# Project Duration Forecasting
## …a comparison of Earned Value Management methods to Earned Schedule

Walt Lipke
Member of Oklahoma City Chapter
Project Management Institute (USA)

**Abstract:** *Earned Value Management (EVM) methods for forecasting project duration have been taught in training courses and used by project managers for four decades. These EVM methods are generally considered to be accepted practice, yet they have not been well studied and researched as to their predictive capability. Using real project data, this paper examines and compares the duration forecasts from four EVM methods to the Earned Schedule prediction technique.*

During the spring of 2003 the concept of Earned Schedule (ES) was introduced, demonstrating the possibility of describing schedule performance in units of time [1]. ES facilitates time-based analysis of the schedule employing uniquely the Earned Value Management (EVM) measures of cost. One year subsequent to the publication of ES, the concept was extended to include project duration forecasting [2]. This paper put forth two equations for forecasting the final duration for a project; one of which is used in this study.

In the time span from year 2004 through 2007, two independent papers were published investigating the capability of the ES forecasting method. One paper written by Lew Hecht describes, positively, the usefulness of ES in a case study of a single US Navy project [3]. The second paper is a comprehensive examination of the capability of ES. The research team of Vanhoucke and Vandevoorde applied a simulation method for assessing the performance of two EVM-based methods and ES in forecasting project duration [4]. A portion of the Vanhoucke and Vandevoorde paper has been updated and published in the Winter 2007-2008 issue of <u>The Measurable News</u> (MN) [5]. The conclusion from the MN paper and its parent indicates "The results …confirm…that the Earned Schedule method outperforms, on average, the other forecasting methods."

Although the results of the research performed by Vanhoucke and Vandevoorde are well regarded, there remains the question of whether the simulation technique is truly representative of real project circumstances. Likewise, the case study testimonial, while strongly supportive of the use of ES indicators and forecasting, is inconclusive in broadly validating the concept. Beyond the recognized shortcomings of the aforementioned studies, it has recently been recognized that four frequently used EVM-based methods of duration forecasting have not been compared to ES. This research is focused to overcome the gaps identified. Real data from 16 projects is used to analyze the respective forecasting capabilities of the overlooked EVM methods along with ES.

The paper begins by defining the pertinent elements of the EVM and ES methods. Building on this foundation, the forecasting equations are presented. Next, the hypothesis of the analysis is described. Then the computations, needed to perform the analysis and evaluation, are outlined. The project data is characterized and results from the computations and analysis are discussed. Finally, conclusions are drawn.

## Earned Value Management Duration Forecasting

An understanding of EVM and its terminology is assumed in this paper. For convenience, the EVM terminology used to portray project status and forecast final duration is tabulated below:

| | |
|---|---|
| PV | Planned Value |
| EV | Earned Value |
| BAC | Budget at Completion (the planned cost of the project) |
| PMB | Performance Measurement Baseline (the cumulative PV over time) |
| IEAC(t) | Independent Estimate at Completion (the forecast final duration) |

Four EVM duration forecasting techniques have been commonly applied over the last 40 years to predict project completion dates. These methods have the following basic form:

Duration Forecast = Elapsed Time + Forecast for Work Remaining
$$IEAC(t) = AT + (BAC - EV) / \text{Work Rate}$$

where 
AT = Actual Time (the duration elapsed to the time at which PV and EV are measured)
BAC – EV is commonly termed the work remaining
Work Rate is a factor which converts the work remaining to time, the duration forecast for the remaining work

The four Work Rates commonly applied are defined below:
1) Average Planned Value: $PV_{av} = PV_{cum} / n$
2) Average Earned Value: $EV_{av} = EV_{cum} / n$
3) Current Period Planned Value: $PV_{lp}$
4) Current Period Earned Value: $EV_{lp}$

where 
$PV_{cum}$ = cumulative value of PV
$EV_{cum}$ = cumulative value of EV
n = total number of periodic time increments of project execution within AT

The EVM forecasts of final duration, IEAC(t), are associated with the Work Rate employed and identified in the remainder of the paper as follows:
1) $PV_{av}$: $IEAC(t)_{PVav}$
2) $EV_{av}$: $IEAC(t)_{EVav}$
3) $PV_{lp}$: $IEAC(t)_{PVlp}$

4)  EVIp: $IEAC(t)_{EVIp}$

## Earned Schedule Duration Forecasting

A recent extension to EVM, ES, has emerged which provides reliable, useful schedule performance management information. In brief, the method yields time-based indicators, unlike the cost-based indicators for schedule performance offered by EVM.

