



centre for security
and defence studies

A Stochastic Salvo Model Analysis of the Battle of the Coral Sea

Michael J. Armstrong
Michael B. Powell

CSDS Working Paper 03
April 2005

centre for security and defence studies
norman paterson school of international affairs
carleton university | ottawa, canada
www.carleton.ca/csds

CSDS WORKING PAPERS

Centre for Security and Defence Studies (CSDS)
Norman Paterson School of International Affairs (NPSIA) | Carleton University
1401 Dunton Tower | 1125 Colonel By Dr
Ottawa, Ontario K1V8H8 | Canada
(t) 613.520.6655 | (f) 613.520.2889
www.carleton.ca/csds/publications.html

ABOUT CSDS WORKING PAPERS

The *CSDS Working Papers* series provides an outlet for the relatively quick dissemination of research in progress by Centre Faculty and Research Associates, Pre- and Post-Doctoral Fellows, and students on topics related to Canadian and international security and defence issues. Working Papers are published electronically on an occasional basis and are available for download at www.carleton.ca/csds/working_papers.htm. Working Papers are reviewed internally. The views expressed are those of individual authors and not necessarily those of CSDS or the Department of National Defence. The *CSDS Working Papers* series continues the *NPSIA Occasional Papers* series.

ABOUT THE CENTRE FOR SECURITY AND DEFENCE STUDIES

The Centre for Security and Defence Studies (CSDS) is one of twelve independent national academic Centres of Expertise supported through the Security and Defence Forum (SDF) of the Department of National Defence (DND). Housed at the Norman Paterson School of International Affairs, Carleton University, CSDS brings together scholars and students from across Carleton and the Ottawa area who are engaged in research and teaching on Canadian and international security, defence and foreign policy issues. More information on the Centre and its research, teaching, and outreach activities can be found at www.carleton.ca/csds.

CSDS DIRECTOR David Mendeloff
CSDS DEPUTY DIRECTOR Chris Penny
WORKING PAPERS EDITOR Grant Dawson

© 2005 Centre for Security and Defence Studies.

A Stochastic Salvo Model Analysis of the Battle of the Coral Sea

Michael J. Armstrong
Michael B. Powell

Abstract

In this work we study the Battle of the Coral Sea using a stochastic version of the salvo combat model. We begin by estimating the range of probable alternative results for the battle, given the forces employed; i.e., if the battle were to be "re-fought", how likely are outcomes other than what historically transpired? Our analysis suggests that a wide range of results was indeed possible, even without any change in forces on either side. We then estimate the impact of hypothetical but plausible changes in the American forces employed. Our analysis suggests that a material advantage could have been obtained by committing extra aircraft carriers to the battle or by dispersing the carriers that were already deployed; on the other hand, equipping each carrier with more fighters but fewer bombers would have yielded a net disadvantage.

About the Authors

Michael J. Armstrong is Associate Professor, Sprott School of Business, and Faculty Associate of the Centre for Security and Defence Studies. Prof. Armstrong teaches courses in Operations Management and in Quality Management. His research focuses on the modeling of systems involving risk or uncertainty, including topics such as missile combat, reliability & quality, and services management. Prior to joining Carleton in 2001, he taught at the Royal Military College of Canada. *Michael B. Powell* is Ph.D. candidate in history at Carleton University. His research interests include Canadian resource policy, Canada's involvement in international organizations, and federal-provincial relations.

Funding for this study was provided by the Sprott School of Business and the Centre for Security and Defence Studies (CSDS), Carleton University. Comments on this paper can be sent to Michael Armstrong at <michael_armstrong@carleton.ca>.

INTRODUCTION

In May 1942 the Japanese military were preparing to invade Port Moresby on the south shore of Papua New Guinea, from where they would have been able to threaten Allied ship movements northeast of Australia. To protect their invasion forces from interference, the Imperial Japanese Navy (IJN) dispatched a task force built around two aircraft carriers (CVs). The United States Navy (USN) responded to this threat by sending two CVs of its own into the area.

After some initial skirmishes involving secondary forces, the opposing carrier groups engaged in battle on 8 May 1942. Each side launched a wave of air strikes against the other at roughly the same time. As a result of these attacks the IJN carrier *Shokaku* suffered heavy damage and was put out of action, while the *Zuikaku* was unharmed. The losses on the USN side were somewhat greater, with the *Yorktown* being temporarily disabled and the *Lexington* eventually sinking. Tactically therefore the battle could be considered a USN defeat; strategically however it was an Allied victory because it saved Port Moresby from Japanese invasion (Lundstrom, 1984: 278).

The Battle of the Coral Sea is an interesting one for study because of both its importance and its simplicity. It was important not only because it marked the first rebuff of IJN advances, but also because it was the world's first carrier duel: both sides' attacks were executed by carrier aircraft, while the ships themselves never sighted the enemy. It was relatively simple in that the main battle involved only a single carrier force and a single air strike for each side.

