Assessment of Uncertainty in Ballistic Response Estimates Obtained from Ballistic Limit Testing

Michael A. Riley, Kirk D. Rice, and Amanda L. Forster

Law Enforcement Standards Office, NIST, 100 Bureau Dr., Gaithersburg, MD, USA michael.riley@nist.gov

Abstract. Ballistic limit tests are widely used to assess armor performance, particularly to determine the velocity at which half of all projectile impacts will perforate the armor, a velocity typically described as the V_{50} . For armor designs where the armor's performance transitions quickly from stopping all impacts to being perforated by all impacts as a test projectile's velocity increases, the V_{50} can often be determined with acceptable certainty with only a small number of shots. When the transition is more gradual, such as is common with soft armor designs, significantly more testing may be necessary to determine either the V_{50} , or the performance at other velocities, with reasonable levels of uncertainty.

In the case of armor designs with gradual transitions, the estimated performance may be influenced by a variety of factors, including the number of shots used to assess the performance, the starting velocity of the shot sequence, and the model chosen to represent the armor performance.

This paper describes work to assess the uncertainty in performance estimates obtained from ballistic limit testing. Monte Carlo style simulations of ballistic tests are used to asses the influences of starting velocity, shot sequence length, total shots, and armor performance on the uncertainty of performance estimates.

Based on a better understanding of the performance uncertainty that is inherent in the test methods, ballistic tests can be planned such that the uncertainty is reduced to an acceptable level with a minimal amount of testing.

1. BACKGROUND

Ballistic limit tests are typically used to determine the failure point of an armor, particularly to determine the velocity at which half of all impacts with a specific type of projectile will perforate the armor, a velocity typically described as the V_{50} . Depending on the armor materials and construction, the way the armor performs as the projectile velocity increases can be quite different, and the importance and meaning of the V_{50} can vary.

For certain combinations of projectiles and armor designs, particularly hard armors with consistent thickness and material properties impacted by non-deforming projectiles, the probability of the projectile perforating the armor at a velocity slightly less than the V_{50} can be negligible¹, while the projectile will nearly always perforate the armor when the impact velocity is slightly greater than the V_{50} . This type of response can be modeled with a curve such as the dotted gray line in Figure 1. In such cases the V_{50} represents the failure velocity of the armor design, and it can be determined with reasonable accuracy using a relatively simple up-down shooting method, such as the ones described in MIL–STD–662 [7] or STANAG 2920 [6], with a small number of shots. When the tests are properly performed, the resulting estimate of the V_{50} will be sufficiently accurate, and it is only necessary for the V_{50} to be slightly greater than the expected threat velocity for the armor to be capable of consistently stopping the threat.

Unfortunately, for armor designs that use less consistent materials and for most soft armor designs, such as the concealable vests used by law enforcement, the performance transitions more slowly from the velocity where nearly all impacts are stopped to the velocity where all impacts perforate the armor. Such responses can be modeled with curves such as the solid black, dashed red, or dash-dot blue lines in Figure 1. While the rate of the transition will vary with the specific armor design, it is not unusual for there to be a low, but significant, probability of perforation for impacts that are 100 m/s below the V_{50} , and similarly a small percentage of impacts at velocities as high as 100 m/s above the V_{50} will be stopped. Experimental results have shown that for such armor designs, the small number of shots used in traditional V_{50} tests may yield inconsistent results; moreover, the V_{50} itself has less value for such armors, since it does not provide an indication of the velocity at which the probability of perforation becomes significant.

¹Certain phenomenon, such as a shatter gap may allow low velocity projectiles to perforate such armors; however, this is not addressed in this work.



Figure 1: Typical textile armor responses, modeled using a logistic distribution.

2. MODELING OF ARMOR RESPONSE

2.1 Mean of Perforations and Stops

One of the simplest methods of estimating the V_{50} is to obtain a series of stops and perforations using an updown series of shot velocities, and then to calculate the average velocity of the fastest shots that are stopped and the slowest ones that perforate the armor. This is the approach used in many test methods [7, 6, 4] that estimate the V_{50} from a small number of shots. This approach, however, does not provide an estimation of the uncertainty of the V_{50} estimate, and the estimate can be skewed by small errors or limitations of the test procedures.

