Presented at the 2011 ISPA/SCEA Joint Annual Conference and Training Workshop - www.iceaaonline.com

TECOLOTE RESEARCH, INC.

Bridging Engineering and Economics

Since 1973

Joint Confidence Level of a Parametric Software Cost and Schedule Estimate

Michael A. Ross

Technical Expert

Tecolote Research, Inc.

3601 Aviation Boulevard, Suite 1600

Manhattan Beach, CA 90266-3758

mross@tecolote.com

Santa Barbara Los Angeles Boston Washington, D.C. Chantilly Huntsville Dayton Albuquerque

Montgomery Ogden Cleveland San Diego Colorado Springs Vandenberg AFB Denver Silver Spring Tacoma

Oklahoma City New Orleans Patuxent River Charleston Rosslyn

"The revolutionary idea that defines the boundary between modern times and the past is the mastery of risk: the notion that the future is more than a whim of the gods and that men and women are not passive before nature. Until human beings discovered a way across that boundary, the future was a mirror of the past or the murky domain of oracles and soothsayers who held a monopoly over knowledge of anticipated events."

(Bernstein, 1996)

TECOLOTE RESEARCH, INC.

Outline

- Problem and Background
- Model Summary
- Regressing CDERs
- Probability in CDER Mathematics
- Joint and Conditional Probability Defined
- Examples
- References

TECOLOTE RESEARCH, INC.

Problem

- When cost uncertainty analyses are presented to decision-makers, questions often asked are
  - *"What is the chance the system can be delivered within cost and schedule?"*
  - *"How likely might the point estimate cost be exceeded for a given schedule?"*
  - *"How are cost reserve recommendations affected by schedule risk?"*
- During the past thirty years, techniques from univariate probability theory have been widely applied to provide insight into P(Cost ≤ x1) and P(Schedule ≤ x2).
- Although it has long been recognized that a system's cost and schedule are correlated, little has been applied from multivariate probability theory to study joint cost-schedule distributions.

(Garvey, 2000)

TECOLOTE RESEARCH, INC.

Background

- **USAF (Marvin Sambur / Peter Teets memo of 2004)**
- **NASA Cost Estimating Handbook**
- **Past ISPA/SCEA Conferences**
  - Panel of Experts Discussion
  - Presentations
- **SSCAG Risk Working Group (Tim Anderson, Eric Druker, et. al.)**
  - ***Parametric***
    - ***"Disjoint Cost and Schedule distributions are conflated into a bivariate distribution through the injection of correlation between the two"***
  - Buildup
  - Estimate to Complete Projection
- **Joint Confidence Level (JCL)**
  - NASA
  - USAF AFCAA

TECOLOTE RESEARCH, INC.

Software CDER Regression Equations

CER

$$\left[\boldsymbol{E} = \boldsymbol{b_1 S}^{a_1}\right]_{\text{<data set name>}} \rightarrow \left[\boldsymbol{E}\left(\boldsymbol{b_1 S}^{a_1}\right)^{-1} = 1\right]_{\text{<data set name>}}$$ *power regression*

SER

$$\left[\boldsymbol{T} = \boldsymbol{b_2 E}^{a_2}\right]_{\text{<data set name>}} \rightarrow \left[\boldsymbol{T}\left(\boldsymbol{b_2 E}^{a_2}\right)^{-1} = 1\right]_{\text{<data set name>}}$$ *power regression*

$CER\, g SER \rightarrow CDER$

$$\left[\boldsymbol{E}\left(\boldsymbol{b_1 S}^{a_1}\right)^{-1}\boldsymbol{T}\left(\boldsymbol{b_2 E}^{a_2}\right)^{-1} = 1\right]_{\text{<data set name>}}$$

Letting $\alpha_E \,\mathrm{A}\, \frac{1-a_2}{a_1}$ *and* $\alpha_T \,\mathrm{A}\, \frac{1}{a_1}$ *and* $\boldsymbol{D} \,\mathrm{A}\, \left(\boldsymbol{b_1 b_2}\right)^{\frac{1}{a_1}}$ *yields*

$$\left[\boldsymbol{E}^{\alpha_E}\boldsymbol{T}^{\alpha_T} = \boldsymbol{DS}\right]_{\text{<data set name>}}$$ *factor regression*

(Ross, 2008a), (Ross, 2008b), (Valerdi, et al, 2009), (Ross, 2011b)

Software CDER Graphing Functions

Work Function

$$\left[\boldsymbol{E} = \left(\boldsymbol{DST}^{-\alpha_T} \right)^{1/\alpha_E} \right]_{\text{<dataset name>}}$$

Intensity Function

$$\left[\boldsymbol{E} = \boldsymbol{IT}^{\gamma} \right]_{\text{<dataset name>}}$$

(Ross, 2008a), (Ross, 2008b), (Valerdi, et al, 2009), (Ross, 2011b)

TECOLOTE RESEARCH, INC.

Software CDER Solving for Cost and Schedule

Cost (Effort) Estimating Relationship (CER)

$$\left[\boldsymbol{E} = \left(\boldsymbol{I}^{\alpha_T} \left(\boldsymbol{DS} \right)^{\gamma} \right)^{1/(\gamma\alpha_E + \alpha_T)} \right]_{\text{<dataset name>}}$$

Schedule (Duration) Estimating Relationship (SER)

$$\left[\boldsymbol{T} = \left(\boldsymbol{I}^{-\alpha_E} \boldsymbol{DS} \right)^{1/(\gamma\alpha_E + \alpha_T)} \right]_{\text{<dataset name>}}$$

(Ross, 2008a), (Ross, 2008b), (Valerdi, et al, 2009), (Ross, 2011b)

Software CDER Solving for Intensity

Cost (Effort) Estimating Relationship (CER)

$$\left[I = \left(E^{\gamma\alpha_E + \alpha_T} \left(DS \right)^{-\gamma} \right)^{1/\alpha_t} \right]_{\text{<dataset name>}}$$

Schedule (Duration) Estimating Relationship (SER)

$$\left[I = \left(T^{\gamma\alpha_E + \alpha_T} \left(DS \right)^{-1} \right)^{-1/\alpha_E} \right]_{\text{<dataset name>}}$$

(Ross, 2008a), (Ross, 2008b), (Valerdi, et al, 2009), (Ross, 2011b)

