to be the case in these results.

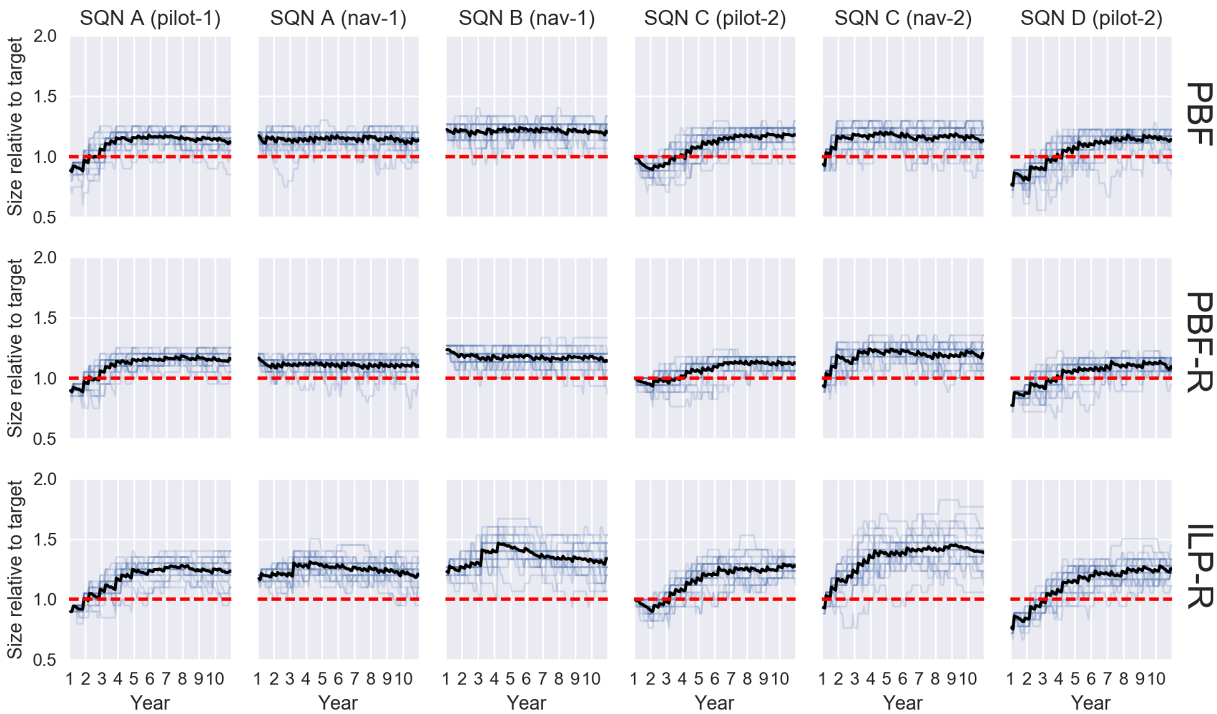

Figure 2: Squadron numbers relative to target for PBF, PBF-R, ILP-R at 10% risk. Thin lines are individual play outs, thick line are the averages, and the thick dotted lines represent the operational targets.

2. The augmenting of the ILP solution with the PBF solution for the earlier region of the graph can invalidate the assumptions of the ILP solution - in particular, the number of recruits available at the midpoint at each time-step may be limited by the number currently in the pipeline, a factor ignored in the ILP model. This may explain some of the increased variability in the ILP-R results.

Furthermore, when testing ILP-R on larger scenarios (eg. 100+ nodes), we found that the time taken to solve the ILP (using CPLEX 12.7 on an i7 machine) varied significantly depending on the initial conditions, with computation times ranging from 5-120 seconds. This cost is incurred multiple times in each year, in each simulation run. In a practical context, this would necessitate an ILP optimality gap tolerance, or time limit, either of which would weaken the optimality of the solution. In contrast, the PBF and PBF-R strategies perform single passes through the training continuum, resulting in an approximately linear increase in computation time as the graph grows.

