25th International Congress on Modelling and Simulation, Darwin, NT, Australia, 9 to 14 July 2023
mssanz.org.au/modsim2023

Markov chain modelling of causality in anomalous and non-anomalous combat event series

J.K.R. Tiong, **T. Bogomolov** and **B.A. Chiera**

UniSA STEM, University of South Australia, Adelaide, South Australia
Email: junia.tiong@unisa.edu.au

Abstract: Combat mission studies are critical for developing an understanding of combat behaviour as part of informed decision-making. Combat simulators are commonly used for this purpose, with the goal of understanding complex relationships, such as evaluating competing options, force design and operational requirements. As part of the simulation process, a collection of event series is recorded to capture the progression of combat over time and provide insight into causal relationships over the combat space. In this work we modelled such event series using Markov chains, to capture the unfolding pattern of events within a combat scenario to establish cause-and-effect relationships. The use of Markov chains is appealing here as this approach can be used not only to capture causal behaviours, but in other applications such as predictive modelling as part of an early warning system, to provide an advantageous tool for combat planning.

A second, complementary focus taken here was to develop Markov chain models in such a way as to allow the identification of potentially anomalous combat event series. In general, anomaly detection helps to uncover unusual patterns in large data sets to provide critical information when there is a departure from typically observed behaviours. In the context of combat simulation and the analysis conducted here, a mechanism for the identification and thus detection of anomalous and non-anomalous combat event series is presented using Markov chains, to provide valuable insight into the differing characteristics of these series.

Specifically, we studied the progression of a collection of simulated combat event series arising from three different combat scenarios: A, B and C. The simulated events were categorised as being in the state of either Movement (M), Detection (D), Shot (S) or Kill (K). A collection of event series comprising these four states resulted from repeated combat simulation, with 201 replications per scenario to yield 201 distinct event series per scenario. Replications for each scenario relied on a single set of non-repeated pseudorandom numbers to seed the simulator, so as to reflect the stochastically varying nature of a combat mission. Markov chain models were developed firstly for the exploration of combat progression from an overall perspective and then extended to consider a team-based perspective in terms of an offensive Blue and defensive Red team, to provide richer exploration of the event state-space. The Akaike Information Criterion and Bayesian Information Criterion were used to determine the appropriate Markov chain model order, with results showing a Markov chain model of order 0 was not appropriate, tacitly suggesting causality is present in the studied event series. A Markov chain model of order 1 was considered most appropriate for each scenario as well as for the overall and team-based combat analyses. To gauge the statistical distance between the resulting models the Kullback–Leibler divergence (KLD) and Earth's Mover distance (EMD) statistics were calculated, indicating whether an event series could be considered anomalous or non-anomalous, in terms of combat progression.

The steady-state transition probabilities governing the Markov chain models and derived from the event series data showed that the state of Movement had the highest probability in a typical combat scenario. The difference in transition between states for Scenarios A, B and C in general is negligible, however the difference in Scenario C is more pronounced. We found the suspected anomalous event series typically produced Markov chain models in which multiple transition probabilities took the value 0. For example, an anomalous event series for Scenario A indicated a total of 23 transitions between states were not reached during simulated combat; similar conclusions were drawn for Scenarios B and C. Interestingly, the suspected anomalous event series for Scenarios B and C overlapped in part, whereas the suspected anomalous event series for Scenario A formed a disjoint set to those for Scenarios B and C, indicating a difference in combat design being realised through simulation. Future work includes strengthening insight into combat scenarios to include the impact of modelling the interarrival time between combat events to augment combat mission understanding and potentially provide another mechanism for anomaly detection, as well as using the Markov chain models as part of an early-warning system in support of informed decision-making.