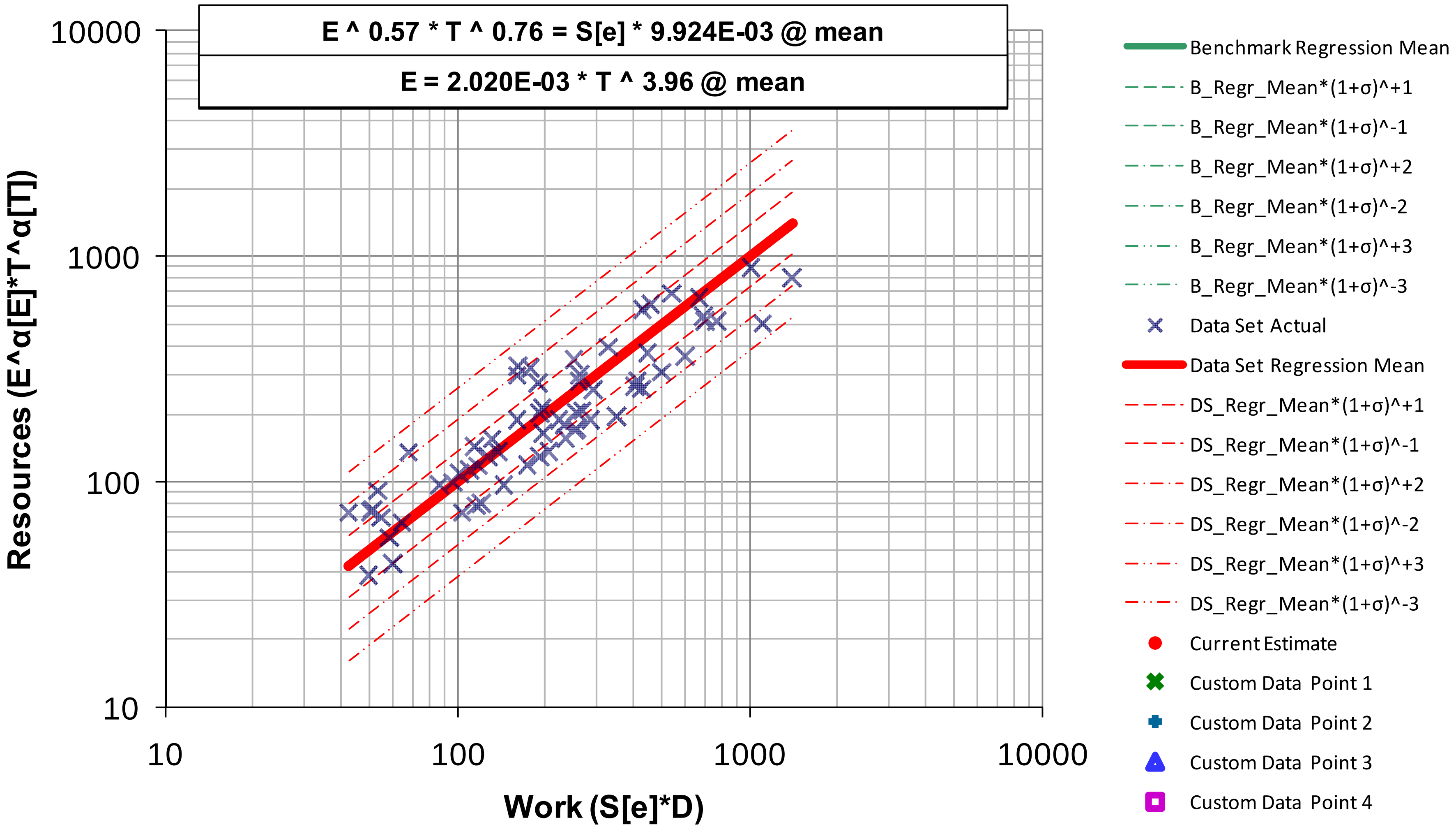
TECOLOTE RESEARCH, INC.

Example CDER Regression Results

Aerospace 2004: Military Mobile Operational

Resources vs Work

n=64 Factor MPE-ZPB [x,y]: y=(1.000E+00)*x SEE=38% BIAS=0% R^2=0.89 PRED(25)=42% MMRE=30%