2.2 Logistic Distribution

An improved approach is to fit the test data to an appropriate statistical model that uses all of the test data to estimate the armor performance. The *logistic distribution* is one such model, which is appropriate for binary data [2, 1, 3] and has been used to model a wide variety of systems with discrete responses. The shape of this model is shown in Figure 1, and it is defined by the equation [2]:

$$p(v) = \frac{e^{\beta_0 + v\beta_1}}{1 + e^{\beta_0 + v\beta_1}}.$$
(1)

Here, p(v) is the probability of a shot fired at velocity v perforating the armor, the constant, β_0 , is a dimensionless parameter, and the velocity coefficient, β_1 , has units of time/length and defines the steepness of the curve. These parameters can be estimated by the *maximum likelihood* method [2].

Equation 1 can be rearranged such that the velocity, v, is a function of the probability of perforation, p:

$$v(p) = V_p = \frac{\ln\left(\frac{p}{1-p}\right) - \beta_0}{\beta_1}.$$
(2)

For the special case of p = 0.5, the V_{50} , this relationship simplifies to:

$$v(0.50) = V_{50} = \frac{-\beta_0}{\beta_1}.$$
(3)

The variance between the estimated model and the test data can be calculated, and this result is used to calculate confidence intervals for the estimated model. The notation $\hat{\beta}_0$ and $\hat{\beta}_1$ will be used to indicate the *estimated* model coefficients, and the estimated logit, $\hat{g}(v)$, is defined as:

$$\hat{g}(v) = \hat{\beta}_0 + v\hat{\beta}_1. \tag{4}$$

Then, based on a Wald test with a normal distribution, the $100(1-\alpha)\%$ confidence intervals for the estimated logistic response are [2, 3]:

$$CI(v) = \frac{e^{\hat{g}(v) \pm z_{1-\alpha/2}} \sqrt{\operatorname{Var}[\hat{g}(v)]}}{1 + e^{\hat{g}(v) \pm z_{1-\alpha/2}} \sqrt{\operatorname{Var}[\hat{g}(v)]}}.$$
(5)

Here, $z_{1-\alpha/2}$ is the normal distribution critical value that is exceeded with a probability of $1-\alpha/2$, and:

$$\hat{\operatorname{Var}}\left[\hat{g}(v)\right] = \hat{\operatorname{Var}}\left(\hat{\beta}_{0}\right) + v^{2}\hat{\operatorname{Var}}\left(\hat{\beta}_{1}\right) + 2v\hat{\operatorname{Cov}}\left(\hat{\beta}_{0},\hat{\beta}_{1}\right).$$
(6)

The terms $\hat{Var}(\hat{\beta}_0)$, $\hat{Var}(\hat{\beta}_1)$, and $\hat{Cov}(\hat{\beta}_0, \hat{\beta}_1)$ are, respectively, the estimated variance in the constant, $\hat{\beta}_0$, the estimated variance in the velocity coefficient, $\hat{\beta}_1$, and the estimated covariance between these two coefficients.

Equation 5 can be solved to determine the range of probabilities of perforation that can be expected to occur at the selected velocity with a confidence of $100(1-\alpha)\%$ if the test is repeated. These values provide an indication of the quality of the estimation. A small range between the upper and lower confidence intervals indicates that the uncertainty of the estimated probability is small, and repeated tests will produce similar results. A large range, however, indicates that the results are less certain, and repeated tests may produce dissimilar results.