Figure 1 is an illustration for understanding the concept. The ES measure identifies when the amount of EV accrued should have occurred. As depicted by the diagram, this is the point on the PMB where PV equals the EV accrued. The vertical line from the point on the PMB to the time axis determines the "earned" portion of the schedule. The duration from the beginning of the project to the intersection of the time axis is the amount of earned schedule (ES).
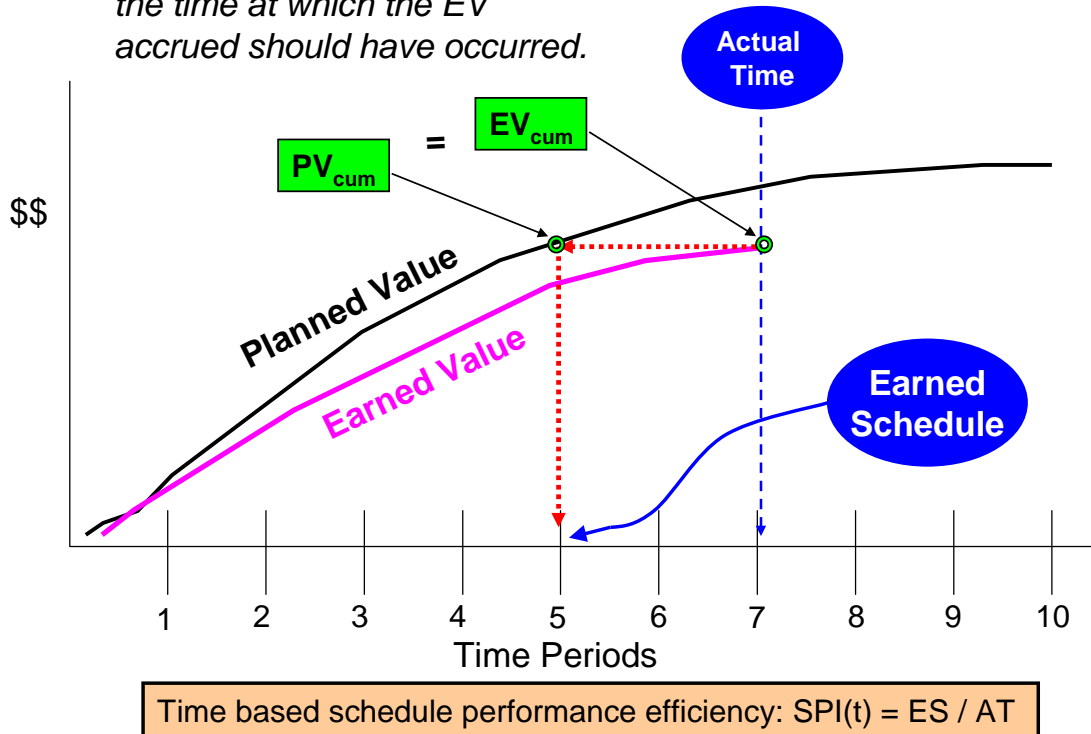
## Earned Schedule Concept



*Figure 1.  Earned Schedule Concept*

With ES and AT defined, the schedule performance efficiency is formulated as depicted in figure 1, SPI(t) = ES / AT. From EVM, final cost may be forecast from the formula, IEAC = BAC / CPI; CPI = EV / AC, where AC is the actual cost. In an analogous manner final duration is forecast from IEAC(t)es = PD / SPI(t), where PD is the planned duration for the project and IEAC(t)es is the ES forecast of final duration.

## Discussion of Forecasting Methods and Study Considerations

The objective of the study is to investigate and understand the forecasting capability of the five methods, four from EVM and one from ES. By inspection, it can be deduced that the EVM Work Rates have mathematical failings which affect their performance.

When the project executes past its planned duration, PVcum is equal to its maximum value, BAC, and is invariant thereafter. Thus the PVav Work Rate becomes PVav = BAC / m, where m is a number larger than the planned number of time periods for the project. Obviously as m becomes larger PVav is decreasingly smaller, thereby causing the work remaining forecast to be longer than its planned time.

The situation for the PVlp Work Rate is more severe. After the planned project duration has passed, there are no periodic values of PV, thereby making the computation of $IEAC(t)_{PVlp}$ indeterminate. These observations are excluded from the study because it may be that $IEAC(t)_{PVlp}$ is a good predictor otherwise. A tenet of the study is to provide each method reasonable opportunity to show well, despite the known limitations.

The two Work Rates, EVav and EVlp, normally do not have indeterminate calculation conditions. There is, however, one exception of when a period elapses with no EV accrued; this condition may occur for smaller projects which assess their status weekly. When EVlp is equal to zero, $IEAC(t)_{EVlp}$ cannot be calculated. Just as for PVlp, the condition is accommodated in the study so as to not discredit the overall forecasting performance of EVlp. When an anomalous instance is encountered, the forecast for the previous valid observation is used.

