Ship Construction Estimates at Completion: A New Technique Using the Weibull Function

Shawn Rudolph

9 June 2010

Presented at the 2010 ISPA/SCEA Joint Annual Conference and Training Workshop - www.iceaaonline.com

Technomics

Outline

A. Introduction

B. EVM Basics

C. Weibull Function Basics

D. Weibull-based Estimate at Completion (EAC) Model

E. Future Work

Introduction

A new method for generating EACs in labor hours.

The characteristics of the method are:

1. Applies to ship construction programs, specifically, follow-ships
   - "Work-arounds" for lead ships
   - Might be applicable to other large-scale commodities
2. Recognizes four, observable construction phases
3. Uses the Weibull function for three of the four phases
4. Organizes historical phase data into logical groupings:
   - Own-ship
   - Own-class
   - Own-shipyard
   - Own-flight
   - Own-contract
5. Uses the historical data to provide constraints
   - Forecasts at the end of each phase
   - Maximum or minimum values within each phase
6. Does not rely on the planned work
7. Provides sensitivity to schedule changes

Four Phases of Ship Construction

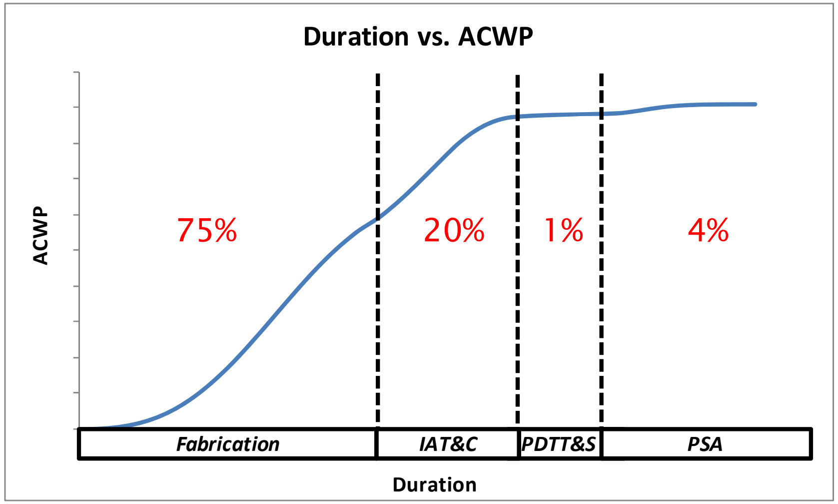
For ship construction programs, we recognize four phases:

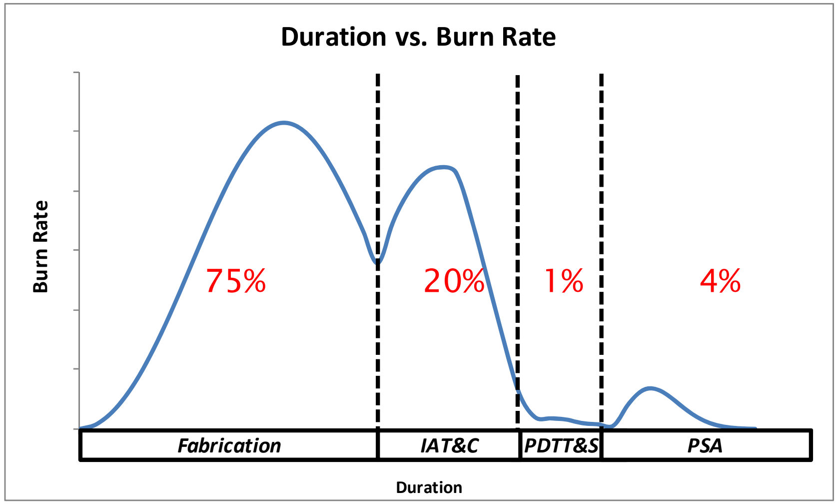
1. Fabrication,
2. Integration, Assembly, Test, & Check-Out (IAT&C),
3. Post Delivery Test and Trials and Shakedown (PDTT&S), and
4. Post Shakedown Availability (PSA).

In our model, we estimate the hours for each phase; the sum of the estimated labor hours = EAC.

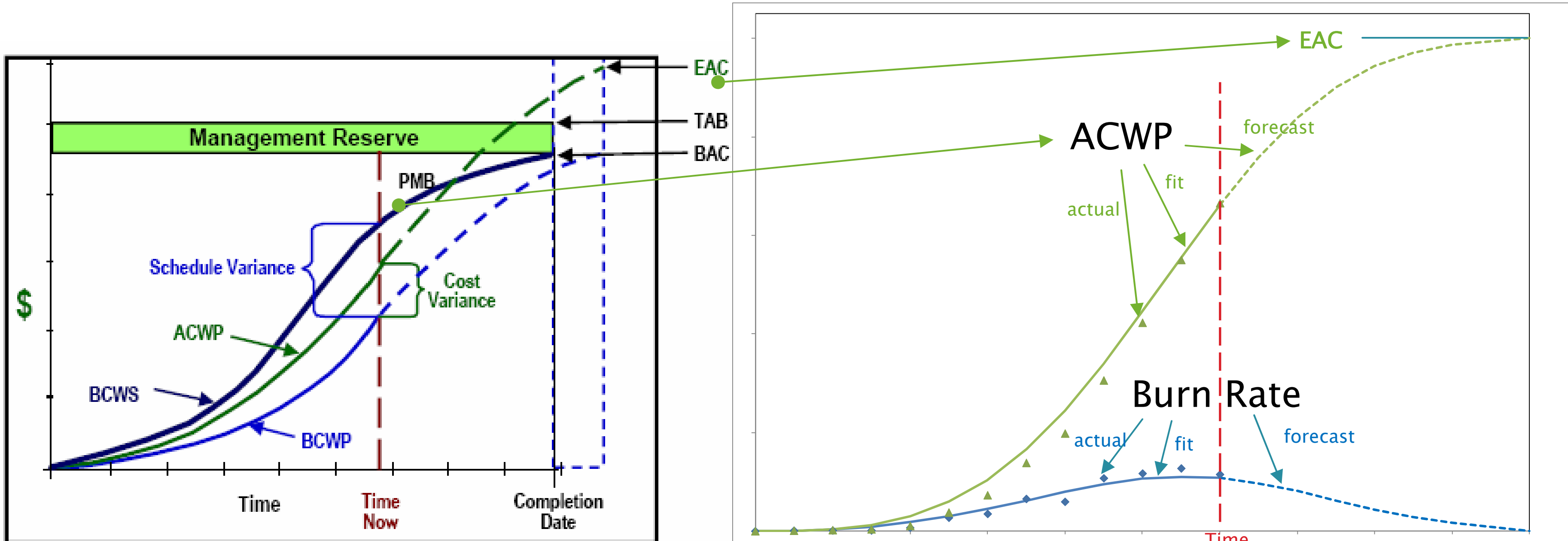
Before I move onto the modeling technique, two items need to be explained:

1. Burn Rate and ACWP
2. Weibull function





EVM Basics: Burn Rate and ACWP



Defense Acquisition University Gold Card

Burn Rate (BR), labor hours per month, is the first difference of ACWP

$$BR(t) = ACWP(t) - ACWP(t-1)$$

Actuals for BR can be "fit" to a functional form via regression.
The resulting model can be used to forecast future periods.

