

A Study on Efficient Modelling in Higher Education Academic Workforce Using Simulation

Nethal K. Jajo and Shelton Peiris

Abstract — This paper shows how the Operations Research (OR) tools (in modelling and simulation) can be modified, applied in planning and their understanding of any long-term impacts due to sudden policy changes. The proposed approach is particularly useful to investigate the movements of the university academics and their impacts on changes in research, research funding, teaching and services as they are the integral parts of the career at any level. We argue that the Discrete Event Simulation (DES) approach can be used to model such dynamics in Higher Education Academic Workforce Model (HEAWM) and show that it can provide a comprehensive projection of future requirements within the context of career progression. Consequently, this HEAWM allows universities to interrogate factors influencing the academic workforce planning as this process often requires the new attributes to be tracked in the model which is difficult with other OR models. It is shown that this approach is easy to apply via DES and creation of the corresponding HEAWM provides better understanding of the factors that will influence the future workforce than the existing results.

Key words — Arena, academic workforce, discrete event simulation, dynamical systems, modelling.

I. INTRODUCTION

OR techniques have been widely used in simulating the dynamics of workforce systems. Since the late 1950s, much of the work has been done on developing models for manpower systems which are being in use for the purpose of planning, see for example [1], [2], [6] and references therein. Markov Chain Models, trend and regression analysis methods can also be used in academic manpower planning in certain occasions, see [2]–[4]. Markov chain modelling of workforce systems including Markov models are linear in the sense that they do not incorporate the feedback mechanism such as the influence of promotion possibility on wastage [5]. However, there is a requirement on the number of people in each class, for example, academic level, age, research area, etc., must be large for the transition probability estimates to be stable. Therefore, classes with staff numbers fewer than 100 are not suitable for aggregate-based techniques, such as Markov chain as discussed in [6].

An alternative way to model workforce systems with feedback mechanisms is to use the supply chain modelling techniques [6]. Higher education academic manpower has no exception and constitutes a crucial sector of human resource development and management in many countries. DES and System Dynamics (SD) are among the techniques that have been most established and widely used OR tools to simulate the dynamics changes in workforces, see [7], [8] for details.

In order to determine a suitable tool, there has been a considerable number of studies conducted and compared as given in [8]–[14]. Reference [15] initiate an alternative tool by combining both tools in forming one. These studies suggest that the selection of the appropriate tool for simulation will depend on the problem under investigation.

An approach given in [7] on modelling higher education academic manpower shows that DES and SD are the most tempting tools in such modelling problem. Many higher education sectors (HES) require their workforce models to be able to accommodate many new attributes and to adopt any future organization policy changes. Adding such new attributes to an SD model will increase the complexity and the computations within the model, while DES model can accommodate many new attributes without the two mentioned issues in SD modelling. Therefore, we expect DES provides a persuasive solution to address most of the issues facing with higher education academic manpower modelling. DES is widely used in modelling the operation of a system as a discrete sequence of events in time. It is assumed that each event occurs at an instant in time and marks a change of its state and no change in the system. However, the simulation time can jump to an occurrence of the next event during the time progression [16].

Over the past few decades, many significant papers on DES have been published with emphasis on various matters of interest with applications. These papers include stress testing, evaluating potential financial investments, and modelling procedures in processes related to various industries. In addition, DES has been used in logistics and supply chain management [17], modelling health-care issues, health economic evaluation and the analysis of human resources and policies, see [18]–[20]. Although there is a considerable diversity in these studies, little attention has been paid in research related to changes in academic policies through DES, especially in the HES. Clearly, such changes have a very significant impact in long-run for the development of a country. As a result, any negative consequences due to delaying such studies can be difficult to correct. Therefore, any such changes to HES should be carefully planned, investigated, and monitored by a dedicated group. In the current literature, this was investigated using other technical methods but not ORs. Reference [21] used a fielded survey of human resources specialists to explore the efforts used in addressing the aging of academic workforce, while [22] uses qualitative study via questionnaire to explore the impact of introducing teaching focus academic role on the reshaping of academic workforce. Reference [23] re-analyzed an existing qualitative data to explore how academic

individuals are adapting to less certain career environment.

Based on the proposed simulation experiments, we investigate the corresponding performance and benefits resulting from changing the policies and providing opportunities for learning outcomes. With that view in mind, this paper is devoted to a study based on modelling and simulation as they are essential for planning and helping to understand relevant long-term effects and impacts.

II. MODELLING PLATFORM

The modelling platform that will be used in this paper is DES using Arena software from Rockwell Automation Technology, Inc. version 16.10. DES works at an individual level, allowing full representation of the academic history of all and interacting with specific individuals. Arena can accommodate discrete, continuous, or hybrid systems and is widely used for workforce and training pipeline modelling by many nations' organizations working in strategic human resources. Selecting which system to apply is straightforward when building the model and the software provides graphical and numerical displays as well as computer animation of the system. It uses icons to represent individual staff members (entities) moving through the system with attributes (such as academic rank, faculty, age, and gender) assigned to each entity using a very simple and straightforward procedure. It simply requires stating the attributes in the definition of the entity in the model along with any code desired to properly indicate the attributes. Unlimited attributes can be tagged to an entity representing different individual identities and the overall entities can be merged or streamed based on their attributes [7]. The animation of personnel flow and on-screen movement in an Arena model can be a valuable tool in providing increased understanding of the flow.

