

F-35 LIGHTNING II

Touch Labor Estimate Modeling

Extrapolation Of Actuals

This model up until now was an idea turned reality with significant contributions from John Dorsett, Michael Yeager, Lyle Davis and others subject matter experts. Michael, and John tossed the idea of tying a few of the ideas in this model together, but up until recently the time and resources escaped them. Lyle arrived on the scene and his excitement about the model contributed materially to the writing of this paper. The benefits of its contribution to acquisition cost estimating are currently being proven in supporting negotiations with multi-billion dollar defense companies around the globe.

To support its mission for the F-35, the government production cost estimating team developed touch labor estimate models which include the flexibility to run production rate effects, loss of learning, commonality adjustments, affordability initiatives and outsourcing impacts, multiple learning curve break point analyzes, and estimates for touch labor in hours or by realization. Given the scrutiny to the F-35 Lightning II by the Department of Defense and Congress detailed and accurate cost modeling allow for better budgeting and more credibility within the services themselves and with the American public. Touch labor modeling with the F-35 Lightning II program is naturally more complex than other acquisition programs not only because of the three different variants, but also because of the international reach of foreign countries participating in the program and in regards to the hundreds of suppliers providing parts and the outsourcing of various assemblies for the jet spread across the globe. The touch labor models developed by the government production cost estimating team create a platform from which a standard has been set on which multiple parameters can be adjusted and changed, without losing a credible method to produce estimates that can be easily replicated, repeated, and followed. This paper is an attempt to present the models and illustrates their functionality which allow for traceable and defensible outputs derived from extrapolation of actuals.

Introduction

The F-35 Lightning II is a 5th generation multi-role fighter manufactured by Lockheed Martin for the Navy, Air Force, Marines and other US allies. Three different variants are being manufactured. The Carrier Variant (CV) for the Navy, the Conventional Take-Off and Landing (CTOL) for the Air Force and the Short Take-Off and Vertical Landing (STOVL) for the Marine Corps. Lockheed Martin's website states, "Aircraft share common parts, support equipment and technical data, yielding greater reliability, and long-term cost savings. And the F-35's strong global partnership and broad industrial base ensures affordability through economies of scale" In this light, the F-35 Lightning II government production cost estimating team intends to support Lockheed Martin in assuring the cost estimates in support of acquiring more jets are as accurate and thorough as possible to ensure eventual long term savings can be achieved. In particular, this paper will address touch labor estimate modeling for the F-35 and the ability of the models to estimate touch labor hours for future production lots incorporated with best fit commonality and standards relationships analyzes.

These models were developed in response to a need to create a standard platform for touch labor estimating from which to maintain continuity in process and logic while allowing for inclusion or exclusion of the various parameters at play specific to the program. These include production rate and loss of learning effects, commonality assumption analysis, affordability initiative and outsourcing impacts, multiple learning curve break point analysis, and estimates for touch labor in hours or by realization.

Use of the models has allowed for convenient, easy to use input options for quick-turnaround drills as well as creating both visual and data driven output for defensible presentations appropriate for the more complex and time consuming estimates.

It should also be mentioned that use of the type of model mentioned here is most appropriate for programs that are past the early stages of development and where actual production data exists in amounts large enough to reasonably draw logical conclusions from their analysis. Any program that does not fit within these parameters would be an unlikely fit for appropriate use with such a model. The use of the model mentioned here is with a program which includes several large past production lots and additional lots currently in production.

Model

The general purpose of the model described here is to predict future touch labor hours for each aircraft by variant in support of contract negotiations for future lot purchases. The functionality of the model can be conceptualized into a few overarching principles.

First, production rate impacts the estimate when coupled with the traditional $Y=Ax^b$ learning curve equation. Instead of $Y=Ax^b$ the equation becomes $Y=Ax^{b*Q^r}$ when production rate is included, where an increase in the production rate may translate into a lower estimate of future hours.

