



LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

Benefit/Cost Ratio in Systems Engineering: Integrated Models, Tests, Design, and Production

Cynthia Nitta, Roger Logan, Steven Chidester,
Mary Frances Foltz

November 1, 2004

Disclaimer

This document was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor the University of California nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or the University of California, and shall not be used for advertising or product endorsement purposes.

BENEFIT / COST RATIO IN SYSTEMS ENGINEERING: INTEGRATED MODELS, TESTS, DESIGN, AND PRODUCTION

*Cynthia K. Nitta, Roger W. Logan, Steven K. Chidester, and M. Frances Foltz
University of California
Lawrence Livermore National Laboratory
Livermore, CA 94551*

ABSTRACT:

We have previously described our methodology for quantification of risk and risk reduction, and the use of risk, quantified as a dollar value, in the Value Engineering and decision tradeoff process. In this work we extend our example theme of the safety of reactive materials during accidental impacts. We have begun to place the validation of our impact safety model into a systems engineering context. In that sense, we have made connections between the data and the trends in the data, our models of the impact safety process, and the implications regarding confidence levels and reliability based on given impact safety requirements. We have folded this information into a quantitative risk assessment, and shown the assessed risk reduction value of developing an even better model, with more model work or more experimental data or both. Since there is a cost incurred for either model improvement or testing, we have used a Benefit / Cost Ratio metric to quantify this, where Benefit is our quantification of assessed risk reduction, and cost is the cost of the new test data, code development, and model validation. This has left us with further questions posed for our evolving system engineering representation for impact safety and its implications. We had concluded that the Benefit / Cost Ratio for more model validation was high, but such improvement could take several paths. We show our progress along two such paths; simple and high fidelity modeling of the impact safety process, and the implications of our knowledge and assumptions of the probability distribution functions involved. At the other end of the systems engineering scale, we discuss the implications of our linkage from model validation to risk on our production plant operations. Naturally, the nature of most such methodologies is still evolving, and this work represents the views of the authors and not necessarily the views of Lawrence Livermore National Laboratory.

INTRODUCTION:

Continuous assessment of safety processes must be made and accommodated in our production operations, especially those involving hazardous materials such as explosives with respect to their potential to react when handled in different manufacturing or functional environments. Over-conservatism hampers productivity, and yet under-conservatism leads to

safety risks. Previously¹, we discussed a process from Requirements to Model Verification & Validation (V&V) and eventually to the investment strategy shown in Figure 1. The V&V process leads to models with Confidence [C] bounds on Uncertainty [U]. We then obtain margins [M] and reliability [R] equivalents, and Quantified Reliability at Confidence (QRC)². The assessed lower-bound QRC value converts to Likelihood and forms one axis of our Risk Diagram. With a model, V&V, and linkages to reliability and risk, we have begun to realize that we have most of the ingredients to view our problem in a systems engineering context, and to show how model V&V fits within that context.

Developing our impact model in a systems engineering context will enable us to balance both of these ends of the scale, and defend our determination of this balance. However, as we show in the analysis, the balancing point depends on the details and fidelity of the data and model validation status.

The key to closure of our system engineering cycle is an investment strategy process linked back to V&V. Our metric for continued investment or closure of the issue is the Benefit/Cost Ratio (BCR):

$$\text{Net BCR} = (\$B - \$C) / \$C \quad [1]$$

Benefit (\$B) is taken to be proportional to assessed Risk Reduction, expressed as Likelihood (a term proportional to reliability and confidence from our V&V assessments) times Consequence. For simplicity we will restrict the current discussion from nonlinear risk-consequence relationships. Then the net "BCR" is (Benefit-Cost)/Cost. The BCR is well accepted as a key to the investment strategy process³; its linkage back to V&V is provided by the QRC analysis.

The dollar benefit of V&V as assessed risk mitigation can be explained using a Risk=Likelihood*Consequence Matrix as shown in Figure 2. V&V plays a quantified role, one that is now directly proportional to Risk Reduction and Value Engineering quantities. Previously, we discussed each aspect of Figure 1 and Figure 2 in turn, relating them to an example of impact testing of explosives to quantify

and bound the reaction thresholds¹.

We can relate our goals and V&V quantification process for the impact safety example to the following system engineering attributes as described by Hsu⁴:

- Requirements Management
 - (Handling the explosive safely)
- Functional Analysis and Allocation
 - (In this paper we will focus on the role of V&V in quantifying impact safety)
- Functional and System Architecture
 - (We consider the system as the physical plant where the explosive material operations take place, the operations sequence and risk mitigation measures, and in this case focus on the impact safety models and V&V that quantify our assessed risks for tradeoff and benefit/cost)
- Integrated Master Plan and Schedule
 - (Beyond the scope of this paper, but our BCR for “what to do next” is part of the basis for developing and revising the plan and schedule)
- Risk Management
 - (During each iteration of this cycle, the V&V assessment of reliability and confidence quantifies our assessed risk and risk mitigation potential)
- Decision Analysis (Trade Study)
 - (Potential for further risk mitigation is balanced against cost via the BCR. This determines the next iteration of system engineering, if any).
- Technical Performance Measurement, TPM
 - (In general, our TPM is process flow at the production plant with minimal risk. Here, we will focus on the TPM for our model V&V, which *enables* lower *assessed* risk, in turn enabling us to meet the production plant TPM).
- Interface Management
 - (We view this as the feedback loop between the requirement for safety, risk level, impact test data, model development, model V&V, and the cost of improving each of these)
- Verification and Validation (V&V)
 - (System Verification here would be viewed as, “did you address the

safety problem correctly?” and System Validation would be “did you address the right safety problem?” In this work we will focus on Model V&V as it feeds into assessed risk; that is, “did we code and run the impact model right”, and “is this the right impact model, compared to experimental data”. It is important to distinguish *System* V&V from *Model* V&V.)