A. Modelling and Design Principles

The primary focus of this work is to construct HEAWM to

represent all significant transformations to an individual member at a certain academic level. To begin the process, consider the initial academic population when inputted into HEAWM, including all relevant attributes of each member. When the model progresses, some of the members are projected as leaving the university, and new "generated" individuals are included (see Fig. 1 **Error! Reference source not found.**). All changes to members are simulated on a yearly cycle and each member is examined (one by one) as the model progresses through the simulation with changes in one year. This process is repeated to project the number of publications, number of citations, amount of funds (at University/Faculty) and academic levels (with number of years of service). Since the model changes are randomly dependent, several independent replicate runs are carried out to validate the model outputs statistically.

New entries to the model are generated to accommodate new recruits. These entries will be inputted into the model via an Excel sheet including their academic level, faculty, and any other attributes. The new entries' number of publications, citations and funds will be read into the model from the same Excel sheet using statistical distributions.

In the HEAWM, Fig. 2, schematically describes the transformations or changes that can occur to individuals at a given academic level. Some individuals will be promoted to the next higher academic level, some others will be leaving the university (voluntary/involuntary separation) and the rest will be staying at their academic level. Typically, the separations are balanced by individuals who are promoted into an academic level from one academic level lower (in most cases) or who enter from the generated population. Moreover, the increment in the number of publications, citations and funds, time at the academic level, time at the university, and age for each academic will be applied for each year.

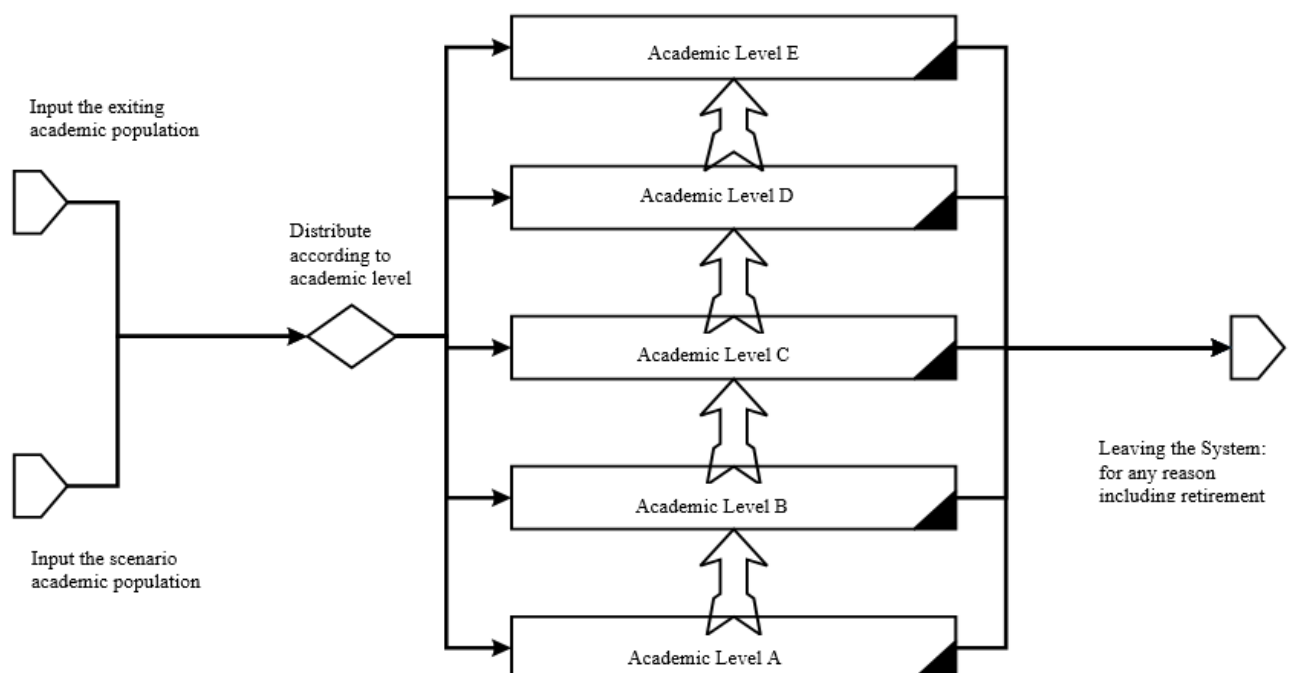


Fig. 1. Overview of the HEAWM system.

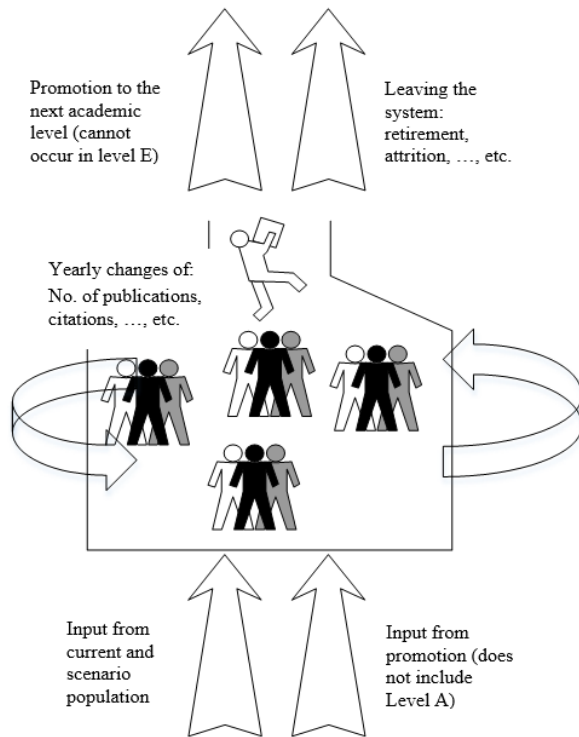


Fig. 2. Changes at a given academic level.