Keywords: *Markov chain models, causality, anomaly detection, defence, simulation*

1. INTRODUCTION

The research on modelling a series of events of interest, captured at measured time instants, has been long studied (Gregory et al., 1993; Tealab, 2018). Understanding such event series data, with respect to the mechanisms that govern transitions between events, has wide application across a number of areas including intrusion detection in cyber security (Haque, DeLucia & Baseman, 2017), weather prediction (Kaneriya et al. 2018), credit card fraud detection (Srivastava et al., 2008) and health care management (Kulick et al., 2020). Identifying and understanding causality within such a series gives insight into factors that lead to differing sets of outcomes; such insight can be used to develop predictive models as well as optimise interventions to achieve desired results (Holland, 1998). For example, Kulick et al. (2020) inferred from causal analysis that long-term exposure to air pollution led to cognitive decline in older adults. Moreover, causal analysis of event series can be refined to separate typical, i.e. non-anomalous, patterns of transition between events, versus those which deviate from typical, i.e. anomalous transition patterns, to provide a richer understanding of the progression of an events series under different conditions.

One application area where the study of causal relationships and identification of anomalous patterns is particularly valuable is in military applications. The insights enabled by not only identifying, but understanding, anomalous event series, gives the ability to recognise atypical patterns within a military exercise and implement corrective action if necessary. For example, the probability of an unsuccessful combat mission could be reduced if army personnel are able to identify and understand atypical patterns as they unfold during combat progression. Moreover, auxiliary knowledge in cause-and-effect relationships between actions taken, and associated outcomes during combat progression, equip personnel with the knowledge of what corrective actions should be taken for a more desirable final mission outcome, based on the current trajectory. Such analysis is attractive as it is able to provide a realistic perspective within combat progression to construct a narrative of what makes for a successful combat mission.

In this paper we explore the dependencies within event series capturing combat progression to model potential causal behaviours, via a case study. We develop Markov chain models of event series produced by a high-resolution, closed-loop, stochastic, discrete event entity-based land-force tactical combat simulator (Chau, 2015), with Markov chains providing a framework to mathematically represent these series of events. We focus on characterising events reflecting non-anomalous and anomalous behaviours both in terms of overall combat progression, as well as combat progression by military team. The adaptability of Markov chains to provide a long-term view of the system supports understanding of how combat missions typically, and atypically, unfold.

This paper is outlined as follows. In Section 2 we introduce a case study in which simulated combat event series are modelled using Markov chains, before investigating appropriate model parameters using maximum likelihood estimation via information criteria. In addition, we also introduce parameters to identify suspect anomalous replications within the case study. Section 3 presents the results of Markov chain modelling, including filtering combat events to accentuate causal relationships in terms of anomalous and non-anomalous behaviours, as well as by combat team. Finally in Section 4 we give our conclusions and future directions.

2. A MARKOV CHAIN MODEL OF COMBAT SIMULATION

2.1. Markov Chain Modelling of the Combat Event Series

Combat simulators are widely used to replicate combat progression and inform decision-making in terms of competing options, characterised as design points for the combat simulator (Chau, 2015), to give rise to different combat scenarios. A design point, or scenario, is the selection of competing options such as the type of armoured vehicle, the protection level of military personnel, engagement range or factors requiring consideration in combat mission, plus other factors of interest such as topological representation and/or weather/lighting, to define the combat environment to be simulated. The simulator used here is a high-resolution, closed-loop, stochastic, discrete event black-box simulator, with specified inputs and measured output performance metrics (Chau, 2015). The simulation is replicated $R = 201$ times, seeded with a set of non-repeated pseudorandom numbers to reflect the stochastic nature of the combat mission. The same set of pseudorandom numbers were used for each scenario to facilitate comparison across scenarios. In this study, simulation was performed on three scenarios, A, B and C, aimed to reflect the different competing options for comparison as part of informed decision-making.

Each simulated combat is recorded as a finite set of events. As time is recorded only when an event occurs, the event series is inhomogeneous over time. The combat event series can be denoted as being in state $s \in \mathcal{S}$, with state-space $\mathcal{S} = \{M, D, S, K\}$ for a given simulation time point $t \in \mathbb{N}$ and the four events *Movement*

(M), *Detection* (D), *Shot* (S) and *Kill* (K). If the observed states are discrete, such that $X_n \epsilon \{1, \dots, K\}$, a discrete-time stochastic process $\{X_n\}$ can be described as Markov chain viz.

