What happened had to happen? An Approximate Bayesian Computation estimation of the Battles of the Coral Sea and Midway

Anelí Bongers[a] and José L. Torres[b]

[a]Department of Economics and Economic History, University of Malaga, Malaga, Spain;
[b]Department of Economics and Economic History, University of Malaga, Malaga, Spain

ARTICLE HISTORY

Compiled February 20, 2022

ABSTRACT

This paper applies a Bayesian technique (the Approximate Bayesian Computation) to estimate a salvo combat model. Approximate Bayesian Computation is a relatively new free-likelihood Bayesian estimation method, particularly useful when data is very limited, and that can be applied to the estimation of a large variety of models. The paper applies this Bayesian technique to estimate a stochastic version of the salvo combat model in which the variable to be explained in the number of hits on carriers in the Battle of the Coral Sea and the Battle of Midway during Second World War. Given the estimation of the parameters, the model is simulated and compared with the historical observed figures. We show that this estimation technique can generate reliable estimates of the model parameters which account for the uncertainty behind observed figures, even when a very limited number of data observation is available.

KEYWORDS

Dynamical Systems; Approximate Bayesian Computation; Salvo combat model

Corresponding author email: jtorres@uma.es

1. Introduction

A widely used mathematical tool, and one of the first models developed in the Operational Research literature, to study and simulate the outcome of battles or the competition between species is the Lanchester-type dynamic combat model (Lanchester, 1916). Estimation of Lanchester-type combat models have proven to be an important analytically tool to study the result of conflicts between two opposing forces and they have been applied to analyze a larger number of battles (see Braken, Kress and Rosenthal, 1995). Several extensions of the basic attrition combat models proposed by Lanchester (1916) have been developed in the literature, including guerrilla warfare (Deitchman, 1962), stochastic environment (Hausken and Moxnes, 2002), space competition (Shanakan and Sen, 2011), three-way combat (Kress *et al.*, 2018), bottleneck restrictions (Bongers and Torres, 2019), etc.

A related model is the so-called salvo combat model developed by Hughes (1995) for modeling modern naval missile warfare. The salvo combat model consists in a two difference-equations system intended to describe a combat among ships where active defense is considered. The salvo combat model replicates some characteristics related to the pulse nature of naval combat with offensive and defensive firepower. Although the salvo combat model has been developed to describe modern naval warfare using missiles, it can also be successfully applied to naval battles between aircraft carriers, such as the ones which occurred in the Pacific theater during World War II (see Hughes 2000).

However, applications of Lanchester's attrition or salvo combat models to real scenarios have encountered the difficulty of estimating the parameters of the model with very limited data availability, which makes difficult the application of most existing estimation techniques. The standard strategy followed by the literature has been the calibration of the parameters of the model to the observed figures. However, recent developments of likelihood-free estimation techniques allows estimation of complex models with limited data. This papers applies a Bayesian technique (the Approximate Bayesian Computation, ABC) to estimate a salvo combat model. ABC method is a relatively new Bayesian free-likelihood estimation approach with a rapid expansion in a number of disciplines. The first application has been in genetics by Tavaré *et al.* (1997), but these techniques rapidly expanded to other sciences, as biology (Beaumont *et al.*, 2002), psychology (Turner and Van Zandt, 2012), epidemiology (Toni *et al.*, 2009), phylogeography (Csilléry *et al.*, 2010), history (MacKay *et al.*, 2016), etc. The main characteristic of this estimation approach is that it is a likelihood-free method but relays on the simulation of the model, where the numerical evaluation of the likelihood function is replaced by posterior distributions constructed based on simulations from the model and their comparision with observed data. ABC approach is particularly useful when the likelihood function is unknown or computational intractable, or when only very limited number of data observation is available. Advantages of the ABC approach and the simplicity of this estimation method has led to its rapid expansion in a number of different disciplines, and can be a promising estimation method to be applied to a wide number of Operational Research areas.

Lanchester-type and salvo dynamic models have been applied to the study of different battles to carry out counterfactual experiments. Armstrong and Sodergren (2015) studied several counterfactual scenarios regarding the Pickett's Charge at the Battle of Gettysburg during the American Civil War. Armstrong and Powell (2005) used a stochastic salvo model to conduct a counterfactual analysis of the Battle of the Coral Sea. They considered several scenarios in order to calculate the effects of the dispersion of USN aircraft carriers (CVs), increasing the number of USN CVs, changing the composition of each air wing between fighter and bombers, and improving air defense. They find that the result for the USN would have been better if the two CVs had been dispersed. Armstrong (2014) applied a sequential version of the stochastic salvo model to the Battle of the Coral Sea to demonstrate that attacking first would have given the American force a larger advantage than that provided by an extra aircraft carrier. Connors, Armstrong and Bonnett (2015) used a stochastic salvo model to conduct a counterfactual analysis of the Charge of the Light Brigade in the Battle of Balaclava during the Crimean War. Finally, Bongers and Torres (2020) calibrate a salvo combat model and carry out several counterfactual experiments for the Battle of Midway, finding that the probability of the Japanese winning was very low and actually close to zero. Even in the most favorable scenario for the Japanese, the Battle of Midway remains an American victory.

However, a critical caveat of counterfactual analysis using statistical or mathematical models is the fact that the model's parameters are calibrated or estimated using observed figures from a particular event, without taking into account the probability of such event to occur. That is, models' parameters are calibrated or estimated to match observed figures but the observed outcome could be far away from the average expected result or even be a very rare and improbable event. Therefore, calibrated model could not be representative of the event under study and some alternative simulated outcomes would be either impossible or very unlikely to occur. If the parameters of the model are calibrated to the observed figures but the observed outcome is far away from the average expected result, hence, the counterfactual analysis using that calibrated model is misleading. In practice, the validity of counterfactual analyses is conditioned to the assumption that the observed outcome was the expected, something that could not be true due to random events.

In the literature, several statistical and econometric techniques have been used to estimate combat models, including Bayesian techniques (see, for instance, Pettit, Wiper and Young, 2003). However, in most of the cases studied data availability is very limited, and even in a large number of cases data is restricted to only one observation which make it impractical to use standard estimation procedures and instead the model is directed calibrated to the observed figures. However, direct calibration of these models to the observed figures does not take into account the uncertainty of such an event ocurring . That is, models' parameters are calibrated to match observed figures but the observed outcome could be far away from the average expected result or even be a very rare and improbable event. If the parameters of the model are calibrated to the observed figures but the observed outcome is far away from the average expected result, the calibrated model can lead to misleading conclusions. In order to circumvent this problem, MacKay, Price and Wood (2016) applied a Bayesian method for the case in which the likelihood function is unknown, the so-called Approximate Bayesian Computation (ABC) method, to the Battle of the Dogger Bank (a WWI naval battle between British and German battlecruisers) to study if the historical outcome was closed or not to the expected one. They show that the ABC method can be useful to estimate the likelihood of the result of a particular battle and an adequate method to estimate the "true" parameters of the model. In their application to real data, they conclude that the result in the Battle of the Dogger Bank was an anomalous result.

This paper applies the ABC technique to study the likelihood of the results of the Battle of the Coral Sea and the Battle of Midway in the Pacific theater of the Second World War, in order to determine whether the historical outcome from these two battles was unlikely or even improbable or, by the contrary, the outcome of these two battles was close to the expected result. The paper focuses on these two battles because they occurred with a difference in time very short (less than one month of difference), and no change in tactics, skill, technology, and equipment occurred between the two engagements. We expect that the "true" parameters of the model would be the same for these two events. These two battles are of special importance. The Battle of the Coral Sea was the first battle between aircraft carriers in history. The Battle of Midway is widely considered as the turning point in the Pacific theater of World War II and the most important defeat of the Imperial Japanese Navy.