In this paper we analyze the battle of 8 May 1942 using a stochastic version (Armstrong, 2005) of the salvo combat model developed by Hughes (1995). Our discussion begins in the next section with a brief review of the salvo model. We then use two versions of this model to "re-fight" the battle as it originally occurred. The first version is an aggregated approach that uses squadrons of aircraft as the unit of analysis; we follow that with a higher-resolution version involving individual aircraft. In both cases our objective is to determine what ranges of ship losses could reasonably have resulted from the battle, given the historical match-up of forces. In other words, we want to know whether the outcome of the battle could have been much different if either side had been slightly "luckier". Our analysis indicates that this was indeed the case, with wide ranges of results for each side being probable.

In subsequent sections of the paper we consider several alternative history or "what-if" scenarios in which the American force is changed in some way.

- What if the USN CVs had traveled separately instead of sailing in a single task force?
- What if the USN had committed a third CV to the battle?
- What if the USN had been equipped with more fighter aircraft but fewer dive bombers or torpedo bombers?
- What if the USN air defense system had been better coordinated and more experienced in the use of radar and combat air patrol tactics?

To study each of these alternatives, we adjust the inputs to our model to reflect the assumed change in forces, and then examine the model outputs to determine the likely impact of that change. Our analysis suggests that having an extra CV or dispersing the carriers into separate task forces would have been mostly beneficial to the USN. The USN would also have benefited from the effects of a better-coordinated air defense, but it appears this benefit would have been insignificant by itself. On the other hand, changing the aircraft mix carried to include a larger proportion of fighters would have provided a net disadvantage for the USN, as the slight improvement in defense would have been offset by a large decrease in offensive power. Overall these results can be interpreted as supporting the proposition (Hughes, 2001: 107) that in 1942 naval warfare the offense was inherently superior to the defense.

Our paper concludes with a discussion of our results and of the principle limitations of our method of analysis. An appendix describes the details of our data and modeling assumptions.

SALVO COMBAT MODELS

The salvo combat model was developed by Hughes (1995) primarily to model missile combat between modern warships, although it has also been used to study aircraft carrier battles (see e.g. Hughes, 2000: 90-116; Johns et al, 2001: 23-28). In this model, combat is assumed to proceed as follows:

- The ships on our side fire a salvo (wave) of offensive anti-shiping missiles at the opposing force. Simultaneously, the opposing force fires their own anti-shiping missiles at our ships;
- Each side uses its missile defense systems to attempt the interception and destruction of as many incoming anti-shiping missiles as possible;
- Any anti-shiping missiles that are not intercepted then strike their targets and cause damage. Ships that take sufficient damage become firepower kills: they are knocked-out of the battle, though not necessarily sunk.

The output of this model is an estimate of the number of surviving ships for each side after one exchange of salvos. The interested reader is referred to Hughes (1995) for a more detailed description of the model and its assumptions.

One limitation of the basic salvo model is its deterministic nature: it does not take into account the inherent variability or randomness in combat, e.g. the uncertainty of whether a particular missile will or will not hit its target. The more recent work by Armstrong (2005) incorporates this variability in a relatively simple way so as to derive a stochastic version of the salvo model. This newer version estimates both the average and the variance of the number of surviving ships on each side, and consequently can provide the probability of specific outcomes or give interval estimates for the losses on each side.

ANALYSIS OF THE HISTORICAL SITUATION

In this section we make use of the stochastic form of the salvo combat model to study the Battle of the Coral Sea as it was actually fought; that is, without making any change to the forces or tactics used by either side. We perform our analysis twice: once at a relatively low resolution, similar in style to that done by Hughes (2000), and then again at a higher resolution that is more in line with the analysis of Johns et al (2001). By using two versions of the model, each with its own data set, we obtain a degree of assurance that our results are not merely artifacts of any particular modeling choice.

Analysis at the Squadron Level

Our first analysis is done with an aggregated treatment of the battle, in which squadrons (typically composed of about 18 aircraft each) are the basic unit of combat power. In terms of the salvo model, we treat each squadron of dive bombers or torpedo bombers as being equivalent to one offensive "missile", which if not intercepted has a probability of successfully attacking the enemy force. We likewise treat each fighter squadron as if it were a defensive "missile" that has a probability of successfully intercepting one attacking squadron. Bomber squadrons that are not intercepted are assumed to attack and cause a variable loss that averages one enemy CV.

The numbers of attack and fighter squadrons carried by each side are obtained from historical accounts of the battle (Johns et al, 2001; Lundstrom, 1984). The probabilities to go with these squadron numbers are derived using the assessment made by Hughes (2000: 100) during his study of the carrier battles of 1942 that "One carrier air wing could on balance sink or inflict crippling damage on one carrier...". Taking this as the desired end result, we work backwards to calculate the probabilities needed for our model. A list of the resulting model parameters is shown in Table 1, while the appendix provides more detail regarding how these figures are derived.