Before we proceed with the V&V process, we have to know the requirements our product or system will have to meet, and which of these our model is to address. After that, depending on both the fiscal and scientific ability to perform a certain level of V&V assessment, we proceed with various degrees of qualitative and (ideally) quantitative validation. In our previous work¹, our focus was on the quantitative validation of our model for impact safety. The model we used was a simple zero-dimensional (0D) regression model. We concluded by showing that our model validation had enabled a quantification of risk and potential risk reduction. Further model development, experimentation, and model validation was warranted and we estimated a high BCR for doing so. However, given the simple regression nature of our simple “0D” impact model, our first step toward improvement was to develop a finite-element model with the potential for physics-based improvements in V&V.

In this work, we show how the development and quantitative validation of a 3D finite-element model for impact safety has enabled us to lower our assessed risk based on model V&V. We also show that given the issues and implications of plant operation, further development and hierarchical validation of our model are warranted, and we estimate a good BCR or return on investment for this further impact model development and validation.

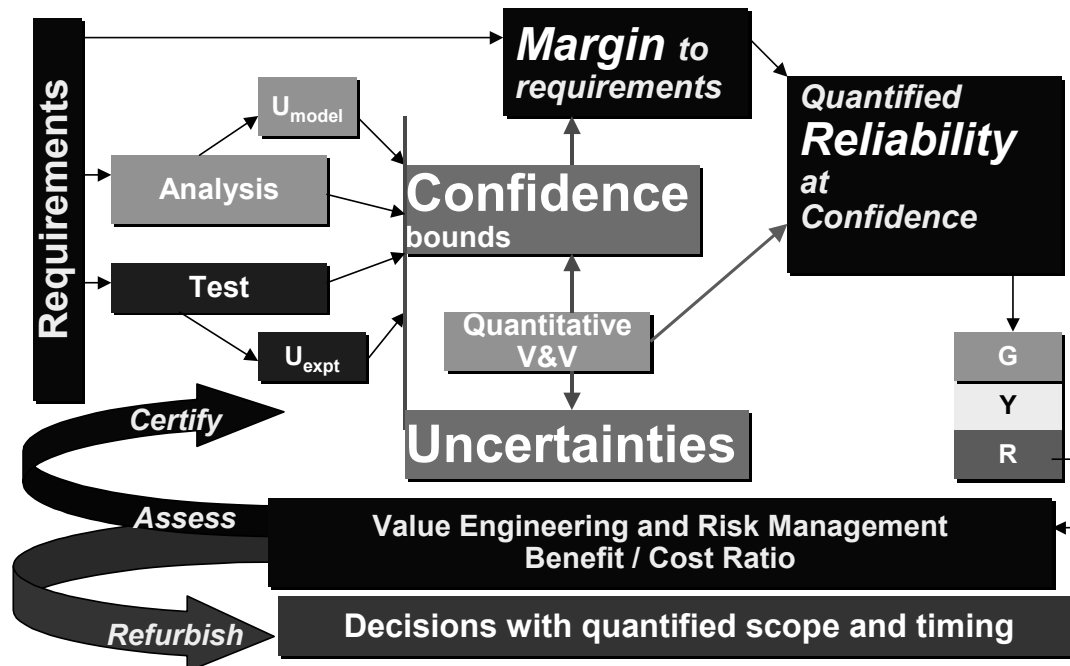


Figure 1. Flow diagram from system Requirements through V&V, through uncertainty quantification and margins; onward through reliability at confidence, then Value Engineering, and Risk Management via Benefit/Cost Ratio BCR. This process is described in detail by Nitta and Logan².

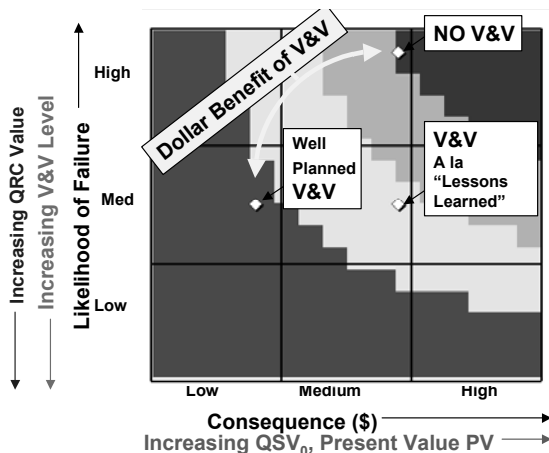


Figure 2. Dollar Benefit of V&V and Quantitative Certification, expressed as a standard Risk=Likelihood*Consequence Matrix. For the sake of simplicity of analysis demonstration we use a symmetric risk matrix. Likelihood becomes analogous to assessed (1-QRC); Consequence is typically expressed in dollar terms.

QUANTITATIVE VALIDATION STATEMENT:

In the V&V portion of this work, we will focus on a *quantitative validation process*, using 4 steps "ABCD" as described previously. These four steps are, briefly:

- Planning and Assessment:** Assess the problem to be modeled, the product and model requirements, available validation data (the referent) and codes. See how well a fit to the data can be obtained with the model. (Note that this last step is calibration, not validation; it shows only that the model has the ability to generate a high fidelity match to the available data. However, since we have not yet generated confidence or prediction intervals on the model to distinguish model validation from model calibration, Step "A" tells us nothing about confidence, reliability, or predictive capability.
- Solution Verification:** For models that are discretized in the spatial, temporal, and/or iterative domains, discretization studies must be done to assess convergence. Solution Verification should quantify uncertainty and error estimates due to the use of models that are not fully converged.
- Validation to the Referent:** This is the quantitative comparison of the model with the referent data. The comparison must be done with a statistical or other quantitative uncertainty method that enables a quantified confidence bound to be established.
- Predictive Capability and Adequacy:** An assessment of the predictive capability, or confidence bound, outside the domain of the

referent is assessed. It is rare that we would take the time to develop a model if we did not intend to extrapolate from the referent data in one or more dimensions of application space. The rigor of the validation methods used in Step C, plus the maximum employment of expert knowledge, determine the credibility of the extrapolation confidence bound estimates from Step D.