We estimate a version of the salvo combat model in which the variable to be predicted is the number of hits suffered by aircraft carriers. Estimation of the parameters of the model is done using a prior beta distribution with mean and variance calculated using the observed values for the two battles. The use of beta distributions is justified given that the parameters of the model can be interpreted as probabilities. Two estimation methods are used: The simple ACB-Rejection method and the ABC-MCMC method. Both methods yield similar results but computational efficiency of the former is very low compared to that of the later. Parameters estimation is done conditioned to the restriction that they must be equal for both battles. Results indicate that in the Battle of the Coral Sea the number of hits suffered by both American and Japanese carriers was lower than the expected average figure. By contrast, in the Battle of Midway the Japanese were lucky in sinking the *Yorktown*. Overall, our results suggest that Hughes (2000) analysis was correct for these two battles and that the Coral Sea was an anomalous result but the outcome in Midway was much closer to the "expected" result.

The structure of the remainder of the paper is as follows. Section 2 describes briefly the main characteristics of the Approximate Bayesian Computation approach. Section 3 describes the stochastic version of the salvo combat type model developed to represent the two battles. Section 4 presents the historical background and data for the construction of prior distributions. Estimation results and their implications are discussed in Section 5. Finally, Section 6 presents some conclusions and summarizes the main findings of the paper.

2. The Approximate Bayesian Computation (ABC) method

Standard Bayesian estimation techniques are based in the combination of prior distributions with the likelihood function by using the Bayes' rule. Bayesian estimation approach works as follows. Let θ be a parameter vector to be estimated. Given a proposed prior distribution $\pi(\theta)$, the goal is to approximate the posterior distribution $p(\theta \mid X^d)$, where X^d are the observed data. In the standard Bayesian framework, the estimation of the parameters is done by the combination of the information obtained from the data, X^d, via the likelihood $p(X^d \mid \theta)$, with some prior information, via the prior distribution $\pi(\theta)$. The likelihood function describes the density of the observed data given the model and its parameters. The posterior density $p(\theta \mid X^d)$ is what we are looking for. All these information are summarized in the posterior distribution, which it is derived using the Bayes rule:

$$p(\theta \mid X^d) = \frac{p(X^d \mid \theta)\pi(\theta)}{\int p(X^d \mid \theta)\pi(\theta)d\theta} \tag{1}$$

where the term $\int p(X^d \mid \theta)\pi(\theta)d\theta$ is a normalizing constant.

As in general we have no closed-form solution for the posterior distribution, the standard procedure uses computer simulation and Monte Carlo techniques to sample from the posterior distribution. A standard approach is the use of Markov Chain Monte Carlo (MCMC) methods, applying, for instance, the Metropolis-Hastings algorithm.

Bayesian estimation methods have several advantages for estimating a model with respect to traditional econometrics. However, standard Bayesian approach is not without problems. Likelihood can peak in regions of the parameter space that are contradictory with plausible parameters values. Furthermore, computing the likelihood can be a hard task; for some problems the likelihood can be impossible to compute or be too expensive computationally. To overcome these problems, the literature has proposed a variety of so-called likelihood-free estimation methods (see, for instance, Didelot *et al.*, 2011). One particular interesting likelihood-free estimation method is the Approximate Bayesian Computation (ABC) technique.

The idea of the ABC approach is very simple. It consists in replacing the likelihood computation by a simulation of the model. That is, the values for the parameters are estimated such that the simulated data from the model are close to the observed data. The main advantage of ABC methods is that they can be used to evaluate posterior distributions without having to calculate the likelihood previously. The simplest algorithm is the ABC-Rejection, first used by Tavaré *et al.* (1997). The pseudo algorithm for the ABC-Rejection method is presented in Table 1.

Table 1. Pseudo-code for ABC-Rejection method.

1. Generate θ' from $\pi(\theta)$
2. Simulate X^s from the model M with parameters θ^*
3. Accept θ' if $\rho(X^d, X^s) \leq \varepsilon$ and return to step 1.

The working of the algorithm is as follows. First, from a priori distribution, a set of parameters is generated. Second, given the set of parameters the model is then simulated. Third, simulated data in the previous step are compared with observed data. If simulate data is equal to observed data, then the proposed set of parameters is accepted, and the algorithm starts a new iteration. Otherwise, the proposed set of parameters is rejected. Step 3 is implemented using some measure of distance between observed and simulated data as the probability of acceptance can be too small, and then, an approximate method must be used.

In practice, the algorithm compares the simulated dataset, X^s, with the observed data, X^d, using a distance function, ρ, and a tolerance level, $\varepsilon \geq 0$. If $\rho(X^d, X^s) \leq \varepsilon$, then accept the candidate parameter vector θ'. That is, the output of the ABC algorithm is a sample of parameters from a distribution $\pi(\theta' \mid \rho(X^d, X^s) \leq \varepsilon)$. If ε is small enough (close to zero), then, that is a good approximation for the posterior distribution $\pi(\theta \mid X^d)$. This is the key characteristic of ABC methods as by simulation of the model and the comparison with observed data through a distance function, it is possible to evaluate the posterior distribution without having to calculate the likelihood.

For empirical implementation, two key elements related to the ABC approach are needed to be chosen: The distance function, and the tolerance level. First, the definition of the distance function can be problematic depending on the dataset characteristics and may imply the definition of some summary statistics. Election of the summary statistics is crucial for the application of ABC. Assuming that a large number of summary statistics is available, Joyce and Marjoram (2008) propose the sequential inclusion of those statistics into the ABC algorithm. In our context, the distance is defined as the difference, in absolute terms, between the observed value (number of hits suffered by each side) and the simulated value. Second, a tolerance level, $\varepsilon \geq 0$, must be fixed. The tolerance level should be kept as small as possible and it is related to computational power. A tolerance value of zero implies that simulated values must be equal to observed values which impose a high computational cost and a low acceptance rate. Indeed, a disadvantage of the ABC rejection method is that can be computationally inefficient as the acceptance rate would be very low when the prior distribution is very different from the posterior distribution. In practice, rather than choosing explicitly a value for ε, it is chosen a number for the percentage of accepted simulations (the acceptance rate), see Beaumont *et al.* (2002).

One disadvantage of the ABC-Rejection sampler is that it is computationally inefficient and thus the acceptance rate is low when the prior distribution is very different from the posterior distribution. The literature has proposed several modifications of the ABC-Rejection method to overcome that problem. Marjoram *et al.* (2003) proposed a combination of the ABC and the MCMC methods using the Metropolis-Hastings algorithm. Table 2 shows the pseudo algorithm for the ABC-MCMC method. ABC-MCMC increases computational efficiency of simple Rejection ABC.

Table 2. Pseudo-code for ABC-MCMC method.

1. If now at θ propose to move to θ' according to transition kernel $q(\theta \to \theta')$
2. Simulate X^s from the model M with parameter θ'
3. If $\rho(X^d, X^s) \leq \varepsilon$, go to next step, and otherwise stay at θ and return to step 1
4. Calculate $h = h(\theta, \theta') = \min(1, \frac{\pi(\theta')q(\theta' \to \theta)}{\pi(\theta')q(\theta \to \theta')})$
5. Accept θ' with probability h and otherwise stay at θ, then return to step 1.