ACWP, labor hours, is the cumulative of Burn Rate

$$ACWP(t+1) = ACWP(t) + BR(t+1)$$

Accumulated BR forecasts through completion, added to ACWP , yield an EAC

Weibull Function Basics

1. Accumulated expenditures over time (ACWP) typically follow an "S-curve" for most programs
   - Therefore phased expenditures (burn rates) typically follow a "hump shaped" profile, possibly skewed
2. A number of probability distributions have these same general properties
   - So probability distributions provide useful functional forms for mathematically modeling phased expenditures
3. The Rayleigh distribution has been applied to phased expenditure modeling for many years
   - Rayleigh distribution has fixed skew (relative position of mode)
4. The Weibull distribution is a more flexible generalization of the Rayleigh that has come into use more recently
   - Weibull distribution has variable skew

Weibull & Rayleigh Distributions

RAYLEIGH PDF

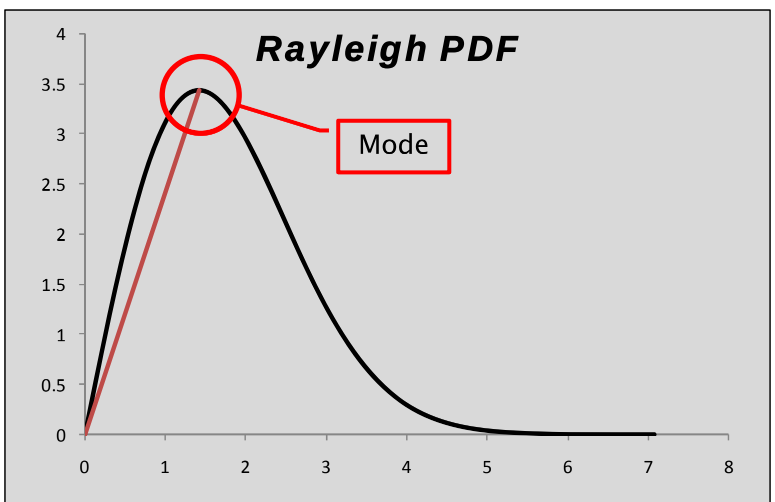
1. One Unknown Parameter Distribution
   - Scale Parameter
   - Shape Parameter = 2

$$F(x) = 1 - \exp\left(-\left(\frac{x}{b}\right)^2\right)$$

$$f(x) = \left(\frac{x}{b}\right) * \exp\left(-\left(\frac{x}{b}\right)^2\right)$$

2. Has no inflection point prior to mode
   - Slow ramp ups are modeled poorly
   - For all distributions, the red line (straight line from 0 to mode) does not intersect the black line on the open interval (0, mode)

3. The area underneath of the curve from 0 to the mode is always 0.393



WEIBULL PDF

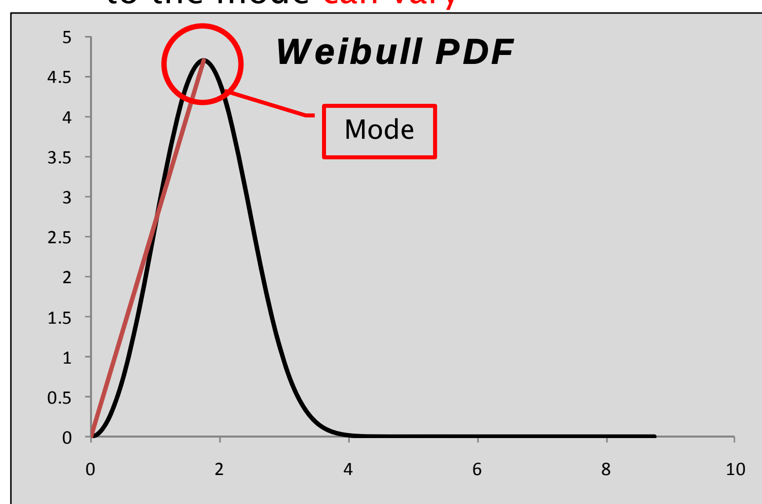
1. Two Unknown Parameter Distribution
   - Scale Parameter
   - Shape Parameter

$$F(x) = 1 - \exp\left(-\left(\frac{x}{b}\right)^a\right)$$

$$f(x) = \left(\frac{x}{b}\right)^{a-1} * \exp\left(-\left(\frac{x}{b}\right)^a\right)$$

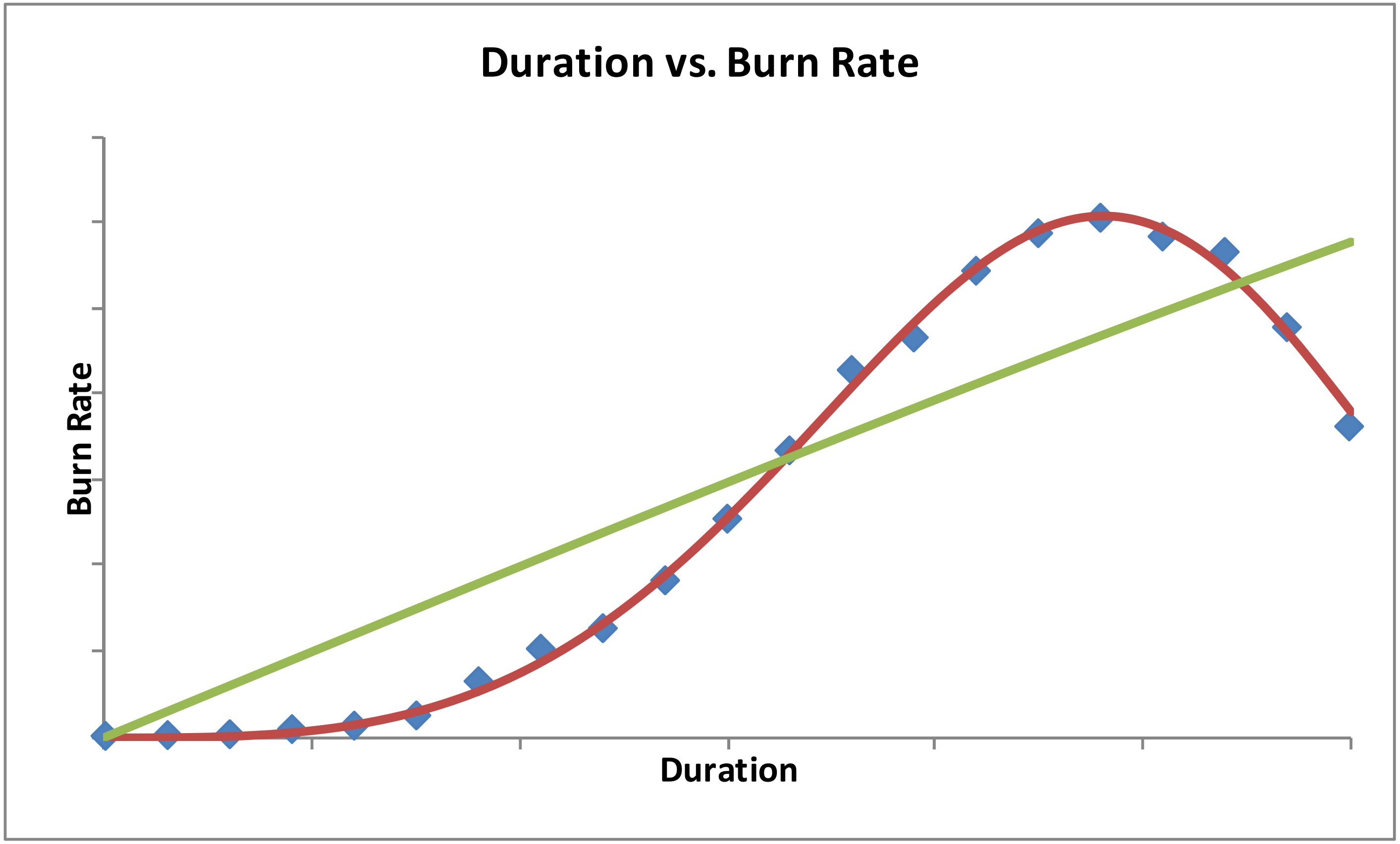
2. Can have inflection point prior to mode
   - Slow ramp up of burn rates are modeled well
   - There exist distributions such that the red line (straight line from 0 to mode) does intersect the black line on the open interval (0, mode)

3. The area underneath of the curve from 0 to the mode can vary



Weibull vs. Rayleigh

The Rayleigh (green) is not a good fit for the actual burn rates (blue). The Weibull (orange) provides a better fit for hulls that exhibit a slow ramp up of hours expended.