Second, lost learning occurs when a gap in the production line occurs. The gap could theoretically affect all three variants or as seen in the recent past be specific to the variants; the wider the gap the greater the effect of lost learning. When a production gap occurs and learning is lost the cumulative unit number is decreased. More on cumulative unit numbers will be discussed later in the discussion of lost learning.

Third, commonality exists where parts between variants share certain features assumed to increase learning to some degree between one or more of the variants in the assembly process. The effect of having common parts is, for example, similar improvement for STOV and CV when a CTOL is assembled when using a STOV, CV, CTOL common part in the assembly process.

Fourth, the goal of an affordability initiative is to reduce the number of hours needed to manufacture an aircraft over time. Affordability initiatives result in reduced hours for jets in the production line starting where the break-in of the initiative occurs and all the following jets manufactured after the break-in into the production line.

Fifth, the prime contractor and its major sub-contractors have a limited capacity to manufacture jets. While in some cases capacity has been increased, as the rate of production begins to significantly increase more of the work must be outsourced. As this continues to happen the effect of learning must follow where the work is performed. This is done by tracking what is being done in-house by the major sub-contractors involved and what is being outsourced. The effect of this outsourcing has resulted in a supply chain for the F-35 aircraft reaching across the globe with many international suppliers. In turn, accurately tracking where learning is actually taking place and where it is not because of outsourcing impacts the rate of the learning. The effect of offloading and its effect on learning are captured in the model presented here.

Sixth, using multiple learning curves allows more accurate modeling of contractor actual performance and creates a clearer picture from which to derive the estimate in the future. This model allows for use of up to four different learning curves. Using multiple learning curves to model actual performance comes into play when significant improvements occur in the production line, or perhaps a strike occurs; one steepening and the other flattening the learning curve. Modeling with multiple curves creates more accuracy and reflects a greater measure of

reality in reflecting past performance and assists in projecting future hours in the estimate. While past performance does not guarantee future performance it does assist in creating a greater measure of fidelity when projecting into the future.

Seventh, the model allows for simple conversion from hours into realization. Where a contractor may speak in terms of realization rather than hours the model allows for a quick conversion to speak on the same terms.

The most basic input into the model is the touch labor actuals and percent complete by tail number received from the contractor. Touch labor actuals represent the sum of actual touch labor hours spent assembling a specific jet as recorded by the contractor. The fidelity of the estimate is dependent upon receiving actuals which are accurate and complete. Also, worth noting is a discussion of standard work expected and planned to occur on a given jet. Standard work is essentially planned work on the jet and the given amount of time to accomplish that work. Non-standard work in theory captures all work that is not standard. This could include time spent repairing damaged parts, scrap, re-working errors or any amount of hours spent on unexpected tasks in the manufacturing process. Non-standard work is normal part of manufacturing; however, clarifying with the contractor how to handle and track the non-standard work is important. Having a detailed discussion about what is included in the touch labor actuals is vital to creating an accurate estimate. Where non-standard work cannot be captured accurately with the standard work it may be best to analyze the non-standard work separately. In terms of the discussion here it is assumed touch labor actuals capture all standard and non-standard work and when a given unit is deemed to be 100% percent complete by the contractor no hours will be added to that given unit from that point on. Be it standard or non-standard work. That said with the touch labor actuals loaded into the model the data can then be normalized.

