Proceedings of the 2020 Winter Simulation Conference
K.-H. Bae, B. Feng, S. Kim, S. Lazarova-Molnar, Z. Zheng, T. Roeder, and R. Thiesing, eds.

WORKFORCE POPULATIONS: EMPIRICAL VERSUS MARKOVIAN DYNAMICS

Robert M. Bryce
Jillian A. Henderson

Director Research Workforce Analytics
Director General Military Personnel Research
Military Personnel Command
Department of National Defence
Ottawa, CANADA

ABSTRACT

Workforce populations are often modeled under a memory-free attrition assumption. This simple theoretical model, which corresponds to an exponential survival time distribution, allows population trajectories to be forecast, for example, by using a discrete time Markov model or the related differential system. However, in practice the distribution of survival times for a given population is often poorly described by an exponential. Here we present a study where different populations in the Canadian Armed Forces are considered. We contrast empirical survival time distributions with the matched exponential, and find distributions ranging from being close to exponential (e.g., Reserve Force) to distinctly non-exponential (e.g., Regular Force). We perform numerical experiments to determine how population dynamics diverge from the assumed Markovian dynamics, finding moderate error levels for the populations studied. On a coarse level the Markovian assumption appears remarkably valid, but with sufficient error (ca. 5–10%) to warrant caution.

1 INTRODUCTION

Markovian approaches have long be used for analysis of workforce populations (Young and Almond 1961; Bartholomew and Forbes 1979) and remains state-of-the-art (Bartholomew, Forbes, and McClean 1991; Nguyen 1997; Wang 2004; Škulj, Vehovar, and Štamfelj 2008; Guerry and De Feyter 2009; Boileau 2012; Zais and Zhang 2016; Jnitova, Elsawah, and Ryan 2017). The distinguishing aspect of Markovian approaches is the attrition process is memoryless, which corresponds to an exponential survival time distribution (Hoel, Port, and Stone 1971) describing the workforce, leading to an analytically tractable approach. However, in practice it is unlikely that real populations will be well described with an exponential survival time distribution and so modeling error is expected. Here we investigate and quantify this modeling error for several sub-populations of the Canadian Armed Forces (CAF), via a methodology readily applied by practitioners to their populations of interest. We find moderate levels of forecasting error, of roughly 5–10%, which is large enough to be operationally significant in some contexts (for example, in the CAF an occupation is labeled "critical" if the Trained Effective Strength is below 10% of the Preferred Manning Level, and labeled "caution" when between 5 and 10%).

2 MARKOVIAN THEORY

We discuss the differential equation associated with a Markovian population (Section 2.1), modify to a specialized form suited for our numerical experiments (Section 2.2), and consider the transient time as the population responds to a sudden impulse (Section 2.3).

2.1 Markovian population dynamics

A simple continuous differential model describing the change of a homogenous population $P(t)$ over time due to attrition at a rate of $\alpha P(t)$ (outflow volume per unit of time) and a constant intake at a rate of *in* at time t (inflow volume per unit of time) is

$$\frac{dP(t)}{dt} = -\alpha P(t) + in,$$

where α is the attrition rate parameter. This ordinary differential equation has the exact solution

$$P(t) = \frac{in}{\alpha} + \left(P_0 - \frac{in}{\alpha}\right) e^{-\alpha t}, \tag{1}$$

where P_0 is the initial population, as can be verified by plugging this solution back into the originating differential equation. As $t \to \infty$ the second term goes to zero and we have the steady state population $P_{SS} = \frac{in}{\alpha}$. This differential model is related to the approximate discrete time Markov model, a standard approach taken in workforce modeling (Bartholomew, Forbes, and McClean 1991); these two Markovian approaches are identical in the limit of diminishing time steps.

It should be noted that a matrix generalization of Equation 1, where flows occur between sub-components, can be obtained (Henderson and Bryce 2019) and exactly solved (Higham 2008) which allows one to consider transitions between military ranks, occupation groups, or other flows between sub-components. Here we focus on single component populations and the differential model in Equation 1 to eliminate complicating factors, allowing us to highlight modeling error accrued due to making a Markovian assumption.

2.2 Markovian population dynamics: specialized form

We want to recast Equation 1 into terms that will map directly to our numerical experiments (as described in Section 3).

We will evolve populations under a given intake level to steady state and then impose a sudden impulse consisting of a step increase in intake, which will, after a transient transition period, lead to a new (higher) steady state. The ratio of final to initial intake is denoted β and captures the magnitude of impulse to which the system is exposed. For example, $\beta = 1.1$ corresponds to a 10% increase in intake and will (eventually) result in the population growing by 10%. The average survival time, $\langle t \rangle$, is inversely related to the attrition rate parameter, α; $\langle t \rangle = \frac{1}{\alpha}$. As P_{SS} defines the natural population scale we will consider the normalized population, and eliminate P_0 by use of β, to get

$$\frac{P(t)}{P_{SS}} = 1 - \left(1 - \frac{1}{\beta}\right) e^{-\frac{t}{\langle t \rangle}}. \tag{2}$$

The $\left(1 - \frac{1}{\beta}\right)$ term is simply the normalized population step increase, as $\frac{(P_{SS} - P_0)}{P_{SS}} = \left(1 - \frac{P_0}{P_{SS}}\right) = \left(1 - \frac{1}{\beta}\right)$.