The “ABCD” process for Quantitative V&V should lead to a Quantitative Validation Statement supporting the definition of Validation given by Cafeo and Roache⁵:

‘Validated Model: A model that has confidence bounds on the output. A validated model output has the following characteristics:

1. The quantity of interest
2. An estimate of the bias (*i.e. confidence bounds are not necessarily centered around the model output*)
3. A set of confidence bounds *drawn at an assessed confidence level*

A validated model is one where evidence supports a *Quantitative Validation Statement* such as⁵:

“I am 90% confident that if I build and measure the quantity of interest, that it will fall within the confidence bands (of uncertainty) shown around the model output.”

Quantitative V&V in Systems Engineering

In previous work¹, we showed a validation study example from a design safety study of velocity reaction thresholds for heated explosives impacted by steel projectiles⁶. We now describe some of the details of the experiments, and then proceed to the validation of the model used to characterize them, including an assessment of its predictive capability.

In the experiments, impact tests were performed on heated energetic material samples to help us develop and validate accident scenario models involving combined thermal and impact conditions. To determine heated thresholds, Steven Test targets (Figure 3) containing explosives PBX 9404 or LX-04 were heated to the range of 150-170 C and impacted at velocities up to 150 m/s by two different projectile heads fired from a gas gun. The projectiles had masses of 1.2 kg and 1.6 kg, and spherical radius heads of 30mm and 6.4mm, respectively.

For the impact tests using the apparatus in Figure 3, the measured and ambient threshold velocities for explosive reaction are shown in Table I and Table II. The following trends were observed:

1. Threshold velocity increases with temperature
2. Threshold increase with temperature is more pronounced for LX-04 (higher binder and lesser explosive content) than for PBX 9404.
3. Threshold velocity increases with projectile head radius.
4. The threshold increase with radius is more pronounced for LX-04, and barely noticeable for PBX 9404.

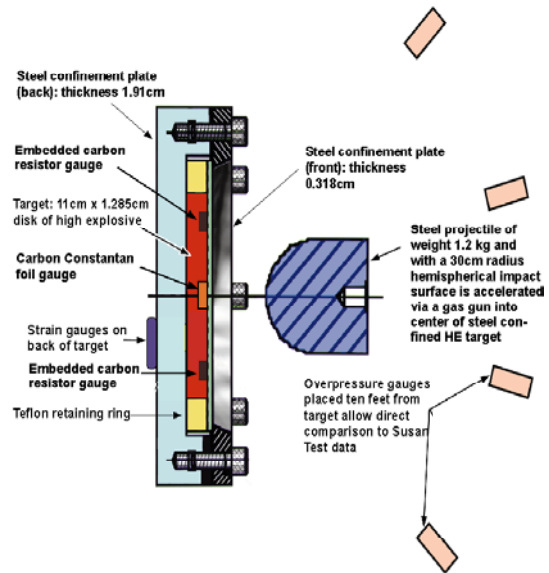


Figure 3. Steven Impact Test: 30mm (Steven, Head #1, shown) and 6mm (Duff R/T, Head #2, not shown) projectiles.

A predictive model for this thermal + impact accident scenario must, as a minimum, capture the four phenomena observed in the data within adequate confidence bounds. In order to avoid undue physics “surprises”, the model should also be physically and geometrically based so as to capture the physical and geometric phenomena during the impact. We had previously used a simple regression “model” for predictive capability, and concluded that we did not meet the system requirements; our benefit / cost ratio analysis supported development of a more physics-based finite element model (FEM) of the impact. The 0D Regression model results are shown in the next to last column of Table I and Table II. The 3D FEM results are shown in the last columns. We can compare the raw values of threshold velocity (meter/second) from the test referent, the 0D model, and the 3D model, in Table I and Table II. However, that is only a comparison that addressed model

fidelity, not model *validation*. In this work, we will compare the quantitative validation process for both the simple 0D Regression Model and the 3D LS-DYNA⁷ Finite Element model (Figure 4), with the validation referent being the impact test data in Table I and Table II.

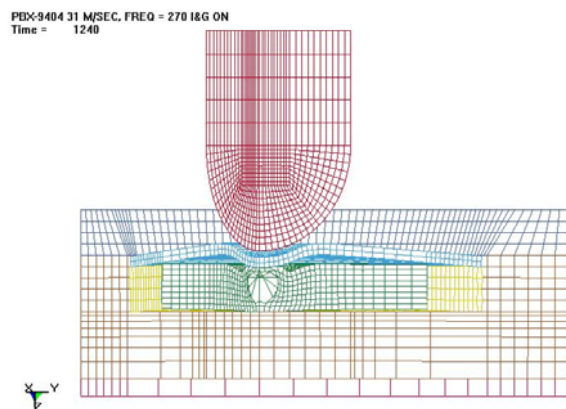


Figure 4. Quarter-symmetry 3D Deformed mesh of the LS-DYNA finite element model of the Steven Test impact. Chemical reaction begins in zones under projectile nose.

Of course, any realistic model validation statement contains caveats and assumptions. Such is obviously the case with a regression-based model, but is also true for the finite element model because its material behavior and kinetics still involve empirical approximations. As stressed in any standard statistics text, extrapolation of empirically based models can be particularly dangerous. Furthermore, several physical and chemical variables discussed in Switzer et al⁶ are ignored for this model validation; we plan to account for these eventually in our mechanical-thermal-chemical finite element model. Nevertheless, the empirical models used here, comparing to nominal and even truncated threshold values, will serve our purpose in demonstrating the method and results of our quantitative validation process, and allow us to carry those results into risk management to quantify the iterations in our systems engineering construct.

We will use these experimental results, and a statistical validation method that does not preclude “tuning dials”, but accounts for them explicitly, to generate validated confidence bounds on our 0D simple regression and 3D finite element models for impacts over the regime of these experimental (validation referent) data.

