Correlation of Flight Test and Analytic *M*-on-*N* Air Combat Exchange Ratios

D. S. Hague*
Aerophysics Research Corporation, Bellevue, Washington

This paper compares analytic and flight test predictions of many-on-many air combat performance capability. It is shown that stochastic effects dominate air combat encounter outcomes. It follows that combat performance estimates should be based on a large sample of encounters. Overall exchange ratios, the number of red losses per blue loss over a series of encounters, predicted by flight test and analysis are within 5% of each other. Trends in exchange ratio with force size obtained by the two methods are also similar. Finally, it is shown that air combat performance is sensitive to both force size ratio and the total number of aircraft engaged and that differences in performance between aircraft types may diminish with increasing aircraft numbers.

Background

A PREVIOUS paper[1] described the multiple tactical aircraft performance analysis code (MULTAC). This code provides a method for analysis of the many-on-many (*M*-on-*N*) air combat encounter; the model is sensitive to initial states, pilot tactics, and aircraft and weapon characteristics. Reference 1 describes MULTAC in some detail and presents some typical results obtained from the code. For convenience the salient features of MULTAC are summarized in the present paper.

In a MULTAC simulation a force of M blue aircraft engage a force of N red aircraft. Initial position and velocity components for each vehicle in the two forces are input to the simulation. Based on these initial state components the relative state of each red player with respect to each blue player, and vice-versa, is computed. The relative states are then converted into pilot cockpit coordinates; angle-off the nose, target aspect (viewing) angle, and range. A heuristic algorithm determines the most likely target and most serious threat for each player in the engagement. Based on the relative state with respect to these opponents a decision is made to either attack or defend using the methods of Ref. 1. In response to this decision each aircraft will seek to place its opponent on the canopy centerline using bank angle as a control variable. Simultaneously each vehicle adjusts angle of attack to pull maximum allowable load factor consistent with structural/aerodynamic limits and a desire to maintain its velocity close to the corner point where turn rate is maximized. This last objective also requires manipulation of a third control variable, throttle setting. With the available control settings determined in this manner on the basis of relative state a forward integration of the combined equations of motion for $(M+N)$ vehicles commences.

This control selection process has resulted in a "filtering" of available courses of action until a single flight objective is established for each player. The simplification which results from such a filtering process may be illustrated by considering the number of combined targetings available to blue and red in an *M*-on-*N* encounter. In the general case, assuming that every vehicle (player) acquires a target, $(M^N N^M)$ combined targeting combinations are possible. This follows by considering that each blue player independently may select any of N red targets and each red player may independently select any of M blue targets. With a one-on-one encounter this process produces a unique targeting combination, player 1 targets player 2 and vice versa. In a two-on-two encounter there are 16 ways to target, a three-on-three engagement produces 729 possible targeting combinations and a four-on-four 65,536 combinations. Clearly the targeting process involves a serious case of the "curse of dimensionality." Reducing the available targetings to a single combination by a hueristic algorithm involves some degree of approximation. Such an approximation must also be performed on a real time basis by the combat pilot. Our present objective is to construct a method for replication of flight test exchange ratios. Accordingly, the degree of success achieved in the target filtering process may be judged from the results of the present paper.

As the combined equations of motion for $(M+N)$ vehicles are integrated forward in time, Fig. 1, a series of firing opportunities develop for the combatants. At these opportunities weapons are fired at the opponent of the moment subject to realistic time delays between firings and to satisfaction of weapon firing constraints from launch to impact. Since weapons involve stochastic kill mechanisms any given impact may destroy the target, or leave it unscathed with a probability which depends on the weapon system, target, and relative state. If kills are recorded in this "sudden death" manner each encounter examined must be repetitively simulated in Monte-Carlo fashion to determine the mean result. With many sets of initial states to consider the computational burden involved in Monte-Carlo simulation can become excessive.