Equation 2 describes how a Markovian population evolves under an intake step increase of size β. We will be contrasting this Markovian theory against real populations that we characterize by measuring their empirical probability distribution functions (PDFs). The residual between the population trajectories given by theory (i.e., Equation 2) and numerical simulations will allow modeling error to be quantified.

2.3 Transient time

A Markovian population asymptotically approaches steady state; a tolerance for "reaching" steady state must be set in order to fix measurements of transient time. We consider a relative threshold so our results do not depend on specifics of the population size. The left-hand side of Equation 2 suggests using the k-th e-folding time to fix thresholds, and we have

$$\frac{P(t^*)}{P_{SS}} = 1 - e^{-k} = 1 - \left(1 - \frac{1}{\beta}\right) e^{-\frac{t^*}{\langle t \rangle}},$$

where t^* is the transient time. For example, a 3 e-folding time is the time taken for the population to reach $1 - e^{-3} \approx 0.95$ of P_{SS}.

Solving the above equation for normalized t^* we find

$$\frac{t^*}{\langle t \rangle} = k + ln\left(1 - \frac{1}{\beta}\right). \tag{3}$$

Note that as $\beta \to 1$ the logarithmic term goes to $-\infty$. Negative values in Equation 3 are due to P_0 being already above the threshold. For any given k we therefore have a minimum β that can considered, which we can find by setting $t^* = 0$ in Equation 3

$$\beta_{min} = \frac{1}{(1 - e^{-k})}. \tag{4}$$

3 STUDYING MODELING ERROR

We investigate the scale of modeling error associated with assuming an exponential survival distribution for CAF sub-populations with attrition profiles that range from visually near-exponential to distinctly non-exponential (Sections 3.1 and 3.2). We compare simulated population dynamics for growth under attrition behavior, as given by empirical survival time distributions, to "equivalent" exponential fit and Markovian theory, for a range of steady state growth scenarios (Section 3.3). Error is then quantified using two measures: maximum gap during the growth phase between populations with differing attrition behaviors (Section 3.4), and transient time to steady state for various values of e-folding time (Section 3.5).

3.1 Survival time distributions

We use a normalized histogram with annual bins as a simple non-parametric estimator for the empirical PDFs. This choice is pragmatic, selected as it is a simple way to capture and visualize PDFs. One advantage of this choice is per annum time scales correspond to standard institutional planning, and so annual bins and frequency are commonly used, making our empirical bins intuitive and comparable to other studies. For example, the CAF characterizes attrition rates by Years of Service (Serré 2019), and such an annual-centric scale is the standard choice in workforce analysis (Bartholomew, Forbes, and McClean 1991).

One can also fit data to parametric distributions. Here we are interested in the equivalent exponential fit in order to match an empirical population with a Markovian one. However, we do not fit the exponential to the raw data via the Maximum Likelihood Estimator (MLE), a standard approach, for the following reason: there is estimation error associated with using a histogram to capture a PDF and so if we estimate both the histogram and the exponential from the raw data, in general, we will expect a small difference between them. Specifically, the choice of bin sizes and number will affect a histograms expectation value and so, in general, the expectation value determined by the MLE will differ from that determined by the histogram. A difference in the average lifetime between the histogram and exponential will introduce error that will affect trajectories which, for example, will lead to different steady states. The trajectories of the empirical and exponential will differ due to both differing shape (which we want to probe here), as well as differing average lifetime (an artificial estimation error). In order to eliminate this error, and find an

"equivalent" matching exponential, we instead determine the average lifetime by finding the expectation of the normalized histogram. This expected value will correspond to the α parameterizing the exponential. This procedure simplifies interpretation. Essentially we are *defining* the population by the histogram and finding the matched Markovian population, so that any differences in trajectories can be attributed to differences in PDF shape (versus being due to both differences in shape and a shift in central location, which complicates interpretation).

3.2 Canadian Armed Forces sub-populations

The two largest sub-populations of interest within the CAF were chosen for analysis: the Regular Force (Reg F), comprised of military members under full-time contract, and the Primary Reserve Force (P Res) comprised of military members under part-time contract. Figure 1 includes the empirical PDFs for these two groups calculated using historical data on career duration (in that group) for all those who attrited within a 15-year period (1 January, 2004 to 31 December, 2019). The mean survival time, $\langle t \rangle$, was calculated for each PDF (15.98 and 6.08 years for Reg F and P Res respectively) and used to build the equivalent exponential model (dotted curve of Figure 1).