TABLE I. DATA FOR STEVEN/DUFF Tests: Head#1, R=30mm (*caveats in Switzer et al⁶)

Test #	HE type	%Explosive	Test Temp C	Nom. m/s Thres. Vel	0D Regress. Model	3D LS-DYNA Model
1-1	PBX 9404	94	20	34.0	33	34
1-2	PBX 9404	94	20	36.0	33	34
1-3	PBX 9404	94	20	35.7	33	34
1-4	PBX 9404	94	150	50.5	50	49
1-5	PBX 9404	94	*165	47.2	52	49
1-6	LX-04	85	20	45.0	44	42
1-7	LX-04	85	20	43.0	44	42
1-8	LX-04	85	150	*125.7	134	72
1-9	LX-04	85	170	*153.2	148	80
1-10	PBX 9501	95	20	46.4	33	46

TABLE II. DATA FOR STEVEN/DUFF Tests: Head#2, R=6.4mm (*caveats in Switzer et al⁶)

Test #	HE type	%Explosive	Test Temp C	Nom. m/s Thres. Vel	0D Regress. Model	3D LS-DYNA Model
2-1	PBX 9404	94	20	29.1	32	
2-2	PBX 9404	94	150	48.8	47	
2-3	PBX 9404	94	*165	48.2	49	
2-4	LX-04	85	20	30.7	34	30
2-5	LX-04	85	20	30.5	34	30
2-6	LX-04	85	150	64.8	63	
2-7	LX-04	85	170	64.7	67	
2-8	PBX 9501	95	20	28.6	32	31

The plot in Figure 5 is actually the end result of completing the “ABCD” validation process described previously¹ and summarized above. The plot shows numerous small squares of experimental data (N=10+8=18 data points) for this Steven Impact scenario. The scatter depicts the variation in Head Radius, Explosive Type, and Test Temperature. If our intended application of the impact model involves thermal conditions, we wish to quantify reliability and confidence during an impact scenario

where impacts of a known velocity must be tolerated without reaction. We develop a model, providing the middle solid line, to answer this question. Quantitative Validation requires a quantitative assessment of the confidence bounds on this model. This is determined by comparison to a known number of data points (N) as shown in the plot. The model error between the measured output and model output is sometimes too small to be shown in a plot such as Figure 5, but is used, along with experimental error, variability, and assumed probability distribution functions (PDFs) to construct a confidence bound (the outer dotted lines in Figure 5) on our analysis. Any adjustable (calibration) parameters used in the model must be counted as model degrees of freedom (K), so our effective number of data points becomes (N-K). Fewer data points (N) or more model adjustables (K) will result in wider confidence bound lines in the plot. Model adjustables are a fact of life; there is neither the time nor funds to avoid them all. It is simply important to quantitatively account for them in the validation assessment.

FROM VALIDATION TO ADEQUACY

The quantitative V&V process provides us with confidence-bounded uncertainties over the domain of the validation referent data. Since a given application of the model may involve excursions away from some or all of the referent data in one or more parametric dimensions, we must use the validation information to construct a performance chart (Figure 5) for the *quantity* of interest and over the *domain* of interest. In the case of Figure 5, the *quantity* of interest is velocity of the reaction threshold. The *domain* of interest is the HE temperature on impact; given that the HE is in this case PBX 9404, and that the impact is with Head #1 (Steven Projectile).

The model fidelity of the fit to the referent data is seen by comparing how close the model predicted threshold velocity (dots) are to the referent data (squares indicating approximate experimental uncertainty). The fidelity of the fit, good or bad, does not necessarily have anything to do with validation. A high fidelity fit can be obtained with excessive calibration of a model. The validation process should trap instances of excessive calibration, and reward, with a tight confidence bound as shown in Figure 5, *only the combination* of a high fidelity fit with minimal calibration parameters.

Once we have performed the quantitative “ABCD” process for V&V (including uncertainty

quantification), and obtained the model mean, bias corrected mean, and confidence intervals on the mean as shown in Figure 5, we have all the information that V&V can provide to feed back into the System Engineering loop.

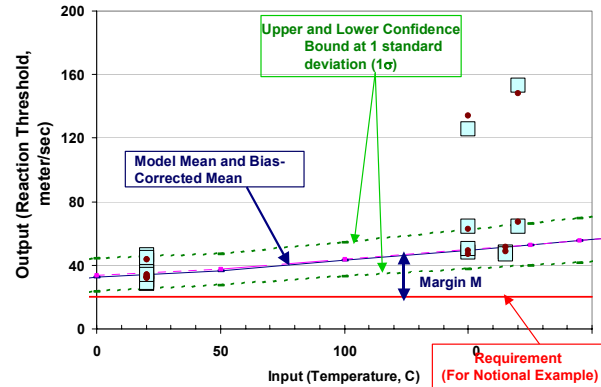


Figure 5. “D” of “ABCD”: Model Validation plot for PBX 9404 impact threshold, using 0D Regression model.

Next, we have to compare our validated assessment of the model for HE impact with the requirement for the system. For this study, we will assume an example requirement that impacts of 20 m/sec or less be endured, at temperatures of 20 °C or above, without encountering the reaction threshold. This is drawn as the “requirement” line on Figure 5. The Margin “M” between the mean bias-corrected model and the requirement line is an interesting quantity to follow, although we will not use it directly in our reliability or risk analyses. Depending on the level of Confidence “C” we choose, we can calculate the model assessed Reliability “R” and maximum product “QRC”, where $QRC = 1000 \cdot R \cdot C$ as discussed previously¹. With a perfect model and abundant referent data, and the only uncertainty remaining was in fact the test-to-test variability in the data, then the quantity $(1 - QRC/1000)$ would in fact be a good approximation to the system failure likelihood, given the scenario and requirement imposed. Since our referent data is usually not abundant but sparse, and our models usually have a quantifiable but far from perfect validation confidence bound, we can only say that the quantity $(1 - QRC/1000)$ is our assessed upper bound likelihood of failure of the *model of the system*, not the system itself. We might be able to infer higher system reliability (lower failure likelihood) for specific application points, but not by using our model applicable over the referent domain. Hence,

we will base our risk assessment and system engineering on the lower bound QRC. Since our definition of assessed failure likelihood of the model of the system is the quantity $(1-QRC/1000)$, the use of the lower bound assessed QRC number is equivalent to use of upper bound failure likelihood.