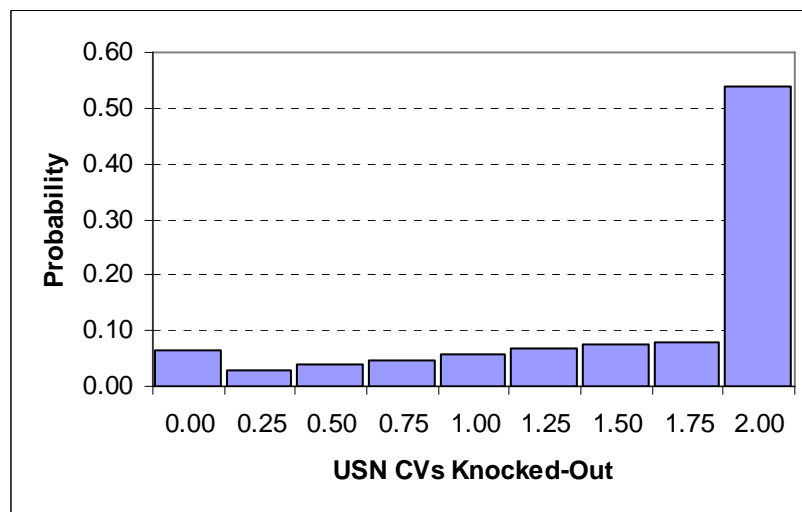
Table 1: Data for Squadron Level Analysis

	USN	IJN
Number of CVs	2	2
Attack squadrons / CV	3	2
Probability of successful attack / squadron	0.4762	0.6429
Fighter squadrons / CV	1	1
Probability of successful intercept / squadron	0.2857	0.4286
Average loss suffered per attack squadron	1	1
Standard deviation of loss / attack squadron	0.3333	0.3333

Given these parameters as inputs, we use a spreadsheet implementation of the stochastic salvo model to estimate the results of the battle. For example, the model predicts average losses (firepower kills) of 1.55 USN CVs and 1.45 IJN CVs. That is, on average each side would lose 1 CV and have the second one moderately damaged. As one might expect, these results are similar to the outcome of the actual battle.

More of interest to us here are the distributions of possible results around these means, as illustrated for the USN in Figure 1. This histogram shows, for example, a 54% probability that the USN could lose the use of both its CVs. Another way to describe the range of "likely" results is with a prediction interval for the losses on each side: if we were to re-fight this battle repeatedly using the same probabilities, a 90% prediction interval is the range in which we would expect the losses to fall 90% of the time. According to our model this interval runs from 0 to 2 ships for both sides, suggesting that "anything was possible", even without changes in how the battle was fought (as in e.g., a surprise attack by one side).

Figure 1: Distribution of losses in squadron level analysis



Analysis at the Aircraft Level

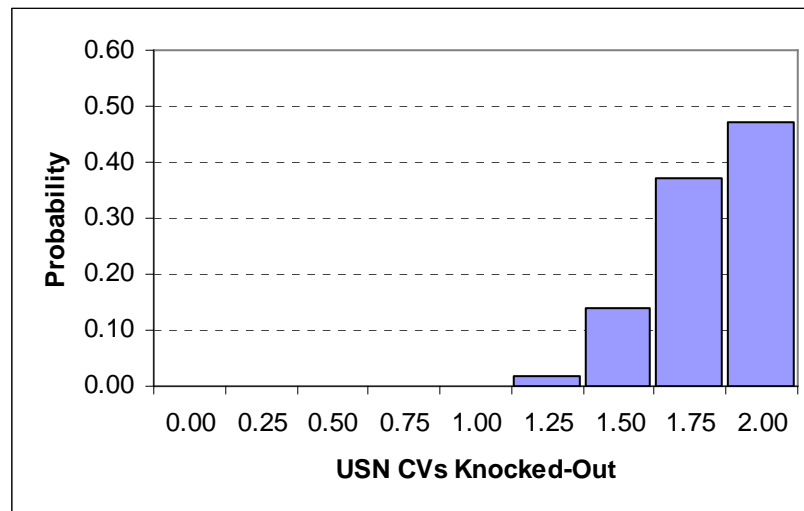
The second version of our study is more detailed in that we model individual attack aircraft as being equivalent to offensive missiles. We likewise treat each fighter aircraft as if it were a defensive missile that has a probability of successfully intercepting one attacking bomber. Attack aircraft that are not intercepted are assumed to attack the enemy task force and cause a variable amount of damage.

Most of the data for this more detailed treatment are derived from the historical account provided by Lundstrom (1984): the actual number of fighter aircraft on each carrier, the number of bomb hits achieved, etc. We treat the actual battle result as being the "average" outcome, and then work backwards to derive the inputs needed by the model. The parameters for the effect of weapon hits on warships, however, are obtained from two studies of bomb and torpedo damage to ships throughout World War II: one by Beall and one by Humphrey, both as reported in Hughes (2000: 157). The resulting model inputs are shown in Table 2; the appendix describes the details of their derivation.

