Mathematical models of the Battle of Britain[1]

Niall J. MacKay

Department of Mathematics, University of York

This talk is, in contrast to the others at this meeting, intended as a mathematical light entertainment, but perhaps with a historical lesson. It began with my setting an undergraduate project on the mathematical modelling of warfare. The precise topic was up to the student, but I intended that it should begin with the simple, deterministic, coupled-ODE models of war due to Lanchester and others. It had struck me that whereas Lanchester had proposed his models in the context of air warfare, the late-20th century industry in fitting them to real battle data all worked with land or all-arms battles. My challenge to my student Ian Johnson was therefore: find me some good data on an *air* campaign, and let's see what we can do with it. The answer, of course, was very close to home: the Battle of Britain. The resulting paper appeared in *Naval Research Logistics*, and we refer the reader to it for all references.[2]

Recall Lanchester's equations, for a battle of attrition and annihilation between $R(t)$ Red and $G(t)$ Green units, beginning at $t = 0$ with $R(0) = R_0$ and $G(t) = G_0$:

$$\frac{dR}{dt} = -gG\,, \qquad \frac{dG}{dt} = -rR \qquad \textbf{aimed fire} \tag{1}$$

—that is, loss-rates are proportional to enemy numbers. Well, the simplistic assumptions here are almost too numerous to mention: no terrain or tactics, no command or control, no randomness, perfect attrition and aggression. But then so too are they in two more models,

$$\frac{dR}{dt} = -gGR\,, \qquad \frac{dG}{dt} = -rRG \qquad \textbf{unaimed fire} \tag{2}$$

and

$$\frac{dR}{dt} = -gN\,, \qquad \frac{dG}{dt} = -rN \qquad \textbf{ancient warfare}\,, \tag{3}$$

where in this last model N is some unknown function of time. Lanchester's point was that there is a crucial distinction between the aimed fire model and the other two. Consider, first, ancient warfare, with hand-to-hand fighting along a fixed battle line. The two sides' prowess may differ, $r \neq g$, but the fight takes place along a fixed battle-line, with $N(t)$ people engaged on each side. Or consider unaimed fire, which causes damage not only in proportion to volume of fire but also to density of enemy units. In either of these models (2,3), divide one equation by the other, separate variables and integrate and we find that $r(R-R_0) = g(G-G_0)$ or $rR - gG = rR_0 - gG_0$. If, say, $rR_0 > gG_0$ then only Red can win, for $R = 0, G > 0$ is

---

[1] Talk given at IMA conference 'Mathematics in Defence 2009', QinetiQ, Farnborough, November 2009, `http://www.ima.org.uk/Conferences/maths_in_defence09.html`

[2] I. R. Johnson and N. J. MacKay, *Lanchester models and the Battle of Britain*, DOI: 10.1002/nav.20328

impossible. This is all quite intuitive: one's power to win the battle, one's *fighting strength*, is the product of number of units by their fighting effectiveness. But do the same in the aimed fire model and something very different emerges, for then

$$\frac{dR}{dG} = \frac{gG}{rR} \qquad \Rightarrow \qquad \int gG\,dG = \int rR\,dR$$

so that now $rR^2 - gG^2 = rR_0^2 - gG_0^2$. This 'square law' has crucial consequences. First, numbers matter more than prowess: for suppose $g = 3r$ and $R_0 = 2G_0$. In the unaimed-fire and ancient models, Green wins (because $3 > 2$), but in the aimed-fire model, Red wins (because $2^2 > 3$ and thus $rR_0^2 > gG_0^2$). Even more important, perhaps, is the value of concentration: for suppose, in this last example, Green were able to split Red into two equal-sized armies and fight them in turn. After the first engagement, between two equal-sized forces, Green has now won, with $G_1 = \sqrt{2/3}G_0$ units remaining to fight the other half of the Red force—which it defeats, for $3r.\frac{2}{3}G_0^2 > rG_0^2$, with $\sqrt{1/3}G_0$ of its original units standing—a complete turnaround! Lanchester's crucial point was that what distinguishes the assumptions of the aimed-fire model is that each unit is capable of long-range aimed projectile fire, with no defence and no shortage of targets—and that this characterized the new conditions of warfare which developed before 1914.