In Table III, the first column of results shows the validated quantities for the impact into PBX 9404 with impact head #1 at 20 C. The peak QRC of $QRC=748$ implies a likelihood of failure for the model of the system of $P_{fail}=1.-748/1000=0.252$. Although this is not to be interpreted as the system failure rate, it is not a very comforting value for likelihood of failure. The margin “M” of 13.9 m/sec looks adequate enough, but with an uncertainty “U” (at 1σ) of 10.5, it is the model confidence bounds that are keeping our validated QRC value so low. And yet, we have gone about as far as we can go with our simple regression model.

SYSTEM ENGINEERING: NEXT ITERATION

The next iteration, as justified by our Benefit / Cost Ratio BCR analysis in our previous work¹, was to develop a more physics-based finite-element model of the Steven Test impact. We applied the LS-DYNA model of DePiero⁸ to this impact condition and geometry. This 3D LS-DYNA finite element model still contains 5 major tunable parameters ($K=5$) which have a large effect and have an unknown probability distribution. None of these parameters or dials “K” have a large effect on the model on their own. For example, consider the LS-DYNA model fit to the PBX 9404 impact data shown in Figure 6. The mean LS-DYNA model (middle line) is a good fit to the 3 data points shown at exactly the conditions modeled. The upper and lower uncertainty bounds are actually 3σ values; that is, the PDF for this parameter (friction coefficient) was actually estimated based on material level tests, and varied $+3\sigma$ and -3σ from its mean value in the model. The effect is quite small. However, these dotted lines in Figure 6 are not 3σ confidence bounds on the LS-DYNA model. They only represent confidence bounds if this single parameter were the only free parameter. Confidence bounds on a model can only be generated by considering all the free parameters in a model and the effect they might have.

This last point is brought home by observing Figure 7. Having applied the 3D FEM model to the PBX 9404 situation of Figure 6, the resulting comparison was encouraging. However, when we applied the model to the LX-04 type HE, the data and 3D FEM model comparison was as shown in Figure 7. There

is a much greater increase in threshold as temperature increases. This is captured by the model qualitatively, but quantitative agreement is poor.

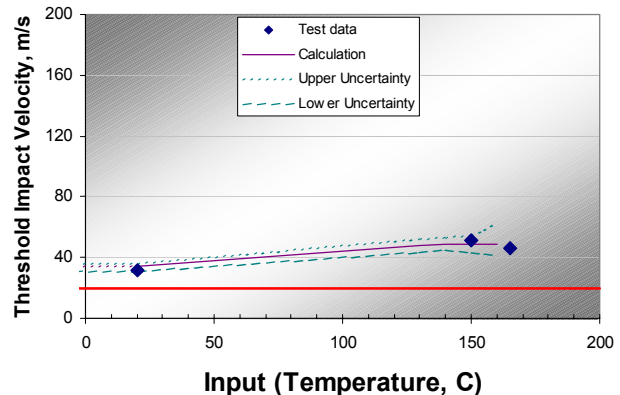


Figure 6. “A” of “ABCD”: Model *planning and calibration* plot for PBX 9404 impact threshold, using 3D LS-DYNA with simple binder material model.

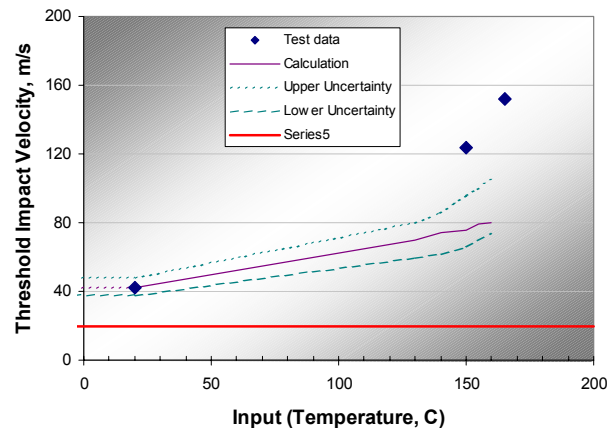


Figure 7. “A” of “ABCD”: Model *planning and calibration* plot for LX-04 impact threshold, using 3D LS-DYNA with simple binder material model.

If we considered the $+3\sigma$ to -3σ effect of the dial “K” for friction alone, and mistakenly called the resulting dotted lines in Figure 7 “confidence bounds”, it would make the comparison, albeit a small sample comparison, very suspect; 2 of the 6 total data points in Figure 6-7 would lie outside our “ 3σ ” limits.

To construct a meaningful confidence bound for a given model of the impact system problem, we must consider all the free parameters in the model, not just one of them. This model uncertainty contribution is often the largest source of uncertainty in V&V and in a system engineering sense. For nonlinear complex physics problems, model uncertainty due to epistemics or free parameters can overwhelm the separate contributions of test variability,

measurement uncertainty, manufacturing or aging variations, or mesh resolution (solution verification) uncertainty. When we have several free parameters ($K=5$ in this FEM model) and we still do not have a high fidelity fit to the data, the resulting confidence bounds can be alarmingly wide. We show such an example in Figure 8. This figure is our first assessment of confidence bounds on our newly built 3D FEM model for the Steven impact test. Figure 8 shows the raw FEM model output for PBX 9404 impact (the solid line), a large bias correction that attempts to correct the model for temperature dependence of threshold (the dashed center line), and 1s (not 3s) confidence bounds on the output of the model. On visual examination, these 1s confidence bounds are now certainly wide enough to account for the discrepancy between our model and the data in Figure 7. Since the 3D FEM model cannot discern between PBX 9404 and LX-04 except by what we enter into the model input deck over the continuum of the domain of application, we must consider the fit (or rather lack thereof) to the LX-04 points, and not just the good fidelity fit we obtained to the PBX 9404 points. It is this latter fact that makes the bias and confidence bounds in Figure 8 so wide. Even though the bias-corrected margin " M "=38.4, is much larger than the 0D model, the uncertainty $U|1\sigma=50.5$, is also much larger than the 0D model. The result is lower reliability and lower confidence at the peak $QRC=581$ at $C=0.5\sigma$ (compared to the 0D model $QRC=748$). This comparison of quantities is shown in Table III, providing the input for subsequent system engineering activities of risk assessment and benefit / cost ratio decision trade analysis.