Table 2: Data for Aircraft Level Analysis

	USN	IJN
Number of CVs	2	2
Attack aircraft / CV	47	33
Probability of successful deployment / aircraft	0.5000	0.7727
Fighter aircraft / CV	17	20.5
Probability of successful intercept / fighter	0.3529	0.04878
Average loss suffered per attack aircraft	0.04766	0.03208
Standard deviation of loss / attack aircraft	0.01589	0.01069

Our individual aircraft version of the stochastic model uses these figures to estimate the average losses at 1.82 USN CVs and 1.44 IJN CVs, so that on average the USN loses 0.38 more CVs than the IJN. It is not surprising that these averages are similar to the actual battle results, since the actual battle is the source of most of our model inputs. The 90% prediction intervals for the losses on each side are likewise calculated to be from 1.48 to 2.00 CVs for the USN and from 1.15 to 1.73 CVs for the IJN. A histogram for USN losses is shown in Figure 2.

Figure 2: Distribution of losses in aircraft level analysis

In comparing these aircraft level results to those from the squadron version of our model, we can see that both versions provide similar estimates of the mean outcomes of the battle, but they differ noticeably in their estimates of the variability of these outcomes: the prediction intervals for the squadron level model are several times wider than those from the aircraft level model. This difference can be explained by the different assumptions that are implied in each model regarding the correlation of aircraft

attacks. The aircraft level model assumes that each aircraft's attack or interception is statistically independent of all other aircraft, whereas the squadron level model implicitly assumes that the aircraft within each squadron are perfectly correlated; e.g. either all of the dive bombers in a squadron are intercepted, or none of them are. The most reasonable estimate probably lies in between those two extremes, as aircraft that were in the same squadron did tend to fly together and so experienced many of the same difficulties (e.g. similar levels of anti-aircraft fire or cloud cover), but not to the extent of them all achieving identical results.

Taken together, the two versions of the model suggest that the results of the original battle could easily have turned out quite differently even without any change in the forces employed. Thus for example it appears that there was a real danger that both USN carriers could have been put out of action, in which case the Japanese military presumably would have carried out its planned invasion of Port Moresby. We believe that the similarity of the results from these two analysis approaches also provides some assurance that our modeling approach is reasonable overall.

ALTERNATIVE SCENARIOS

We now turn our attention to several alternative history or "what-if" scenarios in which we assume that the US force had differed in some way from its historical form. In each scenario we adjust the inputs to our pair of models to reflect this difference, and then use the models to estimate the impact on the battle. The results of each case are summarized in Tables 3 and 4.

Table 3: Results of Squadron Level Analysis

	USN Loss Mean	USN Loss 90% interval	IJN Loss Mean	IJN Loss 90% interval	Mean USN edge
Historical case	1.55	0.00 - 2.00	1.48	0.00 - 2.00	-0.07
USN "what-if" cases:					
Dispersed CVs	0.94	0.35 - 1.00	1.48	0.00 - 2.00	+0.54
Extra CV	1.67	0.00 - 3.00	1.82	0.52 - 2.00	+0.15
More fighters	1.05	0.00 - 2.00	0.67	0.00 - 2.00	-0.38
Better air defense	1.25	0.00 - 2.00	1.48	0.00 - 2.00	+0.23
Extra, more, & better	0.11	0.00 - 0.99	1.08	0.00 - 2.00	+0.97

Table 4: Results of Aircraft Level Analysis

	USN Loss Mean	USN Loss 90% interval	IJN Loss Mean	IJN Loss 90% interval	Mean USN edge
Historical case	1.82	1.48 - 2.00	1.44	1.15 - 1.73	-0.38
USN "what-if" cases:					
Dispersed CVs	1.00	1.00 - 1.00	1.44	1.15 - 1.73	+0.44
Extra CV	1.57	1.17 - 1.98	1.98	1.84 - 2.00	+0.41
More fighters	1.02	0.57 - 1.47	0.64	0.44 - 0.85	-0.38
Better air defense	1.29	0.92 - 1.65	1.44	1.15 - 1.73	+0.15
Extra, more, & better	0.00	0.00 - 0.00	0.99	0.75 - 1.24	+0.99

Dispersed USN carriers

What if the US carriers had been traveling separately, instead of sailing together? This change in tactical dispositions could easily have been ordered by the US force commander, Admiral Fletcher, and our study of it here reflects a debate about the advantages of massing versus dispersal (Hughes, 2000: 96). By traveling together the CVs were able to share their fighter cover and escorts, but this also meant that the discovery of one led to an attack on both. By contrast, at the later Battle of Midway the *Yorktown* was traveling apart from the other CVs; when it was discovered by the IJN it was overwhelmed and sunk, but the other USN carriers remained unspotted and thus unharmed (Lundstrom, 1984: 323).

Herein we suppose that the USN CVs had been moderately far apart: close enough that their strikes against the IJN could still be coordinated, but dispersed enough that their air defenses could be considered independent, so that the IJN attack would have focused on a single CV and missed the other one completely. As we did with the historical scenario in the previous section, we study this alternative scenario with both versions of our model and then compare the results.