An alternative approach to weapon effects simulation involves retaining all vehicles in the analysis but varying their survival probability with time to reflect all weapon impacts

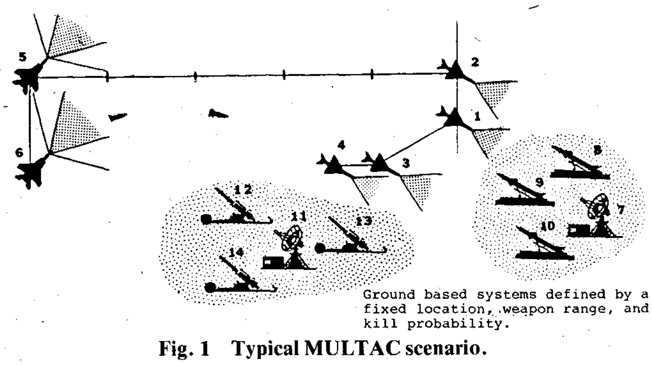

Fig. 1 Typical MULTAC scenario.

Presented as Paper 82-1328 at the AIAA 9th Atmospheric Flight Mechanics Conference, San Diego, Calif., Aug. 9-11, 1982; submitted Aug. 9, 1982; revision received March 7, 1983.

*President. Member AIAA.

received. This is the "co-kill probability approach" introduced in Ref. 2. The essence of the co-kill probability approach is a modification of weapon kill probability at impact time as follows. Let P_{kss} be the weapon single shot kill probability, P_T the target probability of survival at impact, and P_F the survival probability of the firing aircraft at weapon release or impact using active or semiactive weapons, respectively. Then the probability of kill when a weapon impacts is

$$\Delta P_k = P_F P_T P_{kss}$$

reflecting that aircraft which have been destroyed cannot fire a weapon and targets that have been destroyed cannot be destroyed again. Following impact

$$P_T^+ = P_T - \Delta P_k$$

By integrating these difference equations along the path a single simulation of an encounter provides an estimate of the final survival probability of each vehicle. In its simplest form, and in the form used here, these terminal survival probabilities are used to compute an encounter exchange ratio, the number of red losses per blue loss.

During the study reported here MULTAC results are correlated with flight-test encounters flown by experienced military pilots in an air combat maneuver range (ACMR). An existing sample of 253 one-on-one through four-on-four encounters from the available ACMR flight-test data base were simulated in MULTAC and the results compared to flight test. These results were made available to the study by USAF personnel who were responsible for sample selection. Since these flight-test encounters were scored by sudden death modeling they are subject to scatter resulting from stochastic kill mechanisms. Each encounter result is a sample from the spectrum of possible outcomes at each initial state. Since, by definition, we cannot predict the outcome of a single sample an aggregate measure of combat effectiveness must be constructed for correlation purposes. In this paper the cumulative exchange ratio defined as the number of red losses per blue loss summed over a specified number of engagements is employed in this role. The present paper summarizes this work. Overall MULTAC and flight-test exchange ratios based on cumulative results are in essential agreement and at each of seven force sizes flight test analysis showed similar exchange ratio trends.

Flight-Test Correlation

The analysis techniques introduced in the MULTAC code have now been correlated against the series of 253 flight-test encounters supplied by USAF. These were neutral trial encounters flown on an air combat maneuver range. These trials were designated neutral on the basis that both sides have ground control information (GCI) available to the flight-test pilots. The flight tests flown were within the confines of a 30*NM* circle. Test rules required the use of twelve start points. Six of these points were equispaced on the test range circumference. Six more were similarly spaced on a concentric circle having half the test range radius. A precision overflight of designated start points in space and time was imposed on all flights.

In these flights blue force employed either F-14A or F-15A air superiority fighters with enhanced target acquisition capabilities. Red force employed the F5E single seat, high-performance, multipurpose fighter. There were no weapon hardware modifications to these aircraft. Blue weapon loads simulated a full complement of four i.r. and four radar-guided all aspect missiles. In addition, a 20-mm cannon firing 6000 rounds per minute was simulated. Red aircraft carried a load of four i.r. missiles and a 23-mm cannon firing 2400 rounds per minute.