Since the 3D FEM model only permits an assessed $QRC=581$ out of 1000, a value lower than the assessed $QRC=748$ for the 0D model, we should be asking right about now, if it was worth developing the 3D FEM model at all. Our reasoning is yes, from a system engineering point of view. An assessed $QRC=748$ from the 0D model is not high enough to meet system requirements; the model assessed likelihood of exceeding reaction threshold will still be too high and will support a high benefit / cost ratio BCR to do more model development, more testing, and more V&V. However, the 0D model had reached its predictive capability limit; we had no means to obtain more data and constrain or eliminate the free parameters in this simple regression model. With the 3D FEM model, we now have the hope of improving the physics in the model, and obtaining material data to constrain some or all of the free parameters.

For example, our 3D FEM model shows a good fit to

all of the data for the PBX 9404 and PBX 9501 HE⁸ types, and a poor fit only to the LX-04 HE type at high temperature. One of the most significant differences in these HE types is the Viton binder used in the LX-04. We suspect (and hope) that the use of a better temperature and rate dependent model for LX-04 will greatly improve the fit at high temperatures. We also suspect (and hope) that we can obtain these model improvements without tuning the model or adding free parameters, but rather by simply characterizing the actual mechanical behavior of the Viton binder and incorporating this into the 3D FEM model.

What could we hope to obtain from an improved model, with newly obtained data on the Viton binder? As we noted, we would expect the major improvements to the data fit for the LX-04 at high temperatures, the conditions with the notably poor fit as shown in Table II. If we could improve the fit at just these two points, and make the 3D FEM model fit those two data points as closely as the 0D model, and yet retain our limit of $K=5$ tunable model parameters, we could obtain V&V confidence bounds shown in Figure 9 instead of those in Figure 8. The quantitative validation plot in Figure 9 is based on the same information and data vs. model fit as Figure 8, except for our hope of improving the model fit to the two shaded points in Table II from our improved Viton mechanical model. We now have quantities (see Table III) of Margin " M "=15.6 m/s, uncertainty $U|1\sigma=7.9$, a remarkable improvement over our first cut 3D FEM model of Figure 8. This results in a peak $QRC=848$, an improvement of $\Delta QRC=+267$ over the FEM model we have now, and an improvement of $\Delta QRC=+100$ over the 0D regression model.

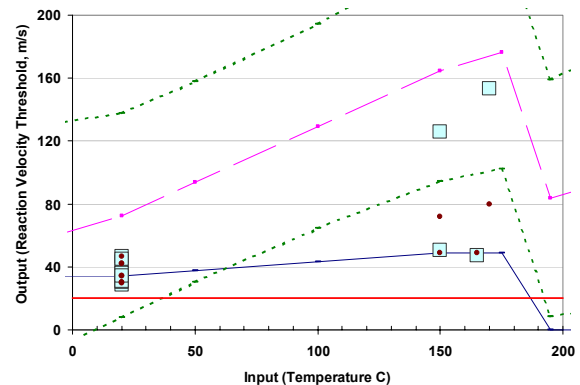


Figure 8. "D" of "ABCD": Model Validation plot for PBX 9404 impact threshold, using 3D LS-DYNA with simple binder model.

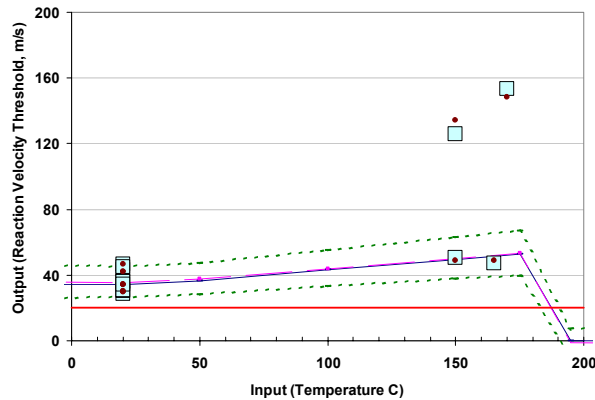


Figure 9. “D” of “ABCD”: Model Validation plot for PBX 9404 impact threshold, using 3D LS-DYNA with *representation* of enhanced binder model.

Better still, we can hope for even more improvement with further V&V, because we now have a 3D FEM model where we can hope to relate remaining discrepancies and free parameter settings to physical quantities and phenomena. Hence, we have a path to improve the model in the future.

TABLE III. Effect of Model Choice (Tuning dials “K”) on Margin, Uncertainty, and peak QRC for tested conditions

Test Condition	0D Regress Model, N=18 K=8	3D FEM, N=13 K=5	3D FEM, est. w/ new Viton Model
Margin M, m/s	13.9	38.4	15.6
Uncertainty U, 1 σ , m/s	10.5	50.5	7.9
Peak QRC (QRC=1000 at R=1 and C=1)	748	581	848
Confidence C at Peak QRC, X σ	0.8	0.5	1.1

V&V Results in Systems Engineering Context

We can use the quantities assessed above for decision trade studies based on inputs from V&V and QRC. We will use a simple Risk=Likelihood*Consequence

construct, with Likelihood=1-QRC/1000 as our model assessed value. Risk reduction can be obtained not only by changing physical scenarios for physical risk mitigation, but also by improving the model, with more integral or hierarchical data, better physics, leading to improved V&V resulting in a higher value of our lower bound model assessed QRC value. The latter is important because it allows us to attach a direct dollar benefit $\Delta\$B = \text{Assessed Risk Reduction}$ to the V&V process. Of course either improving the physical product (tighter manufacturing tolerances, etc.) or improving the model (V&V, model or code improvements, etc.) will cost a dollar amount “ $\Delta\$C$ ”. We can use a Benefit/Cost Ratio:

$$BCR = (\Delta\$B - \Delta\$C) / \Delta\$C \quad [1]$$

The BCR, computed here for product (or model) improvement, gives us a quantity to help answer the question, “was our product or model improvement process worth the cost?” We link the V&V level for particular simulation capabilities (including validation experiments) to the value of products and product decisions made under budget and schedule constraints.