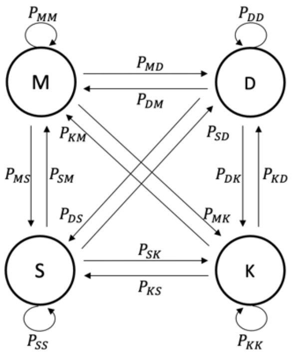

Figure 1: Directed graph depicting the transition between events in the combat state-space.

$$P = \{X_{n+1} = j | X_n = i, X_{n-1} = i_{n-1}, \dots, X_0 = i_0) = P(X_{n+1} = j | X_n = i)$$

with accompanying transition matrix

$$P_{ij} = \mathbb{P}(X_{t+1} = j | X_t = i),\ i, j \in S, t \,\epsilon\, \mathbb{N}$$

where each element of $P_{i,j}$ is the maximum likelihood estimate for the probability event j directly following event i. For each $i \in \mathcal{S}, \sum_{j \in \mathcal{S}} P_{ij} = 1$, with the characteristics of a Markov chain stipulating that over large value of n, all transition probabilities will converge to the stationary distribution $\pi = \pi P$ (Anderson & Goodman, 1957).

Figure 1 shows the general Markov chain for this state-space, illustrating the transition between the events $\mathcal{S} = \{M, D, S, K\}$, and their associated transition probabilities $P_{i,j},\ i, j \in \{M, D, S, K\}$ for each combat simulator replication. From Figure 1, it can be established that the transition of combat events from the combat simulator results in an ergodic Markov Chain in that the event transitions are both irreducible and aperiodic, where every event can be reached from every other event at irregular time intervals. Ergodic Markov chains are particularly attractive here as they converge to a unique steady-state probability in the long-run to support the stable analysis of the event series (Grinstead & Snell, 1997).

2.2. Information Criterion

A necessary step when building Markov chain models is to determine the model order to ensure the appropriate series history length is used during the modelling process. The trade-off between model order is important; a lower-order Markov chain requires fewer past observations whereas a higher-order Markov chain model has the potential to provide a more precise predictive value (Narlikar et al., 2013).

Determining the order of a Markov chain is a well-studied area in the literature. Tong (1975) adapted the Akaike Information Criteria (AIC) developed by Akaike (1974) to allow comparison between different Markov chain models to determine the best fit. The AIC $(\hat{k}_{AIC})$ estimator of the Markov chain order is chosen such that

$$AIC(\hat{k}_{AIC}) = \underset{0 \le k < m}{min} AIC(k)$$

where

$$AIC(k) = \eta_{k,m} - 2(s^m\text{-}s^k)(s-1).$$

Here η is the likelihood ratio statistic, k is the k^{th}-order chain ($k = 0,1, \dots, m-1$) and m the m^{th}-order chain with s possible states. While widely used, the application of the AIC procedure is known to potentially overestimate the true order of Markov chain models regardless of sample size (Katz, 1981). For this reason, it is beneficial to consider a second information criterion when identifying the appropriate order of the Markov chain model. A widely-used alternative is the Bayesian Information Criteria (BIC) estimator of order $(\hat{k}_{BIC})$, chosen as (Katz ,1981)

$$BIC(\hat{k}_{BIC}) = \underset{0 \le k < m}{min} BIC(k)$$

where

$$BIC(k) = \eta_{k,m} - (s^m - s^k)(s-1)\ln n.$$

The estimators used in BIC are identical to the AIC method with n factoring for sample size. A smaller AIC or BIC value is considered a better model fit, as these models penalise additional parameters.