The difference between the survival time profiles of these two groups is stark, with the P Res appearing very close to exponential and the Reg F with clearly non-exponential structure. The exponential-like profile of the P Res suggests a constant year-over-year hazard for separation (attrition), which is consistent with the separation at will Terms of Service (TOS) for the Reserves in Canada. The more complex structure for the Reg F is reflective of a more complex TOS regime in the Reg F, with various fixed term contracts and a historic 20 year pension point and current 25 year pension point (Serré 2019).

A third group, the Reg F sub-population of the Naval Warfare Officer (NWO) occupation (middle panel of Figure 1), was chosen as the Royal Canadian Navy (RCN) has been studying this occupation during the current period of fleet modernization and recapitalization. An on-going program of NWO population modeling began in 2017 and so we are interested in further understanding of this occupation. We find the NWO group presents as a middle-ground between these two extreme cases, both in terms of the distribution structure and it's average survival time of $\langle t \rangle = 10.14$ years.

It should be noted that the Reg F and P Res data sets are comprised of all records of attrition, including releases and transfers between the two components, for the time period stated, indiscriminate of rank or employment state, with sizes on the order of 10^5. The analogous NWO data set additionally takes into account occupation transfers and is comprised of officer ranks only with an approximate size of 3000.

Recall that we have chosen to define the sub-populations by their histograms in order to focus on the effect of PDF shape on growth trajectories (Section 3.1). To get a sense of the difference between $\langle t \rangle$, as detemined by the MLE ($\langle t \rangle_{MLE}$) versus that derived from the histogram ($\langle t \rangle_{hist}$), consider the measured values and their ratio for each sub-population: P Res: $\langle t \rangle_{hist} / \langle t \rangle_{MLE} = 6.082/6.079 = 1.0006$. NWO: $\langle t \rangle_{hist} / \langle t \rangle_{MLE} = 10.140/10.049 = 1.009$. Reg F: $\langle t \rangle_{hist} / \langle t \rangle_{MLE} = 15.977/15.878 = 1.006$. Use of the MLE gives results that are only marginally different from our chosen approach, for our sub-populations. We eliminate this effect, on principle. For smaller populations, where estimation error is greater, this effect could lead to practical consequences.

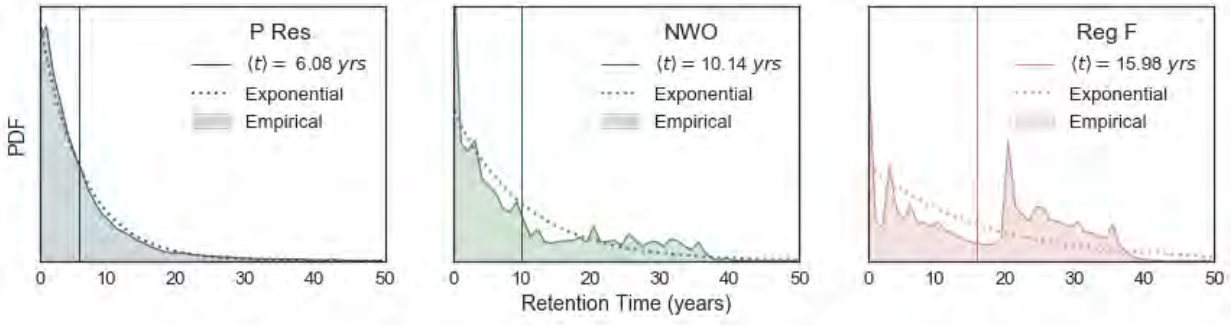

Figure 1: Empirical survival time PDFs (filled areas) and equivalent exponential models (dotted curves) for three groups of military personnel: (*left panel*) Primary Reserve Force (P Res), (*middle panel*) Regular Forces - Naval Warfare Officer (NWO), and (*right panel*) Regular Force (Reg F). Mean survival times, $\langle t \rangle$, for the PDFs are shown for comparison (solid curves).

3.3 Population dynamics

Each sub-population under study was grown under a constant intake corresponding to the steady state solution of Equation 1 and career durations drawn randomly from the empirical and exponential PDFs of Figure 1. Simulated populations were first grown to a steady state population size of $P_0 = 1000$ individuals over a simulation time of 300 years before applying a growth impulse to a new P^{SS}. A range of population growth scenarios were created by increasing P_0 by 1% up to 1000%, corresponding to an impulse set $\beta = (1.01, 1.05, 1.1, 1.25, 1.5, 2, 10)$. Typical personnel increases could be in the range of 5-25% (i.e., $\beta = 1.05 - 1.25$). For example, *Strong, Secure, Engaged*, Canada's Defence Policy (circa 2017) details increases to a target that is $\sim 7\%$ higher than the 2016/2017 year end size for for the Reg F and $\sim 13\%$ for the P Res. Under the same Defence Policy, Canada's Special Forces Command is to grow by $\sim 30\%$, a larger increase underscoring the special capability and utility of that command. For practical considerations we are interested in β values in the ballpark of this range, and therefore consider larger values (i.e., $\beta > 2$) to ensure we bracket such values. An additional advantage accrued by looking at larger impulses is that stochastic noise concerns are suppressed at large β values enabling the underlying trends being studied to be more clearly examined.