B. HEAWM

The modelled academic individuals in the HEAWM are represented as “entities” and characterized by several attributes including the faculty, academic level, age, gender, and years at university, time at current academic level, number of publications, citations, and funds. The top-level of the HEAWM is illustrated in Fig. 3, which contains seven sub models: the two to the left are the current and the recruitment populations and the five sub models to the right are the sub models related to each of the five academic levels.

The model will start by reading the current academic population, assign attributes and variables to each academic and then will distribute the population to the academic level submodels using Decide Block according to the academic level of each member.

Members can be thought of as flowing through this hierarchy as they progress through their careers. Each of the academic level submodels contains many blocks. These blocks are Arena modules modelled within each of the academic level submodel, for example, see Fig. 4 for the SENIOR LECTURER submodel. The changes at a given academic level are occurring within the five sub models.

The bottom portion of Fig. 3, represents the yearly sequence of events and signals. As the simulation proceeds from year to year, these yearly sequences of events change the status of members (e.g., promotion, attrition and recruiting) according to rules which are defined by career progression policies and the occupational organizational structure.

As each “year” is modelled in a simulation run, there is first an “initialization” submodel which resets the values needed to reflect the beginning of a new model year. In this submodel and in the first step, the model will read initial information, write the headers for the tables in Arena output text files and produce a specific number of people to stand by as new entries to the workforce. In the second step, the model will read information about the existing population members, which includes their attributes. The leaving patterns by the faculty, academic level, and age, arbitrarily assumed to represent the faculties’ historical trends which are used to sample those members who leave due to voluntary attrition. Seven age groups were considered for each academic level and faculty. These groups are: 22-29, 30-39, 40-49, 50-54, 55-59, 60-67, more than 68 and the retirement age is set as 67+.

Higher Education Academic Workforce Model (HEAWM)

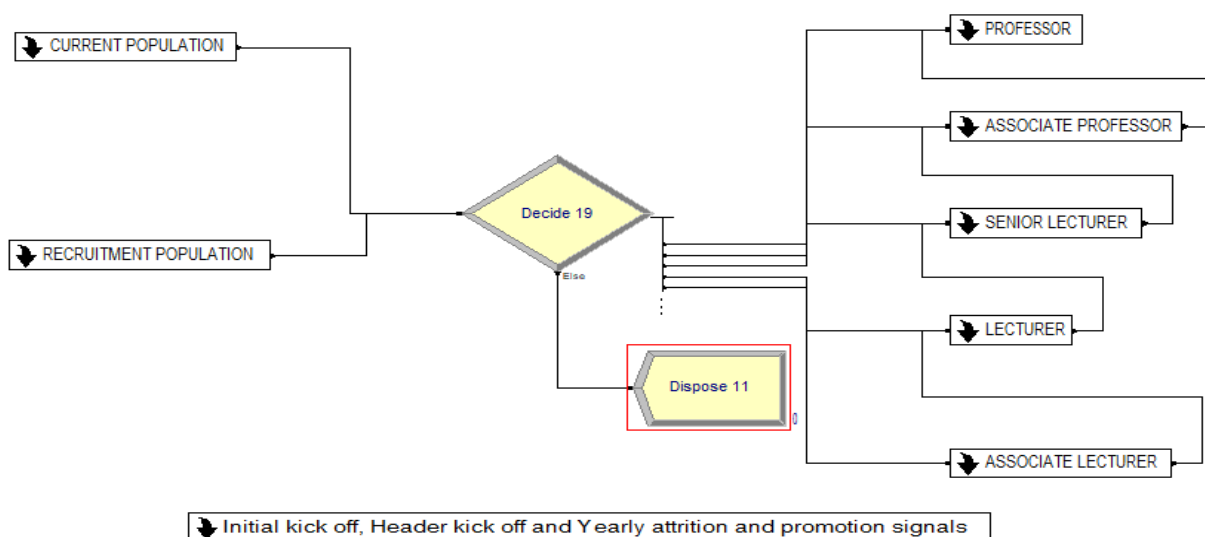


Fig. 3. HEAWM Top-level.

