

Killing Captain Hindsight

Quantifying Chance in Military History



Niall MacKay

THE UNIVERSITY *of York*

HAAC, September 2022

York Historical Warfare Analysis Group

Niall MacKay and



Jamie Wood



Chris Price



Ian Horwood



CAPTAIN HINDSIGHT!



THE HERO OF THE MODERN AGE

History

History happened once, but

History happened once, but
takes place in an envelope of possibilities which only when realized
become the historical record.

History happened once, but takes place in an envelope of possibilities which only when realized become the historical record.

Counterfactual history

History happened once, but takes place in an envelope of possibilities which only when realized become the historical record.

Counterfactual history has a bad name, but

History happened once, but takes place in an envelope of possibilities which only when realized become the historical record.

Counterfactual history has a bad name, but 'restrained' counterfactual history does no more than restore contingency to the historical narrative and agency to historical actors.

Allan Megill, 'The new counterfactualists', *Historically Speaking* (2004)

History happened once, but takes place in an envelope of possibilities which only when realized become the historical record.

Counterfactual history has a bad name, but 'restrained' counterfactual history does no more than restore contingency to the historical narrative and agency to historical actors.

Allan Megill, 'The new counterfactualists', *Historically Speaking* (2004)

'How *exactly* are we to distinguish probable unrealised alternatives from improbable ones?'

Niall Ferguson, 'Virtual History: Towards a "chaotic" theory of the past', in *Virtual History: Alternatives and Counterfactuals* (1997).



Acton



Carr



Fisher



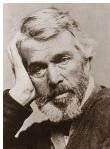
Acton



Carr



Fisher



Carlyle, *On Heroes...*: individuals \Rightarrow history



Acton



Carr



Fisher



Carlyle, *On Heroes...*: individuals \Rightarrow history



Tolstoy: history \Rightarrow individuals



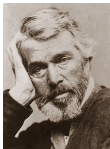
Acton



Carr



Fisher



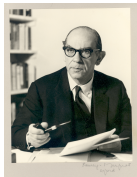
Carlyle, *On Heroes...*: individuals \Rightarrow history



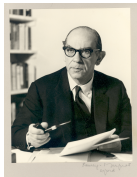
Tolstoy: history \Rightarrow individuals



Buchan, *Causal and Casual in History*

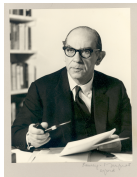


Berlin, *Tolstoy's historical scepticism*



Berlin, *Tolstoy's historical scepticism*



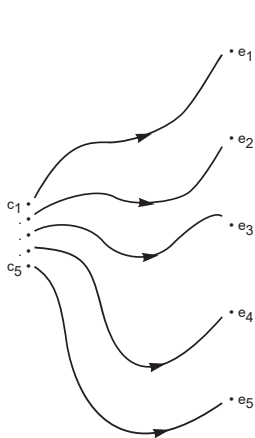


Berlin, *Tolstoy's historical scepticism*

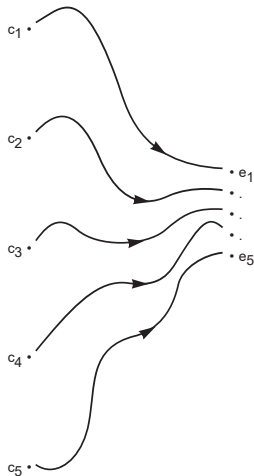


Phil Tetlock, *Good Judgment Project*

two cases in the following way:



Contingency:
Similar causes lead
to different types of effects
High sensitivity to initial
conditions.



Necessity:
Different types of causes
lead to similar effects.
Low sensitivity to initial
conditions.

Critical junctures are much studied by political scientists

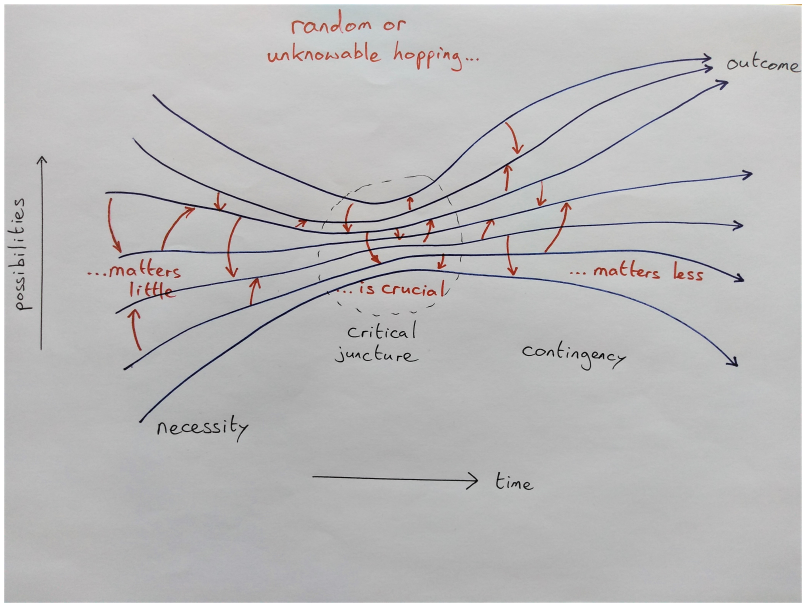
Capoccia & Kelemen. 'The study of critical junctures: theory, narrative, and counterfactuals in historical institutionalism', *World Politics* (2007);

Critical junctures are much studied by political scientists

Capoccia & Kelemen. 'The study of critical junctures: theory, narrative, and counterfactuals in historical institutionalism', *World Politics* (2007);

Geopolitical change cannot be resolved at a timescale finer than decades

Randall Collins, 'The Uses of Counter-Factual History: can there be a theory of historical turning points?' *Amsterdams Sociologisch Tijdschrift* 31 (2004).



Is war declining?



J. R. Statist. Soc. A (2020)
183, Part 3, pp. 909–933

Change point analysis of historical battle deaths

Brennen T. Fagan, Marina I. Knight, Niall J. MacKay and A. Jamie Wood

University of York, UK

[Received April 2019. Revised March 2020]

Summary. It has been claimed and disputed that World War II has been followed by a ‘long peace’: an unprecedented decline of war. We conduct a full change point analysis of well-documented, publicly available battle deaths data sets, using new techniques that enable the robust detection of changes in the statistical properties of such heavy-tailed data. We first test and calibrate these techniques. We then demonstrate the existence of changes, independent of data presentation, in the early to mid-19th century, as the Congress of Vienna system moved towards its collapse, in the early to mid-20th century, bracketing the World Wars, and in the late 20th century, as the world reconfigured around the end of the Cold War. Our analysis provides a methodology for future investigations and an empirical basis for political and historical discussions.

Keywords: Battle deaths; Change point analysis; Correlates of war; Heavy-tailed data; Long peace; Power law distribution

1. Introduction

Is war declining? The record of historical battle deaths surely embodies more human value than any other conceivable data set, for every unit in every data point is a human life violently taken, yet its structure remains poorly understood. Pioneering work was done in the *Journals of the*

Is war declining?



Stephen Pinker: **Yes.**



Nassim Nicholas Taleb: **No.**

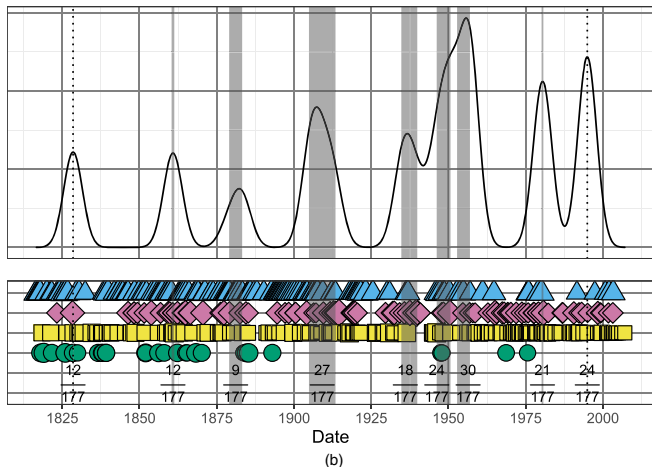


Fig. 12. Results for internal meta-analyses performed on all change points found in any combination of subsets within the data sets (in each plot, there are two images; the lower of each pair of images is a time line of events occurring, sorted by subset; above it is a density estimate of the locations of change points detected; the area under the curve of the estimate is proportional to the probability of finding a change point within that part of the data set; \uparrow , \downarrow , change points, in different locations or the same location respectively, that have been clustered; numbers below the time lines indicate the fraction of identified change points so clustered): (a) COW normalized data (Δ , extrastate war; \diamond , interstate war; \square , intrastate war; \bullet , non-state war; \circ , 10^1 ; \circ , 10^3 ; \circ , 10^5 ; \circ , 10^7); (b) COW normalized data (Δ , extrastate war; \diamond , interstate war; \square , intrastate war; \bullet , non-state war; \circ , 10^{-8} ; \circ , 10^{-6} ; \circ , 10^{-4}) (c) Gleditsch raw data (\square , interstate war; \bullet , civil war; \circ , 10^1 deaths; \circ , 10^3 deaths; \circ , 10^5 deaths; \circ , 10^7 deaths); (d) Gleditsch normalized data (\square , interstate war; \bullet , civil war; \circ , 10^{-8} deaths; \circ , 10^{-6} deaths; \circ , 10^{-4} deaths)

Air Power

Concentration and Asymmetry in Air Combat: Lessons for the Defensive Employment of Air Power

By Ian Horwood, Niall MacKay and Christopher Price

Have air power theorists learned the right lessons from history? We argue that, in the employment of air power to deny air supremacy or defend surface targets, they have not. The classic hypothesized dynamic of air combat is Lanchester's 'Square Law', under which numbers are disproportionately important. Using data from various air campaigns, we demonstrate that air combat does not obey a tactical Square Law. Rather it consistently displays

Lanchester Models and the Battle of Britain

Ian R. Johnson, Niall J. MacKay

Department of Mathematics, University of York, York YO10 5DD, United Kingdom

Received 8 January 2008; revised 1 August 2008; accepted 6 September 2008

DOI 10.1002/nav.20328

Published online 16 December 2008 in Wiley Online Library (wileyonlinelibrary.com).