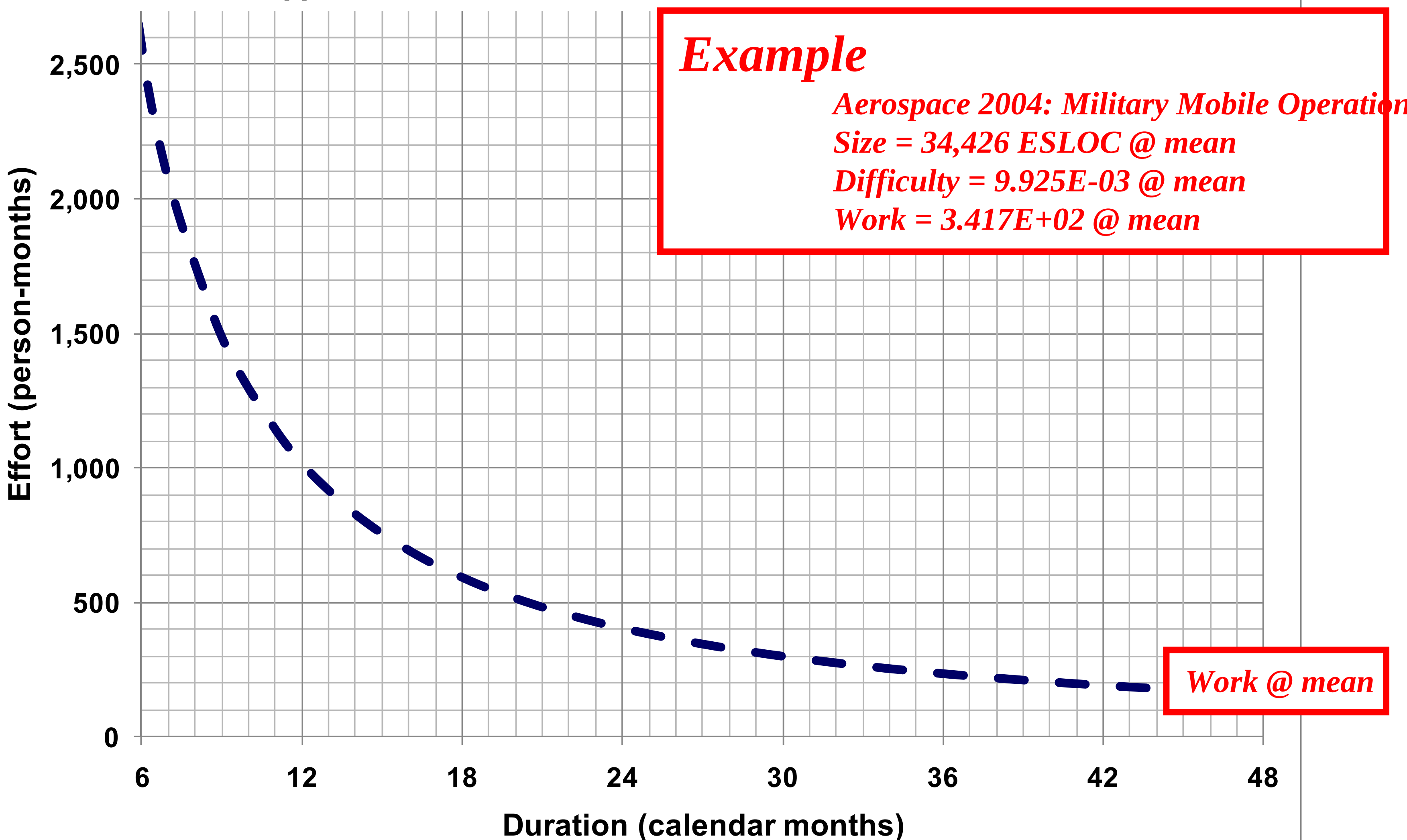


TECOLOTE RESEARCH, INC.

Work Function (CDER) at a Particular Size and Difficulty

Effort versus Duration

S[e] = <SEER PERT, L, M, H> = <SEER PERT, 14,650, 14,650, 92,073>

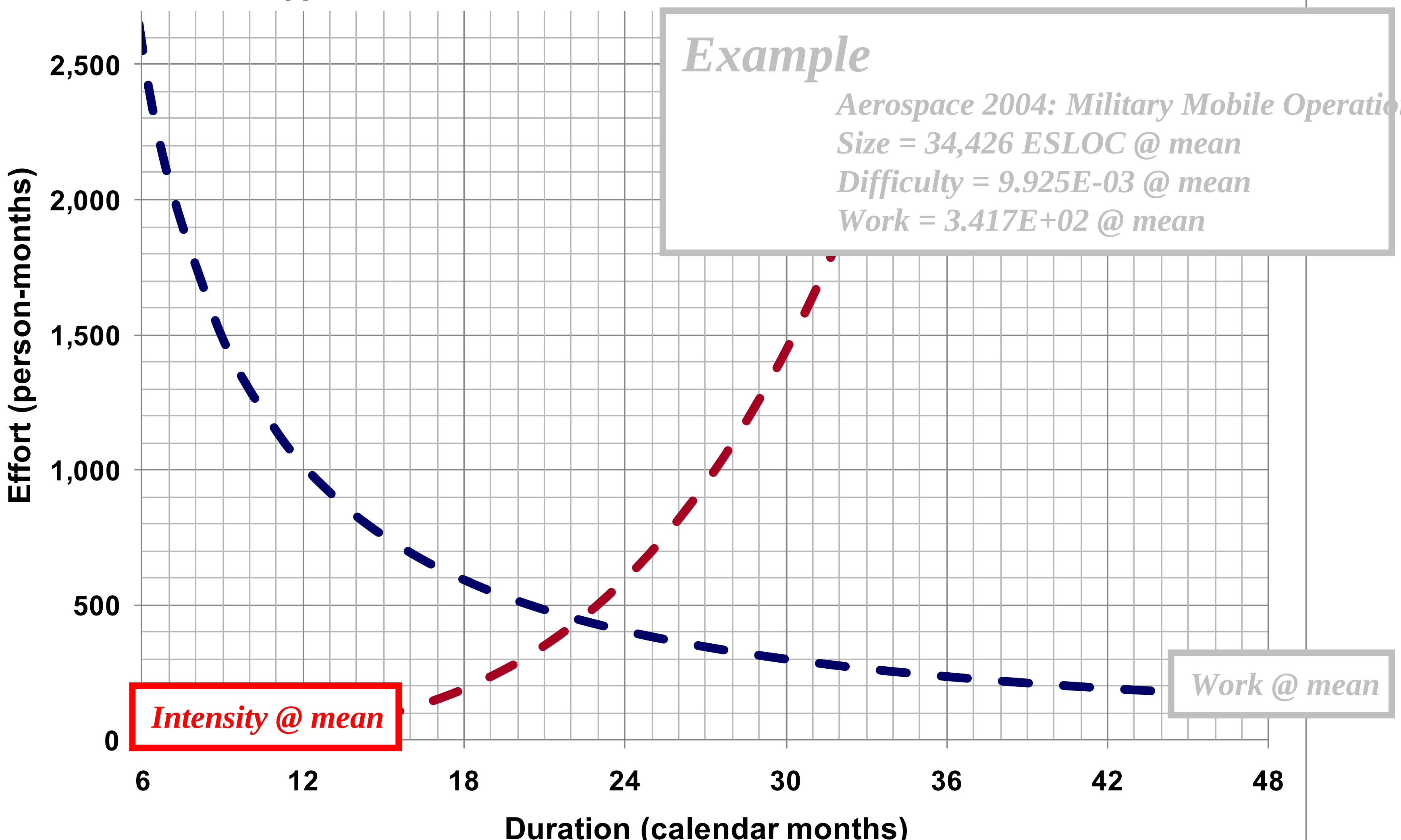


TECOLOTE RESEARCH, INC.