The forecasting from ES does not experience indeterminate calculation conditions. A common positive characteristic of all of the methods with the exception of $IEAC(t)_{PVlp}$ is they converge to the actual duration. The predictive capability of the four EVM-based methods in this study may be superior to the two tested by Vanhoucke and Vandevoorde [4,5]; those methods did not necessarily correctly calculate the actual outcome duration at completion.

## Study Hypothesis and Methodology

The conjecture to be examined in the study is ES provides a better forecasting method of final project duration than the four methods cited previously for EVM. To make a determination concerning this conjecture, the extreme case will be examined and tested. The test is constructed to show that the EVM methods <u>as an aggregate</u> produce better forecasts than does ES. If the EVM methods are shown to be superior to ES it will not be known which one of the EVM methods is better. Thus, if this is the determination, further examination will be necessary to understand the circumstances for selecting the appropriate EVM forecasting method.

The hypothesis from the preceding discussion is formally defined as follows:
Ho: EVM methods produce the better forecast of final project duration
Ha: ES method produces the better forecast of final project duration

where Ho is the null hypothesis (i.e., the statement to be validated) and Ha is the alternate hypothesis [6].

The statistical testing is performed using the Sign Test applied at 0.05 level of significance [7]. Assuming each of the five methods has an equal probability of success, the probability for each trial is 0.8.

Data from 16 projects is used for generating the forecasts from each of the methods. These forecasts are then tested and analyzed. The test statistic for the hypothesis test is computed from the number of times the EVM methods are observed to yield the better forecast. Thus, for each testing condition applied the maximum number of successes for the EVM methods is 16. When the EVM methods successes are fewer than 10, the test statistic has a value in the critical region ($\leq 0.05$). A value in the critical region indicates there is enough evidence to reject the null hypothesis. In clearer language, this test result shows that the EVM methods do not produce duration forecasts better than those from ES. A test statistic value outside of the critical region is the converse; i.e., there is not enough evidence to reject the null hypothesis.

The test statistic is determined from the ranking of the standard deviations for each of the forecasting methods for each project. The standard deviation is calculated from the differences between the forecast values computed at the project status points and the actual final duration as defined below:

$$\sigma_m = [\Sigma\ (FV_m(i) - FD)^2 / (n\text{-}1)]^{0.5}$$

where,      $\sigma_m$ = the standard deviation for forecasting method m
            $FV_m(i)$ = forecast value for method m at status point (i)
            FD = actual final duration
            n = number of status points
            $\Sigma$ = summation over a specific set of status points

The smallest value for the standard deviation indicates the best forecast produced. There are five forecasting methods, thus the ranking will be from 1 to 5; the rank of 1 is associated with the lowest value of standard deviation and 5 the highest. The ranking of the methods is performed for the 16 sets of project data. The number of times the rank of 1 occurs (without ties between the EVM methods) determines the test statistic value. By using the ranking approach, the unit for the periods (e.g., months, weeks) can be different between projects; the ranking of the five methods is performed separately for each project.

To understand whether a particular method is better for early, middle, late or overall forecasting the projects are analyzed and tested for specific regions of performance. Seven groupings are formed using the observations within various percent complete ranges to make the determinations: early (10% - 40%), middle (40% - 70%), late (70% - 100%), overall (10% - 100%). Additionally, other ranges are used to determine if one of the methods converges to the actual final duration more rapidly than the others and thus is better for a portion of the forecast, but is not necessarily superior overall. The ranges used for this purpose are: 25% - 100%, 50% - 100%, and 75% - 100%.

# Data Discussion

A total of sixteen projects are included in the study. Twelve (1 through 12) are from one source with four (13 through 16) from a second. The output of the twelve projects is high technology products. The remaining four projects are associated with information technology (IT) products.

The primary data requirement is that the projects used in the study have not undergone any re-planning. The requirement is necessary to be able to discern the ability of the forecasting methods without having outside influence. All sixteen projects performed from beginning to completion without having baseline changes.

Table 1 illustrates the schedule performance of the projects in the data set. The twelve high technology projects are measured in monthly periods whereas the four IT projects are measured weekly. Two projects completed early, three on time and the remaining eleven delivered later than planned.

| Schedule Performance | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **Project** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| **Planned Duration** | 21m | 32m | 36m | 43m | 24m | 50m | 46m | 29m |
| **Actual Duration** | 24m | 38m | 43m | 47m | 24m | 59m | 54m | 30m |
| **Project** | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 |
| **Planned Duration** | 45m | 44m | 17m | 50m | 81w | 25w | 25w | 19w |
| **Actual Duration** | 55m | 50m | 23m | 50m | 83w | 25w | 22w | 13w |
| **Legend:      m = month      w = week** | | | | | | | | |

*Table 1.  Schedule Performance*

# Results Analysis

To begin the analysis, it is instructive to view the graphs from a single project (Project #13). The first graph, Figure 2, portrays the forecasting performance of all five methods along with the horizontal line for the actual final duration. It is observed that the prediction using the PVav and EVav Work Rates behave in a much less erratic manner than do the forecasts from the current period rates, PVlp and EVlp. The forecast from ES is seen to be much better than any of the EVM predictions, especially after the project completion point of 40 percent.