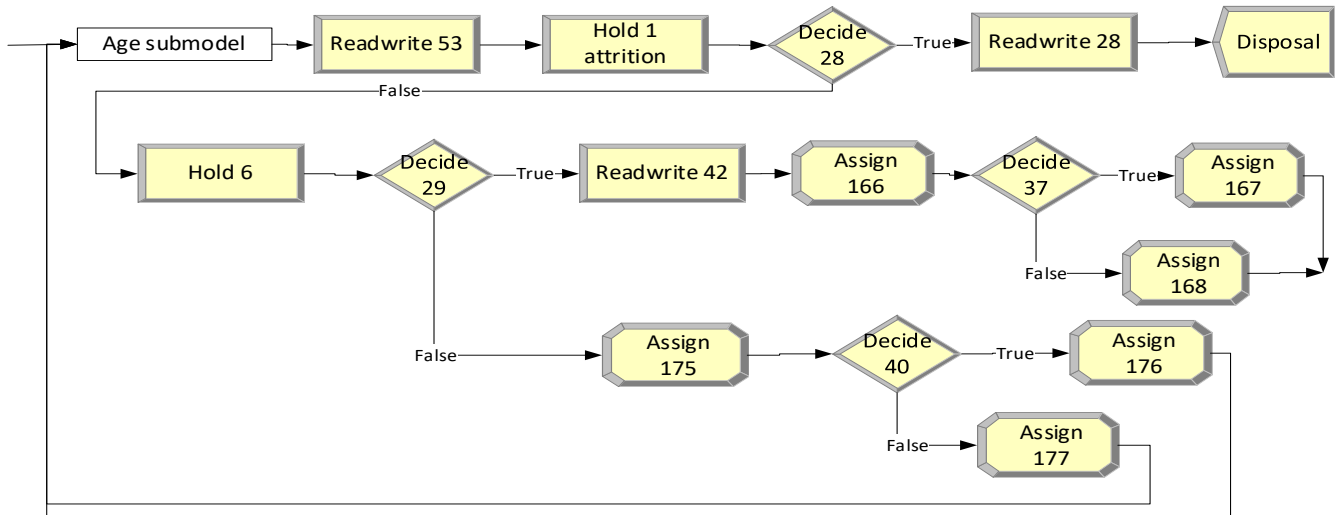


Fig. 4. Senior lecturer submodel.

Promotion criteria includes meeting any specified minimum TiL (Time in Level), number of publications and citations and funding requirements. The values used in this simulation are arbitrarily selected for each academic level and faculty. Minimum TiL values are given in **Error! Reference source not found.**, while the minimum number of publications, citations and funds are given in TABLE II (there is no promotion from the academic level E). An individual academic who completes a specified TiL will be automatically promoted to a higher academic level (sometimes more than one level above the current level) based on specified rules that will be controlled by the user via an Excel sheet which includes number of publications, citations, and funds.

TABLE I: MINIMUM TiL (IN YEARS) REQUIRED FOR PROMOTION

Faculties	Academic Level			
	A	B	C	D
Faculty F1	2.59	4.62	5.02	4.07
Faculty F2	2.04	4.54	5.91	4.21
Faculty F3	1.91	2.55	4.77	4.6
Faculty F4	3.72	4.17	4.78	2.7
Faculty F5	3.03	3.15	4.56	4.68
Faculty F6	2.29	3.37	4.61	4.49

New entries to the model are generated to accommodate new recruits and rehires. The two entry types will be

generated, based on historical statistics for each academic level and faculty, using an Excel sheet.

Moreover, publications, citations, and funds statistics in **Error! Reference source not found.** will be assigned as a tributes to each of the new entries using an Excel sheet.

III. A SIMULATION STUDY

Due to the confidential nature of the academic workforce data, we were unable to acquire the real data. Simulation-generated data were used in this paper with a confirmation that the HEAWM is applicable to real data. An arbitrary example of a middle-size university's academic population of 2000 academics was randomly generated from six Faculties named F1, F2, F3, F4, F5 and F6 and five academic levels A (lowest), B, C, D and E (highest). For each academic level, the age of each member was attached arbitrarily using a triangular distribution, see TABLE III. Production outputs in research volume, number of citations and funds were added randomly and arbitrarily to each member of the current population based on their academic level and faculty (see **Error! Reference source not found.**). Finally, TiL and Years at the University (YaU) were attached to each member randomly, based on academic level and faculty.

TABLE II: ACADEMIC'S PRODUCTIONS BY FACULTY AND ACADEMIC LEVEL

Academic level	Variables	Faculties					
		F1	F2	F3	F4	F5	F6
A	Publications	0	1	1	0	1	1
	Citations	0	0	1	0	10	1
	Funds	966	1913	5764	4895	36885	22067
B	Publications	1	1	3	2	1	2
	Citations	0	0	3	2	2	8
	Funds	6491	8424	89335	21221	38894	153740
C	Publications	2	2	4	5	2	3
	Citations	1	0	4	3	11	7
	Funds	8277	17849	93771	89919	100114	132746
D	Publications	3	2	7	7	5	5
	Citations	1	0	12	0	13	10
	Funds	35765	47368	260954	199184	216799	445529
E	Publications	5	4	14	13	12	11
	Citations	18	5	68	56	135	98
	Funds	97780	187363	865855	470886	1101057	1295243

IV. MODEL VALIDATION

The model validation may be completed using a client's acceptable marginal error. However, due to the use of simulated data, the validation of the model will be executed by testing the sustainability of the model and its robustness.

TABLE III: AVERAGE AGE DISTRIBUTION

Triangular distribution parameters	Academic Level				
	A	B	C	D	E
Minimum	23	28	33	38	43
Average	25	30	35	40	45
Maximum	27	32	37	42	47

Fig. 5 shows the distribution of the population means and standard deviation for the twenty replications. ANOVA table in TABLE IV shows that there are no significant (at 5%) differences in the population means of the replications. This validates the model.