Abstract: We fit deterministic generalized Lanchester models to daily sortie and loss data from the Battle of Britain. The best fit for the period 14th August to 30th October 1940 is $\delta B \sim G^{1.2}$, $\delta G \sim G^{0.9}$, where B and G are RAF Fighter Command and Luftwaffe sortie numbers, and δB and δG are daily loss numbers, respectively. The data naturally divide into two phases, with losses (as a proportion of overall sortie numbers) much reduced after 15th September. Fits were generally better for the first phase than for the second, and for British losses than for German; in every case the dependence on G is stronger than that on B . Days with higher sortie numbers on average favored the Luftwaffe, whereas the loss-ratio was not significantly dependent on the force ratio.
© 2008 Wiley Periodicals, Inc. Naval Research Logistics 58: 210–222, 2011

Keywords: Lanchester equations; military modeling; air power

1. INTRODUCTION

Among the many ideas in the legacy of the great engineer Frederick Lanchester, the simple mathematical model of combat developed in his 1916 book *Aircraft in Warfare* [20] is one of the most remarkable. His key observation was that, with the accurate, long-range projectile weapons then coming into use, forces could be concentrated in a way hitherto impossible. A fighting force with overwhelming numerical advantage could now attack with all its units, rather than being restricted to a narrow battle-line and thus a fixed ratio of engaged units. The startling conclusion emerges that a force's

land warfare. Second, it is not clear (in the second class of attempt) that each day's fighting is truly independent of what has gone before. Finally, Lanchester models are spatially and temporally homogeneous, allowing no variation in unit type, terrain or tactics, command or control, skill or doctrine. Such homogeneities are rarely found in land warfare, and the main re-emphasis in modern modeling is precisely the inclusion of such considerations. Indeed, it has been convincingly argued that the *avoidance* of the "storm of steel" implicit in Lanchester models has been the central tenet of military doctrine since 1917 [2].

In the light of the title of Lanchester's book, it is most



Scaling laws for air combat

Fit loss-rates to powers of own and enemy numbers:

$$\frac{dG}{dt} = -rR^{r_1}G^{g_2} \quad \frac{dR}{dt} = -gG^{g_1}R^{r_2}$$

Scaling laws for air combat

Fit loss-rates to powers of own and enemy numbers:

$$\frac{dG}{dt} = -rR^{r_1}G^{g_2} \quad \frac{dR}{dt} = -gG^{g_1}R^{r_2}$$

Divide, re-arrange, integrate: we find that

$$\frac{r}{\rho}R^\rho - \frac{g}{\gamma}G^\gamma$$

is constant, where $\rho = 1 + r_1 - r_2$ and $\gamma = 1 + g_1 - g_2$

Scaling laws for air combat

Fit loss-rates to powers of own and enemy numbers:

$$\frac{dG}{dt} = -rR^{r_1}G^{g_2} \quad \frac{dR}{dt} = -gG^{g_1}R^{r_2}$$

Divide, re-arrange, integrate: we find that

$$\frac{r}{\rho}R^\rho - \frac{g}{\gamma}G^\gamma$$

is constant, where $\rho = 1 + r_1 - r_2$ and $\gamma = 1 + g_1 - g_2$,
the **exponents**, capture the conditions of battle:

- Green should concentrate its force if $\gamma > 1$, divide if $\gamma < 1$.
- if $\rho > \gamma$ then Green has a defender's advantage, by a factor ρ/γ

Scaling laws for air combat

The crucial tactical dependence is of the **loss ratio** on force numbers:

$$\frac{dG}{dR} = \frac{r}{g} \frac{R^{\rho-1}}{G^{\gamma-1}}.$$

Linear Law:

$\rho = \gamma = 1$, and $\frac{dG}{dR}$ doesn't depend on R or G .

Square Law:

$\rho = \gamma = 2$, and

$$\frac{dG}{dR} \propto \frac{R}{G}.$$

Asymmetric:

$\rho = 1$, $\gamma = 2$, and

$$\frac{dG}{dR} \propto \frac{1}{G}.$$

The Battle of Britain

A battle of attrition and intended annihilation, in which one day's fighting was much like another, the single-seat fighters on each side were well-matched, and all units were seeking engagement.

The Battle of Britain

A battle of attrition and intended annihilation, in which one day's fighting was much like another, the single-seat fighters on each side were well-matched, and all units were seeking engagement.

Take daily loss-rates for RAF (δR) and Luftwaffe (δG) aircraft and fit to RAF (R) and Luftwaffe (G) daily sortie numbers:

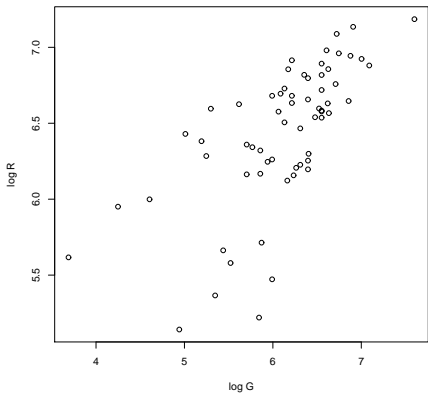
Find the parameters r, r_1, r_2, g, g_1, g_2 for which the data best fit

$$\delta R = g G^{g_1} R^{r_2}, \quad \delta G = r R^{r_1} G^{g_2}$$

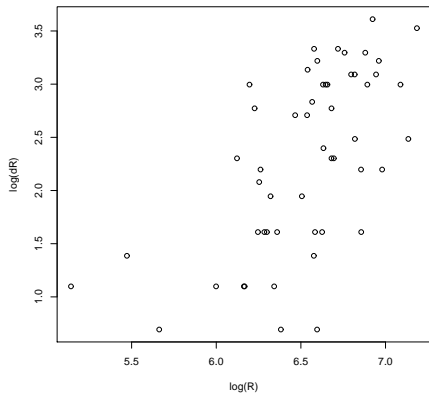
by linear regression onto

$$\log \delta R = \log g + g_1 \log G + r_2 \log R, \quad \log \delta G = \log r + r_1 \log R + g_2 \log G$$

RAF losses



$\log \delta R$ vs $\log G$



$\log \delta R$ vs $\log R$

RAF losses

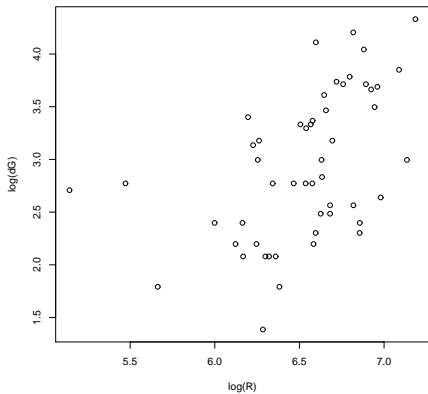
$$\begin{aligned}\frac{dR}{dt} &= -gG^{1.12 \pm 0.17} R^{0.18 \pm 0.25} \\ &= -gG^{1.2} \quad (\Sigma R^2 = 0.66)\end{aligned}$$

RAF losses

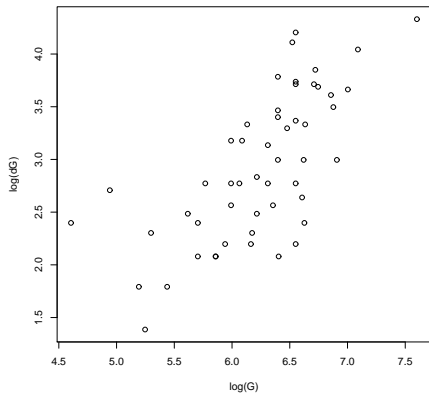
$$\begin{aligned}\frac{dR}{dt} &= -gG^{1.12 \pm 0.17} R^{0.18 \pm 0.25} \\ &= -gG^{1.2} \quad (\Sigma R^2 = 0.66)\end{aligned}$$

Hooray for Lanchester!

Luftwaffe losses



$\log \delta G$ vs $\log R$



$\log \delta G$ vs $\log G$

Luftwaffe losses

$$\begin{aligned}\frac{dG}{dt} &= -rR^{0.00\pm 0.25}G^{0.86\pm 0.18} \\ &= -gG^{0.9} \quad (\Sigma R^2 = 0.49)\end{aligned}$$

Luftwaffe losses

$$\begin{aligned}\frac{dG}{dt} &= -rR^{0.00\pm 0.25}G^{0.86\pm 0.18} \\ &= -gG^{0.9} \quad (\Sigma R^2 = 0.49)\end{aligned}$$

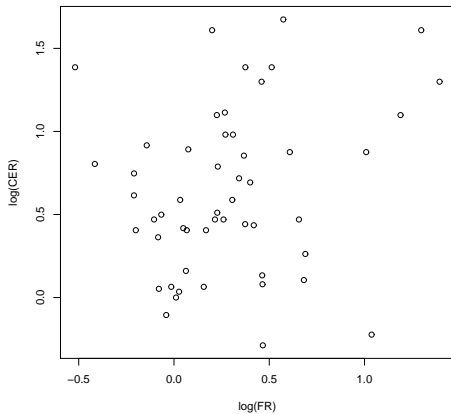
Not so good.