The next graph, Figure 3, portrays similar information. It contains plots of the standard deviation versus percent complete for each of the EVM and ES methods. The behavior seen in figure 2 is amplified by viewing the standard deviation. As described for figure 2, the average work rates are less volatile, while the current rates have large changes from one observation to the next. Again, the ES forecast is observed to be much more stable than any of the other methods. The standard deviation of the ES forecast is noticeably smaller than any of the other methods between 50 and 100 percent complete.
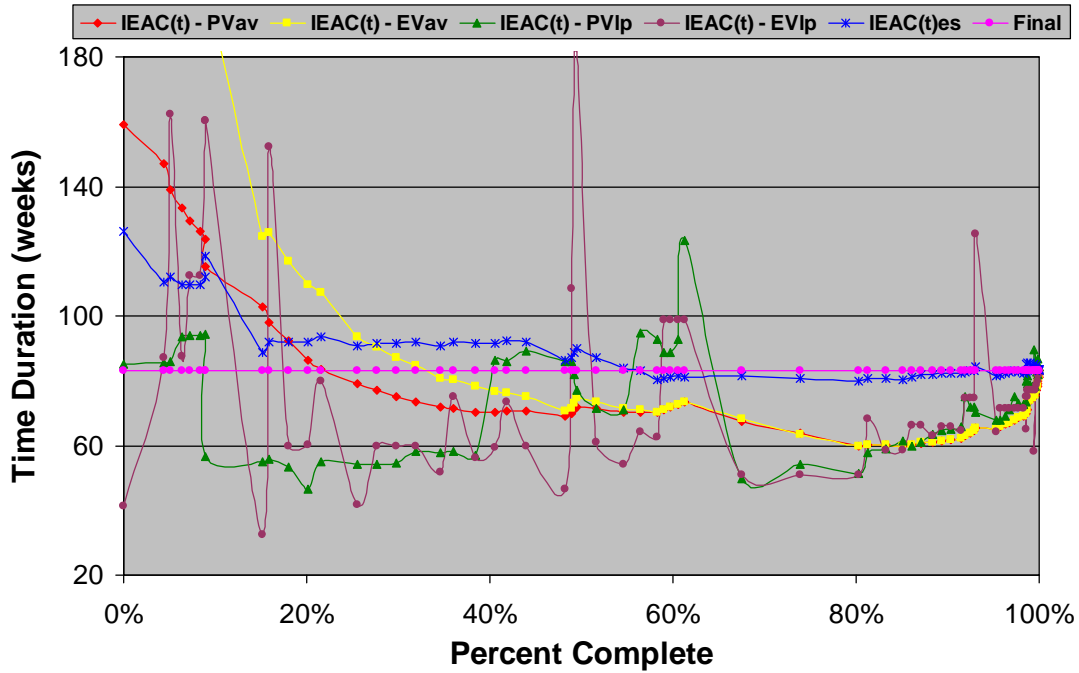
# Final Duration Forecasting Comparisons