The second row of numbers in Table 3 shows the results of the squadron level analysis: IJN losses remain unchanged, while USN losses average 0.94 CVs within an estimated 90% prediction interval of 0.35 to 1.00 ships. This represents a notable decrease in USN losses from the original average of 1.55 ships shown in the top row of the table, and gives the USN an advantage or "edge" in the sense that they now knock out on average 0.54 CVs more than they lose. The aircraft level analysis in Table 4 puts both the mean and the prediction interval for USN losses at 1.00 CV, i.e. the one exposed US CV is almost always lost.

These figures indicate that separation of their CVs would have given the USN significant benefits on defense, by lowering their average loss and also by making their losses less variable and so in some sense less "risky". The main downside to this tactic is that it condemns one of the CVs to destruction, whereas sailing together gives both ships a small chance of survival.

Extra Carriers For The USN

What if the USN had sent an additional carrier to fight in the Coral Sea? Two CVs were involved in the Doolittle raid on Tokyo; a decision to send one or both of these vessels to the Coral Sea could presumably have been made by fleet headquarters, though obviously at the cost of canceling the Tokyo raid.

Let us suppose that the USN had sent 3 CVs to the Coral Sea instead of 2, but that the battle had otherwise been fought in its original form (e.g. the carriers traveled all together, not dispersed as in the previous scenario). In this case our squadron level analysis shows the USN losses actually increasing slightly to 1.67 CVs on average, with

the 90% prediction interval at 0.00 to 3.00 ships. IJN losses in this version show a larger increase, with an average loss of 1.82 ships and a prediction interval of 0.52 to 2.00 ships. The aircraft level analysis on the other hand yields a slightly lower mean USN loss of 1.57 CVs within a prediction interval of 1.17 to 1.98, while the IJN losses average 1.98 ships within a prediction interval of 1.84 to 2.00 ships lost. These figures are shown in the third line of Tables 3 and 4.

Thus it seems that having a third CV in the battle would have provided a significant benefit to the US offense, in that IJN losses would have been higher. The extra CV would not however have helped the US defense much: on the contrary, it would have put more of their ships at risk.

We can carry this investigation a step further by estimating the number of CVs required to essentially guarantee a USN "success". That is:

- How many USN carriers would have been needed to provide at least a 95% probability of a "perfect" offense, i.e. IJN losses = 2 CVs?
- How many USN carriers would have been needed to provide at least a 95% probability of a "perfect" defense, i.e. USN losses = 0 CVs?

First consider the offense. With either the squadron version or the aircraft version of our model, we find that the USN can obtain a 95% probability of achieving the destruction of both IJN carriers by deploying a total of 4 CVs to the battle. This was historically achievable, though it would have meant canceling the Doolittle raid and deploying the entire Pacific carrier force to this one region. Defensive success turns out to be much harder to ensure: to obtain a 95% probability of zero losses, the USN needs a total of 21 CVs according to the squadron model, or 11 CVs according to the aircraft model. Forces of this size were simply not available to the USN in 1942.

These results indicate that while the USN could probably have ensured the destruction of the IJN force by deploying more CVs into the region, they would not have been able to guarantee much about their own survival.

More Fighters And Fewer Bombers

What if the USN had carried more fighters but fewer dive or torpedo bombers? On the day of the battle about 25% of the aircraft on each USN carrier were fighters: 1 squadron with a nominal strength of 18 aircraft. By July 1942 the USN had doubled this allotment to 36 fighters per CV (Lundstrom, 1984: 439) as a result of their experiences at Coral Sea and at Midway, and by 1944 the proportion of fighters had increased to about 65% of total aircraft carried (Hughes, 2000: 107). This ongoing trend gave the carrier air wings a progressively stronger air defense capability, but of necessity it also meant that they contained fewer attack aircraft than they otherwise might have.

To model this scenario we increase the number of fighter aircraft or squadrons on each USN CV to roughly 1944 levels, so that they represent 65% of the total aircraft

carried. We correspondingly reduce the number of attack aircraft to be 35% of the total, so that on a net basis each CV still has the same total number of aircraft. With the stronger defense made possible by these fighters, the squadron level model estimates a drop in USN losses to an average of 1.04 CVs, while the prediction interval remains at 0.00 to 2.00 ships. The losses of the IJN however decrease even more due to the weaker USN offense, as shown by an average loss of 0.67 CVs within a prediction interval of 0.00 to 2.00 ships. The aircraft version of our model shows similar changes: USN losses have a mean of 1.02 CVs within a prediction interval of 0.57 to 1.47 ships, while IJN figures drop to a mean of 0.56 CVs lost within a prediction interval of 0.38 to 0.73.

Based upon these numbers, a switch to an aircraft mix containing more fighters but fewer bombers would have left the USN worse-off overall. The slight improvement in defense would not have been enough to make up for the large drop in offensive power. With more CVs on each side surviving the initial exchange of air strikes, it seems likely that a second round of air strikes would have followed in which the weakened USN would probably have been at a significant disadvantage.