Equation 5 describes the confidence intervals in terms of probability limits at a given velocity, but often it is more useful to understand the range of velocities within which a given probability of perforation, p, can be expected to occur, i.e. CI(v) = p. Setting Equation 5 equal to p and rearranging, yields:

$$0 = \ln\left(\frac{1}{p} - 1\right) + \hat{g}(v) \pm z_{1-\alpha/2}\sqrt{\hat{\operatorname{Var}}[\hat{g}(v)]}.$$
(7)

Expanding this equation with the relationships defined in Equations 4 and 6, using z to represent $z_{1-\alpha/2}$, solving for v, and simplifying the result leads to:

$$v = \frac{-\left[\left(\ln\left(\frac{1}{p}-1\right)+\hat{\beta}_{0}\right)\hat{\beta}_{1}-z^{2}\hat{Cov}\left(\hat{\beta}_{0},\hat{\beta}_{1}\right)\right]}{\left[\hat{\beta}_{1}^{2}-z^{2}\hat{Var}\left(\hat{\beta}_{1}\right)\right]}$$

$$= \frac{\sqrt{\left[\left(\ln\left(\frac{1}{p}-1\right)+\hat{\beta}_{0}\right)\hat{\beta}_{1}-z^{2}\hat{Cov}\left(\hat{\beta}_{0},\hat{\beta}_{1}\right)\right]^{2}}{-\left[\hat{\beta}_{1}^{2}-z^{2}\hat{Var}\left(\hat{\beta}_{1}\right)\right]\left[\left(\ln\left(\frac{1}{p}-1\right)+\hat{\beta}_{0}\right)^{2}-z^{2}\hat{Var}\left(\hat{\beta}_{0}\right)\right]}{\left[\hat{\beta}_{1}^{2}-z^{2}\hat{Var}\left(\hat{\beta}_{1}\right)\right]}.$$
(8)

Again, the V_{50} is of special interest, and setting p = 0.5 causes the logarithmic term to become zero. This simplifies the relationship to:

$$v = \frac{-\left[\hat{\beta}_{0}\hat{\beta}_{1} - z^{2}\hat{Cov}\left(\hat{\beta}_{0},\hat{\beta}_{1}\right)\right]}{\left[\hat{\beta}_{1}^{2} - z^{2}\hat{Var}\left(\hat{\beta}_{1}\right)\right]} \\ \pm \frac{\sqrt{\left[\hat{\beta}_{0}\hat{\beta}_{1} - z^{2}\hat{Cov}\left(\hat{\beta}_{0},\hat{\beta}_{1}\right)\right]^{2} - \left[\hat{\beta}_{1}^{2} - z^{2}\hat{Var}\left(\hat{\beta}_{1}\right)\right]\left[\hat{\beta}_{0}^{2} - z^{2}\hat{Var}\left(\hat{\beta}_{0}\right)\right]}{\left[\hat{\beta}_{1}^{2} - z^{2}\hat{Var}\left(\hat{\beta}_{1}\right)\right]}.$$
(9)

These relationships can be used to estimate the performance of a tested armor and assess how good the performance estimate appears to be. For the simulations described in this paper, the estimated responses are compared to the known response for the purpose of understanding the uncertainty in the test method.

3. SIMULATION OF TEST RESULTS

Past research has shown that ballistic test results can be successfully modeled using the logistic distribution and other, similar, statistical distributions. Extensive testing, however, is often required to provide sufficient information for an acceptable model fit. Such testing can be quite expensive and time consuming; moreover, it can be difficult to separate the uncertainty in the results that is due to variations in the specimens and the limit size of the test data set from the uncertainty due to the test method itself.

In order to assess the uncertainty in the results that is due to the test method, an alternative approach is to use numerical simulations to perform a large number of virtual tests. While these simulations will not provide insight into the performance of armors, they can provide information about the variability and uncertainty inherent in the test methods. Since the performance of simulated armor is defined in the simulation, the performance estimated by the simulation can be compared to the defined performance to gain a better understanding of the quality of the estimate.