Work Function (CDER) at a Particular Intensity

Effort versus Duration

S[e] = <SEER PERT, L, M, H> = <SEER PERT, 14,650, 14,650, 92,073>

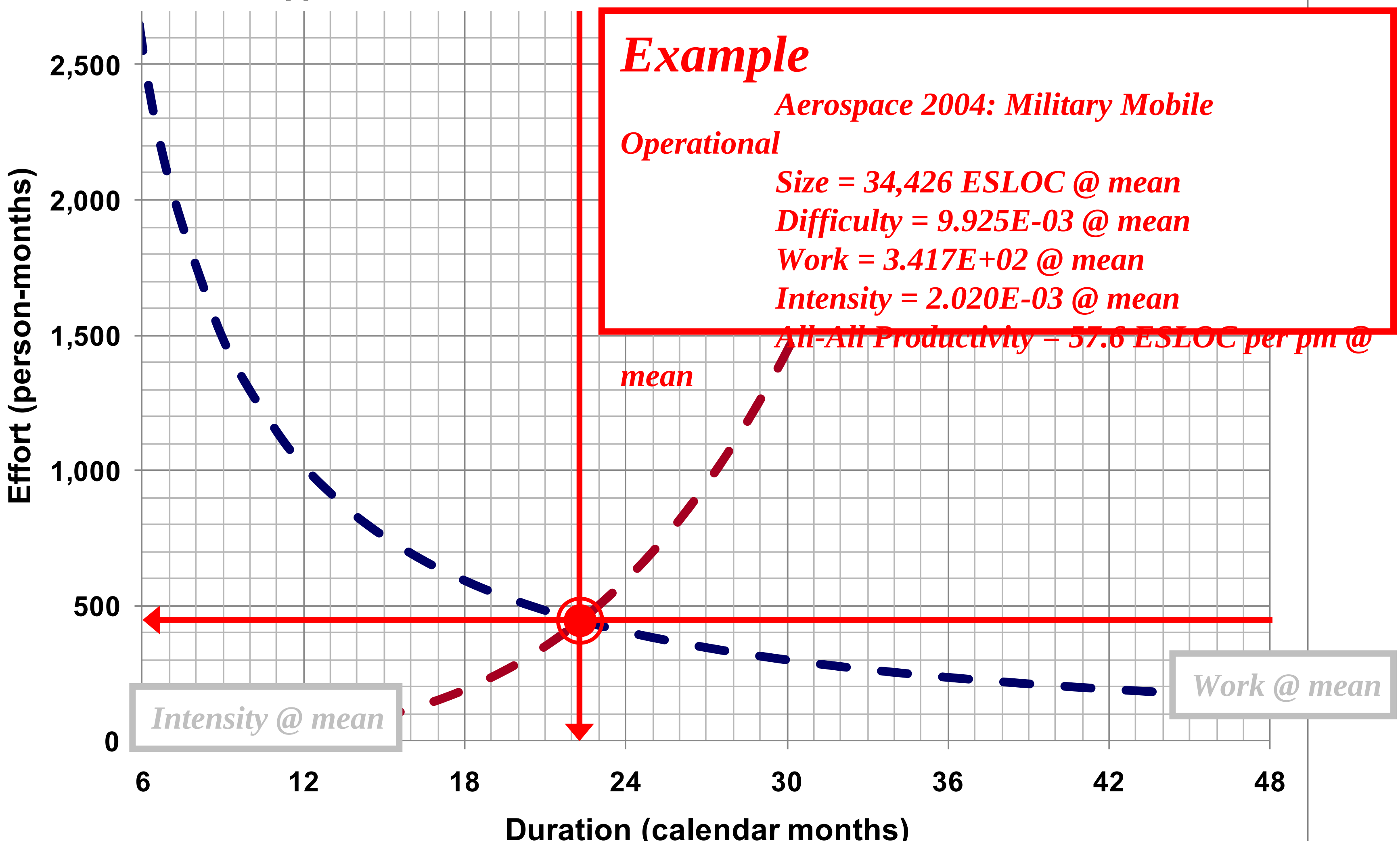


TECOLOTE RESEARCH, INC.

Work Function (CDER) at a Particular Intensity

Effort versus Duration

S[e] = <SEER PERT, L, M, H> = <SEER PERT, 14,650, 14,650, 92,073>



TECOLOTE RESEARCH, INC.

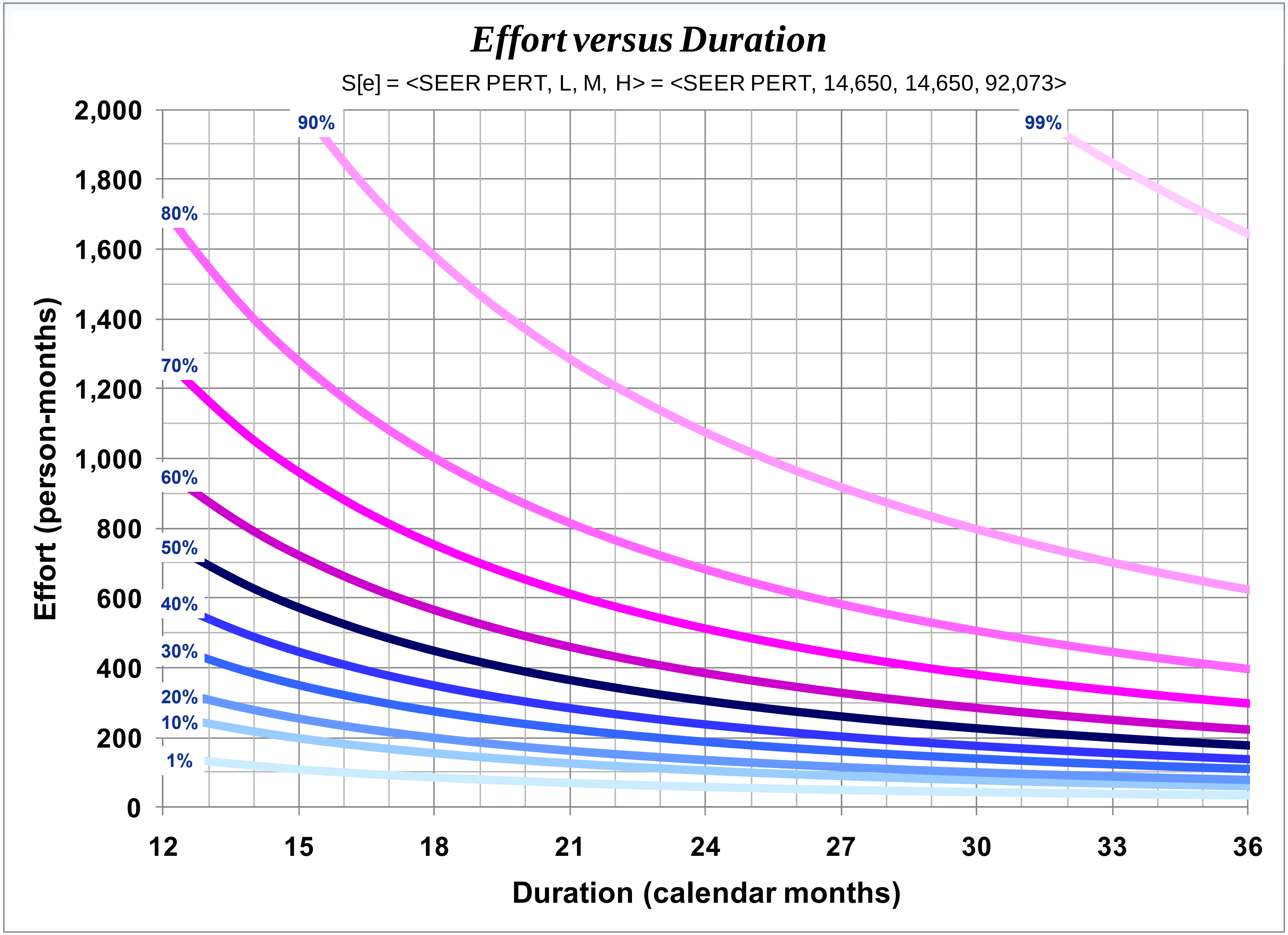
Size * Difficulty Product: a recipe for Monte Carlo