Pilots for both blue and red represented the upper range of aircrew performance in the services. Generally, these pilots were of high caliber and experienced in the air combat arena. Tactics adopted were extremely varied and of a nonstationary nature. Thus, repetitive encounters from similar initial conditions might and did produce wide ranges of differing maneuvers and encounter outcomes. Disengagement was permitted and an aircraft pursuing this maneuver was considered to have reached a sanctuary when escape from the 30*NM* circle was accomplished.

The presence of GCI information was linked to permissible exit modes from combat. With GCI, aircraft were only permitted to exit the test range in the rear semicircle based on their initial headings. Further, exits were only permitted following the detection of an opposing aircraft present. Otherwise, aircraft were required to pass over the ACMR test-range centerpoint before exiting the range. An aircraft was not subject to further fire on escaping the ACMR range outer circumference.

As discussed earlier an air combat encounter involves a stochastic process. Given the same initial conditions, repetitive engagements will produce a spectrum of outcomes rather than a unique result. For example, in a 2-on-2 encounter the exchange ratio, defined as the number of red aircraft lost per blue aircraft loss, can only take on the values 0, ½, 1, 2, and ∞, and it is the relative frequencies with which each possible outcome occurs which determines the average exchange ratio for a particular situation.

Again, it is clear that we cannot predict the result of any single encounter due to the stochastic effects introduced by variable pilot tactics and weapon kill mechanisms. However, if the cumulative exchange ratio over i trials, as defined above, is examined for $i=1,2,...N$, a well-defined average exchange ratio tends to emerge at each force size ratio. It is against this cumulative exchange ratio that MULTAC results will be correlated.

Overall Correlation

An overall comparison between flight-test and MULTAC predicted exchange ratio is presented in Fig. 2, for all 253 available encounters. This figure lumps together results at seven different force sizes in the engagement order of the flight-test data base. Therefore, we are examining predicted exchange ratio over the equivalent of a sizeable airwar. This figure shows the difference between predicted cumulative exchange ratio and flight-test cumulative exchange ratio as a function of the number of encounters considered. It may be seen that cumulative exchange ratio difference starts at −1.5, rises to a high of +2.0 after 25 engagements, and is essentially zero when all engagements are considered. It should be noted that the order in which engagements are flown could affect the convergence behavior from initial to final engagement. However, the final point includes all engagements and, therefore, is independent of engagement order. Again, the

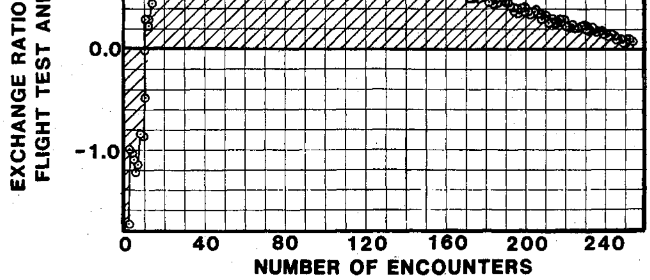

Fig. 2 Cumulative exchange ratio difference, overall.