In the current work a series of Monte Carlo style simulations were used to assess the influence of some test parameters on the estimation of the armor response. The simulations accounted for the influence of many typical experimental limitations, such as the typical levels of random variation in the shot velocity and limitations to the precision of the propellant mass used in the cartridges. For these simulations, a typical load-velocity curve was used to estimate a propellant mass for each simulated shot, the masses were rounded the same increments used our experimental preocedures, and the resulting velocity randomly generated using a normal distribution. This approach created simulated shot sequences that were similar to those that occur during ballistic testing, including the occasional test series that fails to meet the requirements of the test standards². A series of 10,000 simulations were generated for each set of variables considered.

4. RESULTS OF SIMULATIONS

4.1 Influence of Test Starting Velocity

Since the V_{50} of an armor specimen is not known before the specimen is tested, there are a variety of approaches used to determine what velocity should be used for the first test shot. Many test protocols specify that the starting velocity should be at the V_{50} expected from the design or previous tests. Other approaches range from requiring the test to start at a relatively low velocity (e.g. [5, 4]) to some acceptance tests that require a starting velocity that is significantly greater than a required V_{50} (e.g. [7]). If the test method produces the same estimate for the V_{50} no matter what starting velocity is used, then the starting velocity can be considered to be unimportant; however, if different starting velocities lead to different results, a poorly selected starting velocity will lead to an incorrect estimate of the armor performance.

To determine if the starting velocity would influence the estimated V_{50} , a series of simulations were performed with starting velocities that varied from 120 m/s less than the actual V_{50} of the simulated armor to 120 m/s greater than the V_{50} . Simulations were performed for armor models with β_1 coefficients ranging from 0.03 s/m to 0.18 s/m. The simulations considered tests that ranged from four shots to twelve shots³ (plus additional shots that were not used in the V_{50} estimation), which were simulated and analyzed according to the method described in MIL–STD–662 [7]. Velocity steps of 30.4 m/s were used prior to the first reversal, after which the steps were reduced to a minimum step size of 7.6 m/s. For these simulations, the shots used for the V_{50} estimation were required to fall within a range of 45.7 m/s, as required by NIJ Standard–0101.04 [4]. This range is slightly larger that allowed by some other test standards, such as STANAG 2920 [6]. The simulations also considered tests that ranged from twelve to 240 shots, which were simulated and analyzed according to the method described in NIJ Standard–0101.06 [5].

²Simulations that failed to meet the requirements of test standard were discarded and replaced, as such results would be in a test laboratory.

³The number of shots is the total used in the V_{50} calculation, with half required to be stops and half perforations.

Figure 2 contains selected results from the simulations with four to twelve shots. The left side figures (Figures 2a, 2c, and 2e) show the average difference between the estimated V_{50} and the actual V_{50} , while the right side figures (Figures 2b, 2d, and 2f) show the uncertainty in the V_{50} estimates. Each average and uncertainty value in the plots are based on the results of ten thousand simulations.

When the test starting velocity is close to the actual V_{50} , the average estimated V_{50} is quite close to the actual V_{50} ; however, when the test starting velocity is significantly different than the actual V_{50} the average estimated V_{50} is dependent on the response of the armor. When the response transitions quickly from the zone of no perforations to the zone of all perforations (when β_1 is larger), the difference between the starting velocity and the actual V_{50} must be quite large – greater than 50 m/s – before the starting velocity causes a significant bias in the V_{50} estimate. When the transition is slower (when β_1 is relatively small), the difference between the average estimated V_{50} and the actual V_{50} causes a bias in the results, and the difference between the average estimated V_{50} and the actual one can be a large fraction of the starting velocity offset (See Figure 2a). Increasing the number of shots used in the test reduces the bias.

More important is the influence of the starting velocity on the uncertainty of the test results. The expanded uncertainty (94.45 % confidence) of the estimated V_{50} with respect to the actual V_{50} are shown in Figures 2b, 2d, and 2f. The uncertainty varies with the starting velocity offset, the number of shots, and the armor response. The worst cases are tests with a small number of shots, a small β_1 , and a large starting velocity offset. For example, the estimated V_{50} from a four shot test on a armor panel with a β_1 of 0.03 s/m can be expected to vary from the actual V_{50} by 50 m/s if the starting velocity is close to the actual V_{50} , and by more than 80 m/s when the starting velocity offset is 100 m/s.