3. The salvo combat model

Lanchester (1916) developed a variety of dynamic combat models to represent mathematically different types of combats. Lanchester's combat models have been applied not only to the analysis of warfare, but to study a large variety of topics in social sciences, from economics and management to biology. Briefly, Lanchester proposed three alternative specifications: The directed-fire model, the area-fire model and the ancient war model. The aimed- or directed-fire model is intended to represent a situation in which the change in one side is a proportion of the size of the other side, resulting in the so-called Lanchester's square rule, where the number of combatants is more important than their quality (fighting effectiveness). The unaimed- or area-fire model represents a situation in which the fire is unaimed, that is, a battle with area fire weapons as artillery, resulting in the so-called Lanchester's linear rule where the change in one side is a proportion of the products of the number of both sides. A third version considers the case of ancient battles with swords and shields, modeled as a set of individual duels. Lanchester-type combat models are usually defined in continuous-time and they can be applied to land and air battles, as well as to traditional naval battles with gun fire.

A related model designed for naval warfare is the so-called salvo combat model developed by Hughes (1995). The salvo combat model replicates some characteristics related to the pulse nature of naval combat with offensive and defensive firepower, in a discrete time setting. The salvo model presents a number of characteristics different to the Lanchester-type combat models. Whereas Lanchester models are designed to represent gunfire combat where a stream of firepower continuously weakens the enemy force and no defensive firepower exits, the salvo combat models represent combat involving missiles or airstrikes in which defensive firepower is an important characteristic.

Battles between aircraft carriers, without visual contact between ships of both fleets, can be mathematically modeled by the Hughes' salvo combat model (Hughes, 1995). Although the salvo combat model has been developed to describe modern naval warfare using missiles, it can also be applied to naval battles between aircraft carriers, such as the ones that occurred in the Pacific theater during World War II (see Hughes, 2000), as anti-aircraft (AA) fire and fighter defense enable carriers to defend themselves from some level of enemy attack in proportion to their numbers and independent of attackers (Hughes, 2000). In aircraft carrier battles during World War II, attacking firepower is not represented by guns but by dive bombers and torpedo bombers aircraft. Defensive firepower is represented by the combination of AA guns and CAP (Combat Air Patrol) fighters.

Briefly, the salvo combat model consists of the following two dynamic equations for two forces, A and B:

$$\Delta A = -\frac{\alpha B - \beta A}{\gamma} \tag{2}$$

$$\Delta B = -\frac{\delta A - \theta B}{\eta} \tag{3}$$

where $0 \leq -\Delta A \leq A$ and $0 \leq -\Delta B \leq B$. The two equations are structurally identical. The parameter A represents the number of carriers (force strength) on side A at the beginning of the battle. The number of carriers on side B is represented by the parameter B. Without loss of generality, side A represents the USN and side B the IJN. Equation (2) indicates the number of firepower kills (losses) suffered by American aircraft carriers whereas equation (3) does the same for Japanese carriers. Each carrier on side A has an offensive power rating of δ, defined as the probability of attacking aircraft arriving to the enemy carriers, p_S, times the total number of available attacking aircraft, n_S, that is, $\delta = p_S^{USN} \times n_S^{USN}$, where the superscript "USN" indicates the United States Navy. Similarly, offensive power rating of each carrier on side B is represented by $\alpha = p_S^{IJN} \times n_S^{IJN}$, where the superscript "IJN" indicates the Imperial Japanese Navy. Each carrier has a defensive power value, denoted by β and θ, respectively, defined as the probability of defending fighters successfully intercepting and attacking torpedo or dive bomber enemy aircraft before they are able to deliver their ordnance, p_F, times the number of available fighter aircraft for each side, n_F, and . Finally, each carrier also has a staying power value, nF, denoted by γ and η, respectively, that represents the number of hits (bombs or torpedoes) needed to achieve a firepower kill on it, defined as the inverse of the probability of a surviving enemy attacker hitting a carrier, p_H, times the firepower kills per hit suffered by the carrier, a_K, that is, and . For simulation purposes, we treat the parameters representing firepower kills per hit and probabilities as stochastic components of the model.

Equations (2) and (3) are either zero or negative, showing the dynamics of the number of aircraft carriers for both sides during a battle. Figure 1 plots the phase diagram that corresponds to a continuous time representation of the salvo combat model, where the number of IJN carriers is represented in the vertical axis and the number of USN carriers in the horizontal axis. The two lines correspond to values for which equations (1) and (2) are zero. The positive slope for these functions are given by the following expressions:

The two positive zero-change lines (isoclines) split the space in three areas. There are two areas in which the losses of aircraft carriers for one side are zero, which corresponds to a situation in which the defenses are large enough to offset any attack. This occurs when the number of carriers on one side is large enough compared to the number of carriers on the other side and/or when the defense capacity of one side is large enough to completely suppress attacking forces of the other. The area in the upper left represents a combination of the number of carriers on both sides, given the values for the parameters, in which equation (1) is negative and equation (2) is zero. That is, the number of IJN carriers remains constant and the number of USN CVs decreases. Similarly, the area in the lower right represents an initial number of carriers such that equation (2) is zero and equation (3) is negative. Under the assumption that defense parameters are lower than attacking parameters, as indicates by Hughes (2000), we have the following relationship between the zero-solution slopes:

[Insert here Figure 1]

The area between the two zero solution lines reflects a situation with losses for both sides. This is the scenario considered by Hughes (2000) for modeling carrier warfare during WWII.

The basic Hughes' salvo combat model has been extended in several ways. Armstrong (2005) developed a stochastic version of the salvo model in which the number of accurate missiles fired for each ship and the number of interceptions is assumed to be an independent and identically distributed random variable that follows a binomial distribution. Armstrong (2011) analyzed the properties of the stochastic salvo combat model using Monte Carlo simulation on a large number of scenarios and under different assumptions about the distribution of the stochastic components. Armstrong (2013) developed a salvo combat model with area fire in order to study situations in which the enemy location is only approximately known. Armstrong (2014b) developed a sequential salvo combat model for modeling battles in which the exchange of fire is not simultaneous but occurs in different phases.

Salvo combat models, despite their simplicity, have been proven useful in studying different battles, including naval fighting between aircraft carriers, and they have been used to conduct a number of counterfactual experiments. Armstrong and Powell (2005) extended the Hughes' salvo combat model to a stochastic environment and conducted a counterfactual analysis of the Battle of the Coral Sea. They studied several scenarios in order to calculate the effects of the dispersion of USN aircraft carriers (CVs), increasing the number of USN CVs, changing the composition of each air wing between fighter and bombers, and improving air defense. They find that the result for the USN would have been better if the two CVs had been dispersed. Armstrong (2014b) extended the stochastic salvo model to a context in which the exchange of fire is sequential rather than simultaneous. He applied the model to conduct several counterfactual experiments to the Battle of the Coral Sea to demonstrate that attacking first would have given the American force a larger advantage than that provided by an extra aircraft carrier. Salvo combat models have also been applied to battles outside of naval fighting. For example, Armstrong (2014a) used a salvo combat model to study the Israeli Iron Dome's performance during the Operation Pillar of Defense. Finally, Bongers and Torres (2020) uses a stochastic salvo combat model to study the Battle of Midway and to study four alternative counterfactual scenarios: (i) All of the launched American attack aircraft reach the Japanese carriers, ii) the Japanese have one additional carrier, iii) the Japanese do not wait to launch their attack aircraft, and iv) the American carriers are spotted earlier. Contrary to the common wisdom that the result of the Battle was an "incredible" American victory, Monte Carlo simulations show that the probability of the Japanese winning was very low and actually close to zero. Even in the most favourable scenario for the Japanese, the Battle of Midway remains an American victory.