The models can be differentiated based on their relative responses to the risk tolerance. Where PBF ignores the risk tolerance (aiming for a fixed inflation of the notional target), and ILP-R overestimates risk and notional targets (for the reasons given above), PBF-R dynamically chooses the minimum notional targets which are expected to attain the risk tolerance over the horizon, on the assumption that regular re-planning occurs.

The effect of reducing the risk tolerance with PBF-R can be observed in Figure 3, depicting the average numbers for SQN C (nav-2) for risk tolerances of 50%, 10%, 5% and 1%. In general, the more stringent the risk tolerance, the higher the notional target must be to stay above capability target with the specified probability. At 50% risk tolerance we expect the average to be close to the target, which is observed in these results. Lowering the risk tolerance to 10% and 5% result in an average buffer around 25%-30% above the target. At the 1% tolerance level, an average buffer around 40% above the target is maintained. Comparable responses are observed with the other squadron/types. Although a comparison of target vs.

observed risk is required to fully validate the risk estimation of the PBF-R strategy, these results show correct and reasonable adjustments of behavior to a tightening of the risk tolerance.

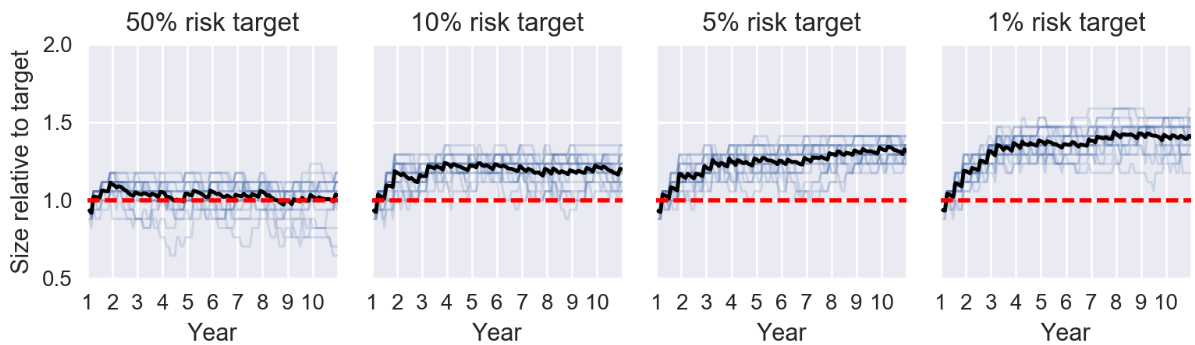

Figure 3: How PBF-R responds to reductions of the risk tolerance (for SQN C, nav-2).

5 SUMMARY

In this paper we have considered the problem of estimating minimum recruitment required to maintain squadron capability with a specified risk tolerance, over a 10 year horizon. Building on lessons from previous research efforts using LP and MDP, we have focused on heuristic and simulation-optimization (ILP) solutions which are capable of scaling up to large problems, while retaining the full set of constraints. The PBF strategy back-propagates expected demand using mean attrition and mean pass rates, with fixed notional targets. PBF-R incorporates a specified risk tolerance by dynamically determining notional targets via risk estimation and Monte Carlo sampling. ILP-R introduces an ILP model to find partial solutions.

Regarding computation time: each of the strategies efficiently computes a solution for the demonstration scenario (PBF: 1-2ms, PBF-R: 1-2 seconds, ILP-R: 2-5 seconds). However, we find significant differences when applied to the full scenario, where PBF and PBF-R computational times increase linearly, and ILP-R computation times become highly variable (PBF: 5-10ms, PBF-R: 5-10 seconds, ILP-R: 5-120 seconds).

Regarding quality of solutions: while we have optimality guarantees for ILP in general, we conclude that the value of this in the ILP-R strategy is undermined by the incompleteness of the solution, plus the significant computation cost preventing re-planning during risk estimation. By contrast, PBF-R provides a complete solution, where, critically, the risk estimation phase accommodates the assumption of re-planning. We observe that PBF and PBF-R successfully maintain steady states which are close to the desired notional targets on average, and that PBF-R effectively incorporates the desired risk tolerance by responding with proportional elevations of the notional targets.