Normalization of Actuals

To regress touch labor actuals the data must be normalized in order to make any comparisons. This process is referred to as normalization. Since not all units may be complete when actuals are taken the data goes through the normalization process to make them comparable. The equation used in this model is as follows:

$$\text{Normalized EAC} = \text{Actual} + [((\text{Actual}/\text{Percent Complete}) - \text{Actual}) * \text{Efficiency Factor}]$$

The formula makes the data comparable for units not 100% complete by estimating the amount of remaining hours it would take to reach 100% complete. The formula takes into consideration the amount of unfinished work and mathematically calculates the number of hours estimated at completion (*Actuals/Percent Complete*). Once all the actuals are run through the

normalization formula they become estimates at completion (EAC) and are comparable with the units already 100% complete. It is important to note actuals are only comparable to actuals for same type of assembly work and are not comparable to non-similar assemblies or sub-assemblies. In addition, the efficiency factor in the formula allows a portion of the EAC to be adjusted for any efficiencies/inefficiencies that may occur during assembly as it nears completion. In particular the efficiency factor allows the unfinished amount of a particular assembly to be adjusted. If, for example, in a particular assembly the last ten percent of work is more difficult than the first ninety percent the efficiency factor allows the unfinished portion to be adjusted upwards to account for the anomaly. It is assumed actuals for units one hundred percent complete reflect all efficiencies/inefficiencies in them. It is worth noting the efficiency factor for this model is not designed to capture non-standard work. The efficiency factor is meant to increase the math EAC where it seems fair and reasonable and has historical precedence that is justifiable.

The topic of percent complete also includes a discussion of the minimum percent complete acceptable in the eyes of the contractor and the government. The GAO Cost Estimating and Assessment Guide suggests actuals used in a learning curve regression should be used at a minimum level of percent completion in order to maintain consistency and a greater degree of fidelity in the estimate. The risk of inaccuracy lies in normalizing an actual to such a large degree that the math estimate at completion (EAC) has become too large or too small a portion of the normalization rather what might occur in reality. The GAO Cost Estimating and Assessment Guide recommends asking the question, 'Was the cost element at least 25% complete before using its data as an extrapolation?', to check for accuracy in estimating techniques. In reality, the contractor and the government program office should work together on what percent complete is acceptable in order to make apples to apples comparisons when regressing actual hours data. In general, within the F-35 program actuals hours data by unit are not included in regression analysis unless they are at a very high percent completion. It may be appropriate to ask such questions as: what is the justification for assuming the last xth percent of the work may include the most hours and complexity, would it be accurate to adjust the math EAC the same as the completed percent of the work, or should the remaining percent in the math EAC include an element that increases the estimated hours? Answering these questions will help in forming a consensus between the contractor and government as estimates are put together. Once the touch labor hours data has been normalized, and a minimum percent complete standard is set the data is included as part of a learning curve regression.

The model described here utilizes part of the ACE IT software estimating package created by Tecelote Research, Inc. called COSTAT. An example of the output of a regression from COSTAT is shown here:

I. Model Form and Equation Table	
Number of Observations Used:	96
Equation in Unit Space:	$UNIT_COST = 6.897e+004 * UNIT_NUM ^ (-0.2725) * 1.006 ^ {cdum} * 1.039 ^ {sdum}$
T1:	68974.8247
Quantity Slope:	82.79%
Dependent Variable:	EAC
Quantity Variable:	First Unit and Last Unit Variables

Shown above in blue is an example of the learning curve equation, the theoretical first unit or T1, and the quantity slope or the rate at which learning is occurring. The learning curve equation above is derived from an example of touch labor hours data loaded into the model. In the output, the T1 above of 68,974.8247 hours represents the CV. The CTOL T1 is calculated as 68,974.8247 x 1.006, which in this example makes it higher than the CV T1. The STOVL T1 would be calculated in the same manner and is 3.9% higher than the CV T1. For this model, the relationship between the T1's in the variants is derived from the actual performance of the contractor as represented in the touch labor hours data and leads to a discussion of standards.