## Project #13



*Figure 2.  Final Duration Forecasting Comparisons*

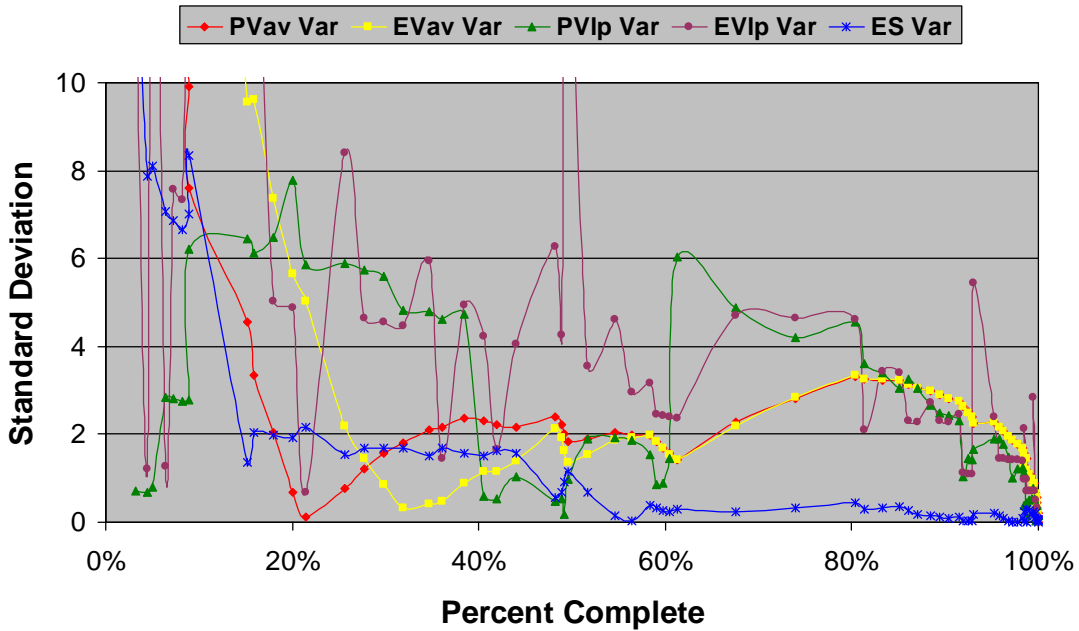# Time Forecasting Std Dev Comparisons

## Project #13



*Figure 3.  Standard Deviation Comparisons*

Figure 4 is a column graph illustrating a view intended for analyzing the forecasting behavior for early, middle, late and overall ranges of project execution. Figure 5 is also a column graph; the ranges applied (25% - 100%, 50% - 100%, 75% - 100%) are used to determine the behavior of the various methods regarding the rate of convergence to the final duration.

For both figures 4 and 5, it is clearly seen that the current period methods are generally more volatile and that the ES method is the better predictor in every range. In fact, for this project, the accuracy of the ES forecasting method is significantly better than the EVM methods.

Recall from the discussion earlier in the paper that, with the exception of the forecast using the PVlp Work Rate, all of the other methods converge to calculate the actual final duration. Because of this characteristic, the expectation is that the standard deviation should decrease as the completion percentage increases. This behavior is observed for ES and EVlp, but the others are nearly invariant between the 25% - 100%, 50% - 100% and 75% - 100% ranges. Looking back at figure 3, the convergence is seen for PVav and EVav but it is not strongly evident until after the project has progressed past 80 percent complete.
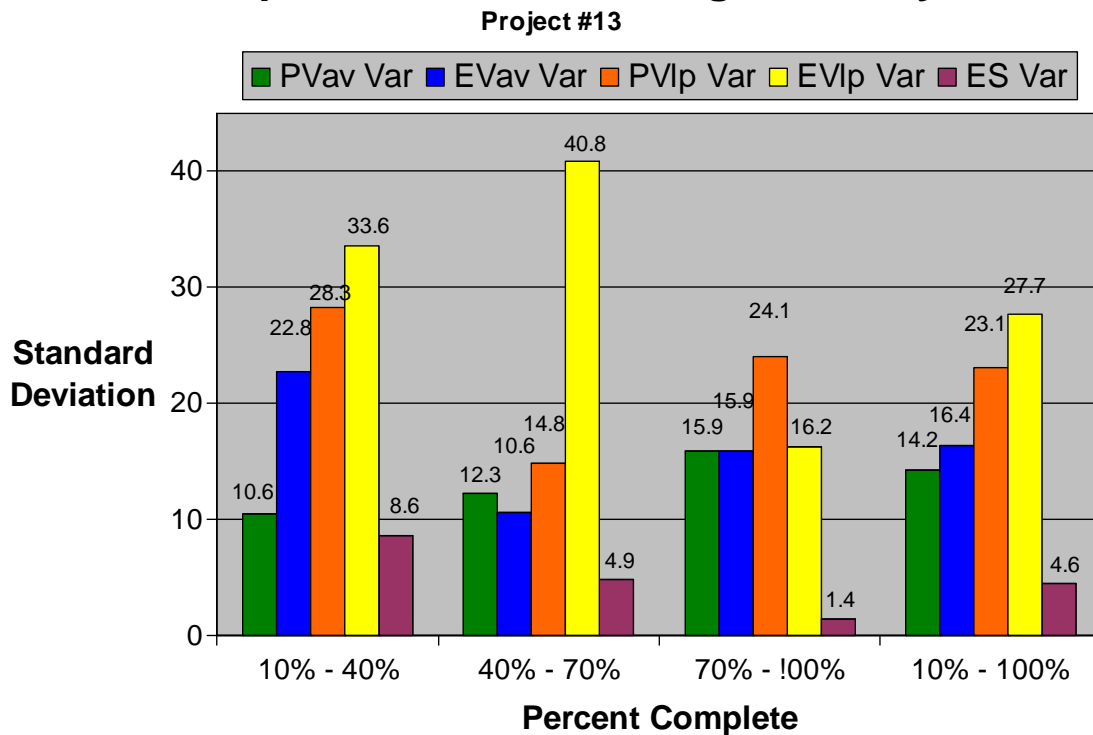
## Comparison of Forecasting Accuracy
### Project #13