In summary, we conclude that PBF-R constitutes a reasonable and efficient heuristic strategy for controlling the flow of recruits through a training continuum, subject to risk tolerances defined for the squadron targets. Finally, we note that if PBF-R is used to pre-compute the optimal notional targets, then it is sufficient to run PBF with those targets fixed, resulting in an extremely efficient and highly scalable optimization strategy for the recruitment problem.

ACKNOWLEDGMENTS

The authors would like to acknowledge Professor Terry Caelli for his insightful comments and editorial feedback.

APPENDIX A: Calculation of Demand Vectors in PBF Algorithm

The calculation of incoming and outgoing demand vectors is dependent on the node type (squadron/course/pool).

`Calculating outgoing demand (squadron)` : Let $\alpha_v = \lceil \mathbf{Y} \cdot O_v \cdot \xi_v \rceil$ approximate the total number of recruits who leave squadron v over the horizon. The departures are spread evenly throughout the outgoing demand vector at times $t_i = \left\lfloor i \frac{\mathbf{T}}{\alpha_{q,r}} \right\rfloor + 1$, for $i \in \{0, \ldots, \alpha_v - 1\}$, so $\psi_t(v) := 1$.

`Calculating incoming demand (squadron)` : Initialize the incoming demand to equal the outgoing demand, $\phi(v) := \psi(v)$. Denote the notional target as $O'_v = O_v + \beta_v$, and denote the difference been current number and notional target as $\Delta_v = O'_v - B_v^{(0)}$. If $\Delta_v > 0$, the squadron is currently below the notional target, and the gap is added to the incoming demand vector: $\phi_1(v) := \phi_1(v) + \Delta_v$. If $\Delta_v < 0$, the squadron is above the notional target, and the incoming demand for the first $|\Delta_v|$ attrition events is reset to 0, satisfied instead by the existing excess of members.

`Calculating arc demand`: When arcs $a_1, \ldots, a_n$ enter node v, then the incoming proportional branching constants π_{a_i} given in configuration determine the demand allocated to each arc: $\lambda_i(a_j) := [\pi_{a_j} \phi_i(v)]$. Some fair integer apportionment logic and random tie breaking is used to avoid fractional amounts without bias. When a node has a single incoming arc, $\pi_{a_i} = 1$, so $\lambda(a) := \phi(v)$.

`Calculating outgoing demand (course/pool)`: The outgoing demand for courses and pools is simply the sum of the demand vectors of the outgoing arcs: $\psi_i(v) := \sum_{a \in out(v)} \lambda_i(a)$

`Calculating incoming demand (course)`: The incoming demand for a course is computed from the outgoing demand, combined with the current state, with a preference for using existing people in the course over pushing demand back. Each time-step $t = 1 \ldots \mathbf{T}$ is considered in turn, and the outgoing demand $\psi_t(v)$ is met according to the following preference order:

1. Reserve people currently in the waiting buffer
2. Reserve people currently enrolled in ongoing session s, if $\varepsilon_s < t$
3. Enrol people in future session s (where $\varepsilon_s < t$), if s has spare capacity (prefer later sessions)
4. Enrol people in future session s (where $\varepsilon_s \geq t$), if s has spare capacity (prefer early sessions)

For (2), (3), (4), the mean pass rate is used to estimate the number emerging from each session.

(3) and (4) ensure that when there are insufficient people currently in the course to satisfy outgoing demand, we will meet the demand either in time, or, as soon as possible afterwards. Only in cases (3) and (4) is the incoming demand vector modified $\phi_{\sigma_s}(v) := \phi_{\sigma_s}(v) + n$, where n is the number of additional enrollments estimated to satisfy the demand at t. Hence, when demand is back-propagated, it is offset in time depending on the lengths and timings of the selected sessions.

In general, the demand at time t may be satisfied by via a combination of (1)-(4), and by creating enrollments across multiple sessions. If (4) occurs often, or (1)-(4) are exhausted altogether, this indicates a bottleneck in the training continuum, and the scenario may be infeasible.

`Solution`: The actions for $1 \leq t \leq 12$ are extracted from λ as follows: $X_{a,r}^{(t)} := \lambda_t(a), \forall a \in A$

APPENDIX B: ILP Formulation

`Variables`: All decision variables are in $\mathbf{Z}^{\geq}$.