In this paper we use a version of the Hughes' salvo combat model in which the variable to be modeled is the number of hits suffered by ships (aircraft carriers). In particular, the model to be used to describe the Battles of the Coral Sea and Midway consists in the following two structural identical equations for the number of hits in a battle between aircraft-carriers:

$$Hits^{USN} = p_H^{IJN}\left[p_A^{IJN} N_A^{IJN} - p_F^{USN} N_F^{USN}\right] \quad (4)$$

$$Hits^{IJN} = p_H^{USN}\left[p_A^{USN} N_A^{USN} - p_F^{IJN} N_F^{IJN}\right] \quad (5)$$

Equation (6) indicates the number of hits (we do not distinguish between bombs and torpedoes) suffered by American aircraft carriers whereas equation (7) does the same for Japanese carriers. Each carrier fleet has an offensive power rating of $p_A N_A$, defined as the probability of attacking aircraft arriving to the enemy carriers, p_A, times the total number of available attacking aircraft, N_A. Additionally, each carrier fleet has a defensive power value of $p_F N_F$, defined as the probability of defending fighters successfully intercepting enemy attacking torpedo or dive bombers aircraft before they are able to deliver their ordinance, p_F, times the total number of available fighter aircraft, N_F. Although defensive firepower also depends on anti-aircraft guns, the model assumes that it is proportional to the number of fighter aircraft. Finally, p_H is the probability of hitting an enemy carrier by surviving attacker.

In summary, the number of parameters to be estimated is three for each side: Probability of a surviving enemy attacker hitting a carrier, p_H, probability of an attack aircraft arriving to enemy target, p_A, and probability of interception of enemy attacking aircraft, p_F. Given the estimation for these parameters, equations (6) and (7) can be used to obtain the expected number of hits and be compared with observed figures.

4. Historical Background and Data

We estimate the salvo combat model described in previous section for two aircraft carriers battles in the Pacific theater of the WWII: The Battle of the Coral Sea and the Battle of Midway. The Battle of the Coral Sea was the first battle between aircraft carriers in history. It took place during 4-8 May 1942 in the vicinity of New Guinea, with the main battle between attacking aircraft-carriers on May 8. USN fleet was integrated by carriers *Yorktown* and *Lexington*. The IJN fleet was composed of another two carriers: the *Shokaku* and the *Zuikaku*. At the end of the day the *Lexington* was sunk hit by 2 torpedoes and 2 bombs, and the *Yorktown* was damage by a bomb. In the Japanese side, the *Shokaku* was heavily damaged with 3 bombs, whereas the *Zuikaku* finished without damage. The Battle of Midway took place on June 4, 1942 in the vicinity of the Midway island. The Japanese attacking fleet was composed of four aircraft-carriers: the *Kaga*, the *Akagi*, the *Hiryu* and the *Soryu* against the USN fleet of three carriers: the *Enterprise*, the *Hornet*, and the urgently repaired *Yorktown*. In that day, Japanese lost all four aircraft-carriers to one carrier lost by the American (the *Yorktown*). IJN carriers were hit by 13 bombs, whereas the *Yorktown* was hit by 3 bombs and 2 torpedoes.

All information is taken from two basic sources: Lundstrom (1984) for the Battle of the Coral Sea, and Parshall and Tully (2005) for the Battle of Midway, completed by information from Naval Staff History (1952). For the battle of the Coral Sea, we use data corresponding to May 8, the day of the battle between the fleet carriers. The number of available USN aircraft in this battle is shown in Table 3. The total number of (operational) aircraft was 110 (33 fighters, 56 dive bombers and 21 torpedo bombers), 55 for each carrier. For the Japanese, the total number of available aircraft (Table 4) was 102 (38 fighters, 35 dive bombers and 29 torpedo bombers).[1]

In the Battle of Midway, the total number of aircraft for the USN was 346 (231 on board the three carriers and 115 based in Midway air base). However, not all aircraft are included in our estimations by different reasons. First, some aircraft are

[1] These figures are lower than the ones reported by Armstrong and Powell (2005) who studied the Battle of the Coral Sea using data from Lundstrom (1984). The difference is explained by aircraft losses in the days before (from 4 to 7 May), as our analysis only consider the main battle on May 8.

Table 3. Battle of the Coral Sea: Number of USN aircraft operational at May 8, 1942.

Aircraft	USN CVs		
	Yorktown	*Lexington*	Total
F4F Wildcat	14	19	33
SBD Dauntless	32	24	56
TBD Devastators	9	12	21
Total	55	55	110

Table 4. Battle of the Coral Sea: Number of IJN aircraft operational at May 8, 1942.

Aircraft	IJN CVs		
	Shokaku	*Zuikaku*	Total
A6M Zero	18	20	38
D3A "Val"	17	18	35
B5N "Kate"	14	15	29
Total	49	53	102

level bombers (USAF B-17s) not suited for naval attack, whereas other are patrol aircraft (USN PBYs), and hence, they are excluded from the analysis. Second, some attack aircraft did not attack IJN carriers, but other secondary targets, and hence they are also excluded to avoid miscalculation of model parameters in the estimation of the number of hits on aircraft carriers.[2] The total number of attack aircraft for the USN was 186 (152 from the three USN CVs and 34 from Midway Air Base, see Table 5). The total number of aircraft for the Japanese was 248 on-board the four IJN carriers, excluding all aircraft on-board battleships and cruisers, as they were all patrol aircraft. Total attack aircraft for the Japanese were 153, as the two D4Y "Judy" on-board Soryu were a patrol version, and hence, excluded from the analysis.[3] The total number of fighter for the US was 107 (79 from the three USN CVs and 28 from the Midway Air Base). For the Japanese, the total fighter was 93 (Table 6).

Three parameters have to be estimated for each side: the probability of an attacking aircraft arriving to the target; the probability of a successful intercept of attacking enemy aircraft; and the probability of hitting the target. As all parameters are interpreted

[2]This was the case of six SB2Us in the first American attack and 16 SBDs in the American counterattack. In the fist case, the six SB2Us decided to attack the battleship *Haruna* as a secondary target given the heavy protection of the IJN CVs by fighters. In the second case, 16 SBDs from the wave attaking the *Hiryu* decided to attack two IJN cruisers as a secondary target as they evaluated that damage to the last IJN carrier inflicted previously by the other 24 SBDs was fatal.

[3]Similarly, all scout aircraft on-board IJN battleships and cruisers are also excluded.

as probabilities, their values fall in the range between 0 and 1. The main difference between both battles is that in the case of the Coral Sea, the attack over carriers was simultaneous, whereas in the case of Midway the battle was conducted in a sequential mode. In order to be comparable both battles, we model the Battle of Midway as a three sequential (dependent) sub-battles (First American attack, Japanese counterattack, and second American attack), whereas the battle of the Coral Sea is modeled as two independent simultaneous battles (American attack and Japanese attack).

Table 5. Battle of Midway: Number of USN aircraft.

	USN CVs				Total	Attack	Fighter	Excluded
Aircraft	*Yorktown*	*Enterprise*	*Hornet*	Midway				
F4F Wildcat	25	27	27	7	86	-	86	0
SBD Dauntless	37	37	35	18	127	127	-	0
TBD Devastators	14	14	15		43	43	-	0
F2A Buffalo				21	21	-	21	0
TBF Avenger				6	6	6	-	0
SB2U Vindicators				12	12	6	-	6
B-17 Fortress				17	17	0	-	17
PBY Catalina				30	30	-	-	30
B-26 Marauder				4	4	4	-	0
Total	76	78	77	115	346	186	107	53

Table 6. Battle of Midway: Number of IJN aircraft.