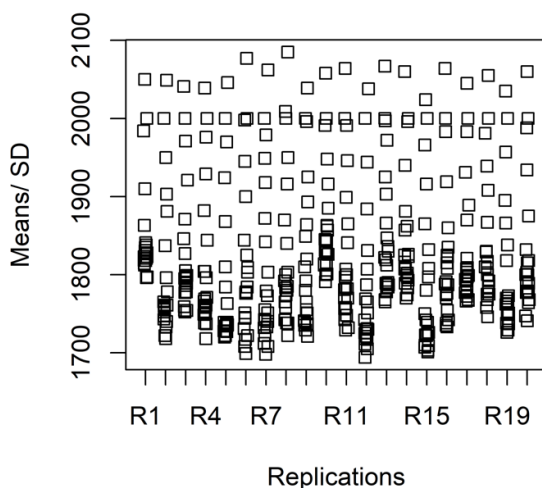


Fig. 5. Mean and SDs of the population for the 20 replications.

TABLE IV: ANOVA TABLE FOR COMPARING REPLICATIONS' MEANS

	Df	SumSq	Mean Sq	F value	Pr(>F)
Replications	19	191460	10076.9	1.1544	0.2941
Residuals	400	3491524	8728.8		

V. MODEL OUTPUTS AND ANALYSIS

The HEAWM's output consists of detailed results (as text files) produced during the simulation run, encapsulating low level specifics of numerous events of various kinds. These data are generated once and will be used multiple times in the analysis. Fig. 6 shows the academic average production in research publications, citations, and funds by academic levels for Faculty F3 as an example.

Meanwhile, exploration of the academic population movement using the same example, Faculty F3, shows that average departures of academics in level B and C (E is not considered because of the retirement age affecting the departures) are a higher than other levels, see Fig. 7, and this may be caused by long TiL before promotion to the next level.

To explore this issue, for the same example, in details, Fig. 8 shows the average TiL for level C is about 6 years prior to being promoted to level D. This may cause those academics who spent this long in their academic level selecting to depart from the university to another university but with a higher academic level. In addition, the correlation matrix between

average number of publications, citations, and funds for the same example shows the highest correlation between average citations and funds (see Fig. 9).

Moreover, even though the data used in this analysis are simulated, not real, providing some statistical analysis may help in understanding the method available for analysis.

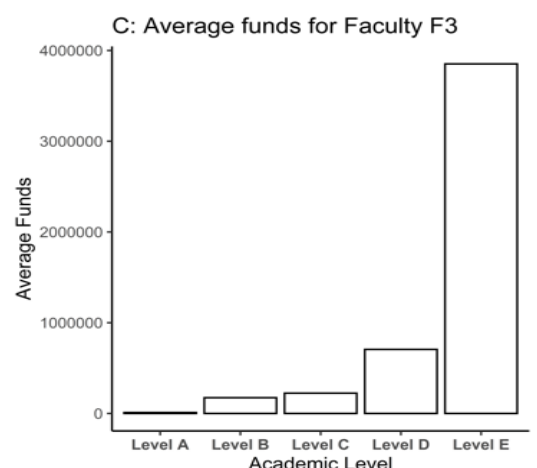
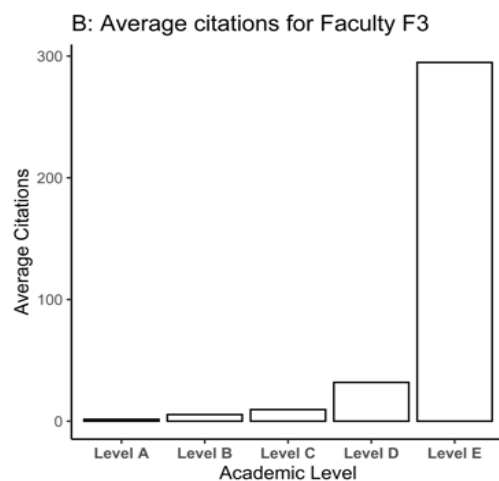
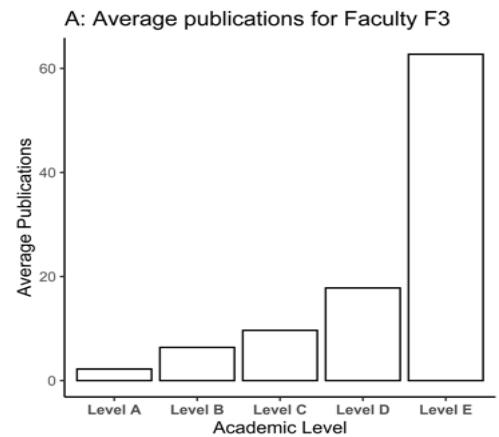


Fig. 6. Academic production for Faculty F3 by academic level.

For example, to investigate the variables that may cause an academic to exit a university, the following methods may apply: In classification or regression problems, we often have many potential explanatory variables.

