Putting the art before the force

A. H. Pincombe[1]   B. M. Pincombe[2]   C. E. M. Pearce [3]

Abstract

We use a dataset from the Battle of Kursk to test three estimators of attrition: linear, quadratic and log dependence on the number of soldiers in each force. Data giving force numbers per day show significant collinearity, so we use force and loss ratios for our tests. We demonstrate that the strongest correlate in the dataset for a sides attrition is its own force strength. This supports the log estimator, and we evaluate the proposition that this counterintuitive connection is a product of the pre-battle art of war, where commanders attempt to balance their forces to their expectations of threat. Thus expected losses generate actual force numbers whereas we seek information on the ways that force numbers generate actual losses, and both processes are based on the same correlation information. We argue that the dataset must still contain information on the mechanisms of attrition, so we widen our search criteria and uncover some remarkable facts.

Contents

1 Introduction

There is reason to believe that the attrition suffered in a battle depends upon the relative strengths of the two forces. Clausewitz [3] states that "superiority of numbers admittedly is the most important factor in the outcome of an engagement". There is also reason to believe that numbers are less important than other factors. Clausewitz [3] also states "To accept superiority of numbers as the one and only rule, and to reduce the whole secret of the art of war to the formula of numerical superiority at a certain time in a certain place was an oversimplification that would not have stood up for a moment against the realities of life". Lanchester equations are based on a belief in the overwhelming importance of numbers. They have been used for more than ninety years and have generated over 700 theoretical developments, yet they have never been validated for complex land battles despite numerous attempts [13].

Lanchester–Osipov models are coupled first order differential equations which describe the evolution of the numerical force strength on each side, given the initial values. Initially there were two models: one for aimed fire (also known as the square law), the other for unaimed fire (known as the linear law). The need to validate was recognised early. Osipov [8] (1915) based

his version of the theory on the outcomes of a large number of historical battles. He developed the square law model for smaller battles and proposed a 3/2 law for larger battles. Importantly, the 3/2 law was based on the observation that, for large forces, the casualties suffered by a side seemed to be related to the square root of their own initial force size. In the square and linear laws, the casualties for a side are related in some way to the numerical force strength of the opposing side. Older historical battle data was simply the beginning and end numerical strengths for the two sides, so results of one battle would equally support the square and linear laws and the results of many battles were needed. This required the astounding assumption that all things, other than force sizes, remained equal. Validation was initially aimed at deciding which of the square and linear laws was better. Outcomes were inconsistent, with neither law performing well [11]. Analysis of empirical data led to the discovery of the log law, where the attrition rate for a side was driven by its own force numbers [10]. The significance of the log law has been a matter of speculation and the search for information to explain that significance has been a major driver for this article and its companion [9]. When time phased data, covering the daily numerical force strengths and numerical casualties for each force, became available for the Battles of Kursk and the Ardennes it caused renewed interest in discriminating between Lanchester laws [1, 5, 4, 7, 6]. Bracken [1] (1995) defined a generalised form of Lanchester equations which includes the square, linear and log laws as particular cases, and applied it to the Ardennes dataset. The general form had five parameters: the individual effectiveness for each side, an exponent for the 'shooting force', an exponent for the 'target force', and a 'tactical parameter reflecting which side is defending or attacking'. He chose days 2 to 11 of the Ardennes dataset for analysis and interpreted the overall result as supportive of the linear law. Hartley [5] used the generalised equations but without the tactical parameter, and found support, over the entire 32 days of the Ardennes dataset, for a mixed log–linear law. Fricker [4] also analysed the entire Ardennes dataset, concluding that it supported the log law and he interpreted this as meaning that the probability of kill "is essentially constant over the range of the opponent force sizes given in the data". Lucas and

Dinges [7] developed their FCUD data (about those actually fighting on each day) from the Kursk database [2] and found, for the homogeneous case, that the linear law worked best $R^2 = 0.622$, followed by the log law $R^2 = 0.535$ and the square law $R^2 = 0.289$. Lucas and Turkes [6] (2004) used a weighted sum of people, armoured personnel carriers, tanks and artillery to calculate the force strength for each side and undertook a complex analysis of the data from the Battles of Kursk and Ardennes. For the Ardennes, the force strengths gave R^2 of around 0.3 but with little to distinguish the three models, while a simple homogeneous approach using the numbers of soldiers gave $R^2 < 0.1$. For Kursk data there was no such distinction, with force strength and force numbers both giving R^2 of around 0.1 for all three models. Lucas and Dinges [7] split the data into four phases (for the German side these are attack against prepared defences, attack against hasty defences, defence against counterattack and withdrawal) and found that the overall performance of all three models improved remarkably to R^2 values around 0.8 and that a model estimating attrition from the average attrition in each of the phases also gave an R^2 of 0.8.