Figure 4.  Forecasting Accuracy – Early, Middle, Late, Overall Completion Ranges

# Comparison of Forecasting Convergence
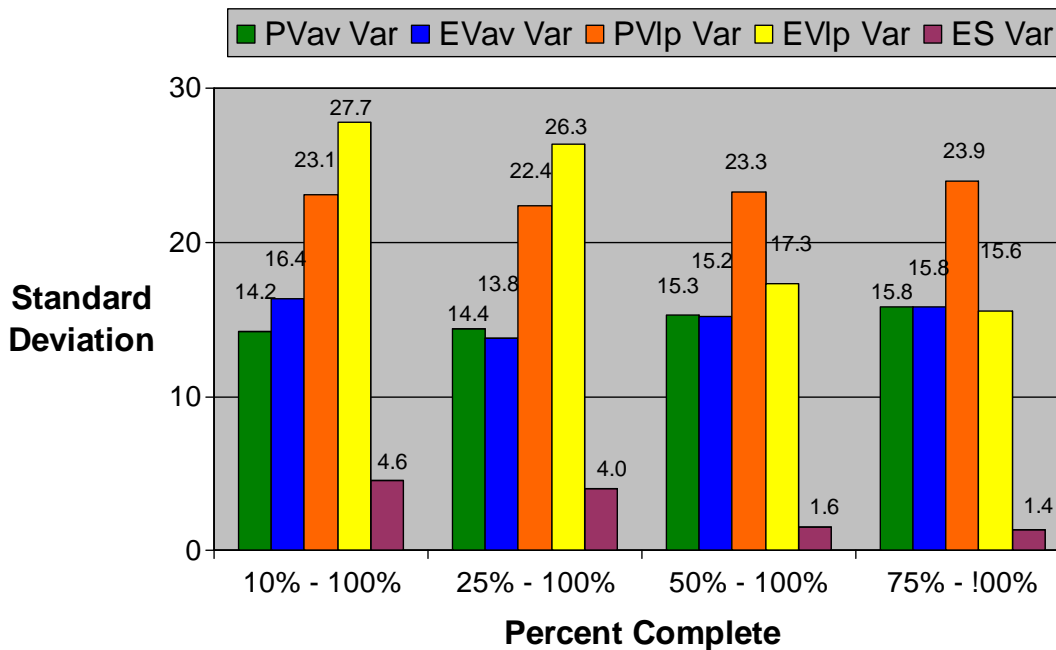## Project #13



*Figure 5. Forecasting Convergence*

Table 2, is an example for the 10% - 40% completion range. It is the tabulation of the computed standard deviations for each forecasting method and their ranking for each project. From reviewing the table, an observation is made that for this completion range, the ES rank is "1" for eleven of the projects. The ES forecasting method provides the best forecasts of final duration for a large majority of the projects. Even so, it does not produce the best forecast results for every project. All seven ranges are analyzed to understand more completely how the various methods perform under different circumstances.

To better understand the goodness of the forecasting methods for the examined completion band, Table 3 was created. It is a condensation of table 2. As can be observed the distribution of the ranking numbers is made between the various forecasting methods. In general, the sum for each of the ranks will equal the number of projects, sixteen. However, when there are ties, as there is for this range, one rank may total more than 16 while an adjacent rank will be equally lower. For table 3, it is noted the sum of the 1s is seventeen while the sum of the 2s is fifteen.

At the bottom of table 3, a weighted average of the ranking distribution is computed for each of the forecasting methods. These weighted averages are then used to rank the methods for the completion range examined. Table 4 is a tabulation of the weighted averages of the rankings for each of the seven completion ranges. For each range, the ES method has the lowest weighted average, indicating that, on average, it is the best predictor of final duration. The only challenge to ES is within the 40% - 70% middle range, where the weighted average of 2.063 for ES is somewhat lower than the 2.500 from PVav.