The results in Table 1 show a model of order 0 for Scenarios A, B, and C resulted in the highest AIC and BIC statistics, indicating a memoryless model is not appropriate here, thus tacitly suggesting there is some level of causality in the event series. Selecting the model order with the lowest AIC and BIC values suggests the data under study is best represented by a Markov Chain of order 1, to provide the best trade-off between the goodness of fit and the complexity of the model, since smaller values of AIC and BIC indicate a better model fit whilst avoiding overfitting.

Table 1. AIC and BIC statistics for Scenarios A, B and C for $\mathcal{S}$. Selected model orders are bolded.

Scenario A			Scenario B			Scenario C		
Order	AIC	BIC	Order	AIC	BIC	Order	AIC	BIC
0	4104154	4104154	0	4238717	4238717	0	4162577	4162577
1	**3595619**	**3595687**	1	**3718050**	**3718118**	1	**3623089**	**3623157**
2	3625160	3625295	2	3722160	3722295	3	3625398	3625601
3	3668553	3668756	3	3723091	3723294	4	3625408	3625679

To construct a more comprehensive understanding of the event series, a team label of *Blue* or *Red* (B or R, respectively) was utilised for additional state-space segmentation. The analysis conducted here aims to create an intricate depiction of interactive events thus by distinguishing between teams it is possible to assign actions and mission outcomes that result in a more accurate representation of combat progression.

To consider the delineation of events by teams, the expanded state-space $\mathcal{S}_{BR}$ was considered:

$$\mathcal{S}_{BR} = \{M_B, M_R, D_B, D_R, S_B, S_R, K_B, K_R\}$$

with the subscripts B and R indicating the Blue and Red teams respectively. The event series displayed a disproportionate number of events of type M_B and M_R, dominating and consequently unduly influencing the Markov chain modelling process, thereby potentially obfuscating the probability of transitions between other events in the state-space. Thus to present a more realistic transition and minimise overinflating transitions between events M_B and M_R in the subsequent analysis, all M_B and M_R events prior to the first detection of any other event type were excluded.

Table 2. AIC and BIC statistics for Scenarios A, B and C for $\mathcal{S}_{BR}$. Selected model orders are bolded.

Scenario A			Scenario B			Scenario C		
Order	AIC	BIC	Order	AIC	BIC	Order	AIC	BIC
0	7389408	7389408	0	7497371	7497371	0	7497371	7497371
1	**6292407**	**6292528**	**1**	**6401049**	**6401171**	**1**	**6401049**	**6401171**
2	6292425	6292668	2	6401067	6401311	2	6401067	6401311
3	6292443	6292808	3	6401085	6401451	3	6401085	6401451

Table 2 presents the AIC and BIC statistics calculated for the expanded state-space $\mathcal{S}_{BR}$. The results indicate that a Markov chain model of order 1 yields the smallest likelihood, although the difference between the first- and higher-order statistics is negligible. Again it is evident that a Markov chain model of order 0 is not suitable for the data, supporting further exploration of causal relationships in event the event series.

2.3. Anomaly Detection

A complementary consideration is the identification and modelling of potential individual anomalous replications for further insight into combat progression. To identify anomalous replications of event series, the Euclidean distance d between all transition probabilities was calculated over all 201 replications. The average distance was then calculated, for each replication, as the deviation between the replication under consideration and all other replications. This process was repeated for each scenario. Figure 2 displays the distances observed in Scenario A, which shows there are potential outlier replications located in the right tail of the distribution. These potential outlier replications exhibit large distances to all other replications, which may indicate the occurrence of anomalous event series in these replications.

An upper cut-off value for the identification of anomalous series was calculated based on the data distributions, such that all distances larger than the upper boundary formed at $Q_4 + 1.5 \times IQR$ were considered anomalous, where Q_4 is the third quartile of the data and IQR the interquartile range. The cut-off value for Scenario A was 0.7845 with event series from 5 replications satisfying the criteria. A similar approach was used for Scenarios B and C with cut-off distances 0.8002 and 0.7957 respectively. For Scenario B, event series from 10 replications were suspected to be anomalous whereas Scenario C yielded 12 replications considered anomalous. For space considerations, the histograms are not included in this paper, although the top 5 replications with the largest deviation from the average for each scenario, are presented in Section 3.