The uncertainty decreases significantly when β_1 is larger, but even with a relatively large β_1 and a reasonably small starting velocity offset the expanded uncertainty was found to range between 13 m/s in a four shot test to 7.6 m/s in a twelve shot test. As β_1 increases, the influence of the starting velocity on the uncertainty of the V_{50} estimate decreases. With a β_1 of 0.18 s/m, the starting velocity offset does not cause the uncertainty to increase until the offset exceeds approximately 75 m/s.

When the number of shots used in the test become significantly large, it becomes practical to model the results with the logistic distribution. Figure 3 contains selected results from the simulations with twelve to 240 shots, with the V_{50} estimated by fitting a logistic model to the simulated test results. Again, the left side figures (Figures 3a, 3c, and 3e) show the average difference between the estimated V_{50} and the actual V_{50} , while the right side figures (Figures 3b, 3d, and 3f) show the expanded uncertainty in the V_{50} estimates.

These results have the same trends that occurred in the simulated tests with four to twelve shots, but the increasing number of shots tends to minimize the differences and uncertainty. The average difference between the estimated and actual V_{50} is small except in the cases where the β_1 coefficient is small, or where the test starting velocity offset is quite large.

The expanded uncertainty of the V_{50} estimates (Shown in Figures 3b, 3d, and 3f) tends to be smaller than in the case of of the four to twelve shot tests, but it is still dependent on the armor response and the starting velocity offset. When the β_1 coefficient is small, the uncertainty is quite large when either the number of shots is relatively small or when the starting velocity offset is large. Even when 120 shots are used and the starting velocity offset is small, the expanded uncertainty is 13.0 m/s. When the β_1 coefficient is larger, the uncertainty is reduced – particularly when a larger number of shots are used. For β_1 coefficients of 0.09 s/m and 0.18 s/m, the expanded uncertainties are less than 6.4 m/s and 3.6 m/s, even when the starting velocity offset is large.

These results indicate that the choice of the test starting velocity can have a significant influence on the resulting V_{50} estimate, but that using a large number of shots will generally improve the quality of the estimate.

4.2 Influence of Armor Response

As discussed above, the simulation results indicate that the goodness of the estimated response are, in part, dependent on the armor response. To clearly show the influence of the armor response on the results, the expanded uncertainty of the V_{50} estimates are plotted versus the coefficient β_1 for the cases of 24 (Figure 4a) and 120 (Figure 4b) shots. With all other simulation parameters held constant, the uncertainty decreases significantly as the coefficient β_1 increases.

This result shows that the uncertainty in the test results will depend on both the test method and the



Figure 2: Differences and Uncertainty in V_{50} Estimated by MIL–STD–662[7] Method.



Figure 3: Differences and Uncertainty in V_{50} Estimated Using Logistic Response Model.



Figure 4: Variation of Uncertainty in V_{50} Estimated due to Armor Response.



Figure 5: Influence of Shot Quantity on V_{50} Estimates ($\beta_1 = 0.090$).

armor response. Increasing the number of shots used in the test and carefully selecting the test starting velocity can reduce the uncertainty, but the uncertainty in the response estimate will still be a function of the β_1 coefficient, and when this response coefficient is small, the uncertainty will remain relatively large.

4.3 Influence of Number of Shots

When determining how to test an armor model, the armor's response is generally unknown, or only poorly known; therefore, a test starting velocity that will minimize the uncertainty cannot always be selected. Moreover, the coefficient β_1 cannot be controlled. This leaves the number of shots used in the test as one parameter that can be easily controlled, and that will have a significant influence on the quality of the estimated response. The mean error in the V_{50} estimate and expanded uncertainty of that estimate are plotted in Figure 5 for an increasing quantity of shots, all for an armor with a response coefficient β_1 of 0.090 s/m. For this armor response the mean error in the V_{50} estimate is small even when the number of shots is small, but the expanded uncertainty is quite high until 48 to 60 shots are used.