- Express each of Size (S) and Difficulty (D) as a distribution
- Create a shuffled list of appropriately-distributed possible outcomes for each of Size and Difficulty
- Multiply each element in the shuffled Size list by its adjacent element in the shuffled Difficulty list to get a list of Size * Difficulty products (Ψ)
- Sort the Size*Difficulty product list in ascending order
- Compute the quantile of each element in the sorted list; the result is the Cumulative Distribution Function (CDF or S-curve) of the Size-Difficulty product random variable Ψ

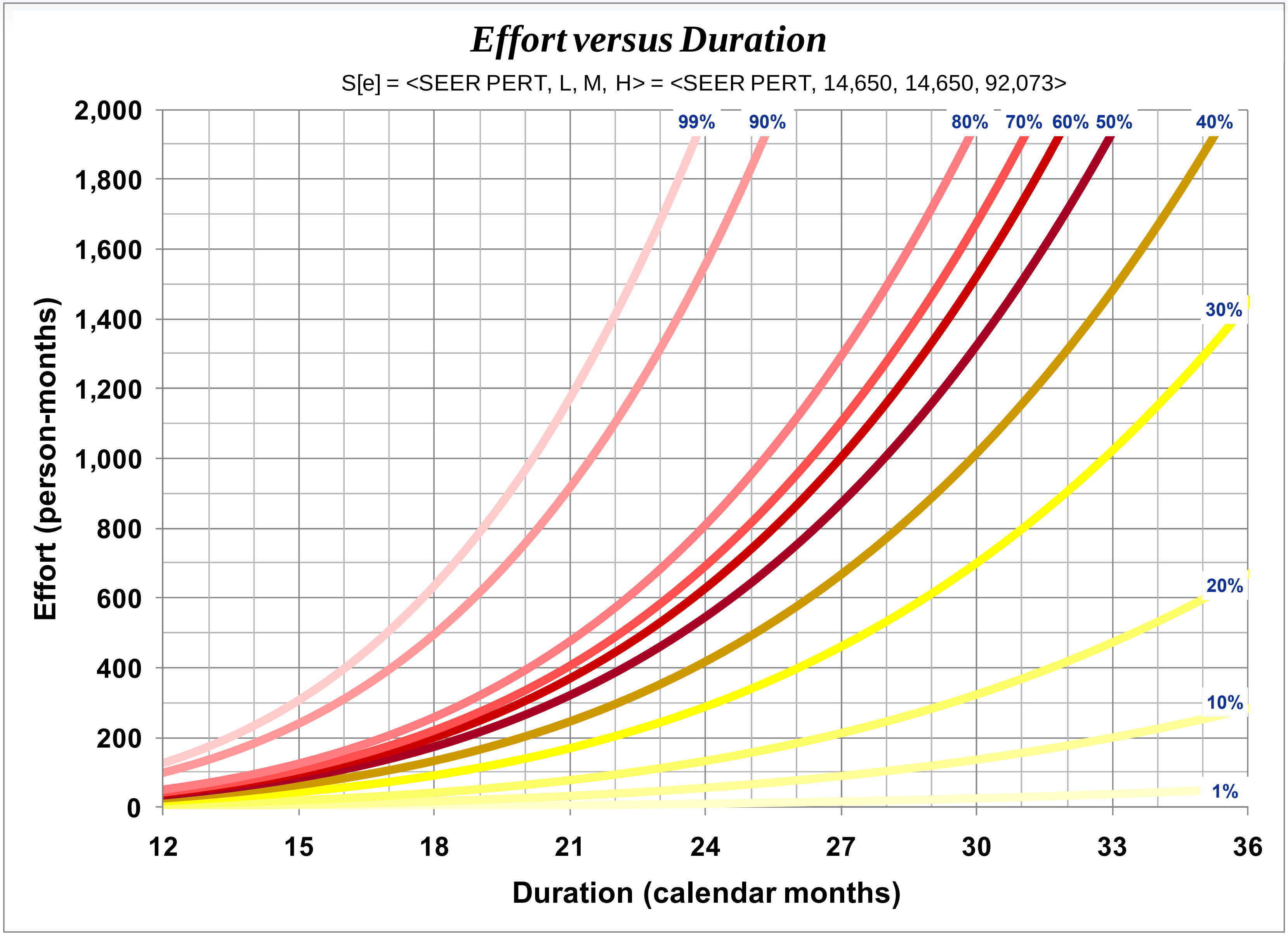
(Ross, 2011b)

TECOLOTE RESEARCH, INC.