## Standard Deviation Results & Ranking for 10% - 40% Completion Group

| | | Project #1 | | Project #2 | | Project #3 | | Project #4 | | Project #5 | | Project #6 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Project ID | | Std Dev | Rank | Std Dev | Rank | Std Dev | Rank | Std Dev | Rank | Std Dev | Rank | Std Dev | Rank |
| Methods | PVav | 14.95 | 5 | 13.01 | 4 | 11.93 | 2 | 25.59 | 2 | 4.38 | 2 | 29.76 | 2 |
| | EVav | 2.65 | 1 | 9.35 | 2 | 8.28 | 1 | 48.68 | 4 | 5.82 | 3 | 42.64 | 4 |
| | PVlp | 5.47 | 2 | 13.62 | 5 | 77.74 | 5 | 42.77 | 3 | 8.67 | 4 | 42.11 | 3 |
| | EVlp | 6.00 | 3 | 12.14 | 3 | 22.38 | 3 | 103.15 | 5 | 9.89 | 5 | 263.03 | 5 |
| | ES | 8.28 | 4 | 4.78 | 1 | 46.76 | 4 | 14.03 | 1 | 1.88 | 1 | 3.57 | 1 |

| | | Project #7 | | Project #8 | | Project #9 | | Project #10 | | Project #11 | | Project #12 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Project ID | | Std Dev | Rank | Std Dev | Rank | Std Dev | Rank | Std Dev | Rank | Std Dev | Rank | Std Dev | Rank |
| Methods | PVav | 9.79 | 3 | 16.16 | 3 | 6.75 | 2 | 9.06 | 1 | 7.66 | 4 | 15.06 | 3 |
| | EVav | 6.00 | 2 | 33.17 | 5 | 15.63 | 3 | 10.55 | 2 | 6.63 | 3 | 30.49 | 5 |
| | PVlp | 17.95 | 5 | 20.69 | 4 | 20.80 | 4 | 39.11 | 4 | 7.70 | 5 | 9.06 | 1 |
| | EVlp | 15.07 | 4 | 5.69 | 2 | 525.62 | 5 | 102.21 | 5 | 6.58 | 2 | 26.86 | 4 |
| | ES | 4.31 | 1 | 5.09 | 1 | 3.74 | 1 | 15.22 | 3 | 4.54 | 1 | 12.49 | 2 |

| | | Project #13 | | Project #14 | | Project #15 | | Project #16 | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Project ID | | Std Dev | Rank | Std Dev | Rank | Std Dev | Rank | Std Dev | Rank | | | | |
| Methods | PVav | 10.57 | 2 | 2.36 | 1 | 15.93 | 3 | 20.18 | 5 | | | | |
| | EVav | 22.78 | 3 | 5.90 | 5 | 18.12 | 5 | 17.10 | 4 | | | | |
| | PVlp | 28.25 | 4 | 2.36 | 1 | 11.24 | 2 | 12.37 | 2 | | | | |
| | EVlp | 33.59 | 5 | 5.49 | 4 | 16.87 | 4 | 16.49 | 3 | | | | |
| | ES | 8.62 | 1 | 4.46 | 3 | 4.45 | 1 | 5.20 | 1 | | | | |

*Table 2. Standard Deviation & Ranking for 10% - 40% Completion Range*

## Rank Count for Data Group 10% - 40%

| | | Methods | | | | |
|---|---|---|---|---|---|---|
| | | PVav | EVav | PVlp | EVlp | ES |
| Count | Nr 1s | 2 | 2 | 2 | 0 | 11 |
| | Nr 2s | 6 | 3 | 3 | 2 | 1 |
| | Nr 3s | 4 | 4 | 2 | 4 | 2 |
| | Nr 4s | 2 | 3 | 5 | 4 | 2 |
| | Nr 5s | 2 | 4 | 4 | 6 | 0 |
| Weighted Average | | 2.750 | 3.250 | 3.375 | 3.875 | 1.688 |
| Composite Rank | | 2 | 3 | 4 | 5 | 1 |

*Table 3. Rank Count for Data Group 10% - 40%*

## Weighted Average of Ranking Results - EVM vs ES Time Forecast

| | ********** Percent Complete Test Bands ********** | | | | | | |
|---|---|---|---|---|---|---|---|
| | 10% - 40% | 40% - 70% | 70% - 100% | 10% - 100% | 25% - 100% | 50% - 100% | 75% - 100% |
| ES | 1.688 | 2.063 | 1.438 | 1.625 | 1.563 | 1.563 | 1.438 |
| PVav | 2.750 | 2.500 | 3.688 | 2.625 | 2.813 | 3.063 | 3.875 |
| EVav | 3.250 | 2.813 | 2.938 | 3.000 | 3.063 | 2.938 | 2.875 |
| PVlp | 3.375 | 3.438 | 3.875 | 3.813 | 3.875 | 3.688 | 3.875 |
| EVlp | 3.875 | 4.188 | 3.063 | 3.938 | 3.688 | 3.750 | 2.938 |