Better Air Defense Coordination

In 1942 the USN was still learning such things as at what altitudes to preposition their fighter coverage and how best to direct these fighters using air search radar (Lundstrom, 1984: 303-304). What if the USN air defenses had been better coordinated in 1942, and had instead been closer to the level of sophistication they reached later in the war (Hughes, 2000: 110)? Admittedly, this alternative history scenario is less plausible than the previous ones we considered, as it presumes the USN would have been able to learn their wartime lessons in peacetime, through e.g. more extensive training exercises.

To model this situation, we simply choose to double the existing probability of successful interception by USN fighters and then see what impact this has on their ship losses; IJN ship losses of course remain unchanged. Our calculations with the squadron level model show USN losses decreasing slightly to an average of 1.25 CVs within the original prediction interval of 0.00 to 2.00 ships. The aircraft level model shows a similar decrease to an average of 1.29 CVs lost by the USN and a prediction interval of 0.92 to 1.65. Thus while the increase in air defense effectiveness does give better results for the USN, its effect on the overall battle is relatively modest; it appears that their fighters are just not numerous enough to make much of an impact on the battle, even when they are better coordinated.

Extra Carriers, More Fighters, and Better Air Defenses

In the previous three subsections we considered the impact of changes in the USN force made one at a time; next we examine the impact of making all three changes at once. Our motivation to study this case comes from the 1944 Battle of the Philippine Sea, where the USN had many more CVs than the IJN, as well as more fighters per CV; the USN also enjoyed a qualitative advantage in its fleet air defenses overall (Hughes

2000: pp 104-107). The combined effect of these advantages allowed the US air wings to dominate the battle. What if the USN had had all these factors in its favor at Coral Sea?

We model this "dream" scenario for the USN by adding 1 extra CV, increasing the proportion of fighters on each of these CVs to 65% of the total aircraft carried, and doubling the intercept probability for these fighters. With these changes made in the squadron level model, the USN losses drop off sharply to an average of just 0.11 CVs, and the 90% prediction interval becomes 0.00 to 0.99 CVs; i.e. at most one CV would have been damaged, leaving at least two still functional. The IJN losses also drop slightly on average, with a mean loss of 1.08 CVs; the prediction interval remains unchanged at 0.00 to 2.00 ships. The aircraft level version of the model suggests an even larger impact: the mean and prediction interval for USN losses are both at zero, i.e. they almost always survive undamaged, while the IJN losses average 0.99 CVs within a prediction interval of 0.75 to 1.24 ships.

It is clear that the USN gains a large advantage from these simultaneous changes, as the combined effect of the increases in both the number and effectiveness of their fighters allows the USN defense to hold back the bulk of the IJN attack, just as they did at the later Battle of the Philippine Sea. The initial exchange of air strikes would likely cause minimal damage to the US force and major (but not complete) losses to the IJN, so that the USN could expect to have a very large advantage in any follow-up air strikes.

CONCLUSIONS & LIMITATIONS

In this paper we made use of the stochastic salvo model of Armstrong (2005) to study the Battle of the Coral Sea, both in historical form and in several alternative or "what-if" scenarios. We analyzed each scenario in two ways: a lower-resolution study in which air power was aggregated into squadrons, and a higher-resolution study in which air power was broken down to the individual aircraft. We found that these two different levels of analysis tended to provide similar estimates for the average number of carriers lost by each side, but that the squadron-level study suggested a much wider range and standard deviation for these losses.

For the base case involving the historical order of battle, our analysis indicated that the battle could easily have had a wide range of outcomes, simply by having one side be slightly more or less "lucky" in its performance. For the alternative history scenarios that we subsequently considered, our findings can be summarized as follows:

- Dispersing their carriers would likely have proved beneficial to the USN, assuming that they remained in close enough communication to coordinate their attacks;
- Deploying extra carriers to the Coral Sea would have provided a significant advantage for the USN offense, but these extra ships would have made little difference to the American defense and would have put themselves at risk;

- Changing the USN aircraft mix to include more fighters and fewer bombers would have been a net disadvantage for the USN, as it would have significantly decreased offensive power in exchange for only a slight increase in defensive capabilities;
- Even if it had been possible to better coordinate and control the USN air defense, this change would by itself likely have had only a small impact on the battle overall.

The validity of these results is limited of course to the extent that our models and their inputs together provide reasonable representations of World War II carrier air battles. The stochastic salvo model we used herein to model naval air strikes was originally developed to model combat involving missiles. One difference between missiles and aircraft is that the latter are inherently more flexible. For example, although fighter aircraft mostly flew combat air patrol in defense of their carriers, they also performed close escort missions to protect their attack aircraft. In a similar manner, USN dive bombers not only attacked enemy shipping but also flew scouting missions to locate enemy targets and even provided secondary air defense of their fleet. Our work implicitly assumed that the proportion of aircraft assigned to these "non-standard" roles was fixed at the levels historically used, whereas a more detailed study might include this mission flexibility as another decision variable.