There is no need to keep the battle symmetric, and a model frequently fitted to data (perhaps only because log-log regression is so tempting) is

$$\frac{dR}{dt} = -gG^{g_1}R^{r_2}\ , \qquad \frac{dG}{dt} = -rR^{r_1}G^{g_2}\ . \tag{4}$$

Rather many authors seem to fit values of the six parameters and stop there. Let us instead once again divide, separate variables and integrate, this time resulting in

$$\frac{rR^\rho}{\rho} - \frac{gG^\gamma}{\gamma} = \text{constant.} \tag{5}$$

The tactical implications of the model are captured by what we shall call the *exponents*, $\rho = 1 + r_1 - r_2$ and $\gamma = 1 + g_1 - g_2$. If an exponent is greater than one, that side benefits from numbers and from concentration (and conversely). The specializations to the aimed- and unaimed-fire models are respectively $g_1 = r_1 = 1, g_2 = r_2 - 0, \rho = \gamma = 2$ and $g_1 = r_1 = 1, g_2 = r_2 - 1, \rho = \gamma = 1$.

But there's also another potentially asymmetric effect, for if $\rho < \gamma$ then Red has what we'll call a 'defender's advantage', by a factor γ/ρ. To see this let's specialize to the case where Green attacks in the open but Red is in concealed positions, so that Red's fire is aimed but Green's is unaimed, $g_1 = r_2 = r_1 = 1, g_2 = 0$, and $\rho = 1, \gamma = 2$. Now, in an asymmetric battle r and g have different dimensions, so one has to be a little careful (and in fact such dimensional issues are crucial whenever $g_1 + r_2 \neq 1$ or $r_1 + g_2 \neq 1$, as we shall see later).

Suppose we rescale g by R_0, so that the model is

$$\frac{dR}{dt} = -gG\frac{R}{R_0}\,, \qquad \frac{dG}{dt} = -rR\,, \tag{6}$$

and the conservation law is

$$rRR_0 - \frac{1}{2}gG^2 = rR_0^2 - \frac{1}{2}gG_0^2. \tag{7}$$

We can see that, compared to an aimed-fire battle with the same initial kill-rates, Green now needs double the fighting strength to win—either double the effectiveness or $\sqrt{2}$ times the numbers. Green benefits from concentration, but Red does not.

The Battle of Britain

So let's see how this fits data from the Battle of Britain. We shall fit the six parameters r, r_1, r_2, g, g_1, g_2 of the model

$$\log\left(-\frac{dR}{dt}\right) = \log g + g_1 \log G + g_2 \log R\,, \qquad \log\left(-\frac{dG}{dt}\right) = \log r + r_1 \log R + r_2 \log G \tag{8}$$

to our 52 daily loss and sortie data from the battle.[3] First, however, we should look for any trends within the battle over time. When one does so, the only striking fact to emerge is captured in Figure 1: it is that the battle naturally splits into a much more intense first phase and a less intense second phase. The natural division is after point 26—which is in fact 15th September 1940, 'Battle of Britain Day'.

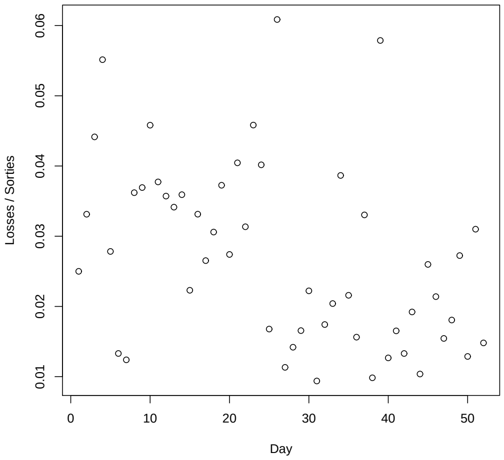


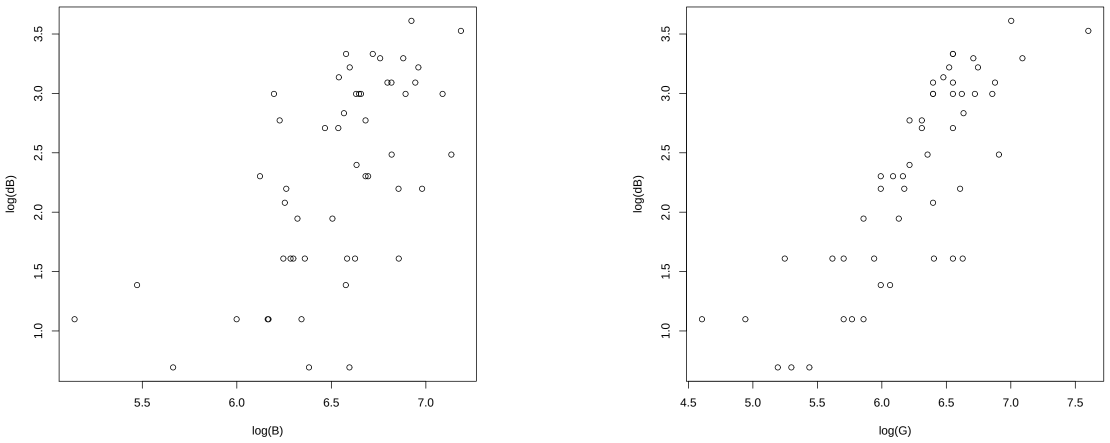
Figure 1: Losses per sortie, $-(dG/dt + dR/dt)/(G+R)$, *vs* day

In what follows we'll look only at the whole battle, but essentially the same conclusions follow for the two phases individually, with better fits in the first phase than the second.