	IJN CVs				Total
Aircraft	*Kaga*	*Akagi*	*Hiryu*	*Soryu*	
A6M Zero	27	24	21	21	93
D3A "Val"	20	18	18	16	72
B5N "Kate"	27	18	18	18	81
D4Y "Judy"				2	2
Total	74	60	57	57	248

Data for constructing prior distributions is as follows. First, we compute the observed probability of an attacking aircraft arriving to the target. For the USN at Coral Sea, a total of 60 dive and torpedo bombers escorted by 16 fighters attacked IJN carriers. Thus, probability of attacking was 60/77=0.7792. For the Japanese, a total of 51 dive and torpedo bombers escorted by 18 fighters attacked USN carriers, that is, the probability of attacking was 51/64=0.7969. In the case of the Battle of Midway, a total of 34 aircraft attacked IJN CVs from Midway Air Base[4] and 90 from USN CVs in the first attack (as all dive bombers from *Hornet*, a total of 35 aircraft, didn't

[4] The Midway aircraft attacks comes in different waves. At 7:10, IJN CVs were attacked by the 4 B-26 armed with torpedoes and the 6 TBFs. At 7:55 dive bombers attack with 18 SBD2. At 8:14 with 16 B-17, and finally at 8:20 with 12 SB2Us. Nevertheless, as indicated in the text, B-17s are level bombers not suitable for attacking ships and therefore they are excluded of the analysis.

spotted IJN carriers as they followed an incorrect heading), and a total of 24 attack aircraft from USN CVs in the second attack.[5] Therefore, for the USN the probability of attacking to enemy CVs during the first attack (phase I) was (34+90)/196=0.6327. Note that total number of launched carrier-based attacking aircraft was 125, but only 90 found and attacked Japanese carriers. In the second American attack (phase III of the battle), a total of 40 attack aircraft were available, of which 24 aircraft attacked the only remaining Japanese carrier and the other 16 aircraft attacked two IJN cruisers. Therefore, probability of attacking was 24/(40-16)=1.00. For the Japanese, in the second phase of the battle, only the *Hiryu* was afloat with 18 dive bombers and 14 torpedo bombers of which 28 aircraft attacked USN carriers. Therefore, probability of attacking USN carriers in Midway-II was 28/32=0.8750.

Next, we compute the probability of a successful intercept of attacking enemy aircraft. This figure must represent the number of attacking aircraft destroyed before they are able to deliver their ordinance (bomb or torpedo) and can be very different from the number of attacking aircraft shooting down as this event can occur after their have launched their ordinance. The probability of a successful intercept is calculated as follows. In the case of the Battle of the Coral Sea, a total of 2 American attacking aircraft were shooting down before launching their ordnance by Japanese defense. Therefore, the probability of successful interception by the Japanese is 2/38=0.0526. Similarly, a total of 6 Japanese attacking aircraft were shooting down by American defense. Thus, probability of successful interception by the American is 6/33=0.1818. In the case of the Battle of Midway, probability of a successful intercept by the Americans is calculated as follows: In the attack to the *Yorktown*, the first wave was composed of 18 Dive-bombers but only 7 where able to avoid USN fighter defense. Therefore, successful intercept in this case was of 11 dive-bombers. In the second wave, 10 torpedo-bombers attacked the *Yorktown* but 5 were intercepted before they could release their torpedoes (although only 4 torpedoes were seen by the *Yorktown*). Therefore, total successful intercepts were 16, that is, a final probability of 16/79=0.2025. For the Japanese, the successful intercept is calculated as follows. First, from the 4 B-26 torpedo bombers, three were able to launch their torpedoes and hence only one was successful inter-

[5] These 24 aircraft corresponds to VB-6 and VS-6. Another 16 dive bombers (VB-8 and VS-8) decided to attack cruisers *Chikuma* and *Tone* after observing damage of the last Japanese carrier.

cepted. From the 6 TBF, two of them were shooting down before they could release their torpedoes. Next, from the 18 SBD2, only 7 were able to release their bombs and so successful interception of this group was of 11 attacking aircraft. Finally, from the 6 SB2Us, three of them were able to release their bombs to the *Kaga*. Next, in the attack from USN CVs, first from the 15 torpedo-bombers from the Hornet, only four where able to release their torpedoes before being shooting down and therefore successful interception of this group were 11. From the group of the *Enterprise*, 6 out of 14 were shot down before launching their torpedoes and from the group of the *Yorktown*, 7 out of 12 were shot down before launching their torpedoes. The *Enterprise* dive bombers groups, a total of 32 Dauntless, attacked the *Kaga* and the *Akagi*. 12 of them were able to drop their bombs. The *Yorktown* dive bomber group, with 17 Dauntless, attacked the *Soryu*, with 13 aircraft being able to drop their bombs. Therefore, total successful interceptions were a total of 65 aircraft, resulting in a probability of successful interception by the Japanese of 65/93=0.6989 for the first phase of the battle. During the second American attack, three dive-bombers were shooting down before being able to drop their bombs. Therefore, probability of successful intercept, given that the *Hiryu* had 25 fighters, were 3/25=0.1200.

Table 7. Observed figures for the parameters of the model.

Param.	Definition	Coral Sea	Midway-I	Midway-II	Midway-III
$Hits^{USN}$	Number of hits received by USN CVs	5	-	5	-
$Hits^{IJN}$	Number of hits received by IJN CVs	3	9	-	4
p_H^{USN}	Probability of hit per shot on IJN CVs	0.0517	0.1525	-	0.1905
p_H^{IJN}	Probability of hit per shot on USN CVs	0.1111	-	0.4167	-
p_A^{USN}	Probability of USN attacking aircraft	0.7792	0.6327	-	1.0000
p_A^{IJN}	Probability of IJN attacking aircraft	0.7969	-	0.8750	-
p_F^{USN}	Probability of intercept by USN fighter	0.1818	-	0.2025	-
p_F^{IJN}	Probability of intercept by IJN fighter	0.0526	0.6989	-	0.1200

Finally, we calculate the probability of hitting a CV. In the battle of the Coral Sea the USN carriers received a total of 5 hits against 3 hits received by the IJN carriers. Excluding intercepted attackers, probability of hitting a Japanese carrier was 3/(60-2)=0.0517, and the probability of hitting an American carrier was 5/(51-6)=0.1111. In the case of the Battle of Midway, USN achieved a total of 13 hits (9 hits during the first attack and other additional four during the second attack), whereas IJN aircraft achieved a total of 5 hits. Excluding intercepted attackers, the probability of hit in Midway-I is 9/(124-65)=0.1525 for the USN, 5/(28-16)=0.4167 for the IJN in Midway-II, and 4/(24-3)=0.1905 for the USN in Midway-III. A summary of the observed parameters values for the two battles used to build prior distributions is presented in Table 7.

5. Estimation results

In this section, we estimate the equations predicting the number of hits on aircraft carriers (equations 2 and 3) using the ABC method. The ABC method implies that prior values of the parameters are updated to obtain posteriors, by comparing the simulated data from the model for a number of battles simulations with the observed figures. For each equation three parameters are estimated: Probability of hit per surviving attacker aircraft, p_H, probability of attacking aircraft arriving to enemy carriers, p_A, and probability of successful interception, p_F. That is, we want to study the joint posterior $p(p_H, p_A, p_F \mid X^d)$ for each side. Given that the parameters of the model can be interpreted as probabilities, we use beta distributions as prior distributions based on the information obtained from both battles. The parameters of the prior beta distributions (α, β), are calculated using prior information (mean and standard deviation) taken from observed values at the Battles of the Coral Sea and Midway. Estimation is done for both battles by assuming that the "true" parameters must be equal. The temporal difference between both battles is very short (less than one month: May 8 and June 4, 1942, respectively), and therefore, no changes in tactics, technology, skills, etc., occurred between the two battles. Differences between estimated posterior distributions and the observed values for each battle are evidence that the observed result deviates with respect to the likely "expected" result.