We use the Kursk dataset and uncover a succession of complications. Multiple regression can be used to infer importances for multiple mechanisms, but it fails here because of significant collinearity between the numbers in the two forces. To reduce the effect of collinearity, we study loss ratios, showing that there is overwhelming support for the log law, where losses appear to have nothing to do with enemy forces. This unexpected and counterintuitive result needs further analysis, so we consider a suggestion [11] that the collinearity in force numbers can result from planning and that the collinearity brings a correlation between the losses for a side and their own force numbers, artificially supporting the log law. Commanders will balance forces, so we should expect collinearity; however, artificial support for the log law can only occur if there is underlying support for the linear or square laws.

When Weiss [12] developed his version of the linear law, he proposed that the attrition rate for a force results from the number of enemy shots through the area occupied by the force and the density of the force in that area. He related

the number of shots to the number of shooters (on the opposing side) and the density of soldiers to the number in the force. We find that the product of the quantity of small arms fire (tonnes per day) for the German force and the force numbers for the Soviet force was strongly correlated (0.89) with Soviet daily attrition, which seems to provide strong support for the linear law. However, when we relate the German small arms usage to the number in the German force, we see that they are completely independent. This is contrary to the Lanchester assumption that the number of shooters in a force can be represented by the number of soldiers in that force, and happens in spite of the restriction of daily force numbers to only those who are involved in the battle.

2   Linear modelling of Kursk data

Multiple regression is the obvious extension to previous [7] simple regression based estimates of the relative importance of the three laws. This failed due to collinearity between the numbers in each of the opposing forces over the fourteen days of the battle ($\rho = 0.6$). To minimise the effect of collinearity, we use loss ratios instead of losses in the analysis. This ratio analysis is based on the state form of the Lanchester equations.

For the square law, the attrition rate equations are

$$\frac{dx}{dt} = -k_y y \quad \text{and} \quad \frac{dy}{dt} = -k_x x\,, \tag{1}$$

where the number of entities (for example, soldiers) in each side (x, y) is the state of the system and k_x and k_y are the 'effectiveness' coefficients. The square law is obtained by integration of the related state equation

$$\frac{dy}{dx} = \frac{k_x}{k_y}\frac{x}{y}\,. \tag{2}$$

Thus the loss ratio $\Delta x / \Delta y$ is driven by y/x, which is the inverse of the related force ratio, and we expect these two ratios to be well correlated if the square law is dominant. The related equations for the linear law are

$$\frac{dx}{dt} = -axy \quad \text{and} \quad \frac{dy}{dt} = -bxy \,, \tag{3}$$

where a and b are rate coefficients, with a state equation

$$\frac{dy}{dx} = \frac{b}{a} \,. \tag{4}$$

In the linear law case, the loss ratio $\Delta x / \Delta y$ should be constant and should not be correlated with either of x and y. The equations for the log law are

$$\frac{dx}{dt} = -c_1 x \quad \text{and} \quad \frac{dy}{dt} = -c_2 y \,, \tag{5}$$

with state equation

$$\frac{dy}{dx} = \frac{c_2}{c_1} \frac{y}{x} \,. \tag{6}$$

Thus, the log law's loss ratio should be driven by the related force ratio.

The appropriateness of the standard Lanchester–Osipov continuous model of a discrete system is discussed in more detail in a companion article [9].