---

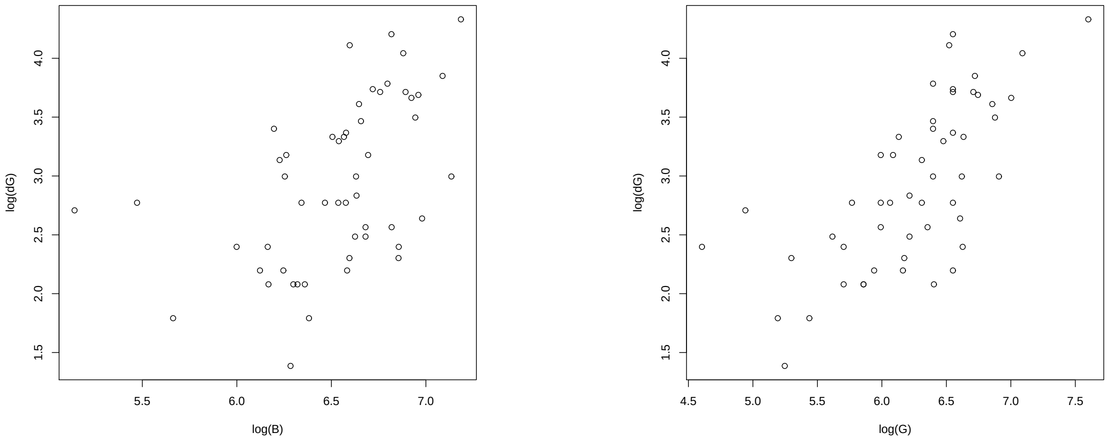
[3] Loss data are precise figures, not claims. German sortie data are, perforce, RAF estimates, for some of the original Luftwaffe data has been lost.

The log-log plots of British and German losses are shown in Figures 2 and 3.



$\log(-dR/dt)$ *vs* $\log R$ $\qquad$ $\log(-dR/dt)$ *vs* $\log G$

Figure 2: British loss rates



$\log(-dG/dt)$ *vs* $\log R$ $\qquad$ $\log(-dG/dt)$ *vs* $\log G$

Figure 3: German loss rates

To cut to the chase, the best-fitting values are

$$\begin{aligned}
\frac{dR}{dt} &= -gG^{1.12\pm0.17}R^{0.18\pm0.25} \qquad (\Sigma R^2 = 0.66), \qquad (9)\\
\frac{dG}{dt} &= -rR^{0.00\pm0.25}G^{0.86\pm0.18} \qquad (\Sigma R^2 = 0.49). \qquad (10)
\end{aligned}$$

(Note that the values of g and r have varying dimensions, are hard to estimate, and do not affect the tactical conclusions anyway—so we do not give them.) Thus, for each side, losses are roughly proportional to the number of German sorties—even though our figures for these are actually RAF estimates. In each case the error-bar for the power of R includes zero, so R has no explanatory power. So Lanchester's insight seems to work for British losses, which are roughly proportional to G, but not for German. But the truth seems quite intuitive—both sides' losses were proportional to the Luftwaffe presence.

Before we think about tactical implications, there are two important subtleties to discuss. The first of these is rather obvious, although papers on the subject often, surprisingly, omit to mention it: it is that R and G are highly correlated (in this case 0.74). The more Luftwaffe sorties there are, the more RAF sorties will follow. The effect of this is that, while the overall powers of numbers on the right-hand sides of (9,10) fit quite well, it remains uncertain whether they are due to R or to G: that is, $g_1 + r_2 (\simeq 1.3)$ and $r_1 + g_2 (\simeq 0.8)$ fit well, but their summands do not.