We investigated the effects of varying the tolerance parameter ε, that is, the difference between the observed and the simulated number of hits suffered by carriers in each side. This parameter affects both the computational efficiency and the accuracy of the inference. With the ABC-Rejection method, the acceptance rate heavily depends on the tolerance value. Unfortunately, with the Markov chain implementation of likelihood-free simulation, one can rarely be both efficient and accurate in the same analysis. The higher the value of ε, the greater the proportion of Metropolis–Hastings steps that are accepted and the faster the sampler moves around the parameter space. However, the fidelity of the posterior distribution to the observed data also becomes reduced. Conversely, a smaller tolerance implies lower acceptance rates, but improved data fidelity. We vary the value for the tolerance from 1 to 3, with little consequence of the estimated posterior distributions.

Gelman, Gilks and Roberts (1995) suggest an acceptance rate of 50% in small dimensions, and an average acceptance rate of 25% in large dimensions. Sherlock (2013).

For the Battle of the Coral Sea, theoretical results predict no survivor, that is, all four carriers (two American and two Japanese) should have been sunk. However, very pour scouting on both sides resulted in only one USN carrier sunk, although two other carriers (one Japanese and the second American carrier) were damaged.

Estimation results from the ABC method for the Battles of the Coral Sea and Midway can be compared with the predicted outcome analysis of aircraft carriers battles by Hughes (2000). Hughes (2000) studied a number of aircraft carrier battles in the Pacific theater of the WWII, suggesting that at this stage of the war between the U.S. and Japan "one carrier air wind could on balance sink or inflict crippling damage of one carrier". For the Battle of the Coral Sea, he obtain that theoretical survivors after the battle is zero for both sides, as all four carriers participating in the battle (two Americans and two Japanese) should have been sunk. However, in this battle only one carrier was sunk (the *Lexington*) and other was heavily damaged (the *Shokaku*). Therefore, the result was very different from the expected by Hughes (2000). For the Battle of Midway, he obtained that, given the initial number of aircraft carriers in both sides and the sequential characteristic of the combats, after the first U.S. attack theoretical survivors are all three American carriers and only one Japanese carrier. After the Japanese counterattack, theoretical survivors are 1 and 2, for the Japanese and the Americans, respectively. Finally, after the U.S. second attack, theoretical survivors reduces to 0 for the Japanese and 2 for the American. That is, for the Battle of Midway the historical result is just the expected result by Hughes (2000). If we consider that the reasoning by Hughes (2000) is correct, the outcome of the Battle

Table 8. Posterior estimation of the parameters.

Parameter	Definition	ABC-Rejection		ABC-MCMC	
		Mean	Median	Mean	Median
p_H^{USN}	Probability of hit per shot on IJN CVs	0.0981	0.0840	0.0810	0.0738
p_H^{IJN}	Probability of hit per shot on USN CVs	0.1522	0.1521	0.1490	0.1486
p_A^{USN}	Probability of USN attacking aircraft	0.6838	0.6918	0.6815	0.6888
p_A^{IJN}	Probability of IJN attacking aircraft	0.8796	0.8816	0.8753	0.8788
p_F^{USN}	Probability of intercept by USN fighter	0.1802	0.1808	0.1764	0.1765
p_F^{IJN}	Probability of intercept by IJN fighter	0.3038	0.1168	0.1904	0.0219

of Midway should have been close to the "average" expected outcome whereas in the case of the Battle of the Coral Sea the outcome should have been far away from the "average" expected result.

Table 8 shows the estimation results of the parameters of the model using both the ABC-Rejection method and the ABC-MCMC algorithm. Differences between both estimation methods are very slight and the main difference comes from the computational efficiency of each method. Figures 1 and 2 plot the estimated posterior distributions for the parameters using both methods, with some important differences in the shape of the posterior distribution for some parameters.

The comparison between observed and estimated parameters gives some interesting results. First, we find that the estimated probability of hit per shot on USN CVs is, in general, larger than the estimated probability of hit per shot on IJN CVs. This means that accuracy by Japanese pilots is supposed to be higher than that of their Americans counterpart. Both posterior distributions are skewed to the right, with the mode close to zero. This is consistent with the observed probability of hit. Posterior mean and median for the probability of hit per shot on IJN carriers are between 0.0738 and 0.0981. Equivalent values for the probability of hit per shot on USN carriers are between 0.1486 and 0.1522. As indicated above, both estimation methods (Rejection and MCMC) yield similar results. The expected probability of hit per shot on USN carriers is higher than the observed figure in the case of the Coral Sea, but much lower than the one observed in Midway. Indeed, for the IJN observed probability of hit per shot was 0.1111 at the Coral Sea and 0.4167 at Midway. For the USN, observed values are 0.0517 at Coral Sea, 0.1385 in the first attack at Midway, and 0.1905 in the second attack at Midway. Consequently, accuracy of IJN pilots was lower than expected in the

Battle of the Coral Sea but surprisingly large at the Battle of Midway, as compared to the expected value. For the USN, probability of hits was lower than expected in the Battle of the Coral Sea, but higher at Midway, especially during the second attack.

The resulting higher expected probability of hit per shot of IJN compared to the USN can be explained by several factors. First, it is argued that at the first stages of the war and before the Battle of Midway, IJN pilots were better trained than USN pilots. This skill gap changed to the opposite direction after Midway. Additionally, a technological factor can also explain such differences, resulting from the poor reliability of American's torpedoes compared to the much more efficient Japanese's torpedoes. All hits suffered by IJN carriers in both battles were with bombs as no American torpedo hit any Japanese carrier. By contrast, Japanese torpedoes hit American carriers in both, Coral Sea and Midway.

Second, estimated posterior distribution for the probability of attacking aircraft also shows an important difference for each side. Posterior mean and median are very similar, as the posterior distribution is hump-shaped. Also we find that estimated probability is significantly higher for the Japanese compare to the Americans. Mean and median estimates are very similar, around 0.68 for the Americans and around 0.87 for the Japanese. Comparing these estimates with the observed values, we find that for the USN, probability of attacking aircraft was higher than expected in the Battle of the Coral Sea but similar to the expected value in the first attack at Midway. For the IJN, the value of the probability of attacking is very close to the expected value in the Battle of Midway. However, at Coral Sea, the observed figure is well below the expected one. In any case, probabilities are large, indicating that the IJN was able to put into air and arrive to the target most of available attacking aircraft. For the USN we find that for the first attack at Midway the value is very close to the expected one, and higher at the Coral Sea. Figure for the second attack at Midway represents a special case, as all available American attacking aircraft were able to arrive to the target but only a fraction (60%) attacked the IJN carriers.

These differences can be explained by the better attack coordination for the IJN compared to that of the USN. In fact, historians agree that American carrier warfare doctrine was superior to that of the Japanese in many aspects, including carrier development, scouting, positioning of carriers in the battle, etc. The only aspect in which Japanese were superior to Americans was in the ability to launch a coordinate strike with both torpedoes and dive bombers from different carriers. Estimated parameters confirm this advantage for the Japanese.

Finally, estimated probability of intercept differs significantly for both sides. In the case of the Americans, posterior mean and median is around 0.18. Estimated values for the Japanese are more imprecise, with a mean of 0.30 for the ABC-Rejection and 0.19 for the ABC-MCMC, and with medians of 0.11 and 0.02, respectively. This is provoked by the fact that estimated posterior distribution for the Japanese is very flat with an extremely high variance. Estimated value for the USN is very close to the observed figures in both battles. Probability of interception was 0.18 at Coral Sea and 0.20 at Midway, values very close to the estimates ones. The observed probability of intercept by IJN fighter was extremely high during the first phase of the Battle of Midway, but very low in the Battle of the Coral Sea, compared to the estimated value. Probability of intercept by USN fighter was lower than expected in both battles.