In fact, fitting to R alone,

$$\frac{dG}{dt} \propto R^{0.87\pm 0.22}$$

explains only $\Sigma R^2 = 0.24$.

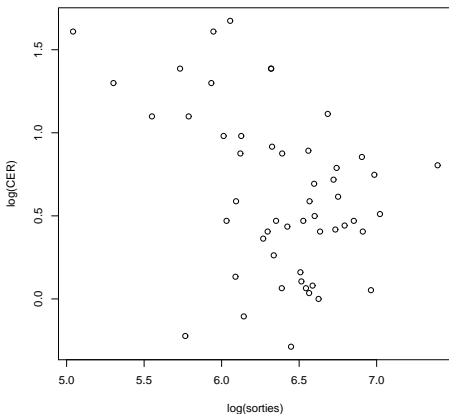
Loss ratio vs Force ratio: Battle of Britain



$\log dG/dR$ vs $\log R/G$

G =Luftwaffe, R =Royal Air Force

Loss ratio vs Sorties: Battle of Britain



$\log dG/dR$ vs $\frac{1}{2} \log(RG)$, G =Luftwaffe, R =Royal Air Force

$$\sum R^2 = 0.17, p = 0.003$$

Scaling laws for air combat

'The dependence of the casualty exchange ratio on the force ratio is not linear; it is exponential'

– Col. John Warden, USAF, *The Air Campaign*

Cites a 1970 study of Korea and WW2.

Scaling laws for air combat

'The dependence of the casualty exchange ratio on the force ratio is not linear; it is exponential'

– Col. John Warden, USAF, *The Air Campaign*

Cites a 1970 study of Korea and WW2.

Well, no.



Is Air Combat Lanchestrian?

Niall J. MacKay, Department of Mathematics, University of York, UK, niall.mackay@york.ac.uk

In his seminal work on the mathematics of warfare, Lanchester's purpose was to make a crucial distinction between what he called "ancient" warfare, in which combat is essentially a set of duels, and "modern" war, in which, with the advent of long-range aimed projectile weapons, combatants with the advantage of numbers could concentrate their fire, many-on-one (Lanchester 1914). When these assumptions are built into the simplest possible dynamical system, ancient war produces a linear law, in which fighting strength is given by units' individual effectiveness multiplied by their numbers, while modern war results in a square law, in which strength is individual effectiveness times numbers squared. Which of these holds for air combat? Lanchester thought the answer was clear: air combat is modern, square-law war, and he included his original articles in a book on *Aircraft in Warfare* (Lanchester 1916).^a The same answer has seemed obvious to modern commentators, too (Warden 1989).

In the last couple of decades, the standard way to perform a Lanchestrian campaign analysis has been to fit the two sides' loss rates to (possibly different) monomial scalings—one writes each side's loss rate as a power of its own numbers multiplied by a power of its opponents' (Bracken 1995,

and Kimball, who number among the fathers of operations research, treat the CER separately from the Lanchester equations (Morse and Kimball 1951, sec. 3.2.1).

To see how this works, let B and R be Blue and Red sortie rates, and $-dB/dt$ and $-dR/dt$ be the corresponding loss rates. Lanchester's aimed-fire equations are

$$\frac{dB}{dt} = -rR, \quad \frac{dR}{dt} = -bB, \quad (1)$$

where r (resp. b) are the losses caused by Red (resp. Blue) per sortie. Upon dividing, we see that the CER

$$\frac{dR}{dB} = \frac{bB}{rR}, \quad (2)$$

which we separate to give $bB dB = rR dR$ and integrate to give the square law $(1/2)rR^2 = (1/2)bB^2 + \text{constant}$. Thus the signature of the square law is that the CER is proportional to the force ratio (FR) B/R , while that of the linear law $rR = bB + \text{constant}$ is that the CER is constant.

One might be tempted to test between linear and square laws by performing a simple linear regression fit of the CER to the FR. However, this would be mistaken, for the conclusions are not invariant under $B \leftrightarrow R$, because the significance of β in fitting $dB/dR = \alpha + \beta(R/B)$ is different from that in fitting $dR/dB = \alpha + \beta(B/R)$. Rather, one

Two Campaigns of World War Two

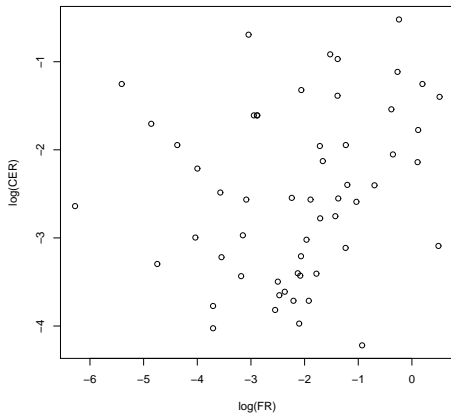
Here we look at two campaigns for which we have finely resolved sortie and loss data, the 1940 Battle of Britain and the 1941–1945 carrier-based Pacific air campaign. For the former, we have daily data, whereas for the latter, the data are at the level of a carrier operation, ranging from a single raid to several weeks' action. Of course, ideal data would be for individual, indivisible engagements. When one sums a number of smaller, independent engagements, the resulting loss rates are pushed toward linear dependence on sortie numbers (owing, at root, to what mathematicians call "Jensen's inequality"). To the extent to which larger loss rates are due to their being the sums of more rather than larger engagements, aggregation effects can obscure evidence for the square law.

The Battle of Britain, 1940

Let B be Royal Air Force (RAF) and R (estimated) Luftwaffe daily sortie numbers, and $\delta B := -dB/dt$ and $\delta R := -dR/dt$ be RAF and Luftwaffe daily loss rates, discretized to $dt = 1$ day (Johnson and MacKay 2011). The logarithm of the CER $= \delta R/\delta B$ is plotted against that of the FR $= R/B$ in Figure 1(a). If we fit

$$\frac{\delta R}{\delta B} = \alpha \left(\frac{B}{R} \right)^\beta, \quad (4)$$

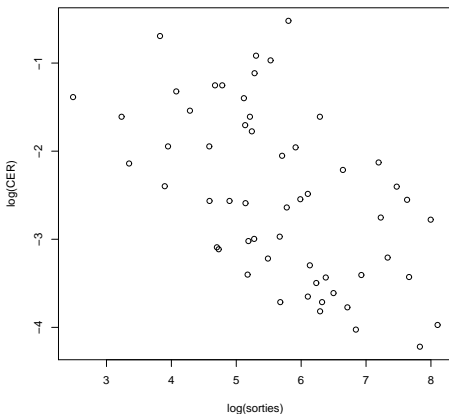
Loss ratio vs Force ratio: Pacific air war



$\log dG/dR$ vs $\log R/G$

G =Americans, R =Japanese

Loss ratio vs Sorties: Pacific air war



$\log dG/dR$ vs $\frac{1}{2} \log(RG)$, G =Americans, R =Japanese

$$\sum R^2 = 0.30, p = 10^{-5}$$

Asymmetry in air combat

What are the exponents for air combat?

Asymmetry in air combat

What are the exponents for air combat?

Battle of Britain: Germans 1.3 , British 0.8

Asymmetry in air combat

What are the exponents for air combat?

Battle of Britain: Germans 1.3 , British 0.8

Pacific air war: Americans 1.3, Japanese 0.9

Asymmetry in air combat

What are the exponents for air combat?

Battle of Britain: Germans 1.3 , British 0.8

Pacific air war: Americans 1.3, Japanese 0.9

Korea: Americans 1.2, North Koreans 0.1

– and these differences are *understated*.

Vietnam 1965-68; Rolling Thunder

Engagement-level data, and a simple linear regression of loss rates against numbers.

Does a sortie lead to a kill, a loss, or neither?

Vietnam 1965-68; Rolling Thunder

Engagement-level data, and a simple linear regression of loss rates against numbers.

Does a sortie lead to a kill, a loss, or neither?

F4 (US fighter) sorties tend to cause NVAF (but not US) losses.

F105 (US bomber) sorties tend to cause neither.

US conclusion: F4s should sortie in numbers.

Vietnam 1965-68; Rolling Thunder

Engagement-level data, and a simple linear regression of loss rates against numbers.

Does a sortie lead to a kill, a loss, or neither?

F4 (US fighter) sorties tend to cause NVAF (but not US) losses.
F105 (US bomber) sorties tend to cause neither.

US conclusion: F4s should sortie in numbers.

NVAF (MiG 17,19,21) sorties tend to cause **own** losses, whether against F4s or F105s.

NVAF conclusion: sortie sparingly, disrupt, avoid engagement.

The Battle of Britain: The Big Wing

Should the RAF's squadrons mass into wings (3 squadrons)
or 'Big Wings' (5 or more) before engaging?

The Battle of Britain: The Big Wing

Should the RAF's squadrons mass into wings (3 squadrons)
or 'Big Wings' (5 or more) before engaging?

Is mere concentration of numbers advantageous for the RAF?

The Battle of Britain: The Big Wing

Should the RAF's squadrons mass into wings (3 squadrons)
or 'Big Wings' (5 or more) before engaging?

Is mere concentration of numbers advantageous for the RAF?

Is the RAF exponent $\rho > 1$?

The Battle of Britain: The Big Wing

Should the RAF's squadrons mass into wings (3 squadrons)
or 'Big Wings' (5 or more) before engaging?

Is mere concentration of numbers advantageous for the RAF?

Is the RAF exponent $\rho > 1$?

No, $\rho \leq 0.8$.

The Battle of Britain: The Big Wing

Should the RAF's squadrons mass into wings (3 squadrons) or 'Big Wings' (5 or more) before engaging?

Is mere concentration of numbers advantageous for the RAF?

Is the RAF exponent $\rho > 1$?

No, $\rho \leq 0.8$.