2.1   Results of ratio analysis

To test each of the models we first need to define the force ratio. Two definitions are possible: either Soviet force numbers divided by German force numbers; or the inverse. Each gives similar results but to avoid confusion we select

$$f_r = \frac{\text{Soviet}}{\text{German}}. \tag{7}$$

The loss ratio must be calculated in the same way, Soviet divided by German.
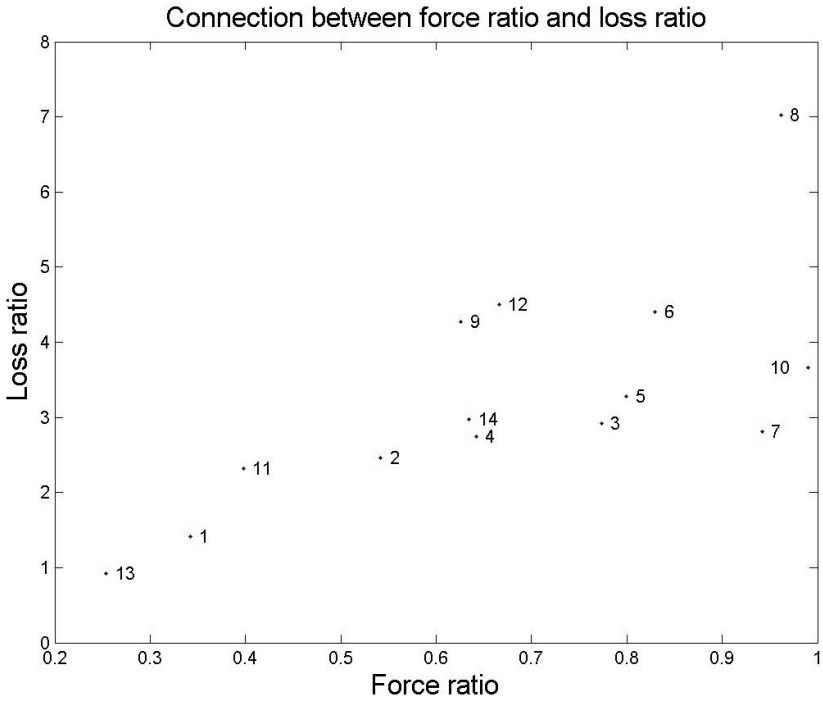
FIGURE 1: Loss ratio as a function of force ratio using daily Kursk data.

The log law will be supported by Kursk data if there is a strong correlation between loss ratio and force ratio. Figure 1 shows a positive linear relationship. This is supported by a moderately strong correlation (0.70) between loss ratio and force ratio. This graph is evidence that the force ratio could be a factor in the expression for the loss ratio.