When assessing the results of a test, the estimated confidence intervals provide insight into the goodness of the estimated model response. When the upper and lower confidence intervals are close together (and both



Figure 6: Influence of Shot Quantity on Mean Confidence Interval Estimates ($\beta_1 = 0.090$).

close to the value of interest) the estimated response is likely to be close to the actual value; however, when the confidence intervals are far apart, the estimated response may be quite different from the actual value. Confidence intervals were calculated for each simulated test, for the purpose of assessing how the number of shots used in the test influenced the confidence limits.

The mean confidence intervals are shown in Figure 6 for both the V_{50} (Figure 6a) and the V_{10} (Figure 6b). As might be expected from the uncertainty analysis, the difference between the mean upper and lower confidence intervals decreases as the number of shots in the test is increased. The mean confidence intervals for V_{50} become reasonably small as the number of shots is increased. The mean confidence intervals for the V_{10} are somewhat larger than the mean confidence intervals for the V_{50} , and they remain larger over the range of test shots considered.

The variability of the individual confidence intervals was also considered. The mean confidences intervals for the V_{50} and V_{10} are shown in Figure 7 with the whiskers indicating the 97.7 and 2.3 percentiles (95.45 % coverage). While both the mean estimates and the variability of the confidence intervals decrease as the number of shots used in the test increase, the variability remains relatively large – with 95 % coverage the variability is about twice the mean value of the confidence limit.

5. SUMMARY AND CONCLUSIONS

These results with simulated ballistic limit tests indicate that the choice of the test starting velocity can have a significant influence on the resulting V_{50} estimate, but that using a large number of shots will generally improve the quality of the estimate. In addition, the uncertainty in the estimated armor response will depend on both the test method and the actual armor response. Increasing the number of shots used in the test and carefully selecting the test starting velocity can reduce the uncertainty, but for armors that can be modeled using the logistic distribution, the uncertainty in the response estimate will be a function of the β_1 coefficient and when this response coefficient is small the uncertainty will remain relatively large.

The simulation results indicate that the uncertainty in the estimated V_{50} will generally remain large when only a small number of test shots are used. Based on these results, 48 to 60 shots are necessary to reduce the uncertainty, and more shots are desirable.

These results provide some insight into the variability and uncertainty associated with standard ballistic limit test methods. Future research will expand this approach to study the influence of alternative test methods on the estimated response, and to determine what test strategies can be used to provide improved performance estimates using minimal quantities of armor specimens and the fewest test shots.



Figure 7: Variability in Confidence Interval Estimates at V_{50} and V_{10} ($\beta_1 = 0.090$).

Acknowledgments

Financial support for this research was provided by the National Institute of Justice under Interagency Agreement Number 2008-DN-R-121. Their support is gratefully acknowledged.

References

- [1] D. R. Cox and E. J. Snell. Analysis of Binary Data. Chapman and Hall, NY, second edition, 1989.
- [2] D. W. Hosmer and S. Lemeshow. Applied Logistic Regression. Wiley-Interscience, NY, second edition, 2000.
- [3] D. Leber. Assessment of body armor characteristics. National Institute of Standards and Technology internal communication, Mar 2005.
- [4] National Institute of Justice. Ballistic Resistance of Personal Body Armor. NIJ Standard–0101.04. U.S. Department of Justice, Office of Justice Programs, Washington, DC, September 2000.
- [5] National Institute of Justice. Ballistic Resistance of Body Armor. NIJ Standard–0101.06. U.S. Department of Justice, Office of Justice Programs, Washington, DC, July 2008.
- [6] North Atlantic Treaty Organization, Military Agency For Standardization. Ballistic Test Method For Personal Armour Materials And Combat Clothing, 2 edition. STANAG 2920.
- [7] US Department of Defense. DoD Test Method Standard, V₅₀ Ballistic Test for Armor, 1997. MIL–STD– 662F.