The RAF's advantages, created and exploited by AVM Keith Park, were of **dispersal** and **parsimony**.



Safety in Numbers: Ideas of Concentration in Royal Air Force Fighter Defence from Lanchester to the Battle of Britain

NIALL MACKAY

University of York

CHRISTOPHER PRICE

York St John University

Abstract

This article examines the Big Wing controversy in the Battle of Britain according to the RAF's understanding of the principle of concentration. RAF doctrine was developed from Fuller's 1916 codification of the principles of war, which in terms of concentration rested on the work of engineer and air-power theorist F. W. Lanchester. We argue that fighter doctrine diverged from its Lanchestrian core between the wars and allowed conflicting interpretations of concentration in British air defence. We conclude that Park's conduct of the Battle of Britain in 11 Group conformed closely to the Lanchester model in concentrating British resources and denying targets to the enemy. Conversely, the Big Wing failed to provide operational concentration and presented the enemy with a massed target it was his mission to destroy. Data suggest greater British losses relative to the enemy when more aircraft were engaged, though damagingly for future policy the mathematics of over-claiming indicated the opposite.

Asymmetry in air combat

The starting hypothesis:

Air combat is a set of duels; random; linear-law.

Asymmetry in air combat

The starting hypothesis:

Air combat is a set of duels; random; linear-law.

To the extent to which it departs from this,

Air combat does not obey the Square Law.

Asymmetry in air combat

The starting hypothesis:

Air combat is a set of duels; random; linear-law.

To the extent to which it departs from this,

Air combat does not obey the Square Law.

Rather,

Air combat is 80% linear-law, 20% asymmetric.

Asymmetry in air combat

The starting hypothesis:

Air combat is a set of duels; random; linear-law.

To the extent to which it departs from this,

Air combat does not obey the Square Law.

Rather,

Air combat is 80% linear-law, 20% asymmetric.

While there appears to be some Lanchestrian advantage in numbers and concentration for the attacker (of surface targets), there is none for the defender – the reverse, even.

The defender needs to disperse, disrupt; be sparing, parsimonious; make minimal, fleeting attacks; utilize information asymmetry; work with ground-based air defence.

Bootstrapping the Battle of Britain



Brennen Fagan, Ian Horwood, Niall MacKay,
Christopher Price, Ed Richards, and A. Jamie Wood*

Abstract

The Battle of Britain (1940) is the focus of much historical controversy. We show here how the statistical technique of *weighted bootstrapping* can be used to create a new quantitative basis to help address such controversies. Bootstrapping facilitates the exploration of alternative campaign possibilities with different tactics. This results in comparative probabilities of “victory” for the actual and various counterfactual campaigns, providing a quantified assessment of the likelihood of German achievement of air superiority, thereby facilitating invasion. We find this more likely had the Luftwaffe targeted airfields more heavily, and greatly more likely had Germany brought forward its air campaign.

1. Introduction

The Battle of Britain (1940) continues to generate fierce controversy amongst historians as its eightieth anniversary approaches. The conflict between critics and

Weighted Bootstrapping for History

Bootstrapping

Weighted Bootstrapping for History

Bootstrapping is re-sampling with replacement.

Weighted Bootstrapping for History

Bootstrapping is re-sampling with replacement.

It is **model-free**,

Weighted Bootstrapping for History

Bootstrapping is re-sampling with replacement.

It is **model-free**, and I would argue that it provides maximal reconciliation between pro- and anti-counterfactualists.

Weighted Bootstrapping for History

Bootstrapping is re-sampling with replacement.

It is **model-free**, and I would argue that it provides maximal reconciliation between pro- and anti-counterfactualists.

Weighted bootstrapping

Weighted Bootstrapping for History

Bootstrapping is re-sampling with replacement.

It is **model-free**, and I would argue that it provides maximal reconciliation between pro- and anti-counterfactualists.

Weighted bootstrapping re-weights the data according to underlying features

Weighted Bootstrapping for History

Bootstrapping is re-sampling with replacement.

It is **model-free**, and I would argue that it provides maximal reconciliation between pro- and anti-counterfactualists.

Weighted bootstrapping re-weights the data according to underlying features and so provides a method for exploring restrained counterfactual (CF) history.

Weighted Bootstrapping for History

Bootstrapping is re-sampling with replacement.

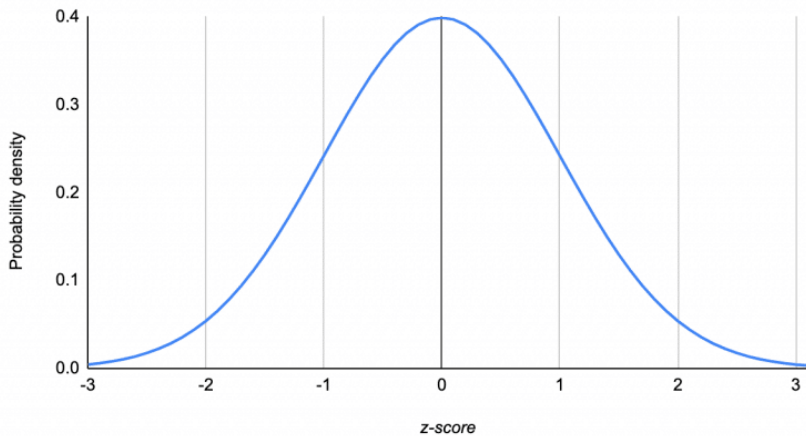
It is **model-free**, and I would argue that it provides maximal reconciliation between pro- and anti-counterfactualists.

Weighted bootstrapping re-weights the data according to underlying features and so provides a method for exploring restrained counterfactual (CF) history.

If we measure the effects of a CF in multiples of the standard deviation, we can quantify by how much it is rational for a historian to change their views (about probabilities of victory) in CF scenarios.

See also Max Little and Reham Badawy, 'Causal bootstrapping', arXiv:1910.09648 (2019).

Standard normal distribution



Weighted Bootstrapping of the Battle of Britain

Weighted Bootstrapping of the Battle of Britain

No switch to attacking London: 2 SD.

Weighted Bootstrapping of the Battle of Britain

No switch to attacking London: 2 SD.

CF Goering has a clear conception of the necessity of defeating Fighter Command.

Luftwaffe targets airfields or London: 3 SD.

Weighted Bootstrapping of the Battle of Britain

No switch to attacking London: 2 SD.

CF Goering has a clear conception of the necessity of defeating Fighter Command.

Luftwaffe targets airfields or London: 3 SD.

CF Hitler has a clear, early conception, before the defeat of France, of the strategic necessity of also defeating Britain.

Battle of Britain starts a month earlier: 3 SD.

Weighted Bootstrapping of the Battle of Britain

No switch to attacking London: 2 SD.

CF Goering has a clear conception of the necessity of defeating Fighter Command.

Luftwaffe targets airfields or London: 3 SD.

CF Hitler has a clear, early conception, before the defeat of France, of the strategic necessity of also defeating Britain.

Battle of Britain starts a month earlier: 3 SD.

CF Hitler and **CF Goering** combined: 6 SD.

Sea Power

Weighing the fog of war: Illustrating the power of Bayesian methods for historical analysis through the Battle of the Dogger Bank

Niall MacKay^a, Christopher Price^b, and A. Jamie Wood^c

^aDepartment of Mathematics, University of York; ^bProgramme of History and American Studies, York St. John University; ^cDepartment of Biology, University of York

ABSTRACT

The application of scientific methods to historical situations is restricted by the existence of a single outcome with no possibility of repetition. However, new computational methods make quantitative historical analysis possible. The authors apply methods of approximate Bayesian computation to simulate a naval engagement of the First World War, the Battle of the Dogger Bank. They demonstrate that the battle's outcome was highly unlikely, with significant implications both for subsequent actions and for historical understanding. Dogger Bank exemplifies the view that Bayesian methods offer historians the tool they need to grapple with the evolving probabilities of historical events, giving a sound scientific basis for counterfactual history and opening up a wealth of possibilities for analysis.

KEYWORDS

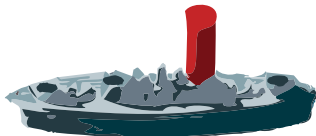
approximate Bayesian computation (ABC); Battle of the Dogger Bank; combat modeling; military simulation; naval warfare

Counterfactual history has provoked an unfortunate response in academia: “Grown-up historians don’t waste time on counterfactuals,” according to renowned historian Michael Howard.¹ In this perspective, historians should confine themselves to what actually happened. Yet, if their fundamental task is to explain why the past developed as it did, they must contend with other possible outcomes: The course of history is complex, contingent, and develops within an envelope of probabilities surrounding the historical narrative. Hindsight promotes the post hoc fallacy,

precisely that of Bayesian probability: to understand historical actors’ implicit prior estimates of chances, and how these changed as events unfolded. Where the evolution of probabilities can be quantified with scientific methods, it is the responsibility of historians to include this in their thinking. Niall Ferguson, a supporter of counterfactual history, has nevertheless asked, “How exactly are we to distinguish probable unrealized alternatives from improbable ones?” (Ferguson 2000, 86). In some cases, Bayesian methods provide the answer (Bayliss et al. 2007).

IN DETAIL

Dogger Bank: Weighing the fog of war



Historical outcomes were at one time only possibilities – but how do we distinguish probable real events from improbable ones? **Niall MacKay**, **Chris Price** and **A. Jamie Wood** use a naval battle of the First World War to explain how Bayesian thinking helps historians reason with the uncertainties of the past

Academic historians reject simple narratives of the past, the “one damn thing after another” cliché of historical study. Yet most still believe human history to be an ordered process, albeit with complex and subtle rules of causation and some element of chance. Thus, for example, the US entry into the Second World War was probable but not inevitable, and the Japanese strike on Pearl Harbor was the culminating event in an intricate play of political and diplomatic interactions.