$B_{v,r}^{(t)}$	recruits in v (session, waiting buffer, squadron) of type r at time t
$X_{a,r}^{(t)}$	recruits moved along arc a of type r at time t
$L_{s,r}$	loss from session s of type r, (i.e. failed students)
$G_{q,r}^{(t)[\delta]}$	δ'th degree gap - the gap between squadron number and notional target, offset by δ

The B variables for courses represent the waiting buffers. The B variables for sessions are defined only for the first time-step of the session, since session enrollment numbers do not change. The following variables are for notational convenience only, aggregating the incoming and outgoing flow respectively:

$$X^{(t)}_{\to v,r} = \sum_{a\in in(v)} X^{(t)}_{a,r}, \qquad X^{(t)}_{v\to,r} = \sum_{a\in out(v)} X^{(t)}_{a,r},$$

`Objective`: The ILP objective is to minimize the number of recruits in the system, calculated as the incoming number, plus the number in waiting buffers, sessions and squadrons, integrated over the horizon:

$$\gamma_r = \sum_{t=1}^{\mathbf{T}} \Big[\sum_{p\in P}\sum_{r\in R(p)} X^{(t)}_{p\to,r} + \sum_{c\in C}\sum_{r\in R(c)} B^{(t)}_{c,r} + \sum_{q\in Q}\sum_{r\in R(q)} B^{(t)}_{q,r} \Big] + \sum_{s\in S}\sum_{r\in R(s)} (\varepsilon_s - \sigma_s + 1) B^{(\sigma_s)}_{s,r} \tag{1}$$

The sum of the gap variables approximates a quadratic function of the actual gap, up to some degree Δ, and a large cost ($k = 1000$) is attached:

$$\gamma_s = \sum_{t=1}^{\mathbf{T}} \sum_{q\in Q}\sum_{r\in R(q)} \sum_{\delta\in\{0\ldots\Delta\}} G^{(t)[\delta]}_{q,r} \tag{2}$$

So the objective function is simply the weighted combination: $\min \gamma_r + k\gamma_s$

`Constraints`: All $B^0_{*,*}$ variables are constrained by equality according to the initial state of the system, given as input, where each pool has a sufficiently large number to meet intake over the horizon (eg. 1000). Movements are defined for $t \in \{1,\ldots,\mathbf{T}\}$ only.

The in/out flow constraint for pools, waiting buffers and squadrons:

$$B^{(t)}_{v,r} = B^{(t-1)}_{v,r} + X^{(t)}_{\to v,r} - X^{(t)}_{v\to,r} \qquad \forall v \in P\cup C\cup Q, r\in R(v), 1\le t\le \mathbf{T} \tag{3}$$

Session enrollment numbers:

$$B^{(\sigma_s)}_{s,r} = X^{(\sigma_s)}_{\to s,r} \qquad \forall s\in S, r\in R(s) \tag{4}$$

The transfer from session s to waiting buffer c incurs a loss of $L_{s,r}$ students, where $L_{s,r}$ is the rounded result of the enrolled number multiplied by the mean failure rate. The parameter $\zeta \in (0,1]$ determines the direction of rounding, and $\zeta = 0.5$ was used in our model for midpoint rounding:

$$(1-\mu_c)B^{(\sigma_s)}_{s,r} + (\zeta - 0.999) \le L_{s,r} \le \zeta + B^{(\sigma_s)}_{s,r}(1-\mu_c) \qquad \forall s\in S, r\in R(s) \tag{5}$$

$$X^{(\varepsilon_s)}_{s\to c,r} = B^{(\sigma_s)}_{s,r} - L_{s,r} \qquad \forall s\in S, r\in R(s) \tag{6}$$

The overall and type-specific session capacities respectively:

$$B^{(\sigma_s)}_{s,r} \le M_{s,r} \qquad \forall s\in S, r\in R(s) \tag{7}$$

$$\sum_{r\in R(s)} B^{(\sigma_s)}_{s,r} \le M_s \qquad \forall s\in S, r\in R(s) \tag{8}$$