Resulting population trajectories, $P_{sim}(t)$, for each sub-population under each survival distribution and impulse scenario, were averaged over 1000 runs and compared with Markovian theory (Equation 2). As a diagnostic to probe finite-sample estimation error we look at the steady state population for each simulated population, P_{sim}^{SS}, which was calculated over the final 100 years of simulation time with a maximum standard error of the mean (SEM) of 0.4 corresponding to the largest impulse $\beta = 10$ for the NWO population. To simplify analysis, we measure $P_{sim}(t)$ with an annual frequency.

In order to concretely illustrate the effect of survival distribution on population dynamics, the population trajectories for $\beta = 1.25$ are shown in Figure 2 where the left-hand panel illustrates the growth phase to P^{SS} over the next 300 simulation years. The left-hand inset focuses on the contrast in growth dynamics between populations of differing survival PDFs. We find that $P_{sim}(t)$ from the exponential fits (dotted curves) follow the Markovian theory (thick solid curves) closely, as expected, while $P_{sim}(t)$ evolving under empirical PDFs differ in trajectory. All of the populations are shown in the left-hand panel, to facilitate comparison.

The right-hand panels showcase the scale of the difference between models during the "arrival" of $P_{sim}(t)$ to P^{SS}. The P Res trajectories (upper panel) closely follow Markovian theory while those for NWO (middle panel) and Reg F (lower panel) increasingly diverge from the Markovian prediction. The scale of these differences, the gap between the simulated trajectory, $P_{sim}(t)$, and the Markovian trajectory, $P_{theory}(t)$,

denoted here as Δ, and the transient time, t^*, to reach the steady state target, will be discussed in Sections 3.4 and 3.5, respectively.

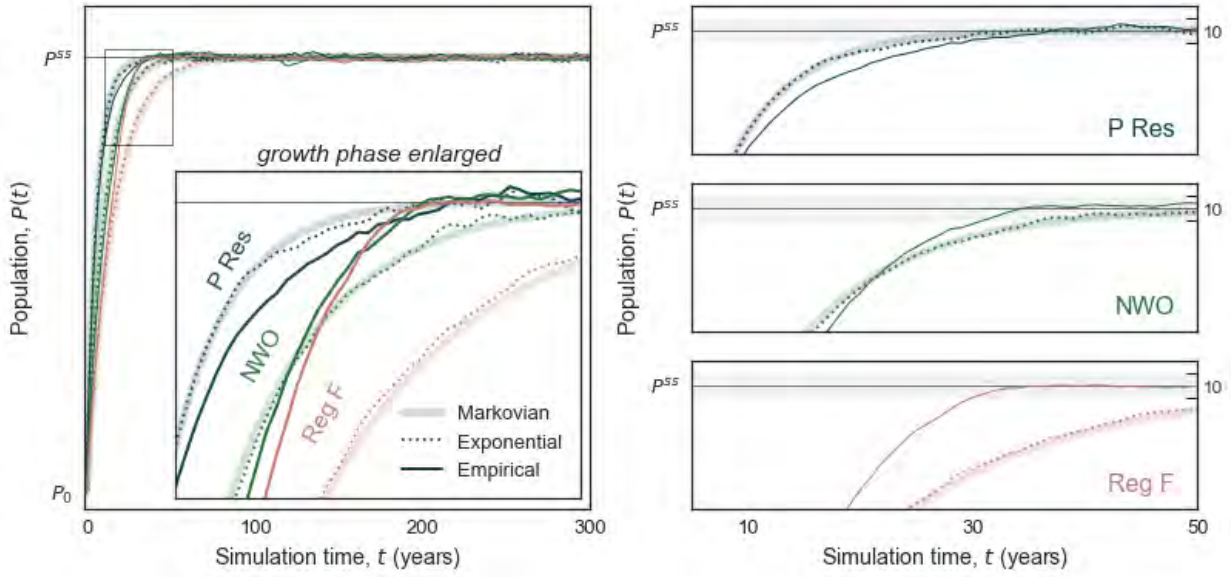

Figure 2: (*Left panel*) Growth phase for a 25% increase in population size. Simulated (normalized) populations defined by the empirical PDF (solid curves) and equivalent exponential model (dotted curves) are compared to Markovian theory (thick solid curve) for P Res (blue), NWO (green), and Reg F (rose). (*Right panels*) Mean population, $P_{sim}(t)$, as it approaches the target steady state population, P^{ss}. A scale of 10 individuals is shown for comparison.