Higher QRC, lower assessed Risk, and the BCR

If QRC is “too low” for predictive adequacy, we are implicitly saying there would be a large Benefit “B” from raising QRC (i.e., a tighter Validation Bound). Since 1-QRC/1000 corresponds to our assessed upper bound model+system failure likelihood, the increased benefit $\Delta\$B$ of raising QRC can be quantified for example in terms of product liability avoidance or increased value to the customer.

Once we know the $\Delta\$B$ we could achieve for a given ΔQRC , we can look for the increased cost $\Delta\$C$ needed to achieve this. We have to examine what we need to do to achieve the ΔQRC . This might include:

- More V&V Analysis
- More system tests (if possible)
- More data (higher “N” for more confidence)
- More accurate data measurements
- Better physics (less tuning dials “K”)
- Better production tolerances
- Bigger computing platform (reduced convergence uncertainty from “B” of “ABCD”; in this analysis, mesh convergence for the 3D FEM model was well established by DePiero⁸, so we expect

minimal gains from better compute platforms in this case)

As an example of this BCR process, let us reconsider the model assessed QRC Numbers of Table III. One of our future goals is to develop a more physics based Viton binder behavior for our 3D finite-element model (FEM) of these impact events and predict threshold velocity. We showed that we could hope to gain increased QRC with an improved binder model in Table III. Now we ask whether this further 3D FEM model development and testing would be worthwhile in terms of benefit / cost.

In Table IV, we reiterate the QRC Numbers from our best 0D regression empirical model (PBX 9404 impact at 20 °C). We then show the new QRC Number in the 2nd row of Table IV; as we noted above, the QRC is not as good as the existing 0D model so the model is not yet ready for use in the risk reduction component of our systems engineering context. However, assuming that our improved hierarchical developments and data for a physics-based Viton binder in the finite-element model of the impact could achieve the same fit of model to data as the 0D model for the two anomalous points shaded in Table 1, we could achieve the model assessed QRC=827 of the 3rd row of Table IV with the 3D FEM model. We estimate the cost of obtaining the Viton data, building this into the 3D FEM model, and redoing the V&V, would bring our investment to a total of $\Delta\$C=\$2.0M$ in the 3D FEM impact model (the 5th column of Table IV). Was the increase a good investment, considering just the impact scenario shown? If we assume a notional Risk Consequence value of \$200 Million for even a mild unplanned explosive reaction accident, we can calculate the Benefit of our model investment as reduction in the risk due from the model+system assessed QRC increase as (in our simplified example), $\Delta B=(\Delta QRC/1000)*\200Million . This value is shown in the 4th column of Table IV. We can then calculate the $BCR=(\Delta\$B-\Delta\$C)/\Delta\$C$ of developing such an improved 3D FEM model. The $BCR=6.9$ value, one for improvement in the *Model V&V* process, is shown in the final column of Table IV.

TABLE IV. Notional process of calculating the assessed Risk Reduction and BCR of development of an improved impact model for each explosive type.

Model	Max QRC	ΔQRC vs prev. model	$\Delta\$B$, Risk Reduction benefit	$\Delta\$C$, Cost to achieve	BCR
0D Model, Base	748				
3D FEM, Current 1 st Cut Model	581	N/A	N/A	\$0.3M Approx.	N/A
3D FEM, New Viton, N=12	827	79	\$15.8M	\$2.0M Approx.	6.9
3D FEM, New Viton, N=13 (add 1 data point)	848	21	\$4.2M	\$0.1M Approx.	42.0

With the cost of each Steven Test and related analyses well under \$0.1M each, we calculate a $BCR=42.0$ for adding to the data referent used for V&V of the model. We have previously noted that in BCR prioritization, items with high BCR, e.g. $BCR=42.0$, should be done first if possible, and this is our next step after the improved Viton model which is needed first to make the 3D FEM model competitive with the 0D regression model. With the 3D FEM model improvements in place to give us the QRC of Table IV (hopefully with high BCR's as we estimate), we are on track for even further improvements in what is now a physics-based 3D FEM model, used in a Systems Engineering construct, with demonstrated ability for reductions in assessed risk.

Naturally, such a series of numbers as in Table 4 cannot replace expert judgment in the decision process. This should be more evident than ever after following all the assumptions and uncertainties in this example validation with calculation of risk reduction and BCR. Furthermore, we stress that in our QRC effort to construct single measure combinations of Reliability and Confidence, we have only shown one such measure in this analysis. The "lower bound" QRC we discuss in this work is essentially a quantification of the assessed model+system Reliability "R" within the stated (and inferred) statistical confidence. Outside the confidence interval, we quantify reliability with what

can be viewed as a “coin flip” or $R=50\%$. There are of course more complex ways to map both R and C onto a symmetric (or asymmetric) risk diagram. However, the analysis process, from V&V to QRC to risk to BCR, can provide a numerical defense of our decision, showing a process with quantitative diligence backed up by expert judgment for a balanced and defensible decision process.

Other Methods of Risk Mitigation: Consequence

In plant operations, we can only work in a System Engineering context using assessed risk; we believe that our real risks have such low likelihoods that we will never accumulate enough data to quantify our actual risk. We can only establish conservative bounds on our model assessed risk. This brings the desire for a high fidelity model (a good fit), followed by a quantitative V&V process (leading to a high QRC and low model assessed failure likelihood). In plant operations, we want even model assessed likelihoods to be of the order 10^n , where “ $n < 0$ ”. For example, $QRC=900$ leads to $n=-1$, $QRC=990$ leads to $n=-2$, etc. Model assessed $QRC > 999$ are obviously desirable in this situation, but may be difficult to obtain in analyses such as the one just shown. While we work on improving our models and V&V confidence bounds, there are two other aspects of the systems engineering cycle that can help as well. One of these is to address the Requirement (e.g. the 20 m/s impact requirement or the 20 °C temperature) used in this analysis. Either of these quantities could be changed in order to raise the model assessed QRC and hence lower the Likelihood term in the model assessed Risk. The other is to mitigate the Consequence of an impact that does achieve threshold.