In general, one would expect a large probability of intercepting incoming attacking enemy aircraft by the USN due to technological factors and better AA defence. American carriers had radar, not available in Japanese carriers, and therefore, they could organize the defence before the arrival of the Japanese attacking aircraft. However, the poor coordination in American strikes, with each squadron arriving over Japanese carriers independent of each other (particularly during the first attack in the Battle of Midway), increased significantly interception by the Japanese. This results in a very imprecise estimation of this parameter for the Japanese side.

Table 9 shows the estimation of the number of hits on CVs given mean and median posterior values for the parameters from the estimated posterior distributions. We found that the observed number of hits on USN carriers in the Battle of the Coral Sea were below the average expected value, whereas the number of hits on IJN carriers was also below although in this case they are close to the expected ones. Estimated number of hits on USN carriers in the Battle of the Coral Sea is in the rank 7 to 9. Therefore, we conclude that American were lucky in the Battle of the Coral Sea as the likelihood number of hits on USN carriers is estimated to be a minimum of 7, at least two more than the historical results, whereas the estimated expected number of hits suffered by IJN carriers is in the range 3 to 5 hits was close to the expected one. Hughes (2000) predicted not survivor for the Battle of the Coral Sea for each side. In reality, one American carrier, the *Lexington*, was sunk and the other, the *Yorktown*, suffered minor damage. In the Japanese side, the *Shokaku* suffered heavy damage, although it was not sunk, whereas the *Zuikaku* was undamaged. Based on data collected by Beall (1990) and Humphrey (1992) regarding damage to warships from bombs and torpedo hits during WWII, one additional hit on the *Shokaku* would have been deadly. Our estimations seem to indicate that Americans should have loss the two carriers and the Japanese one out of two, a closer result to the prediction done by Hughes.

Table 9. Posterior estimation of the number of hits.

		ABC-Rejection		ABC-MCMC	
	Number of hits suffered	Mean	Median	Mean	Median
IJN CVs Coral Sea	3	4.69	3.12	3.85	3.66
USN CVs Coral Sea	5	8.96	6.96	7.49	7.48
IJN CVs Midway-I	9	13.79	9.03	9.99	9.46
USN CVs Midway-II	5	3.67	2.88	2.37	2.39
IJN CVs Midway-II	4	2.38	1.59	2.09	1.93

In the case of the Battle of Midway, we find that in the first American attack, the estimated number of hits on IJN carriers were equal to the expected one or even higher. This result is obtained because although the probability of hit per shot on IJN CV was higher than expected, also the probability of intercept by IJN fighter was higher than expected, cancelling the effects each other over the final number of hits. Expected number of hits in the first American attack is between 9 and 14. This would imply the loss of three carriers, as occurred in reality. We find that for the Japanese counterattack, expected number of hits is lower than that observed. Japanese achieved 5 hits on the *Yorktown* but estimation indicates that the expected number was 2 or 3 hits, evidencing that this counterattack was more effective than expected. Finally, also the number of hits in the second American attack was higher than the expected one. However, this has no consequences as more attacking aircraft were available (an additional 16 dive bombers from VB-8 and VS-8 but they attacked other targets once they observed that the last Japanese carrier was mortally damaged). In summary, these results indicate that the expected losses in the Battle of Midway is all four Japanese carriers and the possibility of no losses for the Americans with only heavy damage carrier, an estimated outcome very close to the observed result and to the prediction by Hughes. Finally, differences in observed figures with respect to estimated expected values introduce an important difference. The *Yorktown* should have been sunk at the Coral Sea. This event would have had important consequences, as this carrier would not have been available at Midway, reducing dramatically the possibility of an American victory in that second battle.

6. Conclusions

Likelihood-free estimation methods are gaining popularity in a large number of research fields (Sisson, Fan and Beaumont, 2018), but still are of limited applications in others. This is the case of Operational Research, where likelihood-free methods can be a promising avenue for estimating complex models where the likelihood is computational intractable. A prominent likelihood-free estimation approach which is gaining a great popularity due to its simplicity is the Approximate Bayesian Computation (ABC) method.

This paper shows that the ABC method can be a very useful technique to obtain reliable estimates of a model's parameters which account for the uncertainty behind observed figures. This is of particular importance for cases where data is very limited. Here, we apply the ABC method to estimate a salvo combat model. Lanchester-type and salvo combat models have been widely used to study battles or competition between species or firms, and for carry out counterfactual experiments. However, observed figures for a particular event can be far away from the average expected outcome or even be a very rate and improbable event, resulting is misleading estimated of the model's parameters. The ABC method can be of value for improving the accuracy of estimates in these situations.

One of the main advantages of the ABC method is that models estimation can be done including when data availability is very limited. This is because the ABC method, contrary to other estimation techniques, relays on simulations of the model and their comparison with observed data. We apply the ABC method to a case where data is very limited, with only five observations corresponding to two aircraft-carrier battles to estimate two equations for a total of six parameters. We show that, using this estimation technique, model estimation can be very different from calibration or direct estimation using observed figures, where uncertainty is not taken into account. Therefore, direct estimation or calibration of a model to conduct counterfactual experiments can be a wrong strategy and lead to erroneous conclusions as they do not account for uncertainty.

Estimation results shows that in the Battle of Coral Sea the number of hits suffered by American carriers was much lower than expected. This result has important implications, not only for this particular battle, but also for the late Battle of Midway, as the expected result indicates the loss of the two carriers, including the *Yorktown.* This implies that the availability of aircraft-carriers for Midway would have been reduced to only two, with dramatic consequences on the result of this battle. The number of hits suffered by IJN carriers was similar but somewhat below the average expected value and could have implied the loss of the *Shokaku.* In the Battle of Midway, we obtain that the expected result in the first American attack is similar to the observed result. However, in the Japanese counterattack, we estimate that the expected number of hits suffered by American carriers is lower than the observed.

Comparing the outcomes of the two battles with Hughes (2000) predictions, we find that he was right in predicting the Battle of Midway result (four Japanese carrier sunk against one American carrier sunk), but he fails in predicting the Battle of the Coral Sea result (only one American carrier sunk and another Japanese carrier heavily damage). Nevertheless, here we show that the historical result at the Coral Sea was far from the expected "likely" result, which it is much closer to Hughes' prediction. In sum, Americans was lucky in the Battle of the Coral Sea, as the expected result implies the loss of the two carriers. On the contrary, the Japanese was lucky in the Battle of Midway sinking the *Yorktown,* as in the Japanese counterattack, observed number of hits on USN carriers as higher than the average expected value.

Acknowledgements

We thank Wayne P. Hughes and Jon Parshall for very useful comments on previous versions of the paper. The authors acknowledge financial support from project financed by the Spanish Ministry of Science and Technology ECO2016-76818-C3-2-P.

References

Armstrong, M. J. (2005). A stochastic salvo model for naval surface combat. *Operations Research,* 53, 830-841.

Armstrong, M. J. (2011). A verification study of the stochastic salvo combat model. *Annals of Operations Research*, 186. 23-38.

Armstrong, M. J. (2013). The salvo combat model with area fire. *Naval Research Logistics*, 60, 653-660.

Armstrong, M. J. (2014a). Modelling short-range ballistic missile defense and Israel's Iron Dome system. *Operations Research*, 62, 1028-1039.

Armstrong, M. J. (2014b). The salvo combat model with a sequential exchange of fire. *Journal of the Operational Research Society*, 65, 1593-1601.

Armstrong, M. J. and Powell, M. B. (2005). A stochastic salvo model analysis of the Battle of the Coral Sea. *Military Operations Research*, 10, 27-37.