This view carries with it the recognition of uncertainty, and the frightening truth that chance in history has shaped our lives. Asking the “what if?” question is unavoidable in historical study. Had Grouchy rather than the Prussians arrived at Waterloo, as many present expected, or had Britain adopted what appeared to be the rational course and made peace with Hitler in 1940, we would now inhabit different worlds. Of course, had the first of these happened, the second almost certainly could not have arisen.

This uncertainty is hard to grapple with. Renowned historian Michael Hirst said once that if there were no historians, death

RIGHT Blücher, one of the German battleships engaged by British forces during the Battle of Dogger Bank

But this approach can only work in very special, uncommon circumstances. Essentially one needs a simple situation governed by a well-founded model with tightly estimated parameters. Then, when the actual historical outcome turns out to be extremely unlikely, one can argue that this must indeed be the case rather than that the model or parameters are wrong.

All these criteria are met in the first all-big-gun naval battle of the dreadnought era. Early in the morning on 24 January 1915, in the first year of the First World War, the battle-cruisers of the British and German navies lined up in the North Sea, 60 miles off the east coast of England, for an engagement in which the fleets would steam in straight lines for several hours, exchanging fire at distance in what immediately became known as the Battle of Dogger Bank.

The battle

Dogger Bank is an excellent case study for this type of analysis. The battle was a simple situation with distinct choices on

The Battle of the Dogger Bank

The first clash of battle cruisers, 24th January 1915



Blücher



Derfflinger



Moltke



Seydlitz



Indomitable



New Zealand



Princess Royal



Tiger

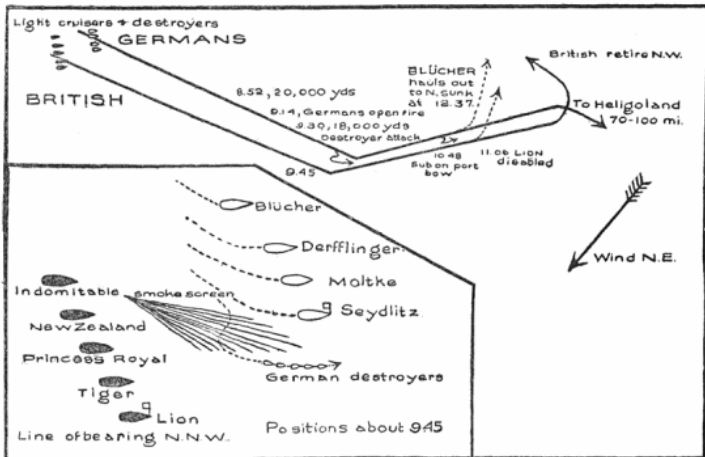


Lion

The Battle of the Dogger Bank

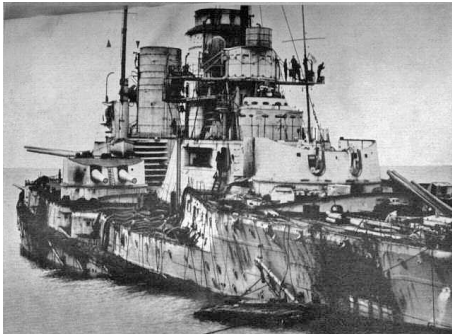


The Battle of the Dogger Bank



First sighting 0705, battlecruisers engage 0852

The Battle of the Dogger Bank



0943 *Seydlitz* damaged

The Battle of the Dogger Bank



1018 *Lion* hit

The Battle of the Dogger Bank

1100-1105

Course North-East

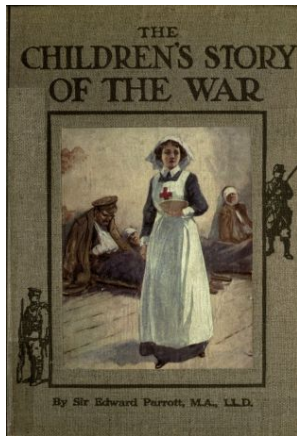
Engage the enemy's rear

The Battle of the Dogger Bank



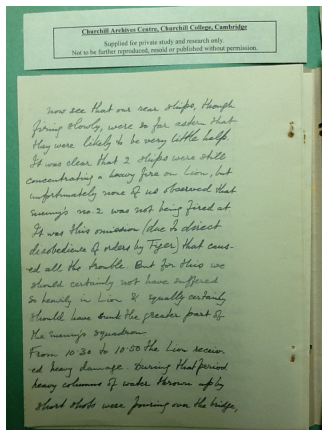
1313 *Blücher* sinks

The Battle of the Dogger Bank



'Only by sheer good luck did any of the German ships escape.'

The Battle of the Dogger Bank



'But for *Tiger's* misdirection of fire we ... certainly should have sunk the greater part of the enemy's squadron'

– diary entry, Cdr (later Vice Adm) Reginald Plunkett-...-Drax

[COPYRIGHTED]

U. S. NAVAL INSTITUTE, ANNAPOLIS, MD.

A PLEA FOR THE BATTLE-CRUISER

By ASSISTANT NAVAL CONSTRUCTOR B. S. BULLARD, U. S. Navy

Prior to the introduction of steam for the propulsive element and of iron or steel for the construction of the hulls of fighting ships, the frigates, fast, comparatively lightly armed sailing vessels, performed all of the multitudinous duties which tactical and strategical considerations of the present day allot to the cruiser. This class of vessel carried on the scouting or despatch service; acted as protectors or destroyers of commerce; took their place in line of battle in concerted fleet action; performed all the duties connected with detached service; and their value as part of any naval establishment which, in time of war, aimed at the control of the sea, was fully recognized and universally understood. However, the true worth of the cruiser was forgotten in the tumultuous wave of enthusiasm which swept over England and France, the leading maritime nations of the period, when wrought iron was introduced in warships for the protection of the ships and of the gun-crews against shell fire. The introduction of armor led to a controversy among the world's leading naval architects upon the relative merits of the casemate ship, where all the guns were grouped in an armored casemate amidships, and the turret ship, where all the guns were isolated in separate armored turrets, but where nearly all of the units comprising a ship's main battery could be brought to bear upon any point of the horizon. These and many other momentous questions connected with the heavily armored first-line or ironclad ships occupied the attention of the naval architects of England and France during this period and the frigate or cruiser's natural course of development, parallel to that of the heavier ships, was cast into the shadow. In spite of this halt in the development of the cruiser, all of the considera-

U. S. NAVAL INSTITUTE, ANNAPOLIS, MD.

THE ARRIVAL OF THE BATTLE-CRUISER

By COMMANDER YATES STIRLING, U. S. Navy

THE BATTLESHIP

"The criterion of warship type will be found in a study of their ultimate service."

The acceptance of this maxim and its application to capital ships has committed the United States to the battleship and to the total exclusion of the battle-cruiser. No fault can be found in the maxim, but, in its application, unfortunately economic considerations have forced the United States Navy to adhere to the pure type in which guns, armor and speed are maintained in proportionate quantities.

"The ultimate test of war is battle. In a fleet action gun power and armor protection are the crowning attributes."

Again the maxim is sound, but have we correctly applied it?

History has shown that in all wars one side will take the initiative and act on the offensive, while the other will surrender the initiative and act upon the defensive. The nation acting upon the offensive does so because it feels itself the stronger. Its fleet, by virtue of its superior strength in type, in material and in morale, will attempt to bring the other fleet to action. The other fleet, the defensive one, will await the attack of its enemy within its own waters and probably behind its own fortifications. This is the condition now existing in the great war in Europe. England, with the stronger fleet, has taken the offensive against the weaker German fleet, which lies within the safety of its strongly fortified harbors of Cuxhaven and Kiel.

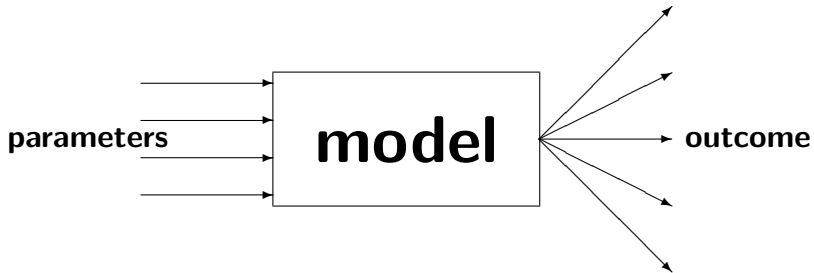
In the application of these two maxims of "ultimate service," there lies the germ of misunderstanding. The idea that the enemy fleet would immediately operate to bring our fleet to action is in itself sound, but have the methods of accomplishment been correctly determined? Once the idea of a decisive battle is considered, attention at once focuses upon being strong at the point

Modeling The Battle of the Dogger Bank

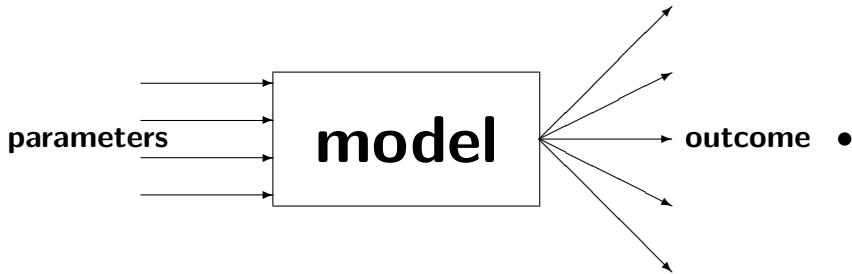
Approximate Bayesian Computation (ABC)

Marjoram, Molitor, Plagnol & Tavaré, Markov chain Monte Carlo without likelihoods, *Proc. Nat. Acad. Sci.* 2003