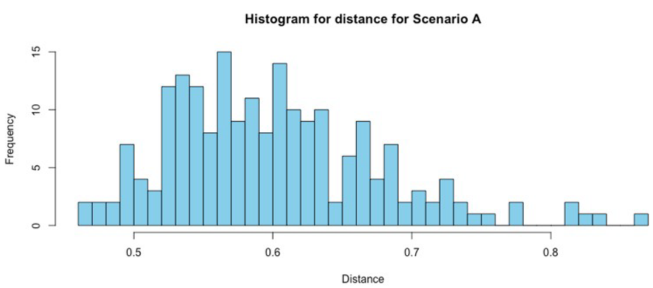

Figure 2: Histogram of the deviation from average distances of each replication for Scenario A

2.4. Kullback–Leibler divergence (KLD) and Earth's Mover distance (EMD)

To determine whether there is a difference in either the combat scenarios or between replications, the Kullback–Leibler divergence (KLD) and Earth's Mover distance (EMD) statistics were calculated. The KLD statistic is a measure of the difference between two probability distributions, indicating how one distribution diverges from the other, which can also be used as an indication of an anomaly (Zeng et al., 2014). For this research, a symmetric KLD statistic is used to compare the transition probability distributions between the three scenarios while an asymmetric KLD statistic is used for the comparison of transition probabilities derived from event series with and without suspected anomalous replications. A small perturbation of $\varepsilon = 1 \times 10^{-10}$ was added to the calculation of the KLD statistic which does not affect the overall results, but circumvents the issue of the existence of zero probabilities. Adding a perturbation in this manner also allows for the possibility of a transition to exist in the physical representation of combat progression, even though it was not observed in the simulated combat space. In contrast, the EMD statistic measures the distance between two probability distributions, accounting for the cost of transforming one distribution into its comparative distribution (Zhu et al., 2014) with large distances indicative of substantially large differences between the two probability transition matrices under consideration.

3. RESULTS OF MARKOV CHAIN MODELS OF SIMULATED COMBAT EVENT SERIES

The resulting steady-state transition probabilities for the Markov chain models for Scenarios A, B and C were empirically determined from the event series, with each scenario representing competing options in combat.

Table 3. Steady-state probabilities for Scenario A, B and C

	Scenario A					Scenario B					Scenario C			
	M	D	S	K		M	D	S	K		M	D	S	K
M	0.9461	0.0394	0.0131	0.0014	M	0.9456	0.0385	0.0146	0.0013	M	0.9469	0.0380	0.0139	0.0013
D	0.6051	0.3590	0.0357	0.0001	D	0.5973	0.3638	0.0388	0.0002	D	0.5923	0.3698	0.0377	0.0002
S	0.6046	0.0925	0.3025	0.0003	S	0.6123	0.0918	0.2951	0.0008	S	0.5926	0.0921	0.3145	0.0009
K	0.5339	0.1882	0.0560	0.2220	K	0.5142	0.2053	0.0648	0.2157	K	0.5335	0.2063	0.0677	0.1925

Table 4. Symmetric KLD and EMD statistics for state-space $\mathcal{S}$ and extended state-space $\mathcal{S}_{BR}$

State-Space, $\mathcal{S}$	Symmetric KLD	EMD
Scenario A vs Scenario B	0.0009	0.0112
Scenario A vs Scenario C	0.0019	0.0184
Scenario B vs Scenario C	0.0010	0.0149

State-Space, $\mathcal{S}_{BR}$	Symmetric KLD	EMD
Scenario A vs Scenario B	0.0028	0.0330
Scenario A vs Scenario C	0.0050	0.0512
Scenario B vs Scenario C	0.0024	0.0365