Rayleigh Regression

Weibull Regression

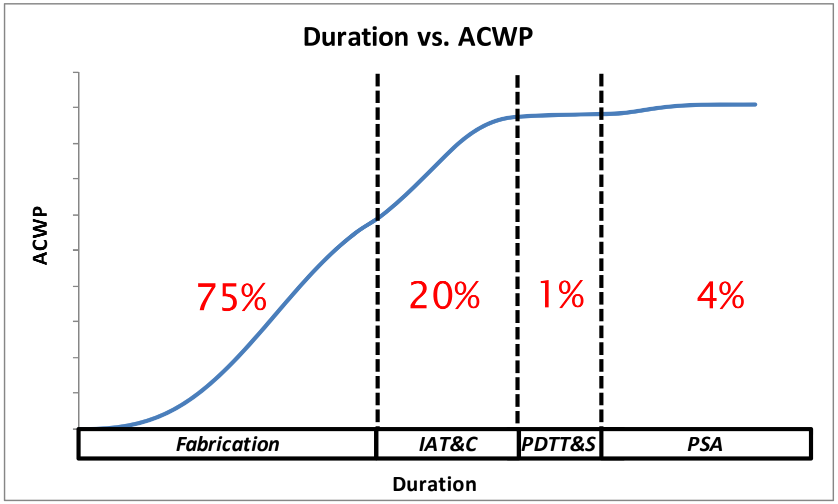
Weibull-based EAC Model

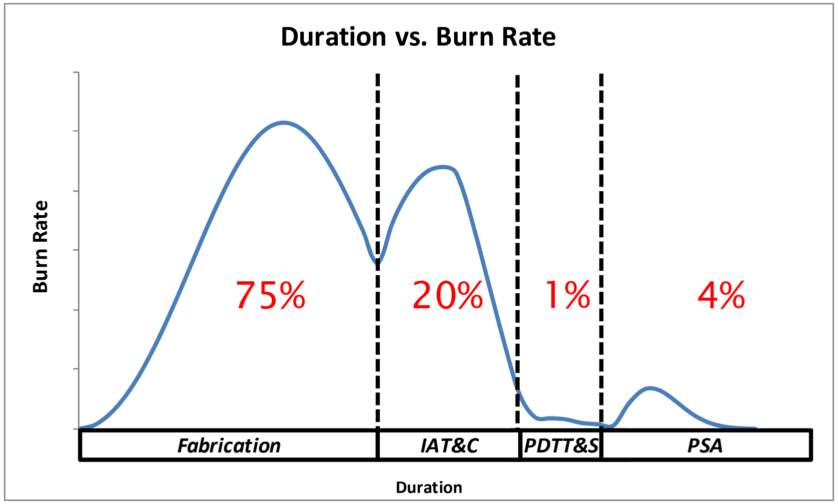
Our four phases of ship construction are:

1. Fabrication,
2. Integration, Assembly, Test, & Check-Out (IAT&C),
3. Post Delivery Test and Trials and Shakedown (PDTT&S), and
4. Post Shakedown Availability (PSA).

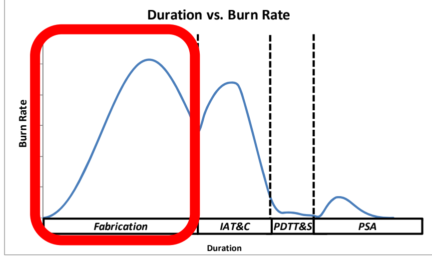
In the next several slides, each phase is discussed.

We also have identified "lack-of-data" as a significant issue; a subsequent section discusses this problem.



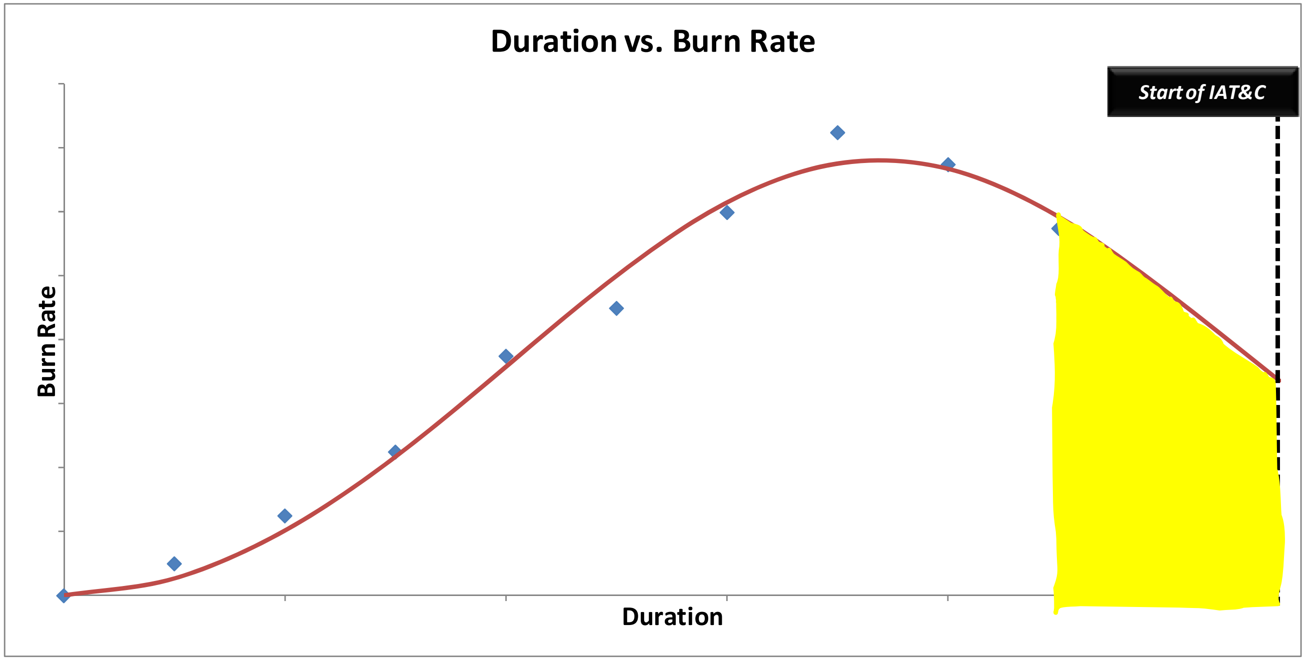


Fabrication Phase



General Properties:

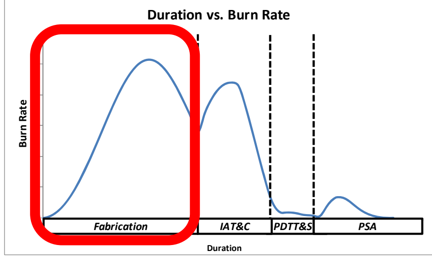
- Consists of hours from Start of Construction to the Start of Integration, Assembly, Test & Checkout
- Accounts for ~75% of total hours expended during ship construction



Estimation Process:

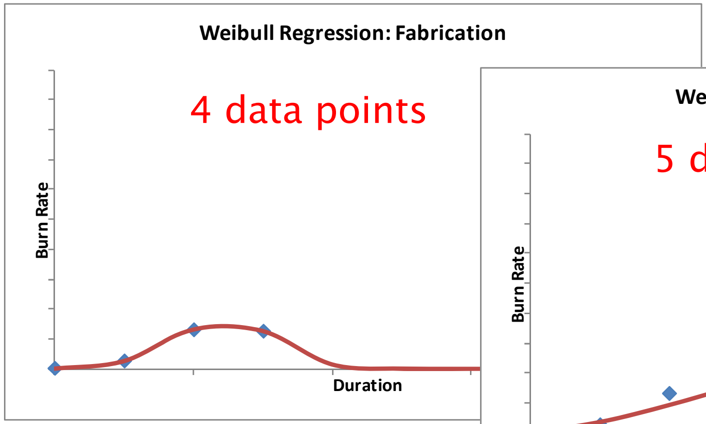
- Estimated by fitting a Weibull probability density function to burn rate data
- The estimated ACWP expended during Construction is the ACWP to date + the area of the yellow region

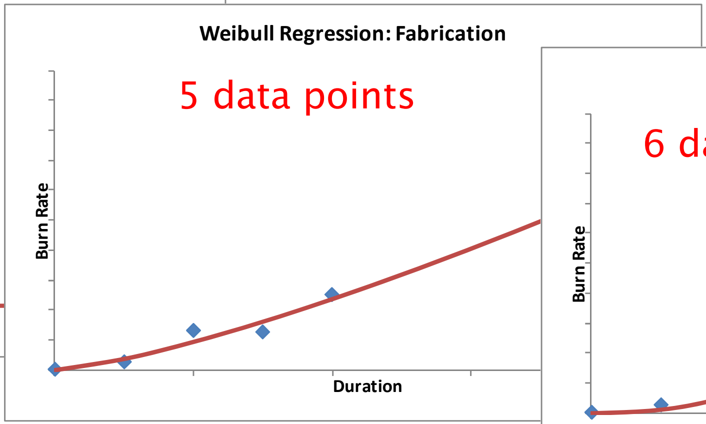
Fabrication Phase Issue

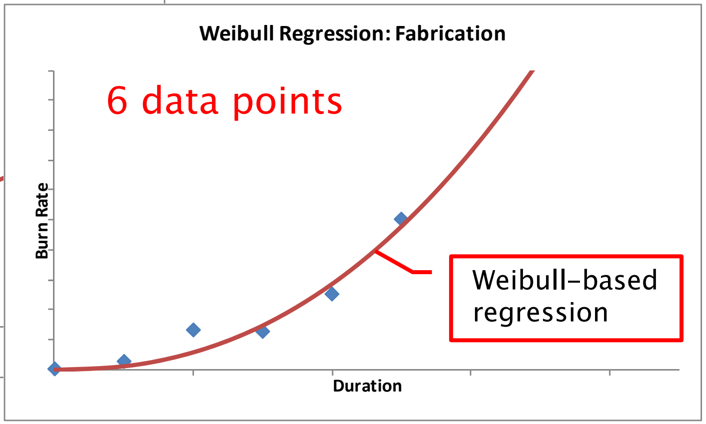


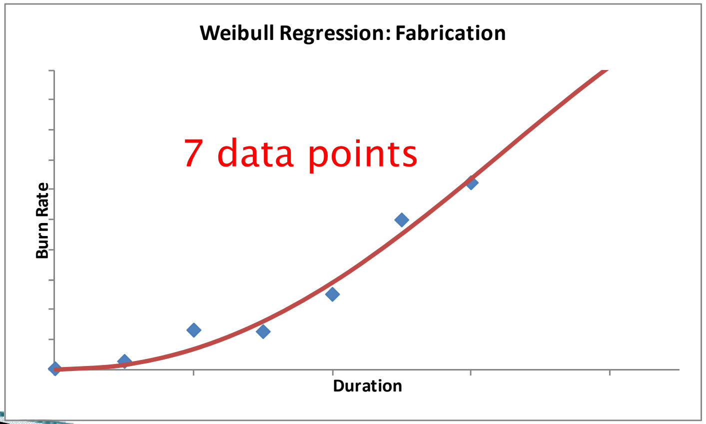
Issue Statement:

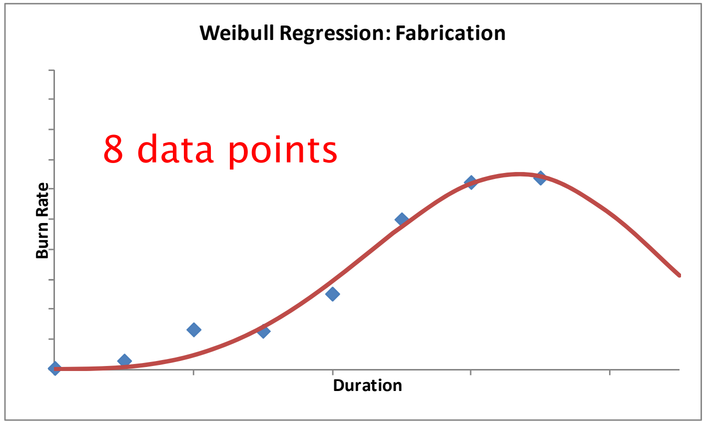
- When insufficient data exist to define the mode, Weibull-based estimates can be unstable
- Weibull PDF stabilizes once the Mode location is established