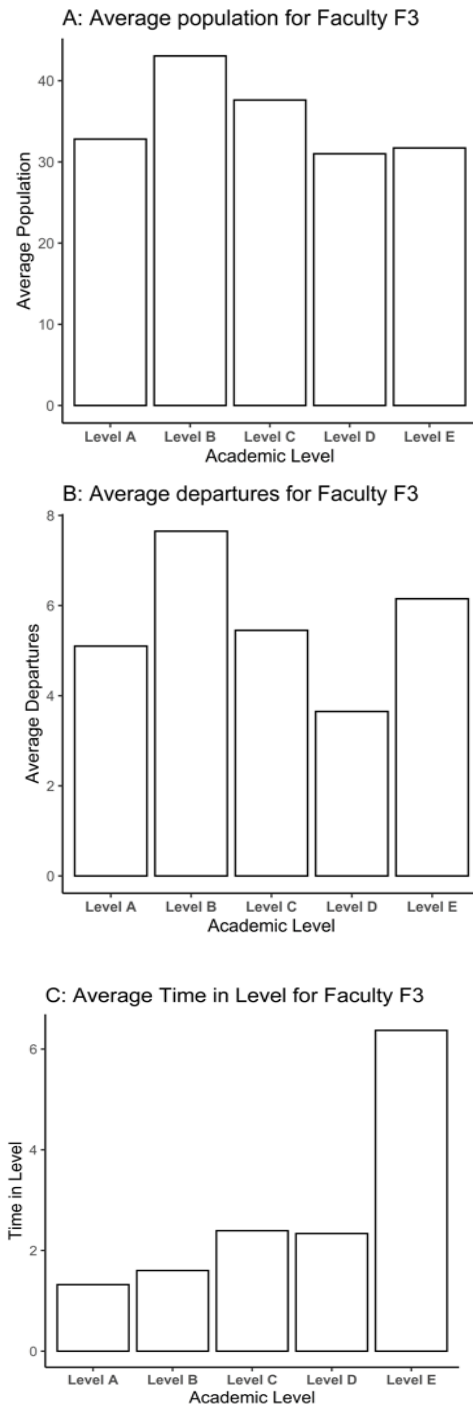


Fig. 7. Academic population movement.

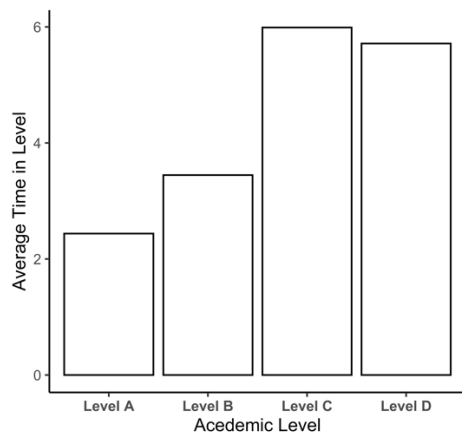


Fig. 8. Average TiL for Faculty F3 prior to promotion.

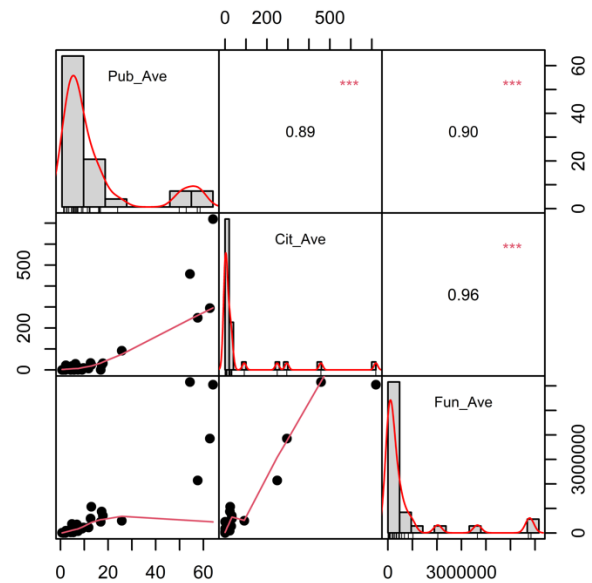


Fig. 9. Correlation matrix for publications, citations, and funds for Faculty F3.

Using all the variables is not likely to be the best solution because combinations of two or more of these variables can be highly correlated [24]. Akaike Information Criterion (AIC) is one of the methods that can be used. This approach can be applied in nonlinear regression, time series, and other settings. Gong was the first to apply the bootstrap as a tool to identify the most important variables to use in a model, see [1] and [25]. In very loose words, the AIC, as summarized by [26], is a way of selecting a model from a set of models. The chosen model is the one that minimizes the Kullback-Leibler distance between the model and the truth. It is based on information theory, but a heuristic way to think about it is as a criterion that seeks a model that has a good fit to the truth but few parameters. In [26] the AIC is given by:

$$AIC = n \times \ln(RSS) + 2K \quad (1)$$

where K is the number of parameters, n is the number of observations, and RSS the residual sum of squares which minimize over K . AIC scores are often shown as ΔAIC scores, or difference between the best model (smallest AIC) and each model (so the best model has the ΔAIC of zero) using stepwise selection process.

TABLE V: ΔAIC FOR DIFFERENT PARAMETERS

Parameters	ΔAIC
Funds	5.02
TiL	3.39
Citations	3.28
Pubs	1.37
YaU	0.97

One way to make the stepwise selection process more reliable is by bootstrapping it. The process can be broken down into the following steps: select a random subset from the data set; run stepwise selection on this subset; record the features that were selected; repeat the previous steps many times. The result is that now we can see how many times a specific feature (predictor variable) gets selected. This

bootstrapping technique creates random datasets from the given dataset (drawing with replacement). We applied bootstrap with AIC for 1000 samples to determine the relative importance variables (see Fig. 11).