Work Function as a Probability Field

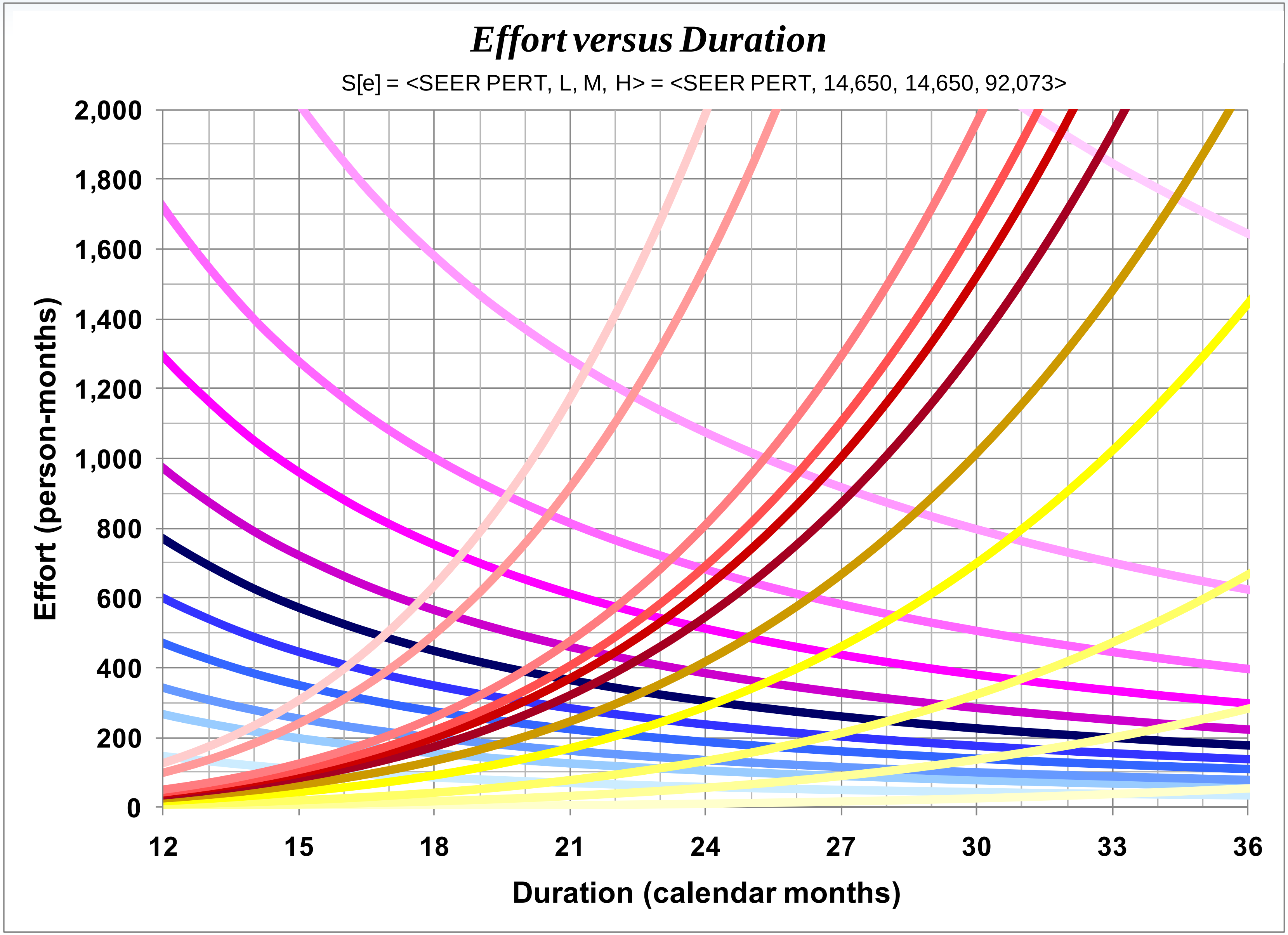


Intensity Function as a Probability Field



TECOLOTE RESEARCH, INC.

Work and Intensity Probability Fields



TECOLOTE RESEARCH, INC.

Satisfying Goals: Solving for Intensity

Finding the Intensity that Satisfies a Cost (Effort) Goal

$$\left[I_{effort\_constraint@p\%} = \left(\frac{E_{effort\_constraint@p\%}^{\gamma\alpha_E+\alpha_T}}{\boldsymbol{F}_{\boldsymbol{\Psi}}^{-1}(p)^{\gamma}} \right)^{1/\alpha_T} \right]_{\text{<data set name>}}$$

Finding the Intensity that Satisfies a Schedule (Duration) Goal

$$\left[I_{duration\_constraint@p\%} = \left(\frac{\boldsymbol{F}_{\boldsymbol{\Psi}}^{-1}(p)}{t_{duration\_constraint@p\%}^{\gamma\alpha_E+\alpha_T}} \right)^{1/\alpha_E} \right]_{\text{<data set name>}}$$

Note: $\boldsymbol{\Psi} \equiv$ size*difficulty product random variable

$=$ convolved ratio of random variables $\boldsymbol{S}$ (size) and $\boldsymbol{D}$ (difficulty)

Note: $\boldsymbol{F}_{\boldsymbol{\Psi}}^{-1}(p) \equiv$ inverse CDF of random variable $\boldsymbol{\Psi}$ at probability p

(Ross, 2008a)

TECOLOTE RESEARCH, INC.

Joint and Conditional Probability Defined

- **Joint**
  - The probability that effort will be less than or equal to some given (goal) value and duration will be less than or equal to some given (goal) value

$$\boldsymbol{P}\left(Effort \le x_1, Duration \le x_2\right)$$

- **Conditional**
  - The probability that effort will be less than or equal to some given (goal) value based on some point estimate of duration being equal to some given (goal) value or vice versa

$$\boldsymbol{P}\left(Effort \le x_1 \mid Duration = x_2\right)$$

or

$$\boldsymbol{P}\left(Duration \le x_1 \mid Effort = x_2\right)$$

(Garvey, 2000)