Standards can be defined as the theoretical amount of hours it would take to accomplish a set task under perfect conditions. Standards are created as a baseline from which to measure and are not necessarily intended to be achieved. However, standards capture what the contractor believes are the differences in the work between the variants. For example, the STOVL lift fan makes it more complex than the other variants and requires more work. Other differences exist between the CTOL and the CV. These differences in work are represented in how the standards are derived by the contractor. Standards represent the amount of touch hours, under perfect circumstances, the aircraft could be assembled and also represent the differences in work content by variant that must be performed to assemble the aircraft. For example, in the table below column 3 represents the contractor proposed amount of standards to assemble each variant. The proposed relationship between the variants is in column 4. The regression equation above defined the relationship differently. In the example below, the government accepts the relationship in column 2, based on past performance, and would project the CTOL as needing more hours to assemble than the contractor, and fewer hours for the STOVL. The difference being the government's view reflects the touch labor hours data, whereas, the contractor position reflects what they propose the relationship is based on how they define the standards. However, for analysis, testing the relationship proposed by the contractor between the standards, while not excluding them, allows the model to reflect what actual performance would define as the relationship between the standards as shown in

columns 1 and 2. This allows for more clarity and fidelity in looking at past performance and theoretically more accuracy for projecting into the future.

	1	2	3	4
	Standards	Gov't	Standards	Contractor
CTOL	1,006	1.006	1,001	1.001
STOVL	1,039	1.039	1,050	1.050
CV	1,000	1.0	1,000	1.000

The COSTAT table below shows the goodness of fit or adjusted R² for the actual hours data loaded into the model.

Goodness-of-Fit Statistics			
Std Error (SE)	R-Squared	R-Squared (Adj)	Pearson's Corr Coef
0.0704	88.66%	88.03%	0.9416

Matching the model and the COSTAT software package allows for a more detailed analysis of the actual hours data than may be typical. The analysis challenges contractor assumptions and should create an opportunity for the government to be more aggressive in cases where significant differences between estimates exist.

The second basic input into the model needed to derive a learning curve is the unit number. The unit number is dependent upon the build profile, commonality among the different variants, lost learning impacts, ground test articles, and any outsourcing of parts and will be discussed more later on. In short, the model tests the commonality as proposed by the contractor, while not excluding it as the best fit, and 17 other different sets of commonality. The best fit is defined as the highest adjusted R² and in affect determines what set of commonality assumptions will be used for projecting future units.

Once the touch labor actuals hours data by unit and unit numbers are loaded into the model and the regression output has been calculated the model is then used to predict the amount of future touch hours by variant. The COSTAT analysis and the adjustments made to the model affect the predicted future hours based on learning curve percentages, commonality assumptions, and T1's.

In the illustration shown above the areas in yellow represent where adjustments in the model would be made after the learning curve regression has been performed. Credit for ground test articles, the efficiency factor, and that minimum percent complete are set before the regressions are run. The following paragraphs will explain the illustration and how it affects the model output.

Production Rate Effects

The models allow the predicted values calculation, $Y=AX^b$, to include the effects of production rate, $Y=AX^b * Q^r$. The effects of a production rate increase will reduce the number of hours predicted by the learning curve formula $Y=AX^b Q^r$ where Q is the quantity to be produced, for the F-35 by lot, and r the rate coefficient. The rate coefficient, r , is calculated as:

$$LN (Rate Curve)/LN(2)$$

In this case r will always be negative and Q always a positive number. Q^r will always be ≤ 1 ; thus, Q^r will always reduce the result of $Y=AX^b$. The one exception being if the cost estimator decides not to include the effects of the production rate; in which case they would set r to 0. Since $Q^0=1$ then $Y=AX^b$ would be multiplied by 1 and eliminate any effects of production rate. The underlying assumption with production rate effects is that an increase in the production rate will decrease costs. Production rate increases are adjusted in the model underneath the four boxes for the learning curves.

Lost learning

According to the GAO Cost Estimating and Assessment Guide lost learning can occur due to personnel learning, supervisory learning, continuity of production, methods, and special tooling. To keep things simple, the output from lost learning calculations result in a setback in the unit number for the variants affected. This model includes the capability to include the setback in the unit numbers. The input screen shown above is modeled to handle a setback for the COTL, STOV, and CV, but can easily include additional setbacks as needed.