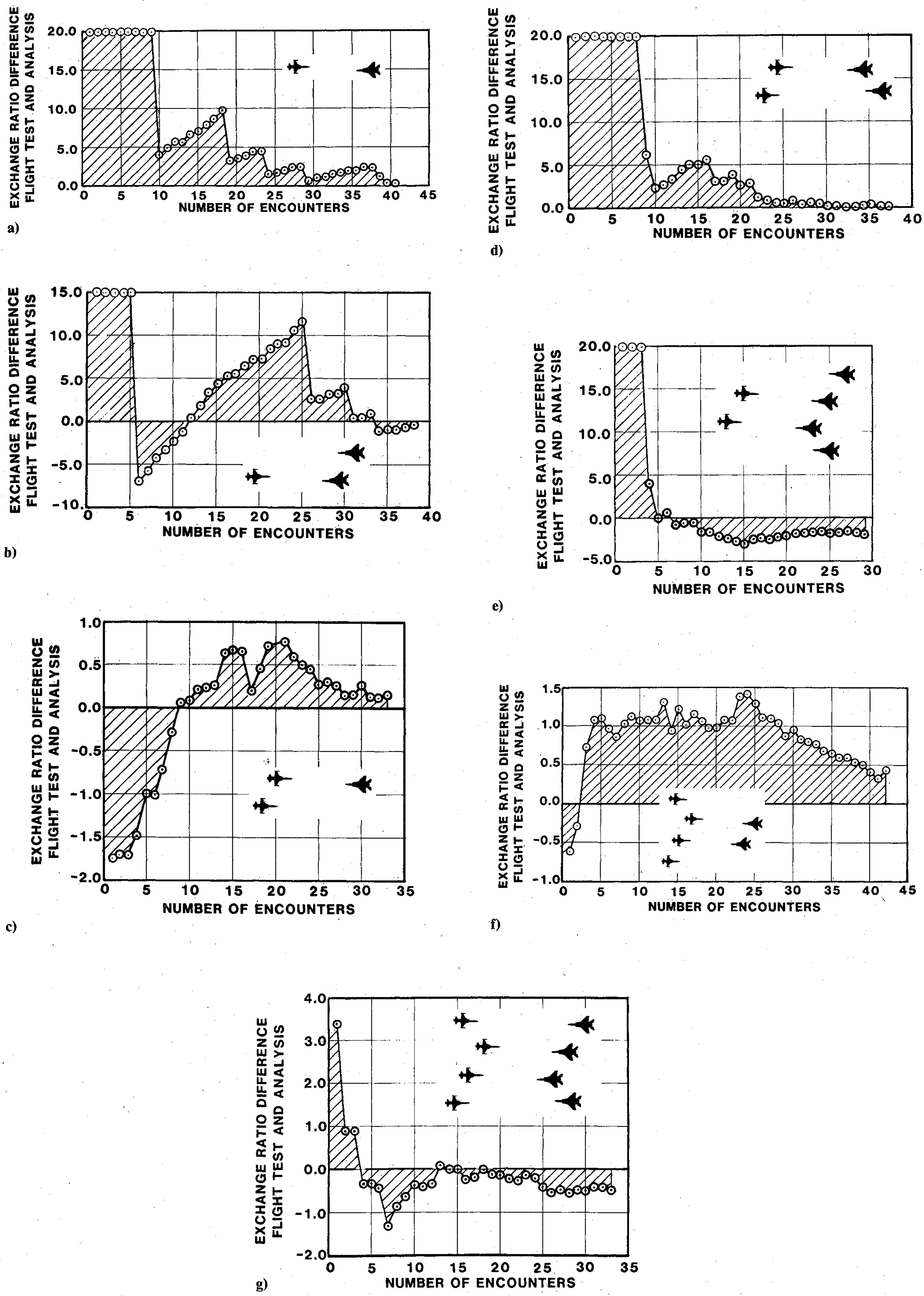

Fig. 3 Cumulative exchange ratio difference. a) 1 red-1 blue, b) 1 red-2 blue, c) 2 red-1 blue, d) 2 red-2 blue, e) 2 red-4 blue, f) 4 red-2 blue, and g) 4 red-4 blue.

particular flight ordering employed is that of the USAF supplied data base rather than a particular ordering selected by the present author. The salient point of Fig. 2 is summarized in the following.

In over 253 one-on-one through four-on-four encounters, the number of red aircraft lost per blue aircraft lost is in essential agreement when a direct comparison is made between overall flight test and MULTAC predictions of the air combat outcome. Thus, it is apparent that the overall outcome of air battles at moderate force size is predictable.