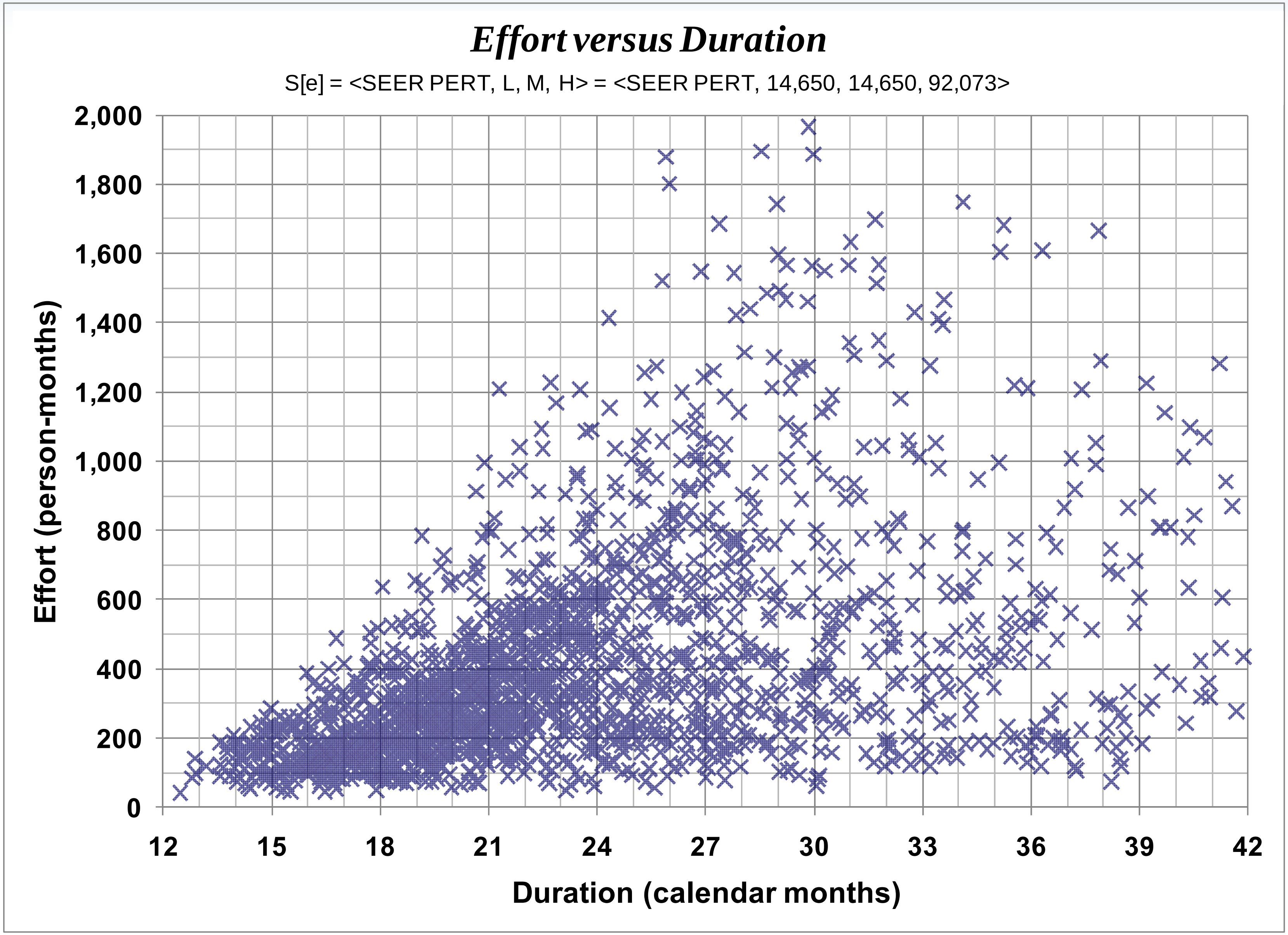
Garvey Question #1

“What is the chance the system can be delivered within cost and schedule?”

(Garvey, 2000)

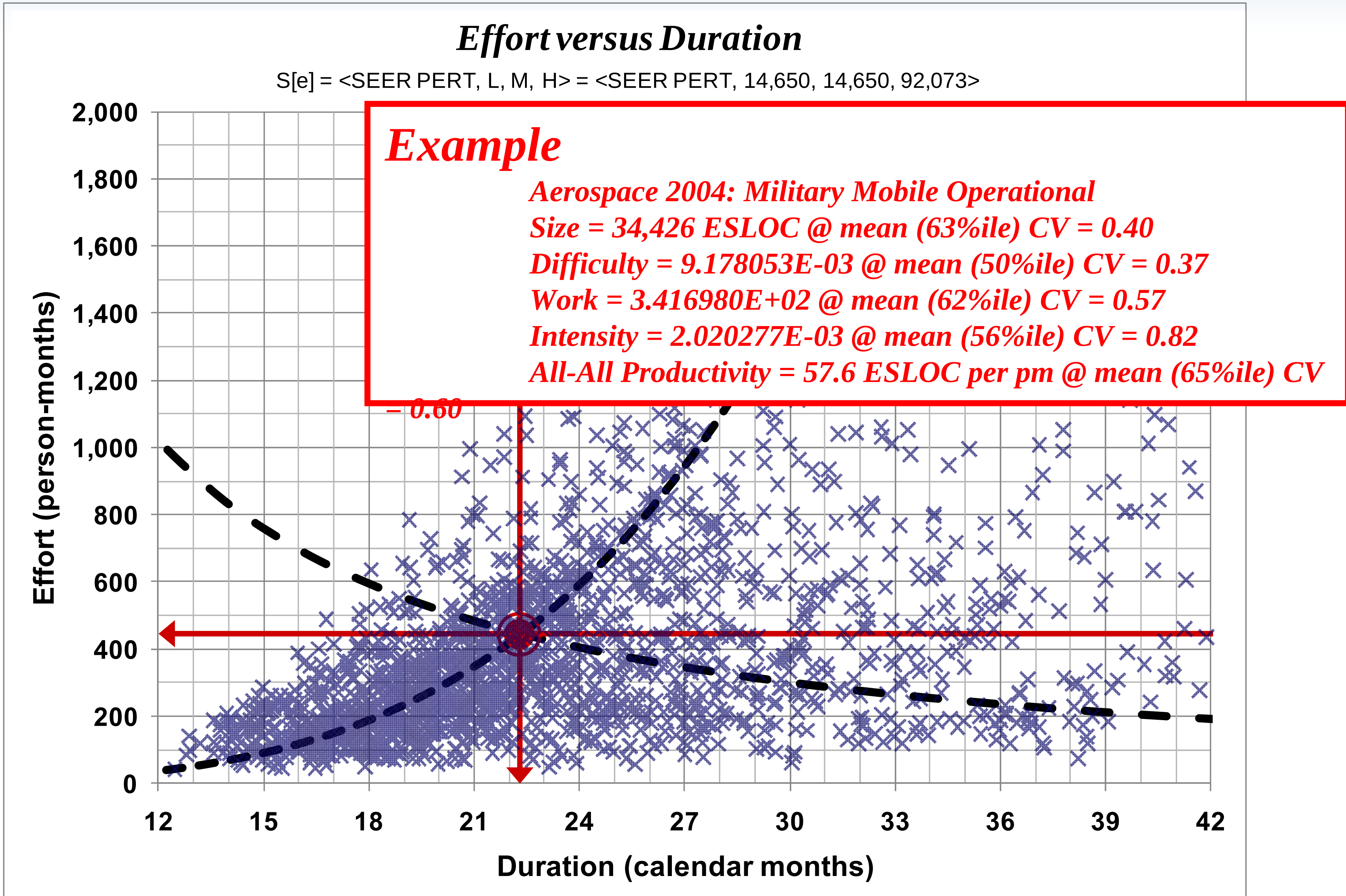
TECOLOTE RESEARCH, INC.