Where a setback has occurred the effects of the lost learning decrease the cumulative unit number and increase the number of hours or cost. The model described here allows adjustments for lost learning by variant and incorporates the loss of learning in the regressions. Lost learning is not entirely uncommon in the aircraft industry; however, when lost learning is proposed by a contractor the government should look for detailed and justifiable analysis of the lost learning effects on production.

Commonality

The idea of commonality for the F-35 suggests some learning can be taken from the variant being manufactured and applied to learning for the other variants. For example, in the table below (reading the table from the left to right) when a CTOL is manufactured, the STOVL receives 30% of a unit, and the CV 33% of unit. When a STOVL is manufactured the CTOL receives 31% of a unit, and the CV 32% and so on.

	CTOL	STOVL	CV
CTOL	1.000	0.300	0.330
STOVL	0.310	1.000	0.320
CV	0.400	0.390	1.000

The values in this table are for illustration purposes only and do not represent actual commonality for the F-35 variants.

In cases where a portion of the work for a unit is outsourced the outsourced work will not increase the cumulative unit number for the part of the work which was performed separately. The model is able to capture what portion of work is done in-house by the contractor being analyzed and what portion of the work is not in-house and calculates the effect of the outsourcing in the unit numbers.

A basic level of commonality explanation suffices for the purposes here. The more detailed explanation of commonality calculated in the model is appropriate at some point, but here where a basic introduction to the model is the purpose a more detailed discussion will be reserved for later discussion.

Several steps are involved to calculate the unit number for each sequential aircraft. Each variant has a separate calculation for unit numbers. For example, by the production of the 20th jet, there could be a mix of CTOL, STOVL, and CV's. A common cumulative unit number must be used to accurately reflect the commonality shared between three different variants. The common cumulative unit number is weighted based on the unit number for each particular variant as follows and the percentage of commonality depends on which variant is being built:

$$[CTOL \text{ unit number} * \% \text{ commonality} + STOVL \text{ unit number} * \% \text{ commonality} + CV \text{ unit number} * \% \text{ commonality}]$$

The formula above represents a combination of several steps taking not only the commonality between all three variants, but also the commonality between the different combinations of just two variants where there may be no commonality between the third, and also the uniqueness of each variant into consideration. A simplistic theoretical explanation

would be as follows: the 15th aircraft is a CTOL, the seventh CTOL produced out of 15. Of the fifteen seven are CTOL, four STOVL, and four CV's. The common cumulative unit number for the 15th unit would look as follows:

$$[7*(100 \% \text{ commonality})+4*(X\% \text{ commonality})+4*(X\% \text{ commonality})]$$

The use of COSTAT in the models analyzes the contractor's assumptions of commonality. COSTAT analyzes the contract commonality assumptions as well as a set range of other iterations of commonality built into the model. The coefficient of determination (r^2) for the set range of commonality assumptions from the model are then compared against each other selecting the regression equation with the highest r^2 .

In cases where this model has been recently used the greater the commonality in existence between the variants, as represented by actual performance, the lower the estimate will be and vice versa. Also, there may be a difference between commonality in existence between the variants in the aircraft design, which in the case of the F-35 is relatively stable at this point, and the commonality as represented by the touch labor actuals. In other words, actual performance may indicate less commonality than the commonality in existence between the variants in the aircraft design itself. For example, repairs needed for a damaged lift fan on a few STOVL jets may create more hours of work than expected. However, the extra hours for the repair work do not represent a change in the aircraft design of the STOVL making it less common, but may represent less commonality in the actual touch labor hours for the STOVL's affected by the repair. In this case, the regression of the actuals may show less commonality for the STOVL, but accurately represents what has occurred in the factory.

Rather than accepting the proposed commonality by the contractor for running regressions, this model tests 17 additional sets of commonality in an attempt to find the best fit for the actual performance data. It does not however, exclude the commonality proposed by the contractor. Where the proposed commonality by the contractor represents the best fit of the touch labor actuals it would be accepted as the best fit for predicting future values.