Now, it should be noted that MULTAC, through the Ref. 1 control laws, has simulated a series of aggressive hard fought battles between the opposing pilots, while the flight tests of ACEVAL simulated encounters from the same set of initial conditions using a mix of tactics performed by pilots experienced in the air combat arena. Nevertheless, the relative losses over what amounts to a sizeable conflict are in agreement by the two approaches. One possible implication of this result is a relative invariance of exchange ratio over a range of flight tactics. This concurs with many large-scale force-on-force models of operations analysis which utilize relative exchange ratios of small-scale engagements to assess directly the outcome of large conflicts. It will be shown that this exchange ratio is a function of the relative force sizes and total number of aircraft engaged, in addition to the aircraft types.

It may also be noted in passing that the sortie loss rates generated by MULTAC, with its aggressive pilot behavior modeling, were very high. In fact, MULTAC sortie loss rates for both blue and red were approximately 65% higher than those of flight tests where pilots employed more diverse tactics. In reality sortie loss rates from both MULTAC and flight test were probably unsustainable in actual air warfare. Sortie loss rate is more sensitive to pilot tactics than exchange ratio. For example, suppose a blue and red force encounter each other during their mission and both elect to turn and flee without engaging. This encounter does not affect exchange rate for it introduced no losses. On the other hand, it directly affects sortie loss rate since it generates sorties but not losses. This proportionality of MULTAC and flight-test results may indicate that exchange ratio is invariant to combat intensity.

Correlation by Force Size

A series of comparisons between flight test and MULTAC predicted exchange ratio differences at each of the seven available force sizes, 1×1, 1×2, 2×1, 2×2, 2×4, 4×2, and 4×4, is presented in Figs. 3a-g. For ease of plotting exchange ratio differences are truncated at plot limits in Fig. 3. The cumulative exchange ratio difference may tend to infinity in the first few encounters between superior blue and inferior red aircraft, for over a limited number of flight-test engagements red may sustain losses while blue does not. The finite cumulative exchange ratio differences displayed after the initial engagements correctly sum all red and blue losses through the engagement in question. Similar behavior to the overall cumulative exchange ratio difference can be observed. As the number of encounters increases, the flight-test and predicted cumulative exchange ratios approach each other. Again, while convergence depends on encounter order in the USAF data base the final point is independent of the order.

The MULTAC predictions employ co-kill probability scoring, and a detailed examination of the actual cumulative exchange ratios typified by Fig. 3 shows that the analytic results converge more quickly to their final value than do flight-test results. This is due to the co-kill probability methods ability to approximate the mean result of each encounter during a single simulation.[1] With flight test, on the other hand, each result is a random sample from the spectrum of possible results from the specified initial state and, hence, such results exhibit a greater variance than the analytic results. It is possible that flight testing might be more efficient if a real time co-kill scoring system were introduced to cut down the required number of flight tests.

Results by force size are illustrated in Fig. 4. Here, results are presented for one red, two reds, and four reds engaging a specified number of blue aircraft. It may be seen that at only one force size, 2 red-4 blue does a significant difference between flight test and MULTAC appear. This may be attributed to blue behavior when the four superior aircraft encounter two inferior ones. There is a tendency for blue to resort to two classical leader/wingman formations in this situation. The wingman on the lookout for other potential opponents in flight test may play a relatively passive role when the additional opponents fail to materialize. Thus, with passive wingmen the 2-on-4 flight-test result is similar to the 2-on-2 result. MULTAC on the other hand, allows each blue to aggressively pursue and shoot at the reds. Thus, with MULTAC the exchange ratio against two red aircraft rises monotonically with the number of blue aircraft. With flight test, on the other hand, the exchange ratio against two aircraft is essentially flat as the number of blue aircraft increases from 2 to 4. It should be noted that in MULTAC flight sizes were known to both sides whereas in flight test opponent force size was not necessarily known.

The impact of force size ratio and total force size is clear from Fig. 4. Thus, compared to the one-on-one result, illustrated by the solid symbols, the exchange ratio would rise approximately 75% if the advanced blue aircraft could meet the red aircraft in a 2 blue-on-1 red condition, thereby enhancing the superior aircrafts performance capability. Conversely, if red could meet blue in a 1 blue-on-2 red condition, then the resulting exchange ratio falls to about 25% of the one-on-one result and the superior performance of the blue aircraft has been neutralized by numbers. It may be noted that an extrapolation of the 2-on-2 result to the 2 red-on-4 blue situation provides a good approximation to flight-test behavior.