3.4 Maximum gap, Δ

As the name implies, the maximum gap is calculated as the largest difference between the simulated population, $P_{sim}(t)$, and the Markovian theoretical population, $P_{theory}(t)$, at a single time step during the growth phase

$$\Delta = max(|P_{sim}(t) - P_{theory}(t)|). \tag{5}$$

Figure 3 shows Δ normalized by P^{SS} and scaled by the normalized step increase $(1 - \frac{1}{\beta})$ as a function of β for each scenario. The amplitude of the max gap curves for the exponential models (dashed curves) indicate the scale of stochastic noise for each β, as divergence from Markovian theory is due to finite sample effects (residual stochastic fluctuations as well as convergence from below, see (Henderson and Bryce 2019)). This "noise floor" drops towards zero as the impulse size increases, as the size of fluctuations becomes small relative to the population step size. On the other hand, the scaled maximum gap for simulated populations undergoing attrition according to the empirical PDFs appear to plateau at values of $\sim 12\%$ for Reg F , $\sim 5\%$ for NWO, and $\sim 3\%$ P Res. We summarize this as $\approx$ 5–10%, in order to distill our findings into a memorable "rule of thumb" which quantifies the general magnitude of error we are observing for CAF populations. Note how all the populations collapse to a common curve as β becomes small, this is the effect of measurements being lost in the noise floor. As β increases we pull out of the noise floor and see a plateau for each population.

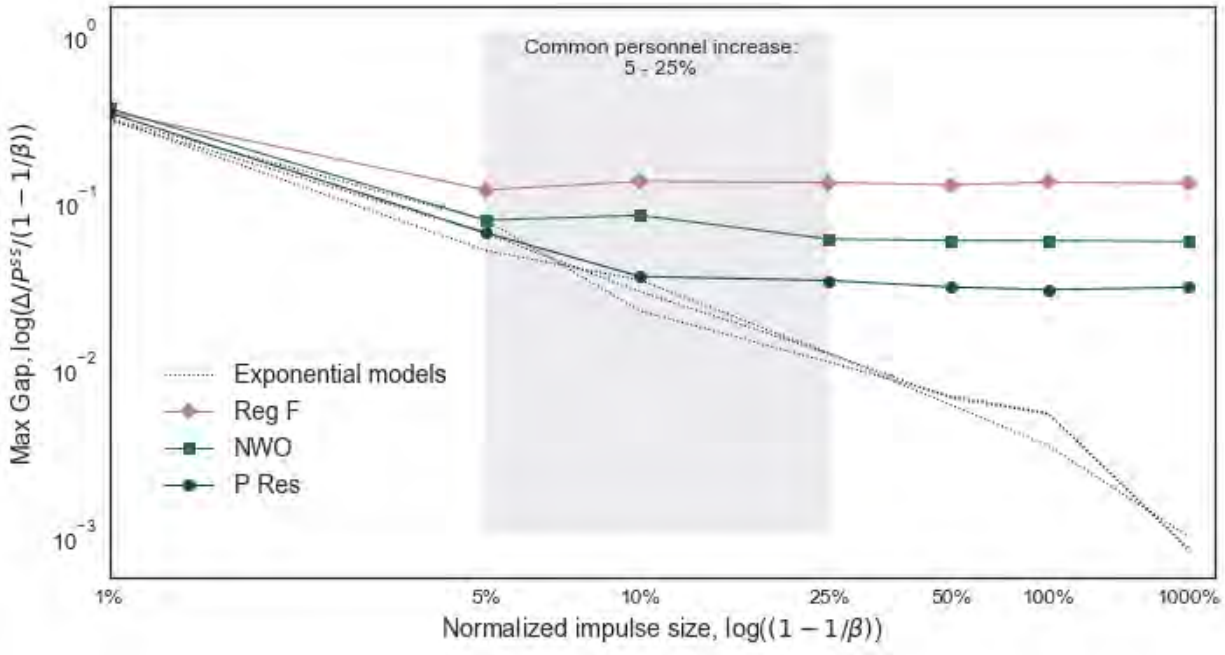

Figure 3: Maximum gap, Δ, between Markovian theory and simulated populations using empirical PDFs (solid marked curves) and equivalent exponential models (dashed curves) for P Res (blue), NWO (green) and Reg F (rose). Note that the maximum gap values shown have been normalized by P^{ss} and scaled by $(1-1/\beta)$. Impulse sizes, β, are labeled as their percent increase from P_0.

3.5 Transient time, t^*

Figure 4 illustrates the time for the simulated population, $P_{sim}(t)$, to reach steady state P^{SS}, as characterized by the transient time t^*, within a particular threshold defined by the e-folding time (Section 2.3). Results are scaled by Markovian theory, t^*_{theory} from Equation 3. Each panel displays the results according to a different e-folding time for $k=4$ (upper), $k=3$ (middle), and $k=2$ (lower). Note that the exponential models follow the Markovian theory as expected, with deviation attributed to residual stochastic noise. The empirical models, however, deviate from the Markovian path, with faster transient times in the case of Reg F and NWO, and slower transient times for P Res. The fractional difference between t^* and t^*_{theory} increases with increased impulse, and in the case of the Reg F up to twice as fast as theory ($t^* = 0.5\ t^*_{theory}$) for $\beta = 10$ at the most stringent threshold of $k=4$. This is also seen in Figure 2 where the Reg F simulated population (rose solid curves in all panels) reaches P^{SS} well before the exponential model and Markovian theory. While less dramatic, this is also seen in the "middle ground" case of the simulated NWO population whose analogous transient time is 1.5 times faster ($t^* = 0.74\ t^*_{theory}$) and the near-exponential case of the P Res which is a fifth slower ($t^* = 1.18\ t^*_{theory}$).