The square law drives the loss ratio in a direction opposite to the observations (see Figure 2). This is supported by a negative correlation (-0.70) between loss ratio and inverse force ratio. The square law, applied to total force numbers, cannot explain the Kursk data.
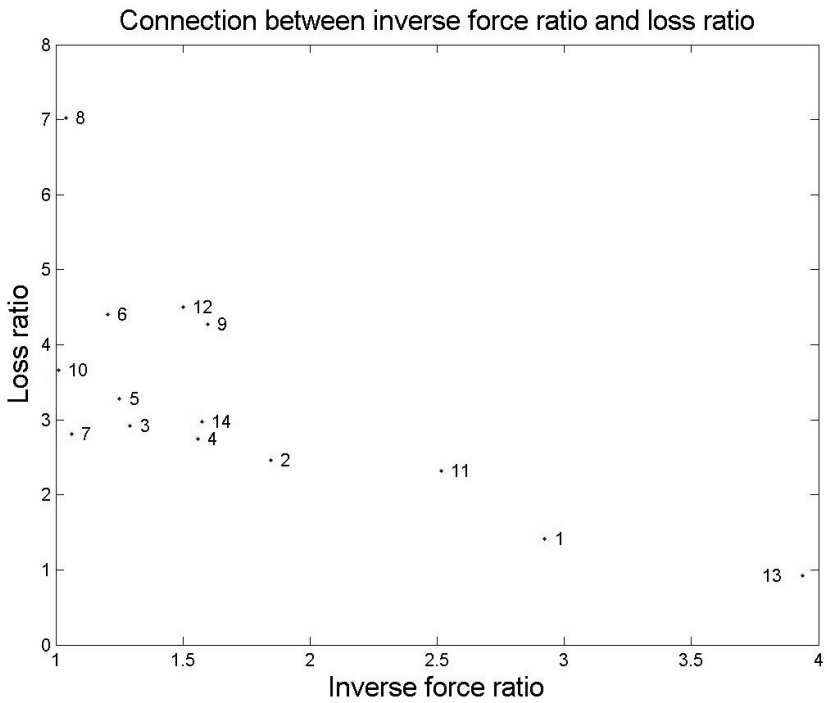
FIGURE 2: Loss ratio as a function of inverse force ratio for daily Kursk data.
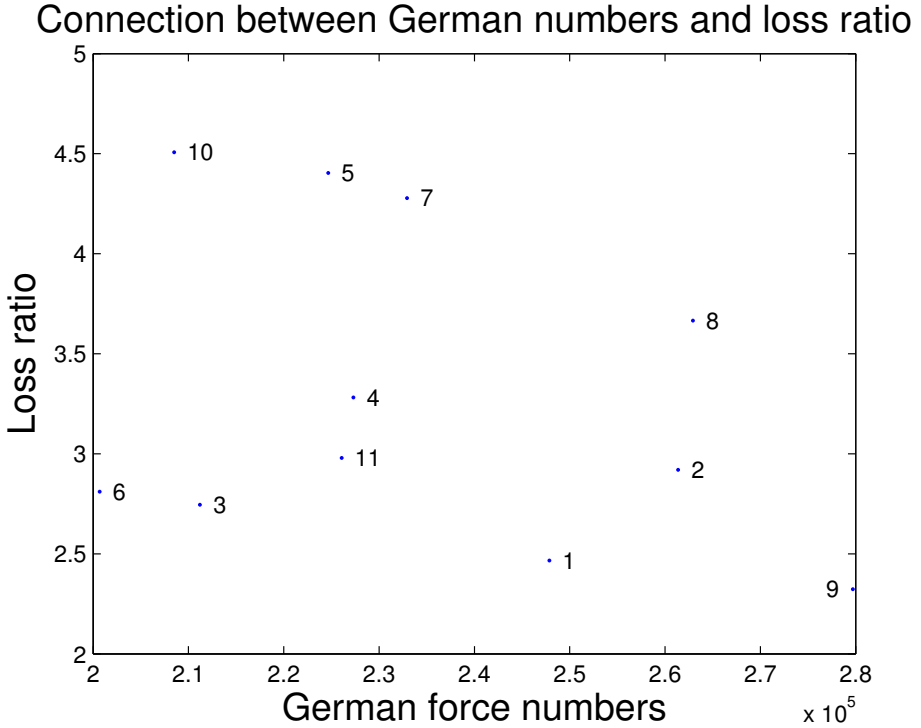
FIGURE 3: Loss ratio as a function of German numerical strength for daily Kursk data.

While there is an obvious connection between loss ratio and force ratio, there is still room for the linear law to be supported. This would require the loss ratio to be independent of the daily force numbers of the two sides.

Figure 3 shows that there is almost no connection between the values of loss ratio and the number of people in the active German force. This is supported by correlation (0.20). Figure 4 shows the equivalent plot for values of loss ratio against the related number of people in the active Soviet force. There is a moderately strong correlation (0.69). The linear connection between loss ratio and Soviet force numbers is sufficient to exclude the the linear law, using
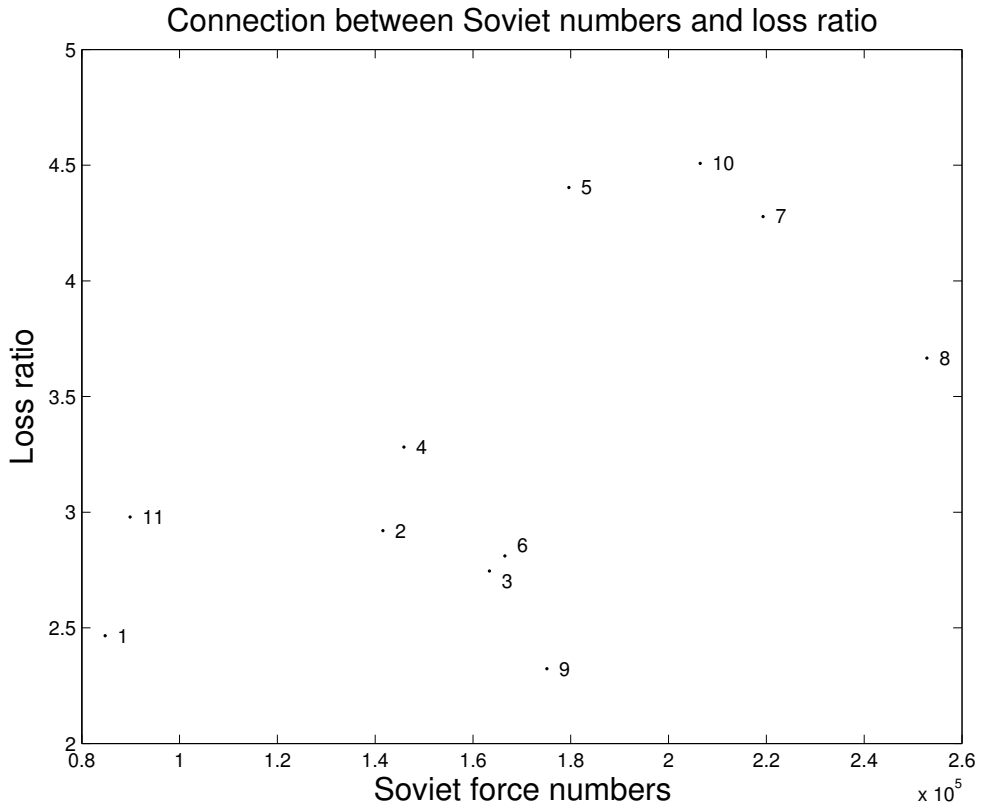
FIGURE 4: Loss ratio as a function of Soviet numerical strength for daily Kursk data.