Table 3 summarises the steady-state probabilities for $R = 201$ replications for Scenarios A, B and C with $\mathcal{S} = \{M, D, S, K\}$ while Table 4 presents the KLD and EMD statistics, which provide a representation of the statistical difference between the three scenarios. In general, the values for the extended state-space $\mathcal{S}_{BR}$ are larger than those in the state-space $\mathcal{S}$, indicating the distances between scenarios are greater in the extended state-space $\mathcal{S}_{BR}$. The larger statistics in extended state-space $\mathcal{S}_{BR}$ is due to larger dimensionality with the resulting symmetric KLD and EMD statistics are generally consistent for the state-space $\mathcal{S}$ and extended state- space $\mathcal{S}_{BR}$.

On the other hand comparison between scenarios shows the distance between Scenarios A and B is smaller than the distance between Scenarios A and C, which in turn is smaller than the distance between Scenarios B and C. This suggests the transition probabilities between states in Scenarios B and C are more alike when compared to those of Scenario A. Additionally, a more prominent dissimilarity in transition probabilities can be observed between Scenarios A and C, which suggests that Scenarios A and C exhibit less similarity in their underlying dynamics as opposed to Scenarios B and C.

The steady-state probabilities in Table 5 for the extended state-space $\mathcal{S}_{BR}$ for Scenario A show there is approximately a 90% chance that if a Blue team member moves, the next state will be another movement by the Blue team ($M_B \rightarrow M_B$). However, the pattern of movement for the Red team contrasts this behaviour, with a 48% chance of transitioning from $M_R \rightarrow M_B$ and a comparative probability (42%) of transitioning from $M_R \rightarrow M_R$. In addition, if a Blue team member detects an opponent in the Red team, the probability the next event in the sequence will be a movement by a Red team member is approximately 6% ($D_B \rightarrow M_R$), in contrast to a detection made by a Red team member, where there is a 57% chance that the next sequence will involve movement by the Blue team ($D_R \rightarrow M_B$). These steady-state probabilities suggest there are differences in combat strategies between the Blue and Red teams.

Table 5. Steady-state probabilities for Scenario A for extended state-space $\mathcal{S}_{BR}$

	M_B	M_R	D_B	D_R	S_B	S_R	K_B	K_R
M_B	0.9038	0.0374	0.0179	0.0254	0.0112	0.0034	9.0E-06	0.0009
M_R	0.4800	0.4174	0.0289	0.0422	0.0106	0.0097	0.0111	0.0001
D_B	0.4939	0.0623	0.3020	0.1030	0.0245	0.0141	0.0001	3.0E-05
D_R	0.5702	0.0738	0.0800	0.2424	0.0116	0.0218	0.0001	0.0001
S_B	0.5764	0.0678	0.0362	0.0255	0.2909	0.0030	0.0002	3.4E-05
S_R	0.4347	0.0807	0.0511	0.1111	0.0053	0.3166	0.0001	0.0004
K_B	0.4712	0.0937	0.1387	0.0507	0.0462	0.0057	0.1938	0.0000
K_R	0.3465	0.1545	0.0641	0.1283	0.0195	0.0401	0.0003	0.2466

Table 6. Distance d, KLD and EMD statistics for suspected anomalous replications for Scenario A, B and C

Scenario A	*d*	KLD	EMD
65	0.8191	0.2502	0.3392
104	0.8171	0.2650	0.4850
124	0.8309	0.2783	0.6411
138	0.8292	0.2308	0.2761
163	0.8661	0.2388	0.3994

Scenario B	*d*	KLD	EMD
51	0.9872	0.4318	0.6770
98	0.9346	0.2834	0.5776
110	0.8301	0.3244	0.3458
189	1.0815	0.4729	0.6465
191	1.1013	0.4471	0.5449

Scenario C	*d*	KLD	EMD
44	0.9101	0.3347	0.3387
51	0.9586	0.4165	0.6580
53	1.0262	0.4043	0.4918
181	0.9571	0.4205	0.4265
189	1.0656	0.4645	0.6160