As discussed in Section 2.3, note that for small values of β and small values of k, measured t^* will be zero as P_0 is above the set threshold. Indeed, a 1% impulse ($\beta = 1.01$) is below β_{min}, for all k considered here according to Equation 4. Note that slightly above β_{min} we expect stochastic noise to drive t^* towards zero; this can intuitively be understood as the population step size will be small relative to fluctuations, and so fluctuations will tend to prematurely push the population above threshold.

The annual frequency at which we measure $P_{sim}(t)$ will introduce a small positive bias of roughly $0.5/\langle t \rangle$ when measuring the transient time. This is a small bias for our sub-populations, as $\langle t \rangle$ is in the 6–15 year range (Figure 1), and will be counteracted by the negative bias due to fluctuations triggering threshold detection early. Looking at simulated results for the exponential models in Figure 4 we see a small

negative bias relative to Markovian theory, which suggest that the stochastic fluctuation bias dominates this frequency bias.

4 DISCUSSION

4.1 Utility and adequacy of a Markovian assumption

The Markovian assumption is attractive as it makes analysis tractable. As examples, in the CAF the steady state rank profile for occupations is determined by solving a Markov model in an Excel based software tool (Boileau 2012) and population trajectory forecasting is performed in another Excel based software tool (Okazawa, Straver, and Arseneau 2018): both tools are facilitated by making a Markovian *assumption*. *Verification* (that software components operate as intended) can be made by comparing output with analytic predictions, as in Henderson and Bryce (2019) where a Discrete Event Simulation developed for modeling RCN occupations is precisely verified. *Validation* (that models adequetly capture reality) requires comparison of empirically rich representations against model predictions—the simple scenario outline here, where a population is subject to a sudden intake increase, allows us to isolate and consider the implications of making a Markovian assumption.

By exposing an empirically described population to shocks (quantified by β in our notation) that are reflective of "*what if*" questions to be asked, one can determine the error associated with making a Markovian assumption. For the specific groups considered here we find moderate errors in population level ($\approx 5-10\%$ as explored in Section 3.4) which, depending on the application, either validate simple Markovian models as being sufficiently accurate or highlight a need for more in-depth and refined models.

We should note that modeling the steady state is much more robust than modeling dynamics, as the average survival time (equivalently, the attrition rate parameter, α), rather than details of the shape of the empirical survival PDF, determines the population levels—thus a properly parameterized Markovian model will capture the correct steady state. See the right-hand panels of Figure 2 where the approach to steady state is not captured by a Markovian model, but the final steady state is. The implication is that (steady state) occupation structure models (e.g., those described in Boileau (2012)) are more robust to assumptions than dynamic models (e.g., Okazawa, Straver, and Arseneau (2018)). Approach to steady state can be faster, or slower, than a Markovian model implies, depending on the details of the empirical survival time distribution, as can be seen in Figure 2 and Figure 4.

If questions of "when" an occupation or target group shortage will be rectified are asked then more nuanced models, such as Discrete Event Simulations, may be required. As the costs of performing nuanced simulations are non-trivial, the approach outlined here is an attractive means of probing when paying the costs involved are warranted; by explicitly testing the modeling error associated with making a Markovian assumption one can ascertain if a simple memoryless model has sufficient fidelity (or not).

4.2 Limitations and intent: validation

We have deliberately ignored many important factors in setting up our simplified numerical experiments. For example, in practice we often are interested in modeling an occupation in the CAF for which ranks and transitions between them are important, while here we consider a single component. We also use a simple histogram with yearly bins to capture the empirical PDF and do not fully consider the modeling and estimation errors associated with that choice. A sudden intake step increase is also highly stylized, as in practice intake fluctuates year-to-year and any increases over time are limited by institutional capacity which will tend to impose slow changes. We purposely strip these details away in order to isolate and identify the modeling error associated with making a Markovian assumption. By isolating this single factor we can probe the adequacy of a Markovian assumption. If population trajectories determined with empirical PDFs do not follow Markovian predictions, within some set error tolerance, then the simplification obtained is not justified. The intent is to allow a basic validation framework, which will highlight when more nuance is required and when Markov models are sufficient, *not* to create a simulation schema. If validation fails,