The following constraints calculate the gap variables. Since the gaps are minimized in the objective function, all gap values will be zero when the squadron is at or above the notional target.

$$B^{(t)}_{q,r} + G^{(t)[0]}_{q,r} \ge O_{q,r} + \beta_{q,r} \qquad \forall q\in Q, r\in R(q), 1\le t\le \mathbf{T} \tag{9}$$

$$G^{(t)[\delta]}_{q,r} \ge G^{(t)[\delta-1]}_{q,r} - 1 \qquad \forall q\in Q, r\in R(q), \delta\in\{1,\ldots,\Delta\}, 1\le t\le \mathbf{T} \tag{10}$$

Finally, a vector $A(q,r)$ of average attrition over the horizon is computed for each squadron/type, identical to the outgoing demand vector $\psi(v)$ of PBF-R. This defines the squadron outflows to a null node $\emptyset$:

$$X^{(t)}_{q\to\emptyset,r} = A_t(q,r) \qquad \forall q \in Q, r \in R(a), 1 \le t \le \mathbf{T} \tag{11}$$

`Solution`: The one year action plan is taken directly from the decision variables, $X^{(t)}_{a,r}$, for $1 \le t \le 12$.

APPENDIX C: Risk Estimation and Determination of Dynamic Notional Target

PBF-R and ILP-R use the same procedure to estimate squadron risk each year, conditional on the strategy:

1. Run $N = 1000$ Monte Carlo runs, each following the strategy
2. For each squadron/type v:
   (a) For each year $y \in \{1,\ldots,\mathbf{Y}\}$, count the number of runs $f^{(y)}_v$ in which the squadron fell below the target for any duration in the year (i.e. failed)
   (b) Estimate the risk of each v in each year as $\hat{\rho}^{(y)}_v = f^{(y)}_v/N$
   (c) Estimate the overall risk of v over the horizon $\hat{\rho}_v$ as the arithmetic mean of the yearly risks
3. Compute the maximum estimated risk of any squadron/type, $\hat{\rho} = \max\{\hat{\rho}_v,\ldots\}$

Note that a solution is considered acceptable when $\hat{\rho} \le \rho^*$. Step (1) can be sped up considerably with halting conditions based on confidence intervals around squadron risk. The choice of arithmetic mean to aggregate the squadron risk over the horizon in step (2c) ensures that high risk during the initial period (often unavoidable given the initial state) does not render the problem infeasible. However, other aggregation functions such as the maximum risk, or weighted linear combinations of the years, may be reasonable.

To determine the ideal target inflation β_v for each squadron/type, an iterative process is used, where risk is estimated, and notional target incremented, until all squadrons satisfy the risk tolerance (or maximum iterations reached - in our demonstration scenario, we set this to 10):

1. Initialize all $\beta_v = 0$
2. Until risk tolerance is attained by all squadrons (or maximum iterations reached):
   (a) Estimate the risk of each squadron, given the current strategy and β_v boosts
   (b) Identify which squadrons are above the risk tolerance - $F = \{v : \hat{\rho}_v > \rho^*\}$
   (c) If $|F| = 0$, return the current plan as the solution
   (d) Otherwise, for each $v \in F$, increment the target boost value: $\beta_v := \beta_v + 1$
3. If maximum iterations reached, select the solution from the iteration with the best risk result

Regarding Step (3): it is not the case in general that the final iteration produces the best risk result. To select among the iterations, we define a risk vector for each iteration containing the ordered risks of each squadron, $[\hat{\rho}_{(1)} \ge \hat{\rho}_{(2)} \ge \ldots \ge \hat{\rho}_{(n)}]$. The iteration with the lexicographically least risk spectrum is selected. However, other approaches may be reasonable such as the minimax risk, the minimum mean risk, etc.

REFERENCES

Akinyele, S. T. 2007. "Determination of the Optimal Manpower Size Using Linear Programming Model". *Research Journal of Business Management* 1(1):30–36.

Azimi, M., R. R. Beheshti, M. Imanzadeh, and Z. Nazari. 2013. "Optimal Allocation of Human Resources by Using Linear Programming in the Beverage Company". *Universal Journal of Management and Social Sciences* 3(5):48–54.