It is interesting to examine exchange ratio variation at constant force size ratio as the total number of aircraft engaged increases. This comparison is made in Figs. 5a, 5b, and 5c for encounters in the force size ratios 1:2, 1:1, and 2:1, respectively. First, consider equal force sizes. It is clear that for the aircraft/weapon combinations involved there is a strong tendency for performance advantage to reduce with increasing force sizes. When the higher performance aircraft also have the numerical advantage this effect is again observed. In contrast, when red has the numerical advantage, exchange ratio is insensitive to total aircraft number.

Reduction in Advantage with Increasing Aircraft Numbers

The flight-test correlation results are not an isolated case of performance difference washout with increasing numbers.

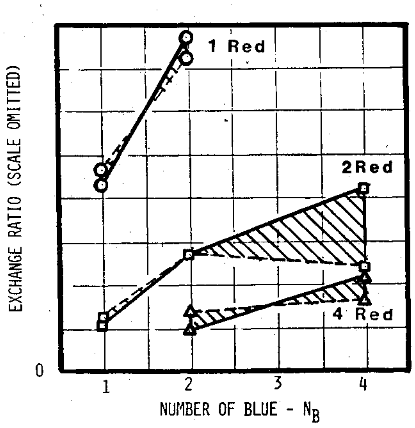

Fig. 4 Final correlations by force size.

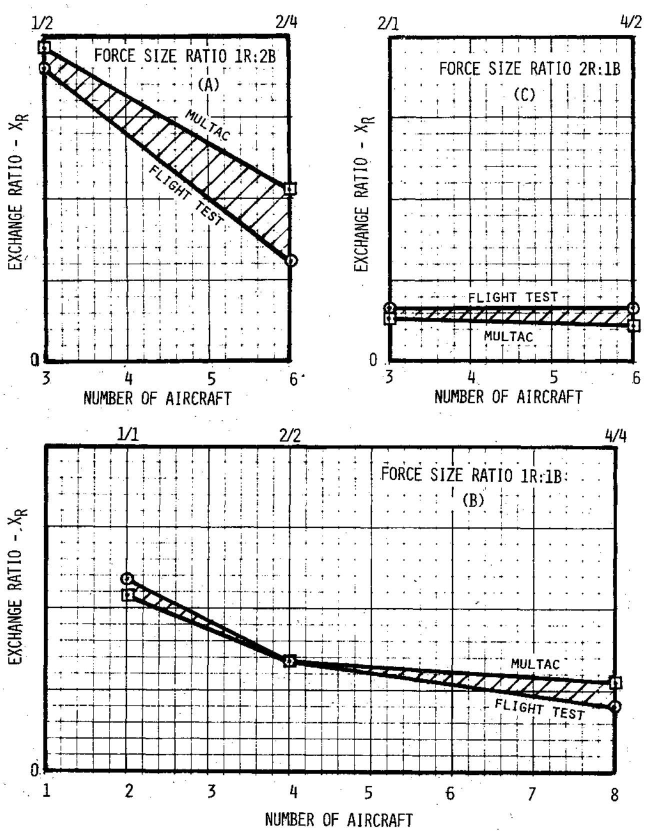

Fig. 5 Effect of numbers on aircraft performance.