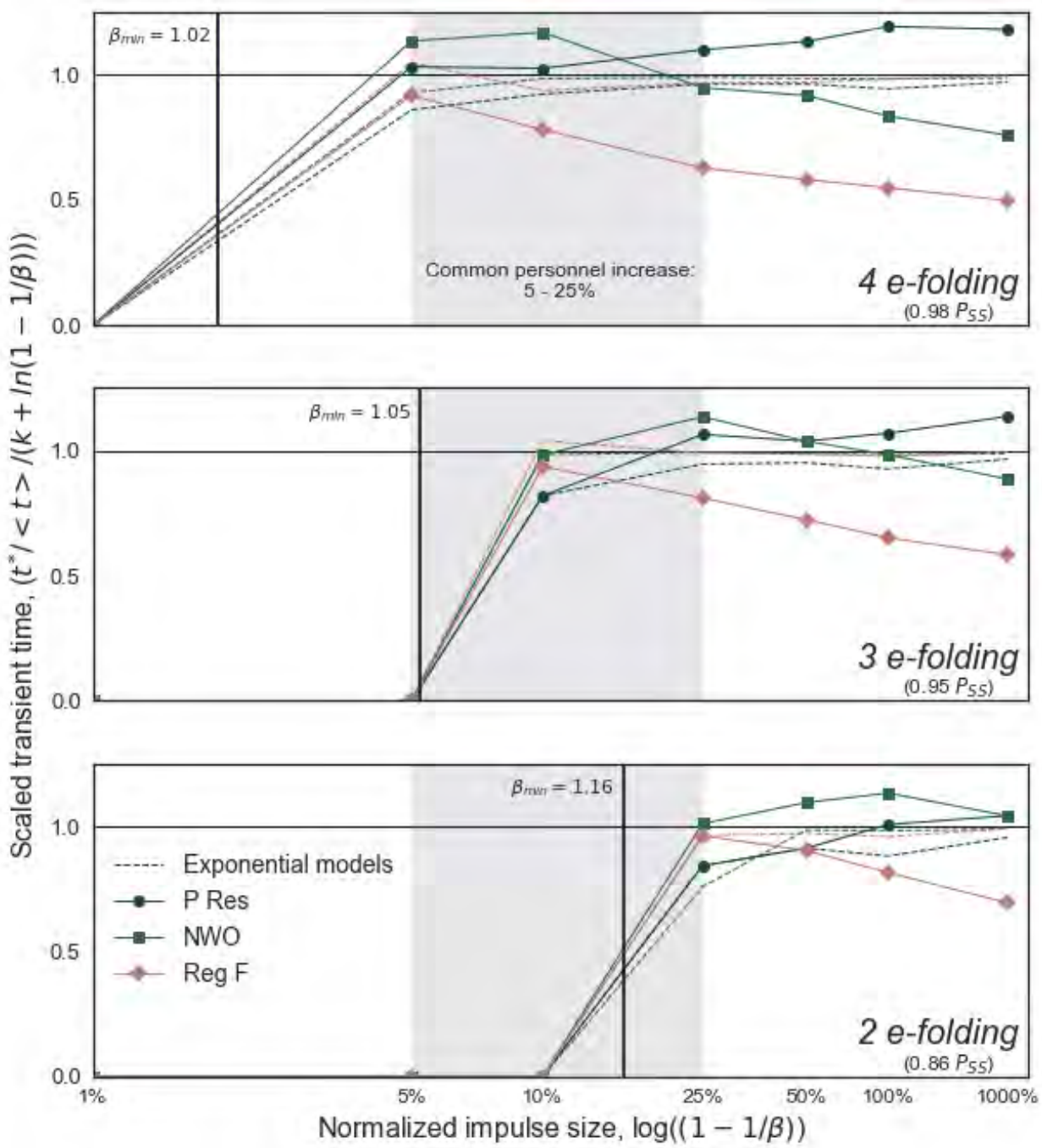

Figure 4: Transient time, t^*, for for several e-folding thresholds for k: (*upper panel*) $k = 4$, (*middle panel*) $k = 3$, and (*lower panel*) $k = 2$. Results shown are from empirical PDFs (solid marked curves) and equivalent exponential models (dashed curves) for P Res (blue), NWO (green) and Reg F (rose). Note that, for each group considered, t^* has been scaled by the Markovian theory, t^*_{theory}, with unity (black curve) and β_{min} (Equation 4) shown for comparison. Impulse sizes, β, are labeled as their percent increase from P_0.

then a simulation schema, such as Discrete Event Simulation, with all the important aspects and nuances must be created.

The goal is to get a practical sense of the error associated with making a Markovian attrition assumption. In one sense, the Markovian model is quite good as the maximum error is moderate, even for PDFs visually far from being exponential (e.g., Regular Force). This can be qualitatively understood from a steady state perspective. At steady state the average survival time, not the full details of the distribution shape, is important. This suggests that "small" perturbations from steady state will have weak effects, with Markovian trajectories being a good representation. As the intake is a linear forcing term which drives the ordinary differential equation describing a population, it is perhaps not surprising that perturbations/forcings appear to have limited ("small") effects on the quality of predictions. That said, the error levels involved are large enough that they can become meaningful. A $5-10\%$ error level is a good "engineering approximation", but for some questions is insufficiently precise. As one example, the time for an occupation to become healthy will be sensitive to such error levels which might justify more fulsome modeling and analysis than a simple Markovian forecast.