As can be seen from TABLE V, Fig. 10 and Fig. 11, based on simulated data, years at a university is the most important factor in the departure of the academics. Not different methods in bootstrap will have different relative important variables for departures. The selection of the best method will depend on the real data. Hopefully, this little extra step 'bootstrap' would give you some insight as to which predictor variables could be most useful.

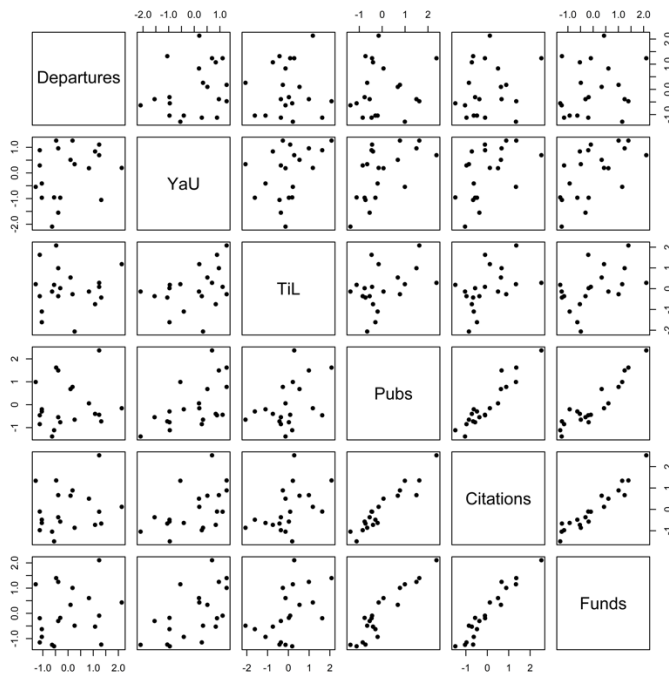
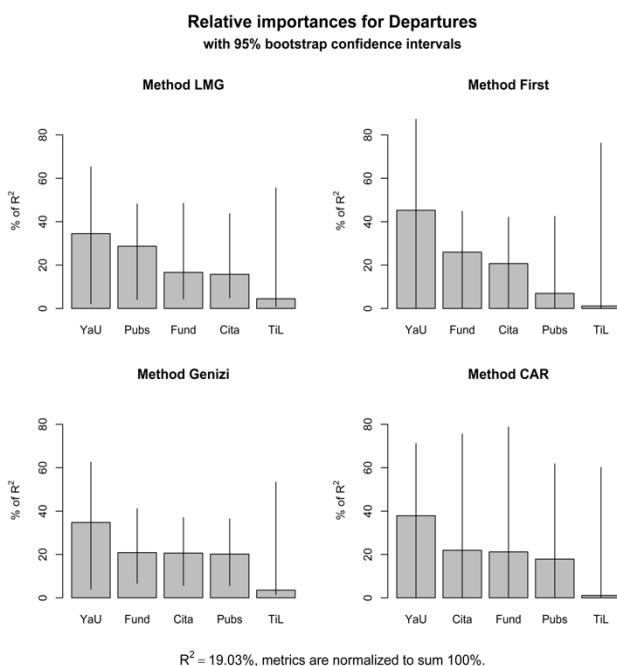


Fig. 10. Pairs plot for departures.



$R^2 = 19.03\%$, metrics are normalized to sum 100%.

Fig. 11. Δ AIC for different parameters.

VI. CONCLUSION

This paper applied the DES methodology to model the HEAWM framework. Using simulated data, the paper has shown that it can provide a comprehensive projection of future requirements within the context of higher education academic career progression with minor refinements to the process. It has been noted that a significant effort is required in HEAWM analyses-based data collection and validation the quality of data. This issue can be resolved by collaborating closely with the Human Resources departments while maintaining the confidentiality within each organization and members.