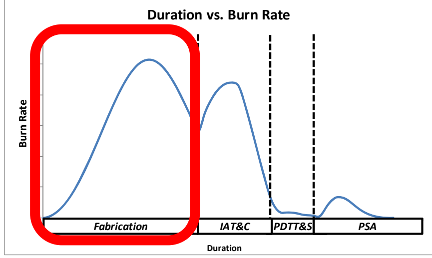






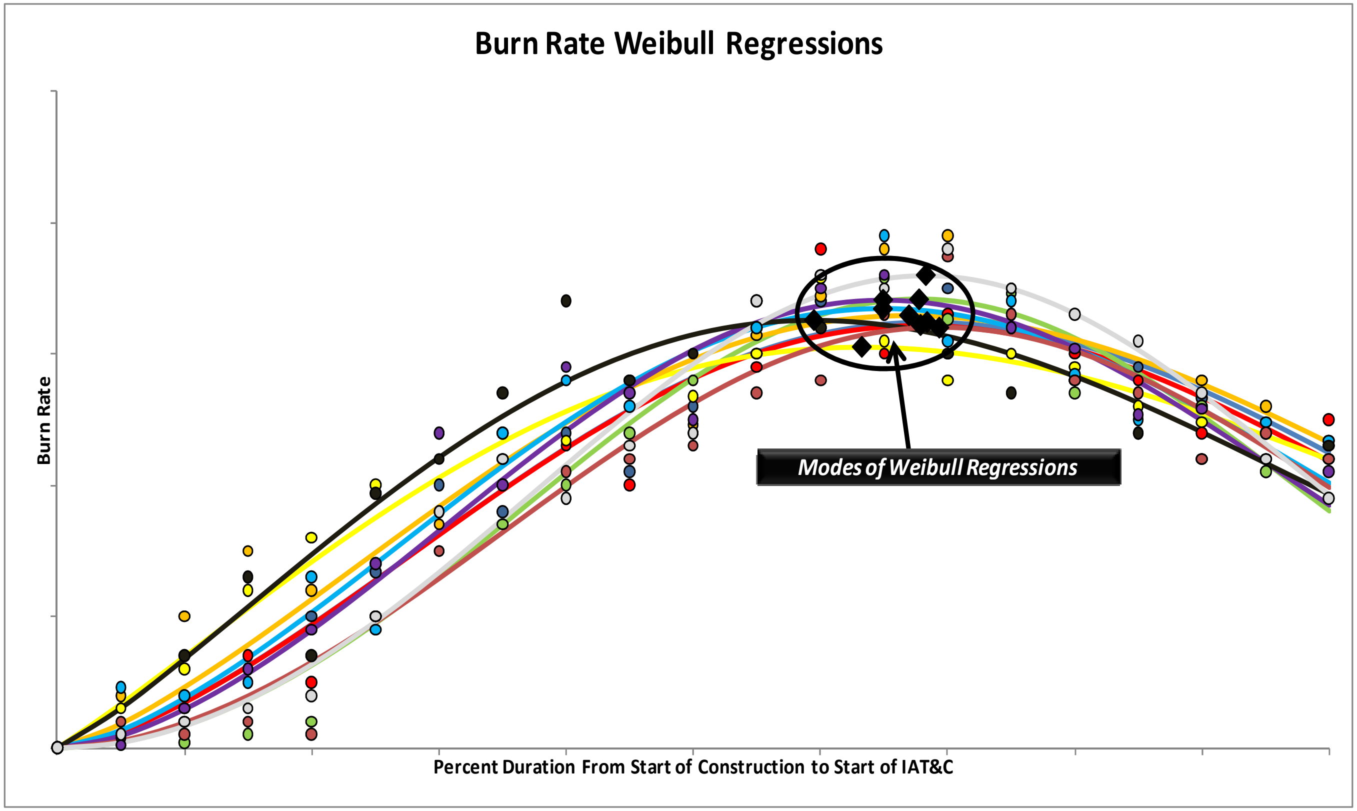


Fabrication Phase Issue

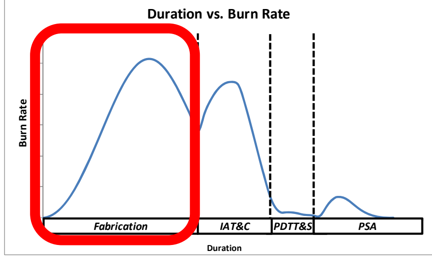


Solution:

- We observe that for all members of the class, the peaks of the Weibull PDFs exhibits similar behavior
  - The modes of all "own-class" Weibull PDFs occur within a tight band (black circle)
- We calculate the values of the mean and standard deviation for the peak Burn Rate and the % Duration
- We can use this information to establish constraints for the Mode

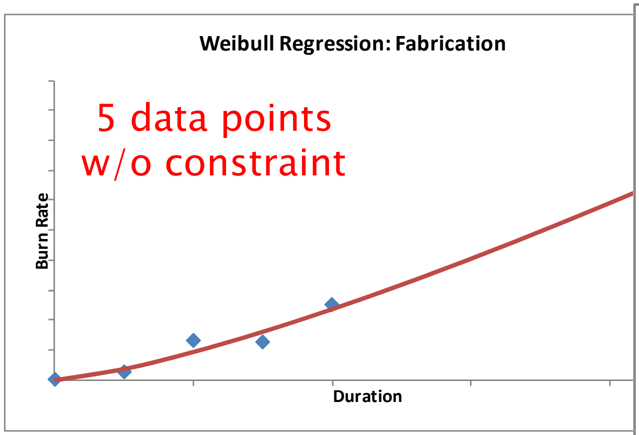


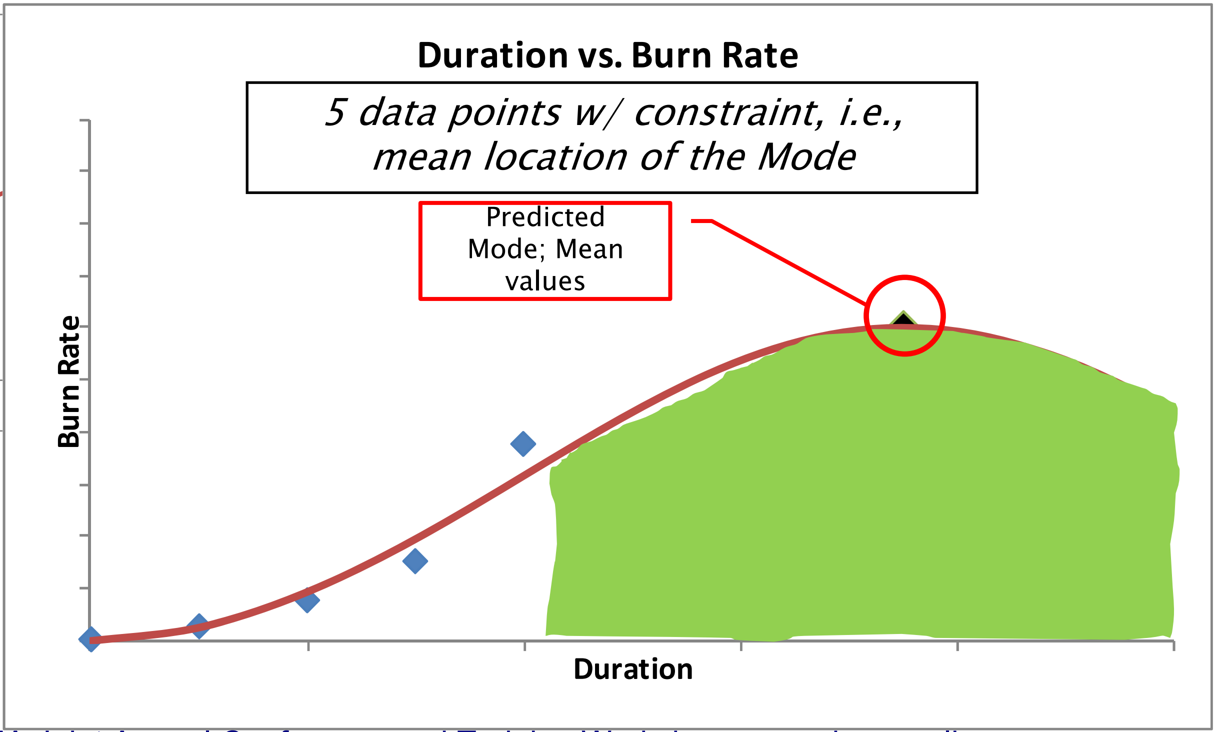
Fabrication Phase Issue



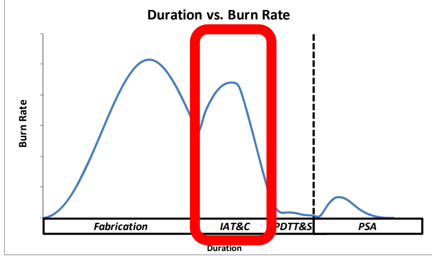
An Example:

1. In the chart below, we have only five Burn Rate data points
2. Assume I have "own-class" data where characteristics of the Mode are known, i.e., the values of the mean and standard deviation for the peak Burn Rate and Duration to the Mode
3. The black diamond represents the mean value-location of the Mode based on "own class" data
4. Now, a Weibull-based regression is performed using the actual data and our predicted location of the Mode
5. Our estimate for the Fabrication phase ACWP is the ACWP at the fifth data point plus our forecast for the remainder of the phase, i.e., the green region