The second, greater subtlety is about aggregation of data points. Recall that ours are aggregated daily data, rather than individual raids or engagements. In fact, what, for these models' purposes, *is* an 'engagement'? In the original aimed fire model, or in any model of the form (4) which has the linearity property $r_1 + g_2 = g_1 + r_2 = 1$, it does not matter. But consider instead the asymmetric model (6). There we divided by R_0, which effectively parametrized Red density. Doubling R and G in this model is not the same as aggregating two equal engagements which take place at different places or times, for in the latter case we must halve the Red density. If this is not done, then the effect of aggregating raids into daily data is to cause problems in (4) whenever the total powers on the right-hand sides, $g_1 + r_2$ and $r_1 + g_2$, differ from one. The upshot is that the curve-fitting yields values of $g_1 + r_2$ and $r_1 + g_2$ which are closer to one than their true values, and with worse fits (*i.e.* lower ΣR^2).

Bearing all this in mind, let us return to our results. On first looking at Figs 2 and 3, we might think (at least of the scaling with $\log G$) 'well, that looks like a clear regression line'. But perhaps this is only with respect to a prejudice that to have *no* such dependence is a reasonable null hypothesis. Of course it is not: it is only to be expected (at least in a consistently aggressive fight) that the more sorties are flown, the more losses there will be. In fact, as we discussed, all tactical implications follow from the exponents, —and recall that we arrived at these via the loss ratio, which is not at all well modelled, as we see from Figure 4. In fact all that can reasonably appear to be said from this is that days of heavy fighting slightly favoured the Luftwaffe, or at least did not favour the RAF.[4]

For the exponents we have $\rho = 1 + r_1 - r_2 \simeq 0.8$ and $\gamma = 1 + g_1 - g_2 \simeq 1.3$. Because of the subtleties discussed above, these are not very accurately known. It is striking, though, that $\gamma > 1$ and $\rho < 1$: that is, that numbers and concentration favoured the Luftwaffe and disfavoured the RAF. This was so in respectively 91% and 86% of cases in a bootstrap analysis[5]

---

[4] In fact this is a good point at which to ask whether time is a confounding factor. Might it be that the RAF did worse earlier on, and the heavy days were mostly in the first phase? But in fact the same conclusion holds for the first phase.

[5] Re-sampling, with replacement, from the known data—effectively a set of plausible alternative realities.

of the first phase. But now notice that $\gamma - \rho = (g_1 + r_2) - (r_1 + g_2) \simeq 0.4$: this quantity is the difference of the two sums of powers, and so is fitted better—and, further, the true values of these sums are probably further from one, so that $\gamma - \rho$ is even larger (and indeed $\gamma > \rho$ in 98% of bootstrap replications of the full battle). Thus numbers and concentration favoured the Luftwaffe more than the RAF.

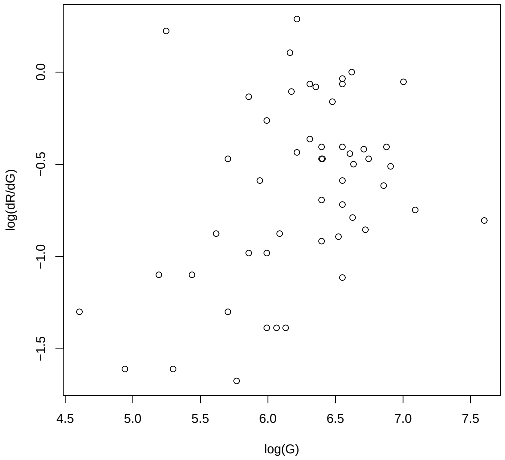


Figure 4: $\log(dR/dG)$ *vs* $\log G$

Of course this bears on the great controversy of the Battle, the use of 'Big Wings'. We could not find any evidence from the data that there was any advantage in mere concentration of numbers for the RAF. Of course the 11 Group commander Keith Park's dilemma was rather that he had to choose, in Forrest's famous dictum, between getting there 'first' and getting there 'with the most', —'there', in this instance, being above and (preferably) up-sun of the Luftwaffe. His view, that 'first' is better, was proved quite correct, by his own later successes in Malta and above all by Leigh-Mallory's disastrous 1941 fighter sweeps.

In terms of the model, to the extent to which $\gamma > \rho$ Park had what we earlier called a 'defender's advantage'. It is impossible to say whether his achievement lay in creating this advantage or exploiting it—whether underlying tactical truths created the model and Park responded, or Park created his own truth.

But we can certainly say that it is a good thing that Lanchester's laws were not acted on, in the sense of being adduced in support of 'Big Wing' strategies, in the Battle of Britain. Of course it is a striking coincidence that the originator of these models should come from the first nation to have to fight an attritional air campaign. It is also true that Lanchester's works influenced the nascent RAF—copies of his book were sent to WWI RFC commanders, and influenced Henderson, probably Trenchard, and above all Fuller. The RAF's Battle of Britain commanders were the product of the culture, developed over a mere twenty years, of this new organization. But that's another story.