Another difference between aircraft and missiles is that aircraft attacks are more complex sets of events. Our study assumed that attack aircraft that were not shot down were able to attempt attacks against ships, while those shot down did not. In reality, fighter interception of an attacking bomber could sometimes distract the attacker enough to ruin its aim, even if it managed to avoid being shot down. On the other hand, even aircraft that were shot down sometimes had time to launch their weapons beforehand. A more definitive model for World War II carrier battles should presumably include these kinds of second-order effects, but this was beyond the scope of our work herein.

REFERENCES

- Armstrong, M.J. 2005. A stochastic salvo model for naval surface combat. *Operations Research* 53, forthcoming.
- Armstrong, M.J. 2004. Effects of lethality in naval combat models. *Naval Research Logistics* 51: 28-43.
- Hughes, W.P. 1996. A salvo model of warships in missile combat used to evaluate their staying power. *Naval Research Logistics* 42: 267-289.
- Hughes, W.P. 2000. *Fleet tactics and coastal combat*, 2nd ed. Annapolis, MD: Naval Institute Press.
- Johns, M.D., S.E. Pilnick, and W.P. Hughes. 2001. *Heterogeneous salvo model for the navy after next*. Technical Report IJWA 2001 0215 065. Monterey, Calif.: Naval Postgraduate School.
- Lundstrom, J.B. 1984. *The first team: Pacific naval air combat from Pearl Harbor to Midway*. Annapolis, MD: Naval Institute Press.

APPENDIX

Data for Squadron Level Analysis

Each CV carried $n_y=n_z=1$ fighter squadron (Johns et al, 2001: 23). The USN CVs each had $n_\alpha=3$ attack squadrons (1 torpedo plus 2 dive bomber), while each IJN CV had $n_\beta=2$ attack squadron (1 torpedo and 1 dive bomber).

To get values for p_α , p_z , and $E[v]$ for the USN offense, begin by setting up a salvo equation in which the expected nominal loss from 1 attacking USN air wing (3 squadrons) is 1 IJN CV, as suggested by Hughes (2000: 100).

$$[3 p_\alpha - 1 p_z] E[v] = 1 \text{ ship lost} \quad (\text{A1})$$

Given just this one equation, we need to make some arbitrary though reasonable choices to obtain 3 parameters. First we choose the mean damage per successful squadron attack to be $E[v]=1$ ship, so that 1 non-intercepted attack squadron on average achieves a firepower kill on 1 CV. We then set the standard deviation of damage to be one third of this value, i.e. $1/3=0.3333$; this amount makes the variation in damage relatively large but keeps the likelihood of a spurious negative value relatively small, as then the mean loss is 3 standard deviations above zero. Some numerical experimentation suggests that our results are not very sensitive to the size of this standard deviation.

The other evaluation we use from Hughes (2000: 268) is his estimate that a relatively unskilled defender required 7 fighter squadrons to neutralize 1 attacking carrier air wing. Since each USN air wing has 3 attack squadrons, we give each IJN fighter squadron a chance of success of $p_z=3/7=0.4286$ to produce this result on average. We can now solve Equation A1 for the probability of a squadron's successful attack to be $p_\alpha=(1/3+1/7)=0.4762$.

Calculations for the IJN offense are similar, except that they have only 2 attack squadrons per CV. The corresponding salvo equation is thus

$$[2 p_\beta - 1 p_y] E[u] = 1 \text{ ship lost} \quad (\text{A2})$$

which leads to $E[u]=1$, $p_y=2/7=0.2857$, and then $p_\beta=(1/2+1/7)=0.6429$.

Data for Aircraft Level Analysis

We get most of the aircraft level data from Lundstrum (1984), plus some from Johns et al (2001). The 2 USN CVs combined carried 72 Dauntless dive bombers (capable of delivering 1000 lb bombs) plus 22 Devastator torpedo bombers, for a total of 94 attack aircraft and thus $n_\alpha=94/2=47$ attackers per CV. They also carried 38 fighters in total, or $n_y=19$ fighters per CV. The IJN carriers had 38 Val dive bombers (with 250 kg bombs) plus 28 Kate torpedo bombers, for a total of 66 attack aircraft; so $n_\beta=33$. They likewise had 41 fighters on board, so $n_z=20.5$ (Lundstrum, 1984: 221).

To obtain the probabilities of success, we assume that the historical results of the battle were "average" and then work backwards to determine what the corresponding model parameters would be. Note that this assumption ignores the likelihood that the historical results were different from their "true" underlying means, and so our prediction intervals presumably underestimate the variation in probable results.

Not all bomber aircraft actually participated in the attack: some did not fly due to mechanical problems, some flew scouting missions, and some were sent on the strike but got lost along the way. To model this we calculate probability p_α as the ratio of aircraft arriving in the target area to aircraft carried. For the USN, 47 attack aircraft made it to the target area, thus $p_\alpha=47/94=0.5000$ (Lundstrum, 1984: pp 225-226, 236, 242). For the IJN the corresponding calculation is $p_\beta=51/66=0.7727$ (Lundstrum, 1984: pp 224, 242).