REFERENCES

- [1] B. Efron and G. Gong, "A leisurely look at the bootstrap, the jackknife, and Cross validation," *Am. Stat.*, 37, pp. 36 – 48, 1983.
- [2] A. Soyihio, "Markov Chain application to academic manpower planning," *Vikalpa*, 9, 1, pp. 27 – 41, 1984.
- [3] G. O. Ogbogbo, G. U. Ebul and C. O. Aronu, "Prediction of academic manpower system of a Polytechnic institution in Nigeria," *Science Journal of Applied Mathematics and Statistics*, 1, 5, pp. 54 – 61, 2013.
- [4] M. Aref and M. Sabah, "Manpower Planning for Demand Forecasting of Faculty Members using Trend Analysis and Regression," *International Journal of Academic Research in Business and Social Sciences*, 5, 2, pp. 11-23, 2015.
- [5] J. Wang, "A review of operations research applications in workforce planning and potential modelling of military training," *DSTO-TR-1688*, 2005.
- [6] J. S. Edwards, "A Survey of Manpower Planning Models and Their Application," *Journal of Operational Research Society*, 34, 11, pp. 1031-1040, 1983.
- [7] N. K. Jajo, "The trade-off between DES and SD in modelling military manpower," *Management Science Letters*, 5, 4, pp. 369-376, doi: 10.5267/j.msl.2015.2.002, 2015.
- [8] A. A. Tako and S. Robinson, "Model development in discrete-event simulation and system dynamics: An empirical study of expert modellers," *European Journal of Operational Research*, 207, pp. 784 – 794, 2010.
- [9] J. Morecroft and S. Robinson, "Explaining puzzling dynamics: Comparing the use of system dynamics and discrete event simulation," *Proceedings of the 23rd International Conference of the System Dynamics Society*, The System Dynamics Society, Boston, USA, 2005.
- [10] J. Morecroft and S. Robinson, "Comparing discrete event simulation and system dynamics: modelling a fishery," *Proceedings of the 2006 OR Society Simulation Workshop SW06*, Leanington, Spa, UK, pp. 137 – 148, 2006.
- [11] A. A. Tako and S. Robinson, "Comparing discrete-event simulation and system dynamics: users' perceptions," *Journal of the operational research society*, 60, 3, pp. 296 -312, 2009.
- [12] K. Chahal and T. Eldabi, "A multi-perspective comparison for selection between system dynamics and discrete event simulation," *International Journal of Business Information Systems*, 6, 1, pp. 4-17, 2010.
- [13] M. Tanha, D. Sajjadi and S. Shamala, "A discrete event simulator for extensive defense mechanism for denial of service attacks analysis," *American Journal of Applied Sciences*, 9, 6, 909 – 916, 2012.
- [14] B. Milczarek, "Review of modelling approaches for healthcare simulation," *Operations Research and Decisions*, 1, pp. 55-72, 2016.
- [15] S. C. Brailsford, S. M. Desai and J. Viana, "Towards the holy grail: combining system dynamics and discrete-event simulation in healthcare," *In Simulation Conference (WSC)}, Proceedings of the 2010 Winter*, pp. 2293-2303, 2010.
- [16] G. S. Fishman, *Discrete Event Simulation: Modeling, Programming and Analysis*, New York: Springer-Verlag, 2001.
- [17] R. B. Detty and J. C. Yingling, "Quantifying benefits of conversion to lean manufacturing with discrete event simulation: A case study," *International Journal of Production Research*, 38, 2, 2000.
- [18] L. P. Baldwin, T. Eldabi and R. J. Paul, "Simulation in healthcare management: a soft approach (MAPIU)," *Simulation Modelling Practice and Theory*, 12, pp. 541-557, 2005.
- [19] S. H. Jacobson, S. N. Hall and J. R. Swisher, *Discrete-Event Simulation of Health Care Systems*. In: Hall R.W. (eds) *Patient Flow: Reducing Delay in Healthcare Delivery*. International Series in Operations Research & Management Science, USA Boston, MA: Springer, 2006.

- [20] M. M. Günal and M. Pidd, "Discrete event simulation for performance modelling in health care: a review of the literature," *Journal of Simulation*, 4, 1, 2010.
- [21] B. Kaskie, M. Walker and M. Andersson, "Efforts to Address the Aging Academic Workforce: Assessing Progress Through a Three-Stage Model of Institutional Change," *Innovation High Education*, 42, pp. 225-237, 2017.
- [22] H. Flavell, L. Roberts, G. Fyfe and M. Broughton, "Shifting goal posts: the impact of academic workforce reshaping and the introduction of teaching academic roles on the Scholarship of Teaching and Learning," *Aust. Educ. Res.*, 45, pp. 179-194, 2018.
- [23] C. Whitchurch, "From a diversifying workforce to the rise of the itinerant academic," *High Educ.*, 77, pp. 679-694, 2019.
- [24] M. R. Chernick and R. A. LaBuddle, *An Introduction to Bootstrap Methods with Applications to R*, Hoboken, New Jersey: John Wiley & Sons, 2011.
- [25] G. Gong, "Cross - validation, the jackknife, and the bootstrap: Excess error in forward logistic regression," *J. Am. Statist. Assoc.*, 81, pp. 108 - 113, 1986.
- [26] K. P. Burnham and D. R. Anderson, *Model selection and multimodal inference: a practical information-theoretic approach*, New York: Springer, 2002.
- [27] M. Moffatt. (Aug. 27, 2020). An Introduction to Akaike's Information Criterion (AIC). <https://www.thoughtco.com/introduction-to-akaikes-information-criterion-1145956>.



Nethal K. Jajo completed his PhD in mathematics from Nankai University in China, in 1999. His major field of study was Mathematical Statistics and Probability Theory.

He is holding three positions within the University of Sydney, Australia: modelling and projection analyst at the DVC-Research, Research Portfolio; Honorary affiliate, School of Mathematics and Statistics; and sessional lecturer at the Discipline of Business analytics, Sydney Business School. He

tutored at The School of Mathematics and Statistics-University of New South Wales and lectured at Western Sydney University, The Northern Consortium of British Universities and Macquarie University. He also worked as mathematics modeler at the Australian Department of Defense.

Dr Jajo is a member of Australian mathematical society and Australian and New Zealand statistical society with more than 35 publications including books, book chapters, conferences and referred articles. He is also at the editorial board of International Journal of Statistics & Economics, International Journal of Ecological Economics and Statistics and American Open Statistic and Mathematics Journal.



Shelton Peiris did his PhD at Monash University, Melbourne, Victoria, Australia. His research interests are on Statistical Analysis of Time Series Applications in Financial Econometrics, topics in Mathematical Statistics and statistics teaching/ education. His recent academic visits include the Department of Statistics and Actuarial Science, Simon Fraser University, Burnaby, Vancouver, Canada (Fall 2019) and the Faculty of Economics, Soka University, Hachioji, Tokyo, Japan (Spring 2015).