Affordability Initiatives

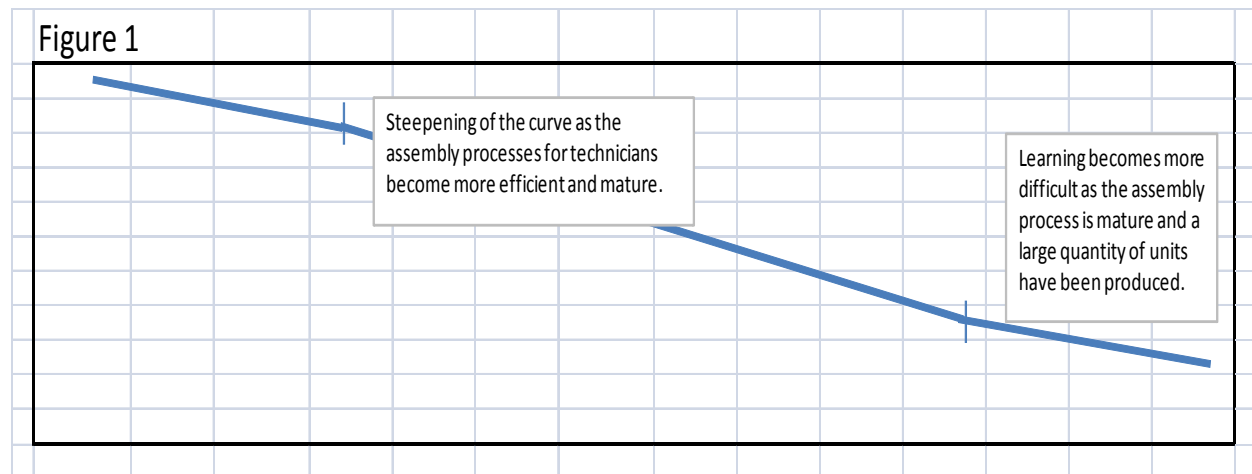
Special initiative projects to reduce touch labor over time in specific areas are also a consideration in the model. These projects where implemented by the contractor show up in the model as credits allotted by aircraft and function as an adjustment to reduce the actuals after they are normalized. The model is able to run regressions with or without the consideration of the special initiatives credits given over time. The intent of such initiatives is to reduce the overall cost of manufacturing the jet over time making more affordable to procure.

Outsourcing Impacts

As mentioned throughout the paper outsourcing or offloading work will continue to increase as the quantity of jets begins to significantly increase. Accurately capturing what is outsourced can be very complex and difficult to model. However, the model here recognizes when an entire unit or a portion of a unit has been offloaded; thereby, making it possible to accurately reflect the work being done in-house by the contractor.

Learning Curve Breakpoints Analysis

In general terms, a learning curve occurs at a constant rate and is considered as a whole from the beginning of production until the end. A breakpoint analysis takes a closer look at the slope of the curve for a given set of consecutively occurring touch labor actuals in the production sequence to see if a break has occurred or in other words if learning is occurring at a higher rate than the previous units. It assumes there may be sections of the curve where a break occurs and the slope becomes either steeper or flatter. For example, after the development of x number of units a regression analysis may show a section of the curve seems significantly steeper than the beginning portion of the curve and then may flatten after a large number of units have been produced. A common learning curve break pattern in the military aircraft world is in the form of a backwards S as illustrated in Figure 1. The few initial aircraft experience a seemingly flat curve, and at some point begin to move down a steeper curve, and then again flattens over time after a large number of aircraft have been produced.



The sequential learning output from COSTAT can be used to better understand exactly where the curve may be breaking and the slope of curve.