*Table 4. Weighted Average of Ranking Results*

Finally, more conclusive evidence of the goodness of the ES forecasting capability is provided from the statistical hypothesis testing. Table 5 provides the compiled results from the testing analysis. In the table the count of the rank of "1" is provided for the aggregate of the EVM methods and for ES. With the exception of one test range, ES shows to be superior to the other methods combined. In one instance, the 40% - 70% range, the number of 1s counted for EVM exceeds the number for ES. However, the value of the test statistic is in the critical region indicating there is enough evidence to reject the null hypothesis that the aggregate of the EVM methods is better than the ES method. Thus, from the results of the Sign Test, ES is indicated to be the better forecasting method regardless of project completion stage, early, middle, late and overall.

| Hypothesis Test Results - EVM vs ES Time Forecast | | | | | | | |
|---|---|---|---|---|---|---|---|
| Significance | ********** | | Percent Complete Test Bands | | | ********** | |
| α = 0.05 | 10% - 40% | 40% - 70% | 70% - 100% | 10% - 100% | 25% - 100% | 50% - 100% | 75% - 100% |
| Test Statistic | 0.0000 | 0.0267 | 0.0000 | 0.0000 | 0.0000 | 0.0002 | 0.0000 |
| Sign Test | Ha | Ha | Ha | Ha | Ha | Ha | Ha |
| Count #1s ES | 11 | 7 | 12 | 11 | 11 | 10 | 12 |
| EVM | 5 | 9 | 4 | 5 | 5 | 6 | 4 |

**Hypothesis Test:** Sign Test at 0.05 level of significance.
**Ho:** The aggregate of EVM forecasts is better / the null hypothesis
**Ha:** ES forecast is better / the alternate hypothesis

*Table 5.  Hypothesis Test Results*

## Summary and Conclusions

Five methods of project duration forecasting were examined in this study, four from EVM and the ES technique.  Performance data from sixteen projects was used to assess the capabilities of the various forecasting methods. The analysis strategy segregated the project data into seven ranges of percent complete in order to isolate possible forecasting characteristics or tendencies among the methods.

Each of the methods was used to create forecasts from the project data. The standard deviation of the forecasts from the actual final duration was computed for each project and each percent complete range studied. The forecasting methods were then ranked from best to worst using the standard deviations.

The tabulation of best forecasts, one of the four EVM methods or ES, for each range was used to calculate the test statistic for the Sign Test. The null hypothesis, EVM provides the better forecast, was rejected for every percent complete range examined.

Conclusively, from among the methods and data set studied, ES is shown to be the best method of forecasting project duration.

## Acknowledgement

**References**

1. Lipke, W. "Schedule is Different," _The Measurable News_, Summer 2003: 31-34
2. Henderson, K. "Further Developments in Earned Schedule," _The Measurable News_, Spring 2004: 15-22
3. Hecht, L. "A Case Study of Earned Schedule to Do Predictions," _The Measurable News_, Winter 2007-2008: 16-18
4. Vanhoucke, M., S. Vandevoorde. "A Simulation and Evaluation of Earned Value Metrics to Forecast Project Duration," _Journal of the Operations Research Society_, October 2007, Vol 58: 1361-1374
5. Vanhoucke, M., S. Vandevoorde. "Measuring the Accuracy of Earned Value/Earned Schedule Forecasting Predictors," _The Measurable News_, Winter 2007-2008: 26-30
6. Wagner, S. _Introduction to Statistics_, Harper Collins, New York 1992
7. National Institute of Standards and Technology Data Plot 2005. "Sign Test." http://www.itl.nist.gov/div898/software/dataplot/refman1/auxillar/signtest.htm

**About the Author**

**Walt Lipke** retired in 2005 as deputy chief of the Software Division at Tinker Air Force Base. He has over 35 years of experience in the development, maintenance, and management of software for automated testing of avionics. During his tenure, the division achieved several software process improvement milestones, including the coveted SEI/IEEE award for Software Process Achievement. Mr. Lipke has published several articles and presented at conferences, internationally, on the benefits of software process improvement and the application of earned value management and statistical methods to software projects. He is the creator of the technique _Earned Schedule_, which extracts schedule information from earned value data. Mr. Lipke is a graduate of the USA DoD course for Program Managers. He is a professional engineer with a master's degree in physics, and is a member of the physics honor society, Sigma Pi Sigma ($\Sigma\Pi\Sigma$). Lipke achieved distinguished academic honors with the selection to Phi Kappa Phi ($\Phi\Kappa\Phi$). During 2007 Mr. Lipke received the PMI Metrics Specific Interest Group Scholar Award and the PMI Eric Jenett Award for Project Management Excellence for his leadership role and contribution to project management resulting from his creation of the Earned Schedule method.

1601 Pembroke Drive
Norman, OK 73072
Phone: (405) 364-1594
E-mail: waltlipke@cox.net