Table 6 summarises the Euclidean distance, d, KLD and EMD statistics of the top 5 suspected anomalous replications for Scenarios A, B and C. For example, replication 124 is suspected to be anomalous as it exhibits the highest KLD statistic for Scenario A, which highlights a significant deviation in transition probabilities produced by the associated event series, compared with the transition probabilities calculated across all other replications of the combat simulator. It is noteworthy that in both Scenarios B and C, replication 189 exhibits the largest KLD statistic and a high EMD statistic, indicating a substantial difference in transition probabilities in this case as well as replication 51. No such overlap was observed for Scenario A.

Next, the transition probabilities of events for suspected anomalous replications were calculated. Table 7 presents the transition probabilities for Scenario A, replication $R = 124$. Notably, zero probabilities occur in 23 transitions, particularly for states K_B and K_R; a feature that sets it apart from other replications. Another example is given in Table *8* for Scenario B, $R = 189$. These transition probabilities were similar to those for replication 189, Scenario C, which was also suspected to be an anomalous replication. An interesting observation here is the transitions between S_B and S_R where it appears that the opposing teams do not shoot one another. Similarly interesting insights were observed for the remaining replications from Table 6, not shown here for space considerations, however indicating differences in causal behaviours across event series.

Table 7. Transition probabilities for $R = 124$ in Scenario A

R = 124	M_B	M_R	D_B	D_R	S_B	S_R	K_B	K_R
M_B	0.9224	0.0415	0.0068	0.0247	0.0012	0.0028	0.0000	0.0007
M_R	0.6257	0.3055	0.0049	0.0508	0.0000	0.0106	0.0025	0.0000
D_B	0.5616	0.0493	0.1675	0.1872	0.0197	0.0148	0.0000	0.0000
D_R	0.6307	0.0755	0.0445	0.2439	0.0013	0.0027	0.0000	0.0013
S_B	0.6364	0.0000	0.0606	0.1212	0.1818	0.0000	0.0000	0.0000
S_R	0.5862	0.1839	0.0230	0.0115	0.0000	0.1954	0.0000	0.0000
K_B	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
K_R	0.5385	0.2308	0.0769	0.0769	0.0000	0.0000	0.0000	0.0769

Table 8. Transition probabilities for $R = 189$ in Scenario B

R = 189	M_B	M_R	D_B	D_R	S_B	S_R	K_B	K_R
M_B	0.9056	0.0405	0.0190	0.0192	0.0126	0.0028	0.0000	0.0001
M_R	0.5106	0.4114	0.0301	0.0306	0.0054	0.0010	0.0109	0.0000
D_B	0.5141	0.0587	0.2946	0.1098	0.0217	0.0011	0.0000	0.0000
D_R	0.6000	0.0861	0.1030	0.1952	0.0109	0.0024	0.0024	0.0000
S_B	0.6104	0.0576	0.0230	0.0096	0.2994	0.0000	0.0000	0.0000
S_R	0.6238	0.0594	0.0297	0.0495	0.0000	0.2376	0.0000	0.0000
K_B	0.4848	0.0000	0.0909	0.0303	0.1212	0.0000	0.2727	0.0000
K_R	0.0000	0.5000	0.5000	0.0000	0.0000	0.0000	0.0000	0.0000

4. CONCLUSION

In this paper we modelled a collection of combat event series using Markov chain models, to provide probabilistic insight into the unfolding of a combat scenario. We considered three different scenarios, A, B and C, representing competing options for a combat mission. Use of the Akaike and Bayesian information criterions suggested causality is present within the data, with a first-order Markov chain model developed for all event series. We augmented the study in two ways: first to include team membership and then to identify suspected anomalous event series, for which we also developed Markov chain models.