It seems prudent to mention the importance of correctly applying regression and learning curve theory. The output from regressing a set of touch labor hours as discussed in this paper will produce T1's for each variant, the adjusted r^2 , the slope, and other various statistics. Once the

T1's have been calculated from regressing the touch labor data learning curve theory does not allow arbitrarily changing or ignoring the T1's associated with that data. Ignoring or arbitrarily changing the T1's or the learning curve percentage while keeping the T1's would be considered outside of learning curve theory and would introduce infidelity into the modeling process.

As mentioned earlier, the learning curve equation output form COSTAT includes a variable for the CTOL, STOVL, and CV. These variables make it possible for the model to use the T1's associated with each variant to more accurately predict what hours will look like in the future. The relationship between the variants as included in the output from COSTAT (shown below) are derived from the actual performance data from the contractor. This relationship between the variants is partially representative of the difference in work content as well.

I. Model Form and Equation Table	
Number of Observations Used:	96
Equation in Unit Space:	$UNIT_COST = 6.897e+004 * UNIT_NUM^{(-0.2725)} * 1.006^{c_{dum}} * 1.039^{s_{dum}}$
T1:	68974.8247
Quantity Slope:	82.79%
Dependent Variable:	EAC
Quantity Variable:	First Unit and Last Unit Variables

Work content or the difference in the aircraft design between the variants that creates differences in the work of assembling a CTOL versus a STOVL versus a CTOL can be characterized by standards and commonality. This model allows the touch labor actuals to define the effects of work content by variant. Differences in modeling has shown some contractor models constrain actual work content by accepting standards and commonality as proposed when considering regressions for estimating rather than accepting how the actual performance, based on touch labor hours data, defines the relationship between the standards and commonality. It becomes a question of whether work content is defined by actual performance of the contractor in estimating based on regressions, or looking at work content based on a relationship a contractor proposed is true, but where actual performance may represent something entirely different than what the contractor proposes is true. This model rests on the side of the fence where running regressions to predict future hours based off of touch labor actual performance data may be more accurate in representing past performance and projecting into the future than otherwise.

Hours or realization

Realization in its simplest form is derived from a simple math equation and relates to the amount of actual hours taken to assemble a part of the aircraft and the amount of standards hours associated with it. Realization is defined as:

$$\text{Actual hours for a set task} / \text{amount of standards hours for the set task}$$

Realization above 1 explains the amount of imperfection existing above the standard hours to perform a task. As more and more learning occurs the realization should move closer and closer to 1. For clarity, this model does not run regressions using realization, but regresses touch labor actual hours. The feature in the model mentioned here is mentioned only in terms of converting a predicted hours position back into realization terms in order to be able speak in the same terms where the contractor may be using realization.

Projection of Future Hours

The projection of future hours in this model is based on the equation $Y=AX^b Q^r$. It uses touch labor hours data to create one or more learning curves to represent as accurately as possible what performance has been in the past and is used to project what future hours may be based on regressions from touch labor actual hours data. The regression output from COSTAT provides a learning curve equation derived from a set of touch labor actual hours data along with separate T1's for each variant and may be used to derive multiple curves and multiple T1's for each variant. It defines the relationships between standards and commonality both defined based on the actual hours performance data from the contractor.

Conclusion

The F-35 Lightning II program office has developed touch labor models that allow for production rate effects, loss of learning, commonality adjustments, affordability initiatives and outsourcing impacts, learning curve break points analysis, and estimating touch labor by hours or realization. Part of the beauty of this model is the ease in which it can be customized for multiple contractors working to support the F-35 Lightning II program. It provides a new platform from which dependable, justifiable, and traceable estimates can be derived; challenges and tests relationships for standards and commonality as proposed by the contractor and includes a much broader perspective than typical by providing a best fit solution.

Works Cited

1. *F-35 Lightning II*. n.d. 9 October 2012. <<http://www.lockheedmartin.com/us/products/f35.html>>.
2. United States Government Accountability Office. "GAO Cost Estimating And Assessment Guide." March 2009. *U.S. GAO*. 16 November 2012.