total force numbers, as an explanation of the Kursk data.

2.2   Putting the art before the force

Speight [11] suggested that "the observed results (are) due to (commanders) calculation of the resources required for mission success". This is a double hypothesis: first that force balancing causes collinearity in force numbers; second that collinearity between force numbers produces an artificial support for the log law. Force balancing can and does occur either prior to or on the day so there is support for the hypothesis that force balancing contributes to collinearity in force numbers. However, we question the second part of the hypothesis.

If the underlying attrition mechanisms follow either the linear or square laws, then losses for the Y force will depend on the number x in the X force. A strong correlation between x and y, the number in the Y force would mean that y could also be used to predict Y force losses, but not quite so well as x. When part of the attrition is via the linear law and the rest is via the square law, it is conceivable that the correlation effect could make it appear that the log law was dominant.

Two events are needed for collinearity to artificially support the log law. First, for the effect to occur at all, Lanchester–Osipov models need to be valid; second, for the log law to dominate, the underlying mechanisms need to be a mixture of the linear and square laws. Our ratio analysis has shown that the dataset provides absolutely no support for the square or linear laws and that it unequivocally supports the log law. In fact the collinearity effect may work in the opposite direction, providing some support for square and linear laws in earlier studies.
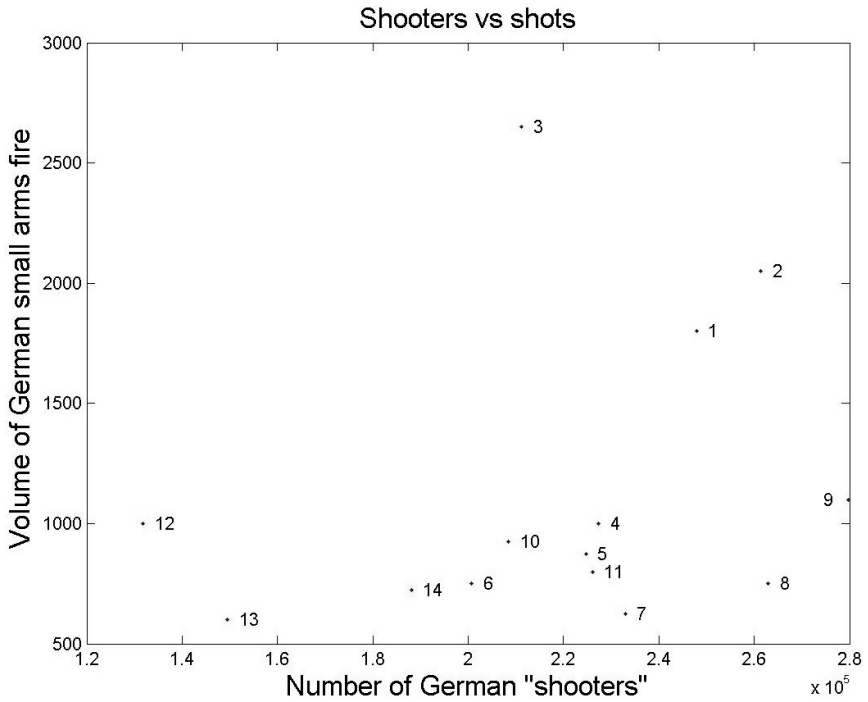
FIGURE 5: Volume of German small arms fire as a function of German numerical strength using daily Kursk data.