Further research can extend upon this initial work. The current study differentiated between combat teams, however further granularity can be added by distinguishing between other combat components, e.g. infantry and vehicles used. A further addition is to include time as part of the modelling process, such as the interarrival time between events, to further enhance the capacities of understanding the progression of a combat over time. Of particular interest would be time-based behaviours in combat action sequences from anomalous and non-anomalous event series. The work presented here thus provides the foundation for a more comprehensive understanding of combat scenarios to ultimately form part of an early warning system to inform effective decision making in combat, as well as in support of the development of military strategies.

ACKNOWLEDGMENTS

The Commonwealth of Australia (represented by the Defence Science and Technology Group) supported this research through a Defence Science Partnership agreement. We would like to thank Dr Maria John for her invaluable input into this research through expert insight.

REFERENCES

Anderson, T.W. & Goodman, L.A. 1957. Statistical inference about Markov chains. The annals of mathematical statistics, pp.89-110.

Chau, W. Gill, A.W. & Grieger, D. 2015. Using Combat Simulation and Sensitivity Analysis to Support Evaluation of Land Combat Vehicle Configurations. In Syme, G., Hatton MacDonald, D., Fulton, B. and Piantadosi, J. (eds) MODSIM2017. 22nd International Congress on Modelling and Simulation. Modelling and Simulation Society of Australia and New Zealand, December 2017, pp. 536-542.

Gregory, J.M. Wigley, T.M.L. and Jones, P.D. 1993. Application of Markov models to area-average daily precipitation series and interannual variability in seasonal totals. Climate dynamics, 8, pp.299-310.

Grinstead, C. & Snell, J.L. 1997. Introduction to Probability. American Mathematical Society.

Haque, A. DeLucia, A. & Baseman, E. 2017. Markov chain modeling for anomaly detection in high performance computing system logs. In Proceedings of the Fourth International Workshop on HPC User Support Tools pp. 1-8.

Holland, J. H. 1998. Emergence: From chaos to order. Oxford University Press.

Kaneriya, S. Tanwar, S. Buddhadev, S. Verma, J.P. Tyagi, S. Kumar, N. and Misra, S. 2018. May. A range-based approach for long-term forecast of weather using probabilistic markov model. In 2018 IEEE international conference on communications workshops (ICC workshops) (pp. 1-6), IEEE.

Katz, R.W. 1981. On some criteria for estimating the order of a Markov chain. Technometrics, 23(3), pp. 243-249.

Kulick, E.R. Elkind, M.S. Boehme, A.K. Joyce, N.R. Schupf, N. Kaufman, J.D. Mayeux, R. Manly, J.J. and Wellenius, G.A. 2020. Long-term exposure to ambient air pollution. APOE-ε4 status, and cognitive decline in a cohort of older adults in northern Manhattan. Environment international, 136, p.105440.

Narlikar, L. Mehta, N. Galande, S. & Arjunwadkar, M. 2013. One size does not fit all: on how Markov model order dictates performance of genomic sequence analyses. Nucleic acids research, 41(3), pp. 1416-1424.

Schütze, H. Manning, C.D. and Raghavan, P. 2008. Introduction to information retrieval. Vol. 39, pp. 234-265). Cambridge: Cambridge University Press.

Srivastava, A. Kundu, A. Sural, S. and Majumdar, A. 2008. Credit card fraud detection using hidden Markov model. IEEE Transactions on dependable and secure computing, 5(1), pp.37-48.

Tealab, A. 2018. Time series forecasting using artificial neural networks methodologies: A systematic review. Future Computing and Informatics Journal, 3(2), pp. 334-340.

Tong, H. 1975. Determination of the order of a Markov chain by Akaike's information criterion. Journal of applied probability, 12(3), pp. 488-497.

Zeng, J. Kruger, U. Geluk, J. Wang, X. & Xie, L. 2014. Detecting abnormal situations using the Kullback–Leibler divergence. Automatica (Oxford), vol. 50, no. 11, pp. 2777–2786.

Zhu, X. Liu, J. Wang, J. Li, C. & Lu, H. 2014. Sparse representation for robust abnormality detection in crowded scenes. Pattern Recognition, 47(5), 1791–1799.