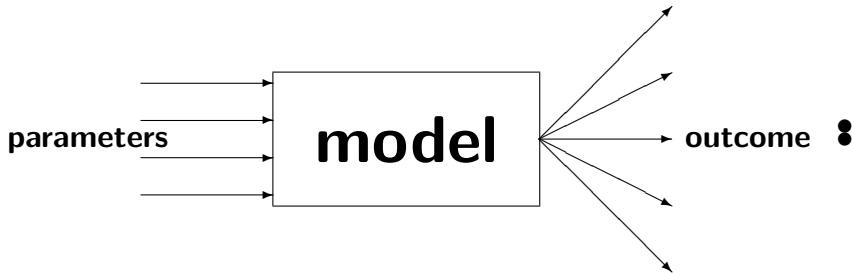
Modeling The Battle of the Dogger Bank



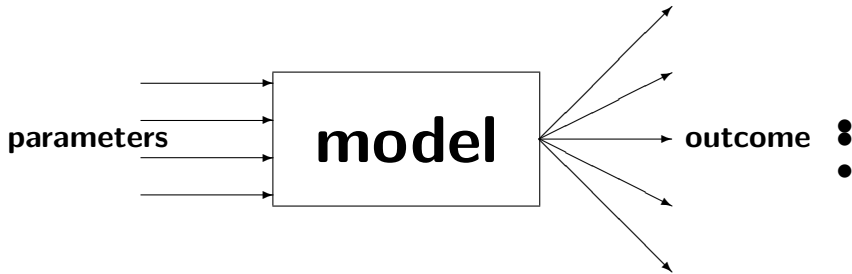
Modeling The Battle of the Dogger Bank



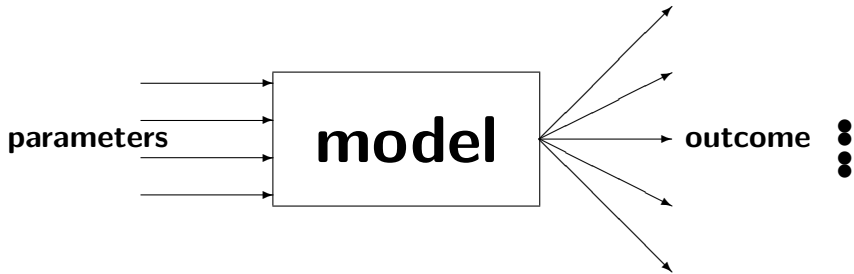
Modeling The Battle of the Dogger Bank



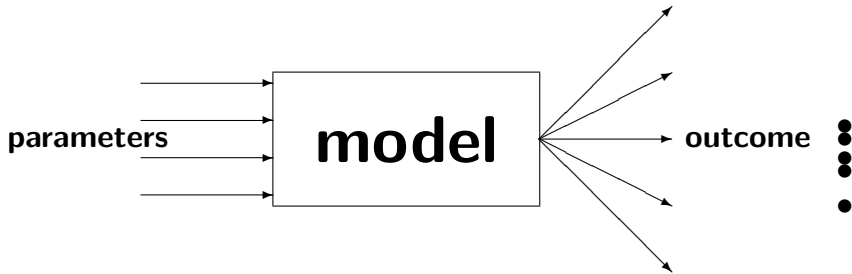
Modeling The Battle of the Dogger Bank



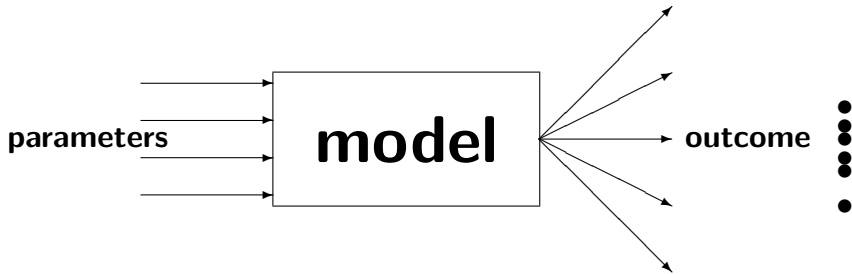
Modeling The Battle of the Dogger Bank



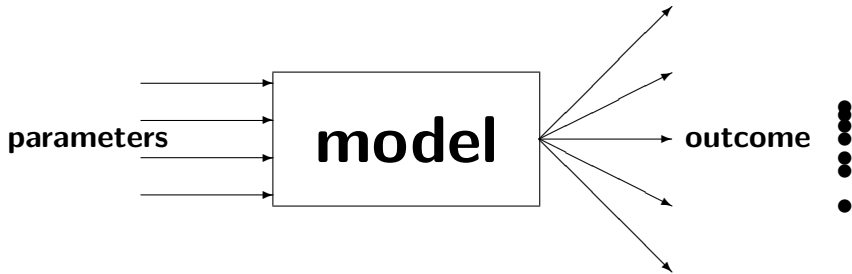
Modeling The Battle of the Dogger Bank



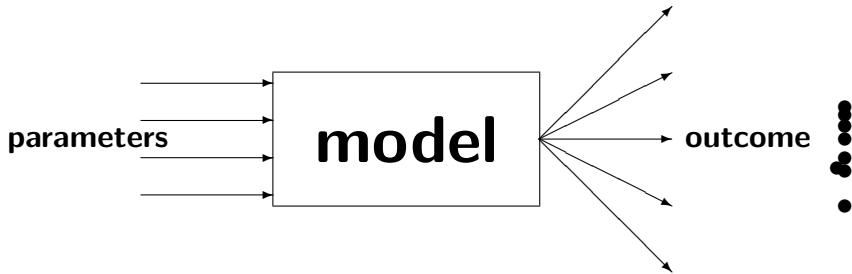
Modeling The Battle of the Dogger Bank



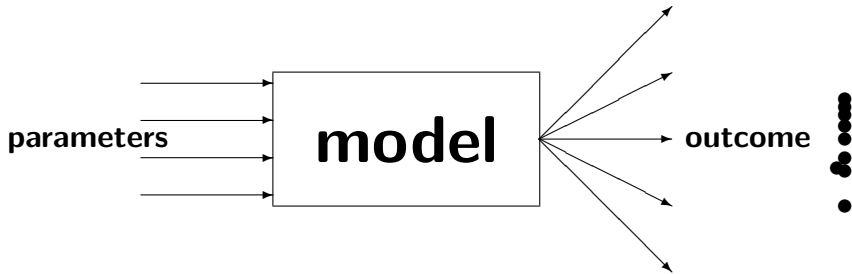
Modeling The Battle of the Dogger Bank



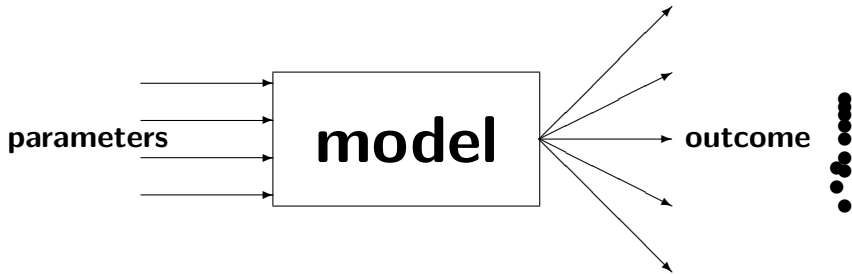
Modeling The Battle of the Dogger Bank



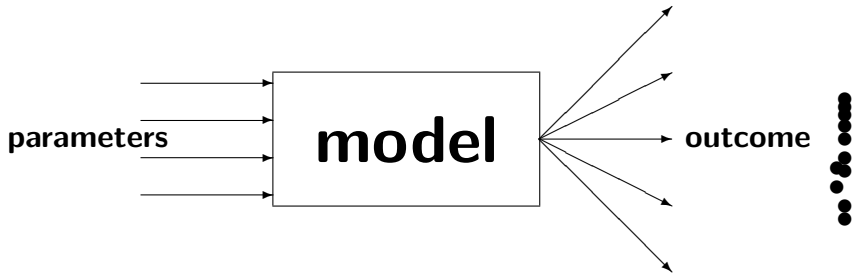
Modeling The Battle of the Dogger Bank



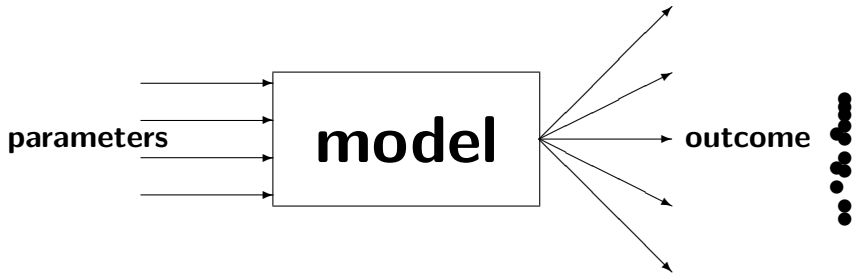
Modeling The Battle of the Dogger Bank



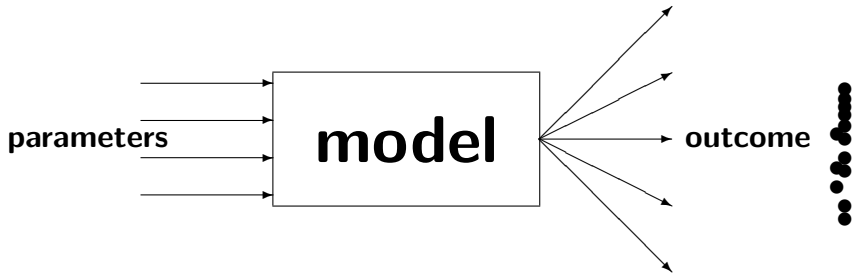
Modeling The Battle of the Dogger Bank



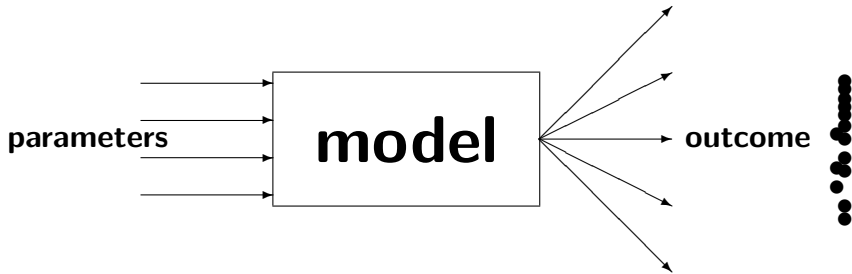
Modeling The Battle of the Dogger Bank



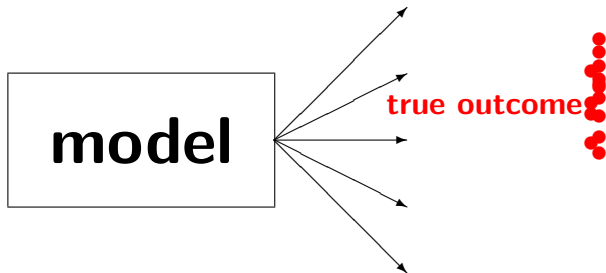
Modeling The Battle of the Dogger Bank



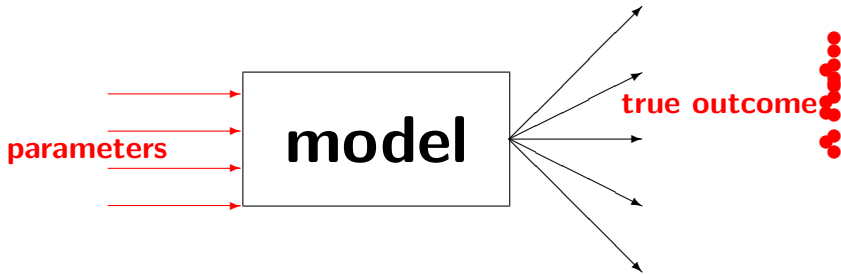
Modeling The Battle of the Dogger Bank



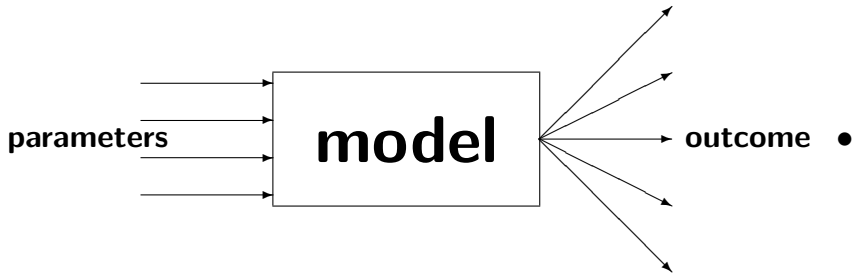
Modeling The Battle of the Dogger Bank



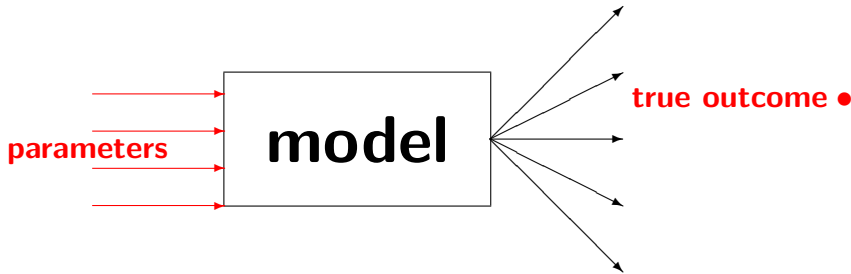
Modeling The Battle of the Dogger Bank



Modeling The Battle of the Dogger Bank

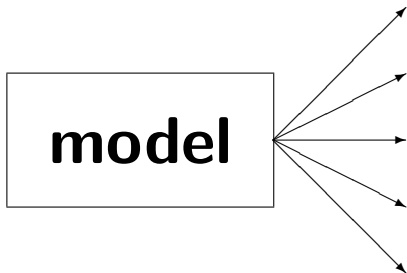


Modeling The Battle of the Dogger Bank



Modeling The Battle of the Dogger Bank

true outcome ●



Modeling the Battle of the Dogger Bank

Either

Modeling the Battle of the Dogger Bank

Either

the model is wrong

or

Modeling the Battle of the Dogger Bank

Either

the model is wrong

or

the parameters are (very) wrong

or

Modeling the Battle of the Dogger Bank

Either

the model is wrong

or

the parameters are (very) wrong

or

something improbable has happened.

The Model

The Model

A force causes damage in proportion to its numbers

Baudry, Chase, Fiske, Lanchester, Osipov (1902-1916)

The Model

A force causes damage in proportion to its numbers

Baudry, Chase, Fiske, Lanchester, Osipov (1902-1916)

'Your N-squared law has become quite famous in the Grand Fleet'
— Jellicoe to Lanchester, 1916

The Model

A force causes damage in proportion to its numbers

Baudry, Chase, Fiske, Lanchester, Osipov (1902-1916)

'Your N-squared law has become quite famous in the Grand Fleet'
— Jellicoe to Lanchester, 1916

Numbers are of the units which stand/fall together:
turrets (Fiske, 1905) or **ships**.

The Model

A force causes damage in proportion to its numbers

Baudry, Chase, Fiske, Lanchester, Osipov (1902-1916)

'Your N-squared law has become quite famous in the Grand Fleet'
— Jellicoe to Lanchester, 1916

Numbers are of the units which stand/fall together:
turrets (Fiske, 1905) or **ships**.

True? Ship concentrations of 2:1, 3:1, 4:1 were given multipliers of 2, 2.5 and 3 in exercises

Gunnery Practices in the Grand Fleet 1914-1918, ADM 137/4822, ADM 186/339, Kew

The Model

A force causes damage in proportion to its numbers