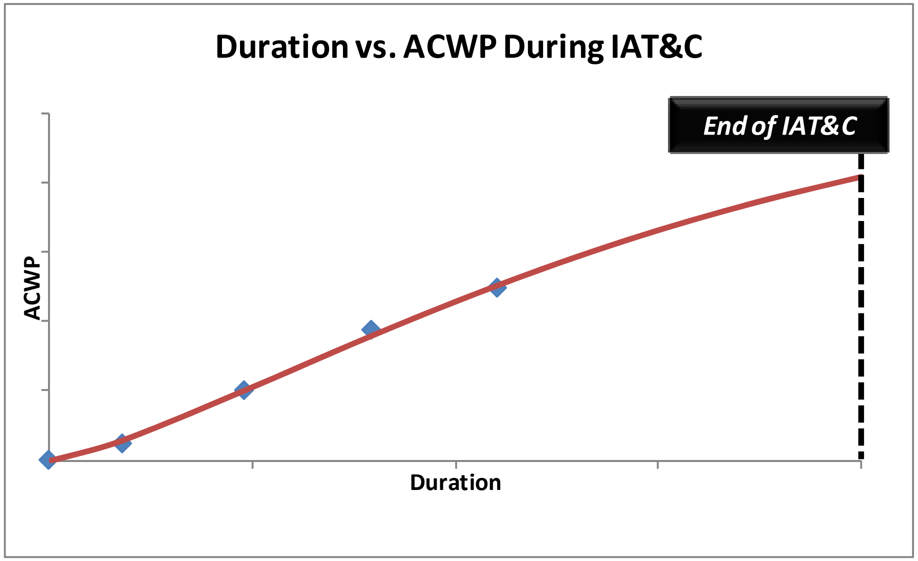


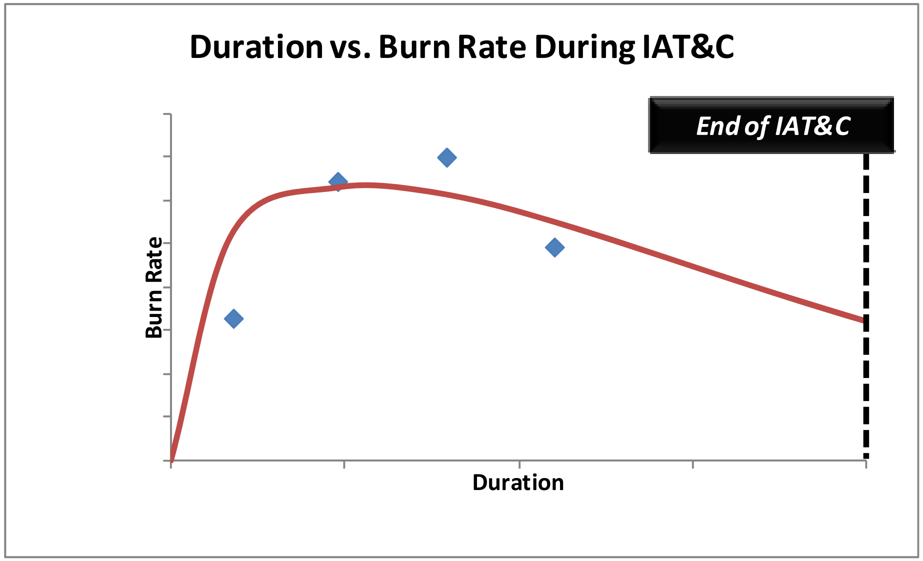
IAT&C Phase



General Properties:

- Consists of hours from the Start of Integration, Assembly, Test & Checkout to the start of Post Delivery Test and Trials and Shakedown
- Accounts for ~20% of total hours expended during ship construction

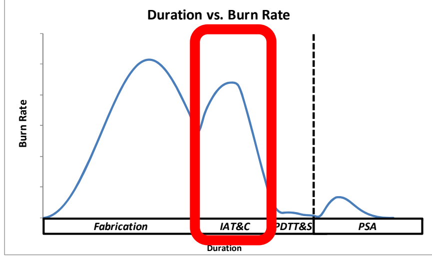




Estimation Process:

- A Weibull cumulative distribution function is fit to the ACWP data
- The estimated ACWP expended during IAT&C is the height of the ACWP forecast at the End of IAT&C

IAT&C Phase Issue



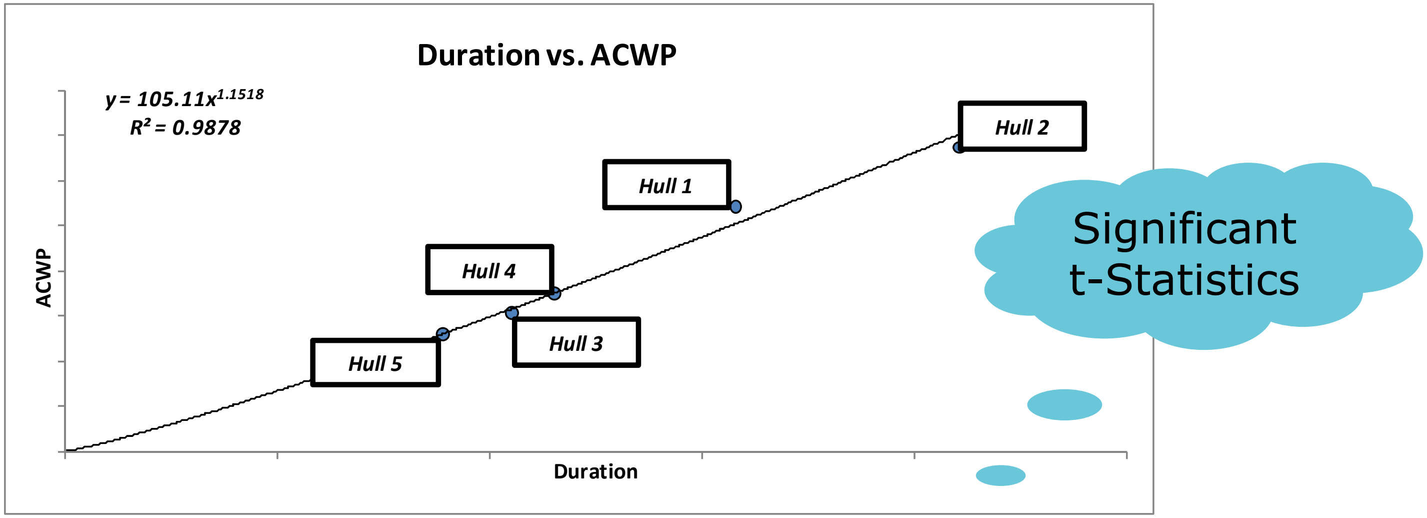
Issue Statement:

- When no data or insufficient data exist to define the mode, Weibull-based estimates can be unstable
- Weibull PDF stabilizes once the Mode location is established

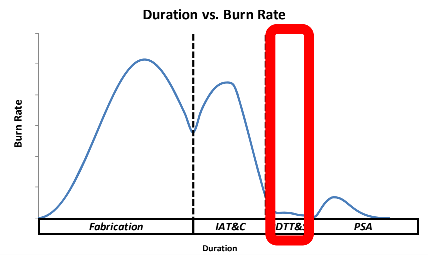
Solution:

- One solution is to use an approach similar to that for the Fabrication phase, i.e., the constrained Weibull
- Another solution is possible if we assume that for all members of the class there is a strong, slightly non-linear correlation between the duration of the IAT&C phase and the ACWP expended during the IAT&C phase
- We can use this observed, statistically sound, "own-class" information to estimate ACWP for the IAT&C phase:

ACWP for IAT&C = 105.11* (Duration of IAT&C)^1.1518



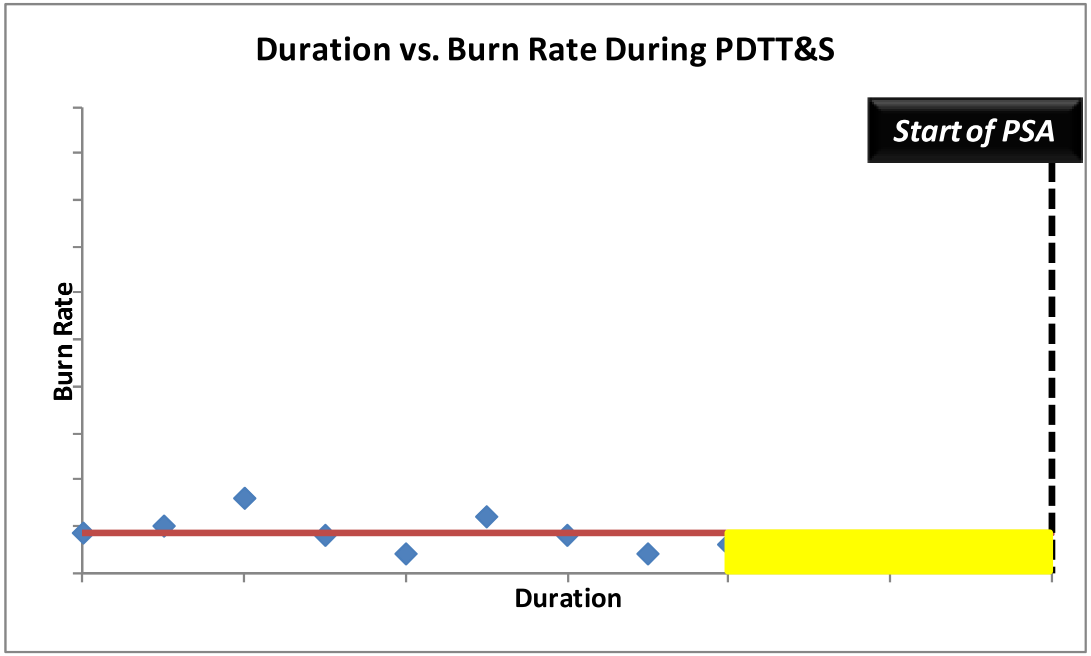
Variable/Term	Coefficient Estimate	Approximate Std Error	Approximate Lower 95% Confidence	Approximate Upper 95% Confidence	T-Statistic	P-Value
a	105.1145	20.0376	41.3147	168.9142	5.2459	0.0135
b	1.1518	0.0740	0.9162	1.3874	15.5649	0.0006



PDTT&S Phase

General Properties:

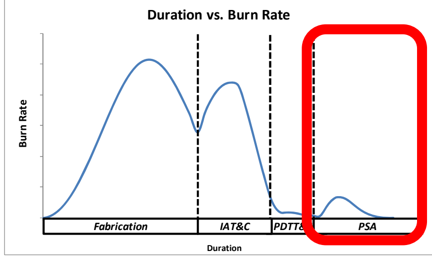
- Consists of hours from start of Post Delivery Test and Trials and Shakedown to the start of Post Shakedown Availability
- Accounts for ~1% of total hours expended during ship construction



Estimation Process:

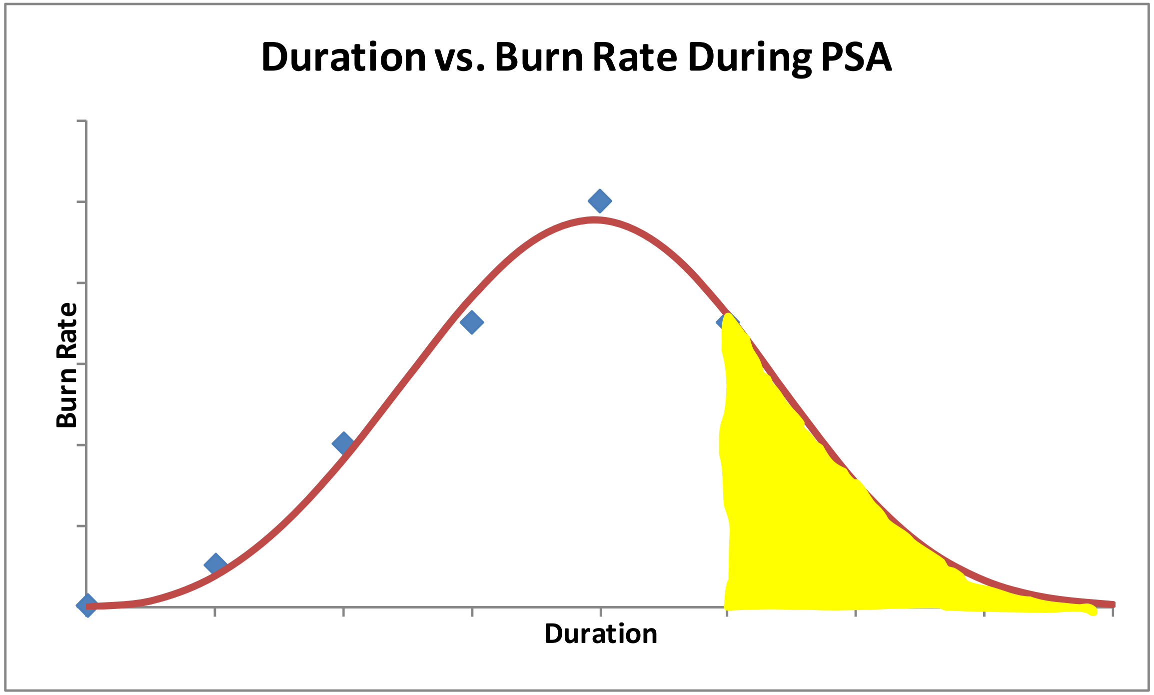
- A straight line is fit to the average burn rate data
- The estimated ACWP expended during PDTT&S is the ACWP to date + the area of the yellow rectangle

PSA Phase



General Properties:

- Consists of hours from start of Post Shakedown Availability to the end of construction
- Accounts for ~4% of total hours expended during ship construction

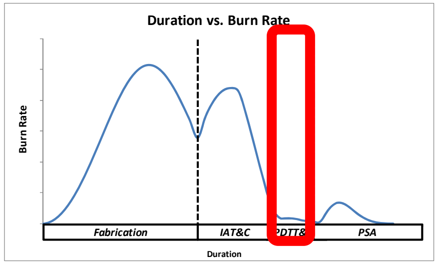


Estimation Process:

- Estimated by fitting Weibull curve to burn rate data
- The estimated ACWP expended during PSA is the ACWP to date + the area of the yellow region

PDTT&S & PSA Phase Issues

PDTT&S Issue Statement:



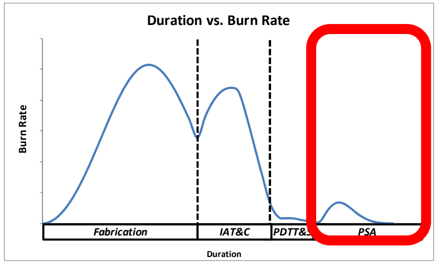
- No data or insufficient data exist for this phase
- Note: ACWP ~ 1% of total

Solution:

- Calculate the “own-class” average Burn Rate for the PDTT&S phase and estimate ACWP using:

ACWP for PDTT&S = (Average Burn Rate for PDTT&S) * (Duration of PDTT&S)

PSA Issue Statement:



- When no data or insufficient data exist to define the mode, Weibull-based estimates can be unstable
- Weibull PDF stabilizes once the Mode location is established
- Note: ACWP ~ 4% of total