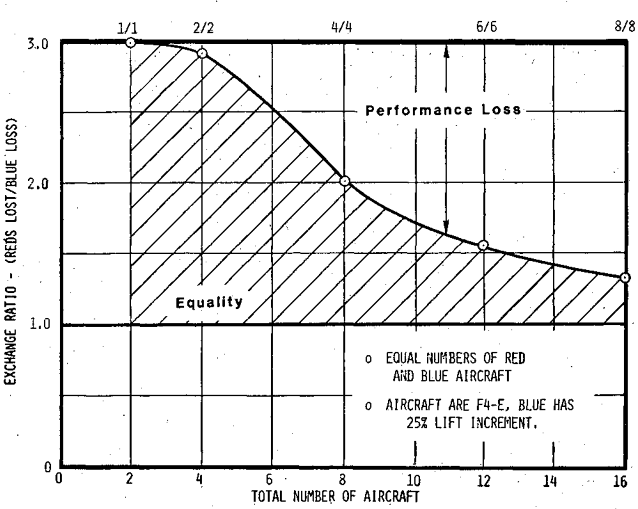

Fig. 6 Combat performance degradation with increasing numbers.

A similar result was predicted analytically in Ref. 1, where the impact of total force size on encounter between standard and a hypothetical advanced F-4 fighter aircraft were examined. The hypothetical advanced F-4 aircraft had an increase in turn capability of 25% without a drag penalty. In these results, summarized in Fig. 6, at equal force sizes the exchange ratio tends to unity as total aircraft involved increases in number. In the one-on-one case MULTAC predicted a 3:1 exchange ratio in favor of the advanced design. Predicted exchange ratio then fell monotonically with increasing number of aircraft. In the largest set of encounters considered, eight-on-eight, the exchange ratio had fallen from 3.0 to 1.3 and the advanced fighter advantage was largely nullified by numbers.

In other studies, reported elsewhere, it has been shown that certain aircraft/weapon combinations do produce exchange ratios which fall toward unity as the total number of aircraft rises. On the other hand, other aircraft/weapon combinations can produce an increase in exchange ratio as force sizes increase, the so-called "target rich" environment. This is particularly true for helicopter vs helicopter engagements where highly maneuverable vehicles are equipped with turreted guns. Thus, it is not possible to generalize on exchange ratio trends with force size ratio and total force size. However, one-on-one results certainly do not provide reasonable exchange ratio estimates for the medium force sizes of this paper. To obtain medium force size exchange ratios we must subject the particular aircraft/weapon combination to a detailed analysis in the M-on-N environment and, thus, establish exchange ratio trends with aircraft numbers. The MULTAC code provides such an analysis technique. It is now being applied to a series of conceptual designs for future aircraft/weapon combinations in both industry and USAF studies. It has also been used to develop comparative rankings of existing U.S. and Soviet fighter aircraft.

Conclusion

The present paper has demonstrated several significant points. First, it is possible to predict the outcome of an M-on-N encounter with reasonable precision in an aggregate sense. That is, while the outcome of any individual encounter cannot be predicted for it is a single sample from a stochastic process, nevertheless, if a large enough sample of flight test and analytic results are compared, the two overall exchange ratios will tend to converge to each other.

Second, in either flight-test or analytic results a sufficiently large sample must be taken or the results will be misleading. This may partially explain why the relative merits of fighter aircraft remains a topic for debate long after the war in which they engaged the enemy is over. Here, the stochastic nature of the encounter is magnified by the difficulty of data collection in the heat of conflict and with relatively uncontrolled initial conditions.

Third, and most significant, it is clear that the use of one-on-one exchange ratios to predict the outcome of large-scale encounters is inappropriate. Thus, there is not a single exchange ratio value between, say, F-15 and F-5 aircraft. Rather, the exchange ratio which holds between these two aircraft types is a strong function of force size and force size ratio, and, other considerations being equal, a superior aircraft will benefit from the advantage of numbers and suffer from a disadvantage in numbers.

Acknowledgment

This work was supported by the Air Force Wright Aeronautical Laboratories under Contract F33615-80-C-3003.

References

[1]Hague, D. S., "Multiple-Tactical Aircraft Combat Performance Evaluation System," *Journal of Aircraft*, Vol. 18, July 1981, pp. 513-520.

[2]Hague, D. S., "An Introduction to Co-Kill Probability Estimation in the M-on-N Encounter," AIAA Paper 79-1729, AIAA Guidance and Control Meeting, Boulder, Colo., Aug. 6-8, 1979.