Bartholomew, D. J. 1971. "The Statistical Approach to Manpower Planning". *The Statistician* 20(1):3–26.

Birge, F. L. 2011. *Introduction to Stochastic Programming*. Springer New York.

Charnes, A., and W. W. Cooper. 1963. "Deterministic Equivalents for Optimizing and Satisficing under Chance Constraints". *Operations Research* 11(1):18–39.

Chow, C.-S., and J. N. Tsitsiklis. 1989. "The Complexity of Dynamic Programming". *Journal of complexity* 5(4):466–488.

Cottle, R. W. 2003. *The Basic George B. Dantzig*. Stanford University Press, Stanford, California.

Dimitris Bertsimas, J. N. T. 1997. *Introduction to Linear Optimization*. Athena scientific Belmont, MA.

Kirszenblat, D., B. Hill, V. Mak-Hau, B. Moran, V. Nguyen, and A. Novak. 2017. "Using Column Generation to Solve an Aircrew Training Timetabling Problem". In *22nd International Congress on Modelling and Simulation*, edited by G. S. et al. Hobart, Tasmania, Australia: MSSANZ.

Mehlmann, A. 1980. "An Approach to Optimal Recruitment and Transition Strategies for Manpower Systems Using Dynamic Programming". *Journal of the Operational Research Society* 31(11):1009–1015.

Nguyen, V., B. Moran, T. Caelli, B. Hill, and A. Novak. 2018. "Dancing Links for Optimal Timetabling". *Journal of Military Operations Research (to appear)*.

Pike, C., B. Moran, D. Kirszenblat, B. Hill, and A. Novak. 2018. "A Stochastic Programming Approach to Optimal Recruitment in Australian Naval Aviation Training". In *Proceedings of the 2018 Winter Simulation Conference*, edited by M. R. et al. Piscataway, New Jersey: IEEE.

Schrijver, A. 2011. *Theory of Linear and Integer Programming*. Wiley Chichester.

Traub, J. F., and H. Woźniakowski. 1982. "Complexity of Linear Programming". *Operations Research Letters* 1(2):59–62.

Udom, A. U. 2013. "A Markov Decision Process Approach to Optimal Control of a Multi-level Hierarchical Manpower System". *CBN Journal of Applied Statistics* 4(2):31–50.

AUTHOR BIOGRAPHIES

BRENDAN HILL is a Researcher at the University of Melbourne specialising in Operations Research. Brendan holds a B.Sc (Mathematics and Statistics) from the University of Melbourne. He has accumulated 18 years experience in commercial software development across many industries, in roles ranging from technical lead to management. Research interests include optimization, simulation, machine learning and artificial intelligence. His website is http://brendanhill.com.au and his email address is brendan.hill@gmail.com.

DAVID KIRSZENBLAT is a mathematics PhD candidate at the University of Melbourne. He received a BA/BSc (Honours) and MPhil in mathematics from the University of Melbourne. His research interests include optimal control theory, geometric network optimization and the computational complexity and applications of integer linear programming, and his email address is dki@student.unimelb.edu.au.

BILL MORAN is a Mathematician/Engineer currently in RMIT School of Engineering and Honorary Professor at the University of Melbourne, Professor Bill Moran is a Fellow of the Australian Academy of Science. He has worked as a both as a Mathematician and as an Engineer across many research areas. Between 2011 and 2013 he was Director of the Defence Science Institute at the University of Melbourne. His current research interests lie in Optimization, Sensor Management, Signal Processing, Radar, and various topics in Mathematics, and his email address is wmoran@unimelb.edu.au.

ANA NOVAK joined DST Group in 2010 as an Operations Research Scientist after four years as a specialist consultant in a boutique demand, inventory and supply chain optimization firm. After immigrating to Australia in 2000, she received her B.Eng. (Honours) in Information Technology and Telecommunications from University of Adelaide in 2003, followed by a Ph.D. degree in Mathematics from The University of Melbourne in 2006 in Queueing Theory. Her research interests include planning and scheduling systems, simulation and modeling, optimization and modeling uncertainty, and her email address is Ana.Novak@dst.defence.gov.au.