Some generalized examples of ways we could adjust Requirements to reduce the Likelihood term, or add mitigators to reduce the Consequence of an event exceeding threshold, include the following:

- Lower the desired production (or rework) rate
 - This would reduce Benefit \$B and increase cost \$C in other ways – so again a trade study would be needed
- Modify or delete hazardous operations (e.g. heavy or sharp tooling, etc)
 - Once again this could adversely affect the productivity BCR though it could enhance the safety BCR
- Add mitigating assembly/disassembly compartments, impact absorption barriers,

etc. to mitigate the consequence of either an impact or potential impact

- Trade study on risk and BCR would again be needed.

In doing so, we must realize that concessions on the Requirements side or the addition of mitigation measures for reduced Consequence, may slow down production operations. This may have a poor or negative “BCR” on the production side, because we slow down production rate (reduced benefit \$B), and achieve production at higher unit cost \$C.

CONCLUSION:

On reflection, our Systems Engineering goal remains that of safe environments for our high explosive components and assemblies, both in production and logistics environments. To assure such safety and provide quantitative evidence, we have chosen a Systems Engineering context that includes:

- Requirements Management
- Functional Analysis
 - Model V&V of HE Impact
- Risk Management
 - Terms from Model V&V combine with Requirements to form Likelihood
 - Consequence values generated separately from V&V
- Decision Analysis
 - Benefit / Cost Ratio BCR provides a quantitative metric as input for what remains a qualitative decision process
- Technical Performance Management
 - In the sense of our work, we can ask “Did we meet our V&V goals with a new model; is uncertainty lower? Are reliability, confidence and other metrics higher?”
 - In the production sense, we can ask, “Has our risk assessment and mitigation impacted production efficiency?”
 - In the financial sense, we can ask “Did we achieve the BCR we had hoped for during our developments?”
- Verification & Validation (V&V)
 - Model V&V fits into the middle portion of the Systems Engineering Context
 - System V&V must close the loop over the domain of cost, schedule,

and performance, all with minimal assessed risk.

We can never do as much experimental work as we would like. Furthermore, experimental costs are rising, while numerical model costs decreasing; cost per compute cycle continues a steep downward trend. Because of this, we have chosen to stress the role of numerical models, and the growing role and flexibility of nonlinear 3D finite element models, in the Systems Engineering context.

However, as part of that same Systems Engineering context, we have attempted to show that models of a complex nonlinear physics problem such as explosive reaction threshold impact velocity, can be very misleading unless rigorous and quantitative model V&V is applied. When model V&V is applied to such problems, we have shown how quantitative linkages can be made directly to assessed reliability and confidence of the model of the product, to risk and risk reduction, and to benefit / cost ratio (BCR) for decision trade studies.

ACKNOWLEDGEMENTS:

This work was performed under the auspices of the U.S. Dept of Energy by UC/Lawrence Livermore National Laboratory under contract W-7405-Eng-48. The authors wish to thank Lori Switzer, Kevin Vandersall, Daniel Greenwood, and Craig Tarver for their help in providing experimental results and theoretical assistance enabling the V&V and risk analyses.

NOMENCLATURE:

$\Delta \$B$	Benefit, usually in \$\$\$
BCR	Benefit / Cost Ratio Δ
$\Delta \$C$	Cost, usually in \$\$\$
C	Confidence, a numerical value
FEM	Finite Element Method
K	Tuning Dials; free parameters
M	Margin, where Factor of Safety = $M+1$
μ	population mean
N	Number of trials as in coin-flipping
PDF	Probability Distribution Function
QRC	Quantitative Reliability at Confidence
R	Reliability
σ	population standard deviation (estimate)
U	Uncertainty, General or "System" [in V&V always at a confidence C]
V&V	Verification & Validation
Z	Standard Normal Distribution Variable for variable X, $Z=(X-\mu)/\sigma$

Z_{qrc} QRC analog to Z, where function Z_{qrc} provides R for a chosen PDF, U, C, and M.

REFERENCES:

- ¹ Logan, R.W., Nitta, C.K., and Chidester, S.K., "Risk Reduction as the Product of Model Assessed Reliability, Confidence, and Consequence", UCRL-AR-200703, Nov 2004.
- ² Nitta, C.K. and Logan, R.W., "Qualitative and Quantitative Linkages from V&V to Adequacy, Certification, Risk, and Benefit / Cost Ratio", LLNL UCRL-TR-205809, Sep 2004.
- ³ Ross, H. et al, "Use of Benefit-Cost Analysis to Develop Roadside Safety Policies and Guidelines", Transportation Research Board Circular No. 362, Aug 1990.
- ⁴ Hsu, J.C., "Introduction to Systems Engineering", Tutorial at 42nd AIAA Aerospace Sciences Meeting, Reno, NV, Jan 2004.
- ⁵ Cafeo, J.A. and P.J. Roache, P.J., private communic. of draft V&V definitions, Apr 2002.
- ⁶ Switzer, L.L., Vandersall, K.S., Chidester, S.K., Greenwood, D.W., and Tarver, C.M., "Threshold Studies of Heated HMX-Based Energetic Material Targets Using the Steven Impact Test", 13th APS Topical Conference on Shock Compression of Condensed Matter, Portland, OR, Jul 2003.
- ⁷ Hallquist, J.O., "LS-DYNA Users Manual", LSTC Corp., Livermore, CA, 2003.
- ⁸ DePiero, A.H., "Modeling of Low Impact Testing of High Explosives", Proc. 5th International LS-DYNA Users Conference, Southfield, MI, Sep 1998.