3   What is the database telling us?

We widen our search for causative factors to cover usage of ammunition. The usage of ammunition is clearly related to the number of shots fired, so, for example following Weiss [12], we could expect that the product of the daily small arms ammunition tonnage for the X force and the daily force numbers for the Y force will represent the effect of the linear law. This product has a correlation with the losses Δy of $r = 0.89$. The rate of fire is usually given by αx where α is the estimated firing rate per shooter and x is taken to be the number of shooters. So why does the linear law not give such a high correlation? Inspection of force numbers x and daily small arms tonnages for the X force, in Figure 5 shows that, for days 4 to 14 the two values are completely independent. The number of bullets fired each day is independent of the number of soldiers, for the Kursk data.

Losses result from the actions of the enemy force, yet, for Kursk, they are independent of the total number in the enemy force. Similarly, moving bullets result from the actions of enemy shooters, yet the number of bullets is independent of the number of soldiers in the enemy force. The problem appears to come from attempts to relate the daily variations in enemy effectiveness with daily variations in enemy force numbers. We investigate this problem in a companion article [9].

References

[1] J. Bracken. Lanchester models of the Ardennes campaign, *Naval Research Logistics*, **42**, 1995, 559–577. C484

[2] W. J. Bauman. *Kursk operation simulation and validation exercise—phase II*. CAA-SR-98-7. US Army Concepts Analysis Agency, 1998. C485

[3] C. von Clausewitz. *On War*. Edited and translated by M. Howard and P. Paret. Princeton University Press, 1984. C483

[4] R. D. Fricker, Attrition models of the Ardennes campaign, *Naval Research Logistics*, **45**, 1998, 1–22. C484

[5] D. S. Hartley. A mathematical model of attrition data. *Naval Research Logistics*, **42**, 1995, 585–607. C484

[6] T. W. Lucas and T. Turkes. Fitting Lanchester equations to the Battles of Kursk and Ardennes. *Naval Research Logistics*, **51**, 2004, 95–116. C484, C485

[7] T. W. Lucas and J. A. Dinges. The effect of battle circumstances on fitting Lanchester equations to the Battle of Kursk. *Military Operations Research*, **9**, 2004, 17–30. C484, C485, C486

[8] M. Osipov. The Influence of the Numerical Strength of Engaged Forces in Their Casualties. *Voennyi Sbornik*, 1915, no. 6 (June) 59–74, no. 7 (July) 25–36, no. 8 (Aug.) 31–40, no. 9 (Sep.) 25–37, no. 10 (Oct.) 93–96. Translated by R. L. Hembold and A. S. Rehm. *Naval Research Logistics*, **42**, 1995, 435–490. C483

[9] A. H. Pincombe, B. M. Pincombe and C. E. M. Pearce. A simple battle model with explanatory power, *ANZIAM Journal*, 51(E), 2010. http://anziamj.austms.org.au/ojs/index.php/ANZIAMJ/article/view/2585 C484, C487, C494

[10] R. H. Peterson. On the 'logarithmic law' of attrition and its application to tank combat, *Operations Research*, **15**, 1967, 557–558. C484

[11] L. R. Speight. Lanchester's equations and the structure of the operational campaign: between campaign effects. *Military Operations Research*, **7**, 15–43, 2002. C484, C485, C492

[12] H. K. Weiss. Lanchester-Type Models of Warfare. *Proceedings of the 1st International Conference on Operational Research*, Amsterdam, 82–99. C485, C494

[13] S. Wrigge, A. Fransén and L. Wigg. *The Lanchester Theory of Combat and Some Related Subjects: A Bibliography 1900 – 1993*. FOA Rapport-D–95–00153-1.1,1–SE. National Defence Research Establishment, 1995. http://handle.dtic.mil/100.2/ADA302237 C483

Author addresses

1. **A. H. Pincombe**, Defence Science and Technology Organisation, Edinburgh, Australia.
   mailto:adrian.pincombe@dsto.defence.gov.au

2. **B. M. Pincombe**, Defence Science and Technology Organisation, Edinburgh, Australia.

3. **C. E. M. Pearce** , University of Adelaide, South Australia, Australia.