Baudry, Chase, Fiske, Lanchester, Osipov (1902-1916)

'Your N-squared law has become quite famous in the Grand Fleet'
— Jellicoe to Lanchester, 1916

Numbers are of the units which stand/fall together:
turrets (Fiske, 1905) or **ships**.

True? Ship concentrations of 2:1, 3:1, 4:1 were given multipliers of 2, 2.5 and 3 in exercises

Gunnery Practices in the Grand Fleet 1914-1918, ADM 137/4822, ADM 186/339, Kew

— more than good enough.

The Parameters

Deduced from Dogger Bank and Jutland, taken together
—valid *unless* they changed greatly between Dogger Bank and
Jutland

The Parameters

Deduced from Dogger Bank and Jutland, taken together
—valid *unless* they changed greatly between Dogger Bank and
Jutland

Shells, ships/guns/armour, gunnery practice,

The Parameters

Deduced from Dogger Bank and Jutland, taken together
—valid *unless* they changed greatly between Dogger Bank and
Jutland

Shells, ships/guns/armour, gunnery practice, *flash fire!*

The Parameters

Deduced from Dogger Bank and Jutland, taken together
—valid *unless* they changed greatly between Dogger Bank and Jutland

Shells, ships/guns/armour, gunnery practice, *flash fire!*

On the German side big changes happened, because of *Seydlitz*

The Parameters

On the British side they probably didn't.

The Parameters

On the British side they probably didn't.

Nicholas Lambert, "Our bloody ships" or "Our bloody system"? Jutland and the loss of the British battle-cruisers, 1916, *JMH* 62 (1998) 29-55

The Parameters

On the British side they probably didn't.

Nicholas Lambert, "Our bloody ships" or "Our bloody system"? Jutland and the loss of the British battle-cruisers, 1916, *JMH* 62 (1998) 29-55

"a mistake was made in firing too slowly during the earlier stages *rapidity of fire* is essential ...rapid fire will be employed by the enemy at 18,000 yds, which must be answered by rapid fire"

The Parameters

On the British side they probably didn't.

Nicholas Lambert, "Our bloody ships" or "Our bloody system"? Jutland and the loss of the British battle-cruisers, 1916, *JMH* 62 (1998) 29-55

"a mistake was made in firing too slowly during the earlier stages *rapidity of fire* is essential ...rapid fire will be employed by the enemy at 18,000 yds, which must be answered by rapid fire" *but* "Plunging fire is a great danger to ammunition anywhere between decks. ... Lids of powder cases should not be removed faster than necessary." – Ernle Chatfield, captain of *Lion*

The Parameters

On the British side they probably didn't.

Nicholas Lambert, "Our bloody ships" or "Our bloody system"? Jutland and the loss of the British battle-cruisers, 1916, *JMH* 62 (1998) 29-55

"a mistake was made in firing too slowly during the earlier stages *rapidity of fire* is essential ...rapid fire will be employed by the enemy at 18,000 yds, which must be answered by rapid fire" *but* "Plunging fire is a great danger to ammunition anywhere between decks. ... Lids of powder cases should not be removed faster than necessary." – Ernle Chatfield, captain of *Lion*

An Admiralty memorandum of February 1915 urges better flash discipline

The Parameters

On the British side they probably didn't.