Next consider the air defense. In the USN strike, 2 attack aircraft were shot down by the IJN (Lundstrum, 1984: 243), thus the probability of a successful intercept (shoot down) was $p_z=2/41=0.04878$ per IJN fighter. The calculation for the USN defense is $p_y=12/34=0.3529$. These calculations implicitly assume that these attack aircraft were shot down only by fighters (rather than by other aircraft types or the defending ships), and that these shoot-downs all occurred before the aircraft had a chance to launch their ordinance (rather than as they withdrew afterwards). Losses of fighter aircraft are ignored in our model.

A total of 3 hits were achieved against the IJN CVs; dividing this by the $47-2=45$ USN attack aircraft that arrived in the target area and were not shot down gives $3/45=0.06667$ ship hits per surviving USN attacker (Lundstrum, 1984: 231, 238). The equivalent calculation for the IJN attack divides 5 hits by $51-12=39$ attackers and so gives $5/39=0.1282$ hits per surviving IJN attacker (Lundstrum, 1984: 253, 256-258).

To calculate the number of ships lost per hit, we used estimates described by Hughes (2000: 157) that were developed by Beall and by Humphrey based on 176 warships that were damaged during World War II; each study estimated the amount of ordinance required to achieve a firepower kill for target ships of different tonnage. For aircraft bombs, the number of thousand pound bomb equivalent (TPBE) hits required to achieve firepower kills were:

- Beall: 1.7 hits for a 15000 ton ship, 2.5 hits for a 45000 ton ship; and,
- Humphrey: 1.6 hits for 15000 tons, 2.7 hits for 45000 tons.

Averaging these two sets of estimates gives 1.65 hits for a ship of 15000 tons and 2.6 hits for one of 45000 tons. By assuming a linear relationship within this range, we derive the following equation.

$$\text{Hits needed} = 1.175 + (0.03167)(\text{Ship displacement in thousands of tons})$$

Since the Lexington displaced 43055 tons and the Yorktown 25484 tons (Johns et al, 2001: 23), the average per USN CV was 34270 tons. Thus we estimate that 1 USN CV needs $1.175+(0.03167)(34.270)=2.260$ TPBE hits for a firepower kill; inverting this gives $1/2.260=0.4424$ USN ships lost per TPBE IJN hit. The tonnage of the IJN carriers *Shokaku* and *Zuikaku* were 32105 tons each, so that the effect of USN attacks would be $1.175+(0.03167)(32.105)=2.192$ TPBE hits per ship, and thus $1/2.192=0.4563$ IJN ships lost per TPBE USN hit.

We followed a similar process for torpedo damage.

- Beall: 1.4 hits for 15000 tons, 2.0 hits for 45000 tons; and,
- Humphrey: 1.3 hits for 15000 tons, 2.2 hits for 45000 tons.

The averages are 1.35 hits for 15000 tons and 2.1 hits for 45000 tons, so that

Hits needed = $0.975 + (0.025)(\text{Ship displacement in thousand tons})$.

Each USN CV needs an average of $0.975+(0.025)(34.270)=1.832$ torpedo hits for a full firepower kill; inverting this gives $1/1.832=0.5459$ USN ships lost per IJN torpedo hit. Calculations for the IJN CVs give $0.975+(0.025)(32.105)=1.778$ torpedo hits per kill and so $1/1.778=0.5625$ IJN ships lost per USN torpedo hit.

Since our model aggregates all of the attack aircraft on each side, we convert specific weaponry into TPBE and then compute a weighted average value for each side. For the USN we proceed as follows.

- 1 Dauntless 1000 lb bomb = 1 TPBE = 0.4563 firepower kills per hit;
- 1 Devastator torpedo = 0.5625 firepower kills per hit.

So on average $[72(0.4563) + 22(0.5625)] / 94 = 0.4811$ IJN ships lost per hit.

We do similar calculations for the IJN weaponry.

- 1 Val 250 kg bomb = 0.55 TPBE = 0.2433 firepower kills per hit;
- 1 Kate torpedo = 0.5459 firepower kills per hit.

So on average $[38(0.2433) + 28(0.5459)] / 66 = 0.3717$ USN ships lost per hit.

Finally, we multiply the number of firepower kills per hit by the number of hits per surviving attacker. For attacks against the USN this gives $E[u]=0.04766$ CVs lost per surviving IJN attacker. When the IJN is the target we get $E[v]=0.03208$ CVs lost per surviving USN attacker. We arbitrarily set the standard deviations at 1/3 of these average amounts for the same reasons as in the squadron level analysis.

Note that the calculated means for the parameters in the aircraft level model indicate a level of *lethality* (as defined in Armstrong, 2004) for this combat situation that

is moderate, though approaching high; that is, an average air strike by 1 CV will damage but not destroy 1 defending CV. On the other hand, the parameter values for the squadron level model indicate a level of lethality that is just barely in the high category: 1 CV's air strike will only just knock out 1 defending CV.