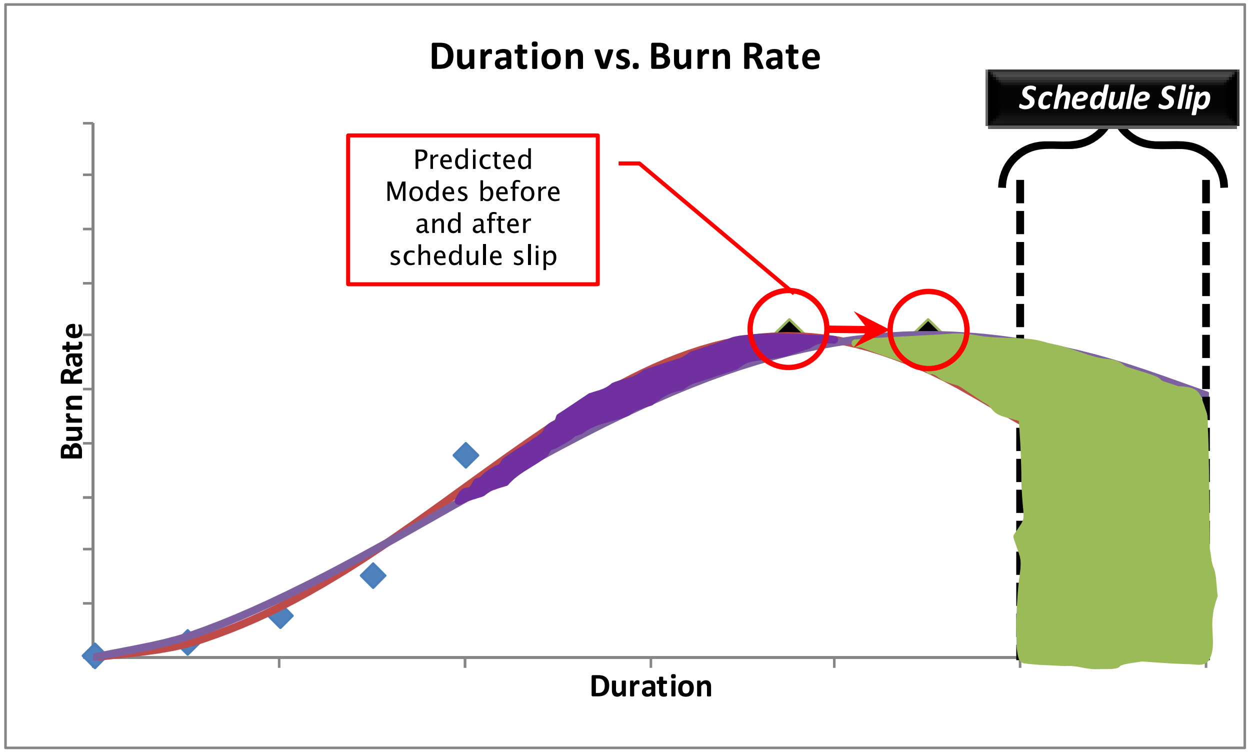
Solution:

- One solution is to use an approach similar to that for the Fabrication phase, i.e., the constrained Weibull
- Another solution is to use an approach similar to that for the IAT&C phase, i.e., “own-class” statistical relationships
- A third solution is to calculate the “own-class” average Burn Rate for the PSA phase and estimate ACWP using:

ACWP for PSA = (Average Burn Rate for PSA) * (Duration of PSA)

Schedule Sensitivity

- Normally, we assume schedules and durations for all phases are defined by an Integrated Master Schedule
- However, the Weibull-based EAC model allows us to perform schedule sensitivity analysis
- For example, let's say the schedule slips
- Since the predicted location of the mode is relative to the end date, a schedule slip would in turn delay the occurrence of the Mode
- So, a new regression is run using actual given data, the new predicted Mode, and the new end date
- The hours added by the schedule slip would approximately equal the area of the green region minus the area of the purple region



Future Work

1. Add prediction intervals to our EACs and understand the effects of our constraint assumptions

2. Consider additional constraints

3. Study the phase-to-phase relationships of earned value data
   - For instance, how is the IAT&C phase related to the Fabrication phase?

4. Study the impact of schedule compression or slippage on the peak burn rate

- For instance, when the schedule is changed, at what point is efficiency beneficially or adversely affected? By how much?

BACK-UP

Derivation of the Mode and Peak

The maximum (mode) and minimum of all continuous, differentiable equations occurs when its derivative is equal to 0 (i.e. its slope is 0). To derive the Mode of the Weibull PDF, we take the derivative of the Weibull PDF, set it equal to 0 and solve for x.

$$f(x) = c * \left(\frac{x}{b}\right)^{a-1} * \exp\left(-\left(\frac{x}{b}\right)^{a}\right)$$

$$f'(x) = \frac{a}{b}\left(\left(\left(\frac{a-1}{b}\right)\left(\frac{x}{b}\right)^{a-2}\exp\left(-\left(\frac{x}{b}\right)^{a}\right)\right) - \left(\left(\frac{x}{b}\right)^{a-1}\left(\frac{a}{b}\right)\left(\frac{x}{b}\right)^{a-1}\exp\left(-\left(\frac{x}{b}\right)^{a}\right)\right)\right) = 0$$

$$0 = \frac{a-1}{a} * \left(\frac{x}{b}\right)^{a-2} - \left(\frac{x}{b}\right)^{a-1} * \left(\frac{x}{b}\right)^{a-1}$$

$$0 = \frac{a-1}{a} * \left(\frac{x}{b}\right)^{a-2} - \left(\frac{x}{b}\right)^{a-2} * \left(\frac{x}{b}\right)^{a}$$

$$\left(\frac{x}{b}\right)^{a} = \frac{a-1}{a}$$

$$x = b * \left(\frac{a-1}{a}\right)^{\frac{1}{a}}$$

The peak burn rate is simply the Weibull PDF evaluated at x = Mode

$$Peak\ Burn\ Rate = c * \left(\frac{Mode}{b}\right)^{a-1} * \exp\left(-\left(\frac{Mode}{b}\right)^{a}\right)$$

$$Peak\ Burn\ Rate = c * \left(\frac{a-1}{a}\right) * \left(\frac{a-1}{a}\right)^{-\frac{1}{a}} * \exp\left(-\left(\frac{a-1}{a}\right)\right)$$

$$Peak\ Burn\ Rate = c * \left(\frac{a-1}{a}\right)^{1-\left(\frac{1}{a}\right)} * \exp\left(-\left(\frac{a-1}{a}\right)\right)$$

Derivation of Constrained Weibull Regression Equation

Therefore, from the previous slide:

$$b = Mode * \left(\frac{a-1}{a}\right)^{-\frac{1}{a}}$$

$$c = PeakBurnRate * \left(\frac{a-1}{a}\right)^{\left(\frac{1}{a}\right)-1} * \exp\left(\frac{a-1}{a}\right)$$

Since b is a function of Mode and a, and c is a function of Peak Burn Rate and a, the Weibull PDF can be modified to reflect the new Constrained Weibull PDF.

$$f(x) = PeakBurnRate * \left(\frac{a-1}{a}\right)^{\left(\frac{1}{a}\right)-1} * \exp\left(\frac{a-1}{a}\right) * \left(\frac{x * \left(\frac{a-1}{a}\right)^{\frac{1}{a}}}{Mode}\right)^{a-1} * \exp\left(-\frac{x * \left(\frac{a-1}{a}\right)^{\frac{1}{a}}}{Mode}\right)^{a}$$

$$f(x) = PeakBurnRate * \left(\frac{a-1}{a}\right)^{-1} * \left(\frac{a-1}{a}\right)^{\frac{1}{a}} * \frac{x^{a-1} * \left(\frac{a-1}{a}\right) * \left(\frac{a-1}{a}\right)^{-\frac{1}{a}}}{Mode^{a-1}} * \exp\left(\left(\frac{a-1}{a}\right) - \frac{x^{a} * \left(\frac{a-1}{a}\right)}{Mode}\right)$$

$$f(x) = PeakBurnRate * \left(\frac{x}{Mode}\right)^{a-1} * \exp\left(\left(\frac{a-1}{a}\right) * \left(1 - \left(\frac{x}{Mode}\right)^{a}\right)\right)$$

Since Peak Burn Rate and Mode are constraints defined based on past ship performance, only one unknown parameter exists (shape parameter a). Therefore, we can regress our data points via OLS to obtain the parameter value for shape parameter a.