5 CONCLUSION

In this work we have aimed to get a practical sense of the modeling error associated with making a Markovian (memoryless) attrition assumption. Real world populations are not expected to be well described by an exponential and so such modeling error is anticipated, the question is not "*if*" but rather "*how much?*". Looking at several different sub-populations in the CAF we highlight examples which range from being visually close to an exponential (Reserve Force), to being quite distinct (Regular Force), and something in-between (Naval Warfare Officers); see Figure 1. By exposing a population to an intake shock we create a simple setup to measure response, allowing us to compare Markovian and empirical population trajectories. We find that, for the groups considered, error is visually distinctive (Figure 2) with empirical trajectories differing from Markovian predictions. In the groups investigated moderate error is found in population levels (roughly in the range of 5–10%; see Figure 3) during the transient response to the intake shock, and we find the measured transient times to be notably different (Figure 4). For example, the Regular Force reaches equilibrium roughly $2\times$ faster than Markovian predictions while the Reserve Force, with survival times visually well described by an exponential distribution, have a meaningful delay of about a fifth of the Markovian time. The simple approach taken here allows probing modeling error via numerical simulations, requiring only the empirical probability distribution of a population's survival times. Two measured parameters, the transient time and the maximum gap between empirical and equivalent Markovian population trajectories, summarize the modeling error associated with a Markovian assumption, allowing validation and a practical understanding of the error.

REFERENCES

Bartholomew, D., and A. Forbes. 1979. *Statistical techniques for manpower planning*. 1st ed. Norfolk, UK: John Wiley & Sons.

Bartholomew, D., A. Forbes, and S. McClean. 1991. *Statistical techniques for manpower planning*. 2nd ed. Chichester, West Sussex: Wiley.

Boileau, M. 2012. "Workforce Modelling Tools used by the Canadian Forces". In *Proceedings of The International Workshop on Applied Modelling & Simulation*. September 24th-27th, Rome, Italy, 18-23.

Guerry, M.-A., and T. De Feyter. 2009. "Markovian Approaches in Modeling Workforce Systems". *Journal of Current Issues in Finance, Business and Economics* 2(3):1–20.

Henderson, J. A., and R. M. Bryce. 2019. "Verification Methodology for Discrete Event Simulation Models of Personnel in the Canadian Armed Forces". In *Proceedings of the 2019 Winter Simulation Conference*, edited by N. Mustafee, K.-H. G. Bae, S. Lazarova-Molnar, M. Rabe, C. Szabo, P. Haas, and Y.-J. Son, 2479–2490. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Higham, N. J. 2008. *Functions of Matrices: Theory and Computation*. Philadelphia: Society for Industrial and Applied Mathematics.

Hoel, P. G., S. C. Port, and C. J. Stone. 1971. *Introduction to probability theory*. Boston: Houghton Mifflin.

Jnitova, V., S. Elsawah, and M. Ryan. 2017. "Review of simulation models in military workforce planning and management context". *The Journal of Defense Modeling and Simulation* 14(4):447–463.

Nguyen, V. 1997. *Analysis of the U.S. Marines Corps' steady state Markov model for forecasting annual first-term enlisted classification requirements*. Ph.D. thesis, Naval Postgraduate School, Monterey, California.

Okazawa, S., M. Straver, and L. Arseneau. 2018. "Occupation Promotion Recruitment and Attrition Model (OPRAM): A Replacement for the Royal Canadian Navy Long-Range Planning Model". Technical Report DRDC-RDDC-2018-R273, Department of National Defence, Ottawa, Canada.

Serré, L. 2019. "Managing the Personnel Resources of a Military Occupation: Attrition Forecasting and Production Planning". In *Royal Canadian Air Force defence economics*, edited by R. Fetterly and B. Soloman, 242–260. Ottawa, Canada: Department of National Defence.

Škulj, D., V. Vehovar, and D. Štamfelj. 2008. "The Modelling of Manpower by Markov Chains—A Case Study of the Slovenian Armed Forces". *Informatica* 32(3):289–291.

Wang, J. 2004. "A review of operations research applications in workforce planning and potential modelling of miltary training". Technical Report DSTO-TR-1688, Defence Science and Technology Organisation, Salisbury, Australia.

Young, A., and G. Almond. 1961. "Predicting Distributions of Staff". *The Computer Journal* 34(4):246–250.

Zais, M., and D. Zhang. 2016. "A Markov chain model of military personnel dynamics". *International Journal of Production Research* 54(6):1863–1885.

AUTHOR BIOGRAPHIES

ROBERT M. BRYCE is a defence scientist at the Department of National Defence. Holding a PhD in Physics from the University of Alberta, Dr. Bryce's research interests lie in explanatory data analysis, using theory, modelling, simulations, machine learning, graphing, and other tools. Contact Dr. Bryce at Robert.Bryce@forces.gc.ca.

JILLIAN A. HENDERSON is a defence scientist at the Department of National Defence. Dr. Henderson holds a PhD in Astronomy from the Universidad Nacional Autonóma de México. Dr. Henderson's research interests lie in numerical modeling and simulation with a current focus on strategic military personnel management. Dr. Henderson can be contacted via email: Jillian.Henderson@forces.gc.ca.