Nicholas Lambert, "Our bloody ships" or "Our bloody system"? Jutland and the loss of the British battle-cruisers, 1916, *JMH* 62 (1998) 29-55

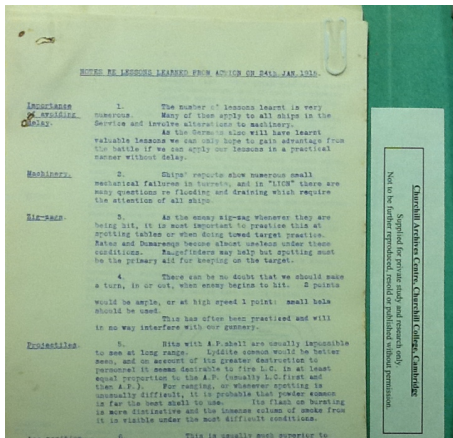
"a mistake was made in firing too slowly during the earlier stages ... *rapidity of fire* is essential ...rapid fire will be employed by the enemy at 18,000 yds, which must be answered by rapid fire" *but* "Plunging fire is a great danger to ammunition anywhere between decks. ... Lids of powder cases should not be removed faster than necessary." – Ernle Chatfield, captain of *Lion*

An Admiralty memorandum of February 1915 urges better flash discipline ... but was not widely acted upon.

The Parameters

Nothing in the 'lessons learned'

although 'German shell, for incendiary effect and damage to personnel, are far inferior to ours. Their only good quality lies in armour penetration and damage to material.'



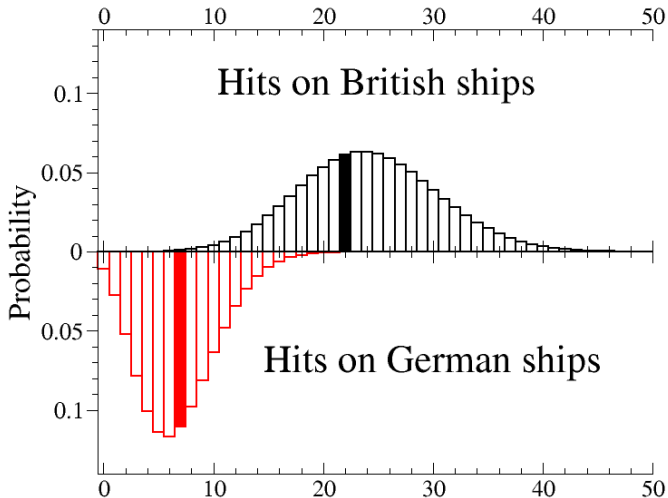
The Results

Using

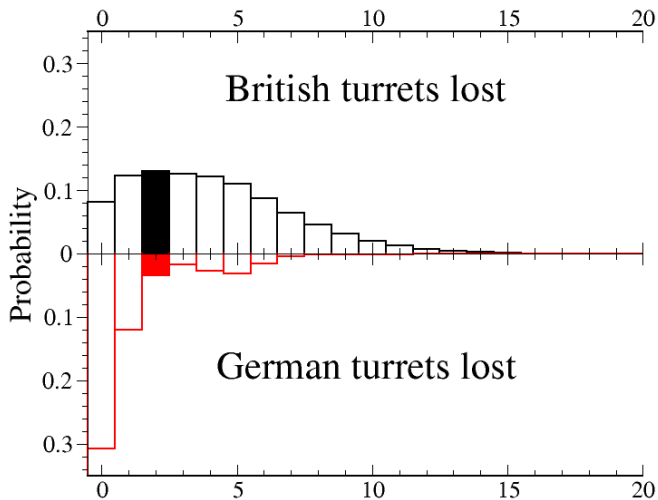
Approximate Bayesian Computation (ABC)

we have a standardized, optimal methodology with which to explore *all* of the parameter space for its capacity to reproduce real results.

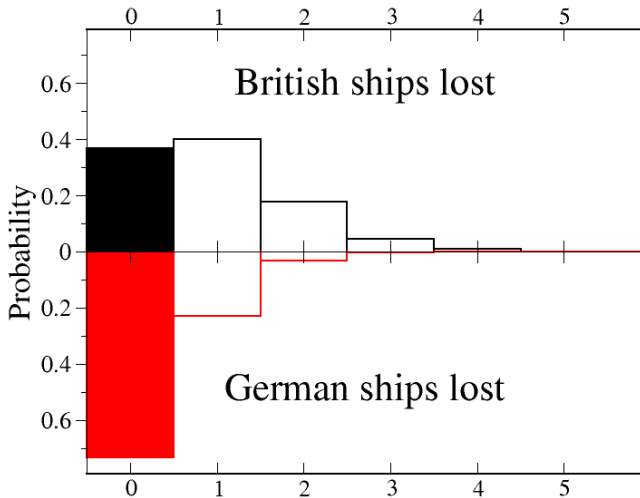
The Results



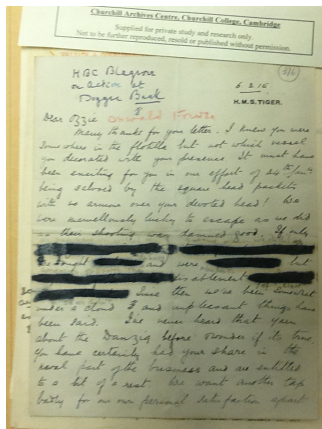
The Results



The Results



The Results



'We were marvellously lucky to escape as we did as their shooting was damned good'

– Lt (later Rear Adm) Henry Blagrove

The Results

Essentially, the British got lucky:

given that they lost 3 battle-cruisers at Jutland (1916) to magazine explosions, they were very lucky not to lose ships at Dogger Bank

The Results

Essentially, the British got lucky:

given that they lost 3 battle-cruisers at Jutland (1916) to magazine explosions, they were very lucky not to lose ships at Dogger Bank — and would almost certainly have done so had the action not been truncated.



Weight of Shell Must Tell: A Lanchestrian Reappraisal of the Battle of Jutland

NIALL MACKAY

University of York

CHRISTOPHER PRICE

York St John University

A. JAMIE WOOD

University of York

Abstract

This article re-analyses the 1916 battle of Jutland (German: Skagerrak), the major naval engagement of the First World War, in the light of the understanding of dreadnought fleet tactics developed over the decade leading up to it. In particular, it considers the interaction of the calculus of Lanchester's Square Law with fleet geometry and the commanders' decisions that determined it, and with the shipbuilding decisions associated with the Lanchestrian trade-off between quality and quantity. There is a re-examination of the behaviour of the commanders in the light of this tactical analysis, and it is concluded that the outcome of Jutland, in spite of apparent British tactical and technological failings, was the culmination of a decade of consistent and professionally insightful decision-making by the Royal Navy, which built and correctly wielded its decisive weapon, the Grand Fleet, to achieve the required strategic victory.

The Battle of Jutland

The Battle of Jutland

Beatty's unstable dynamic creates great variation in outcome

The Battle of Jutland

Beatty's unstable dynamic creates great variation in outcome

Evan-Thomas played a blinder

The Battle of Jutland

Beatty's unstable dynamic creates great variation in outcome

Evan-Thomas played a blinder

Jellicoe played the odds, got them right, and gained his strategic victory

Historical Methodology

'Grown-up historians don't waste time on counterfactuals'

— Michael Howard

Historical Methodology

'How exactly are we to distinguish probable **unrealized alternatives** from improbable ones?' — Niall Ferguson

Historical Methodology

'How exactly are we to distinguish probable **real events** from improbable ones?' — Niall Ferguson

Historical Methodology

The course of history is complex, contingent, and develops within an envelope of probabilities surrounding the historical narrative.

Historical Methodology

The course of history is complex, contingent, and develops within an envelope of probabilities surrounding the historical narrative.

To avoid causation fallacies, the mind-set required by the historian is that of Bayesian probability: to understand historical actors' implicit prior estimates of chances, how these changed as events unfolded, and how they compare with real probabilities, not just realized events.

Phil Tetlock and Dan Gardner, *Superforecasting*, 2015

Current and Future Projects

Superforecasting: wisdom of crowds, time-critical forecasting.

Current and Future Projects

Superforecasting: wisdom of crowds, time-critical forecasting.

Vietnam: attrition, pacification, the Hamlet Evaluation Survey.

Current and Future Projects

Superforecasting: wisdom of crowds, time-critical forecasting.

Vietnam: attrition, pacification, the Hamlet Evaluation Survey.

Birth of Naval Aviation: wargames, fleet exercises, Peleliu.

Current and Future Projects

Superforecasting: wisdom of crowds, time-critical forecasting.

Vietnam: attrition, pacification, the Hamlet Evaluation Survey.

Birth of Naval Aviation: wargames, fleet exercises, Peleliu.

Sea Lion: Could Germany have invaded England in 1940?

Current and Future Projects

Superforecasting: wisdom of crowds, time-critical forecasting.

Vietnam: attrition, pacification, the Hamlet Evaluation Survey.

Birth of Naval Aviation: wargames, fleet exercises, Peleliu.

Sea Lion: Could Germany have invaded England in 1940?

Battle of the Atlantic: U-boat tactics, 'ace' culture, Allied strategy.

Current and Future Projects

Superforecasting: wisdom of crowds, time-critical forecasting.

Vietnam: attrition, pacification, the Hamlet Evaluation Survey.

Birth of Naval Aviation: wargames, fleet exercises, Peleliu.

Sea Lion: Could Germany have invaded England in 1940?

Battle of the Atlantic: U-boat tactics, 'ace' culture, Allied strategy.

Nuclear Deterrence: *Able Archer*, game theory.

Current and Future Projects

Superforecasting: wisdom of crowds, time-critical forecasting.

Vietnam: attrition, pacification, the Hamlet Evaluation Survey.

Birth of Naval Aviation: wargames, fleet exercises, Peleliu.

Sea Lion: Could Germany have invaded England in 1940?

Battle of the Atlantic: U-boat tactics, 'ace' culture, Allied strategy.

Nuclear Deterrence: *Able Archer*, game theory.

Talk to us, give a Zoom seminar, send us students, come and visit!

Thanks for listening