Armstrong, M. J., and Sodergren. S. E. (2015). Refighting Pickett's Charge: Mathematical modeling of the Civil War battlefield. *Social Science Quarterly*, 96, 1153-1168.

Beall, T. M. (1990). *The development of a naval battle model and its validation using historical data*. Master Thesis, Naval Postgraduate School, Monterey, California.

Beaumont, M. A., Zhang, W. and Balding, D. J. (2002). Approximate Bayesian Computation in population genetics. *Genetics*, 162, 2035-2035.

Bongers, A. and Torres, J.L. (2019). A bottleneck combat model: an application to the Battle of Thermopylae. *Operational Research*, doi:10.1007/s12351-019-00513-0.

Bongers, A. and Torres, J.L. (2020). Revisiting the Battle of Midway: A counterfactual analysis. *Military Operations Research*, (forthcoming).

Braken, J., Kress, M. and Rosenthal, R.E. (1995). *Warfare Modeling.* Alexandria, VA: Military Operations Research Society.

Connors, D., Armstrong, M. J. and Bonnett, J. (2015). A counterfactual study of the charge of the Light Brigade. *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, 48, 80-89.

Csilléy, K., Blum, M. G. B., Gaggiotti, O. E. and François, O. (2010). Approximate Bayesian Computation (ABC) in practice. *Trend in Ecology and Evolution*, 25, 410-418.

David, I. (1995). Lanchester modeling and the biblical account of the battles of Gibeah. *Naval Research Logistics*, 42, 579-584.

Deitchman, S. J. (1962). A Lanchester Model of Guerrilla Warfare. *Operations Research*, 10, 818-827.

Didelot, X.,Everitt, R.G., Johansen, A.M. and Lawson, D.J. (2011). Likelihood-free estimation of model evidence. *Bayesian Analysis*, 6, 49–76.

González, E. and Villena, M. (2011). Spatial Lanchester models. *European Journal of Operational Research*, 210, 706-715.

Hughes, W. P. (1995). 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: Naval Institute Press.

Humphrey, R. L. (1992). Comparing damage and sinking data for World War II and recent conflicts. 13th General Working Meeting of The Military Conflict Institute, McLean, Va.

Joyce, P. and Marjoram, P. (2008). Approximately sufficient statistics and bayesian computation. *Statistical Applications in Genetics and Molecular Biology*, 7, Article 26.

(Hausken and Moxnes, 2002)

Kress, M., Caulkins, J. P., Feichtinger, G., Grass, D. and Seidl, A. (2018). Lanchester model for three-way combat. *European Journal of Operational Research*, 264, 46-54.

Lanchester, F. W. (1916). *Aircraft in warfare. The down of the fourth arm.* Constable and Company Limited: London.

Lundstrom, J. B. (1984). *The First Team: Pacific Naval Air Combat from Pearl Harbor to Midway.* Annapolis: Naval Institute Press.

Mackay, N., Price, C. and Wood, A. J. (2016). Weighing the fog of war: Illustrating the power of Bayesian methods for historical analysis through the Battle of the Dogger Bank. *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, 49, 80-91.

Marjoram, P., Molitor, J., Plagnol, V., and Tavaré, S. (2003). Markov chain Monte Carlo without likelihoods. *Procedings National Academic of Science*, 100, 15324-15328.

Naval Staff History (1952). *The Battles of Coral Sea and Midway.* Battle Summaries Nos 45 & 46.

Parshall, J. B. and Tully, A. P. (2005). *Shattered Sword: The Untold Story of the Battle of Midway.* Nebraska: Potomac Books.

Pettit, L. P., Wiper, M. P. and Young, K. D. S. (2003). Bayesian inference for some Lanchester combat laws. *European Journal of Operational Research*, 148, 152-165.

(Shanakan and Sen, 2011),

Sisson, S.A., Fan, Y. and Beaumont, M.A. (2018). *Handbook of Approximate Bayesian Computation.* CRS Press, Boca Raton: Florida.

Tavaré, S., Balding, D., Griffith, R. and Donnelly, P. (1997). Inferring coalescence times from DNA sequence data. *Genetics*, 145, 505-518.

Toni, T., Welch, D., Strelkowa, A., Ipsen, A. and Stumpf, M. P. H. (2009). Approximate Bayesian computation scheme for parameter inference and model selection in dynamical systems. *Journal of the Royal Society Interface*, 6, 187–202.

Turner, B. M. and Van Zandt, T. (2012). A tutorial on approximate Bayesian computation. *Journal of Mathematical Psychology*, 56, 69–85.

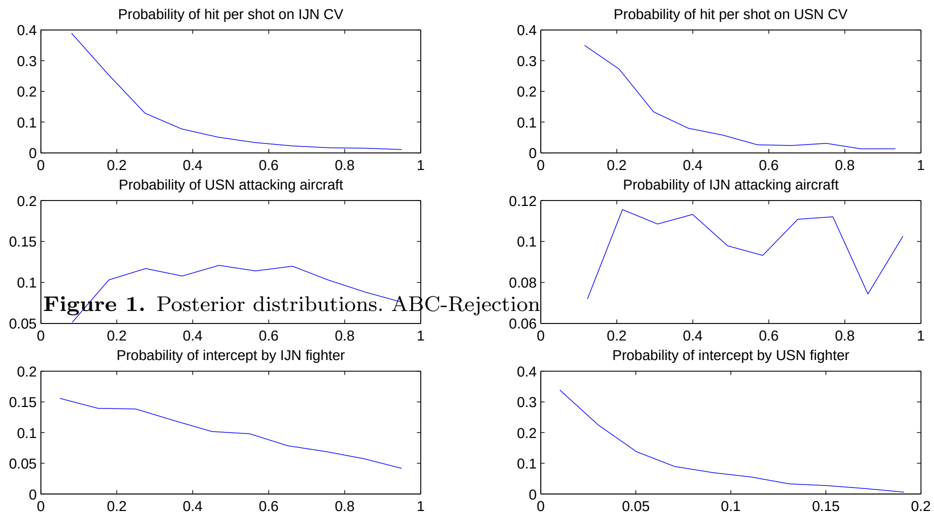

Figure 1. Posterior distributions. ABC-Rejection

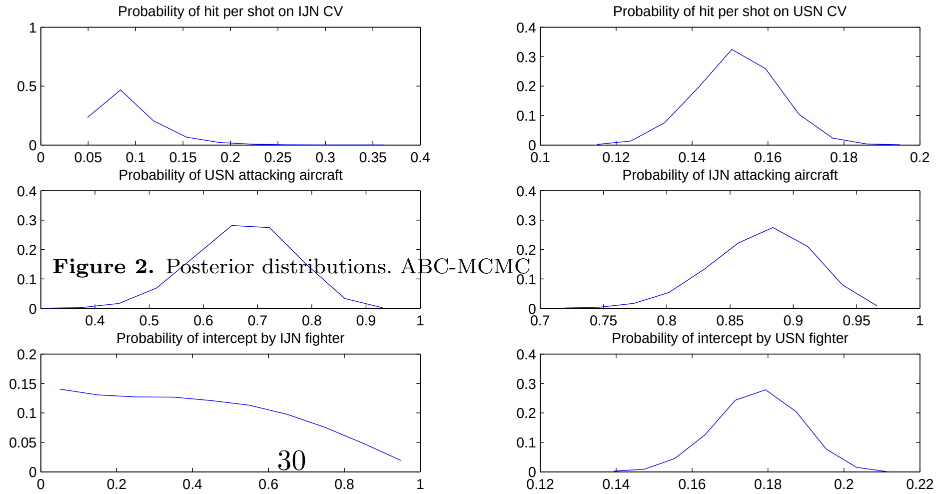

Figure 2. Posterior distributions. ABC-MCMC