Garvey Question #1
Scatter Diagram Approach



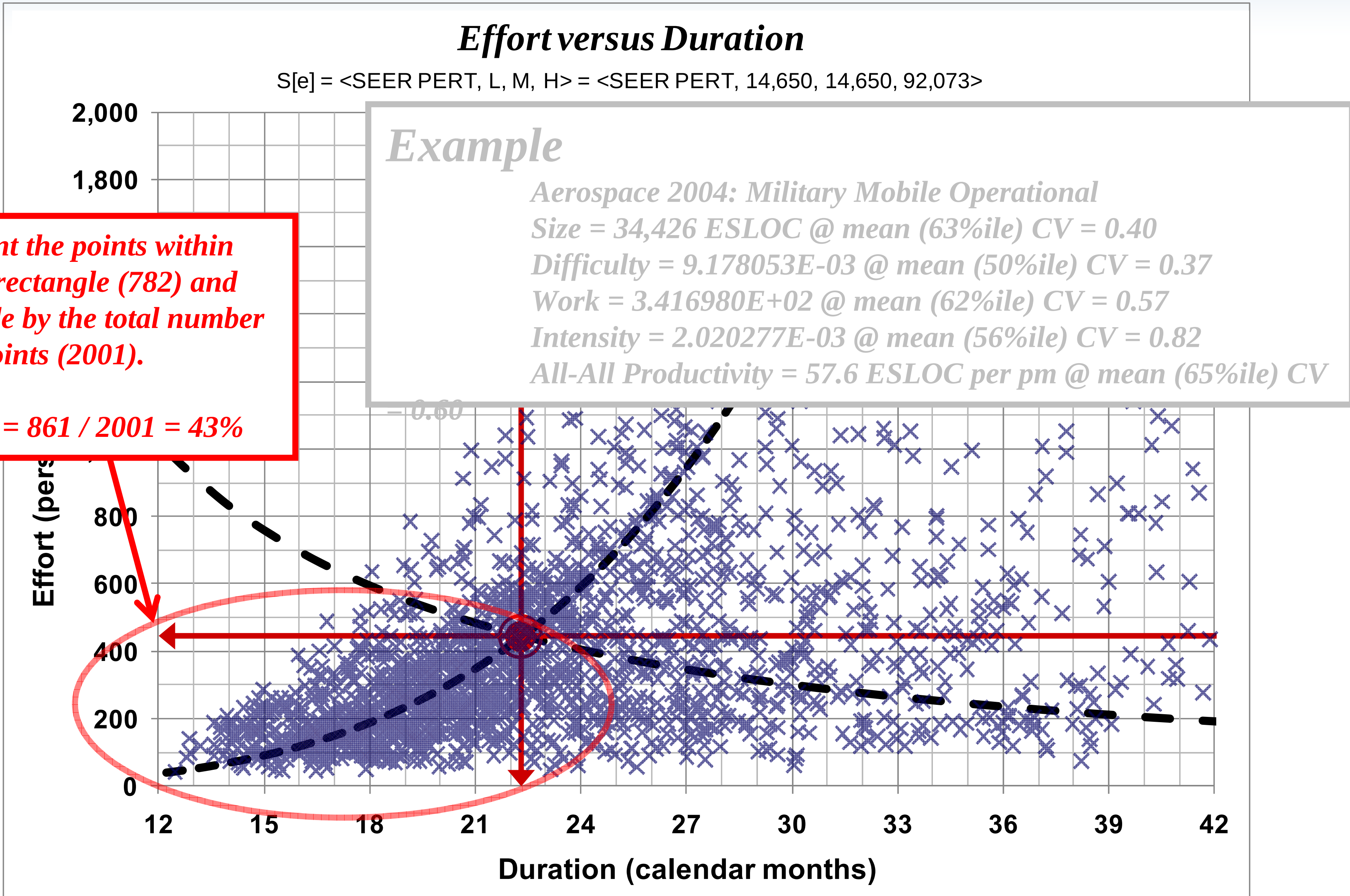
TECOLOTE RESEARCH, INC.

Garvey Question #1
Scatter Diagram Approach



TECOLOTE RESEARCH, INC.

Garvey Question #1
Scatter Diagram Approach



Garvey Question #1
Boolean Random Variable Approach

$$\boldsymbol{J} \equiv \left(\boldsymbol{E} \leq \hat{E}\right) \mathrm{I} \left(\boldsymbol{T} \leq \hat{T}\right)$$

$$F_{\boldsymbol{J}}\left(x \mid \boldsymbol{J}\right) \equiv \begin{cases} 1 - \dfrac{\sum \boldsymbol{J}}{\boldsymbol{count}\left(\boldsymbol{J}\right)} & x = 0 \ \left(\text{FALSE}\right) \\ \dfrac{\sum \boldsymbol{J}}{\boldsymbol{count}\left(\boldsymbol{J}\right)} & x = 1 \ \left(\text{TRUE}\right) \end{cases}$$

$$F_{\boldsymbol{J}}\left(\text{TRUE}\right) = 861/2001 = 43\%$$

$$\therefore \text{Joint Confidence Level} = 43\%$$

Garvey Question #1: Example Answer

"What is the chance the system can be delivered within cost and schedule?"

Answer: Joint Confidence Level = 43%

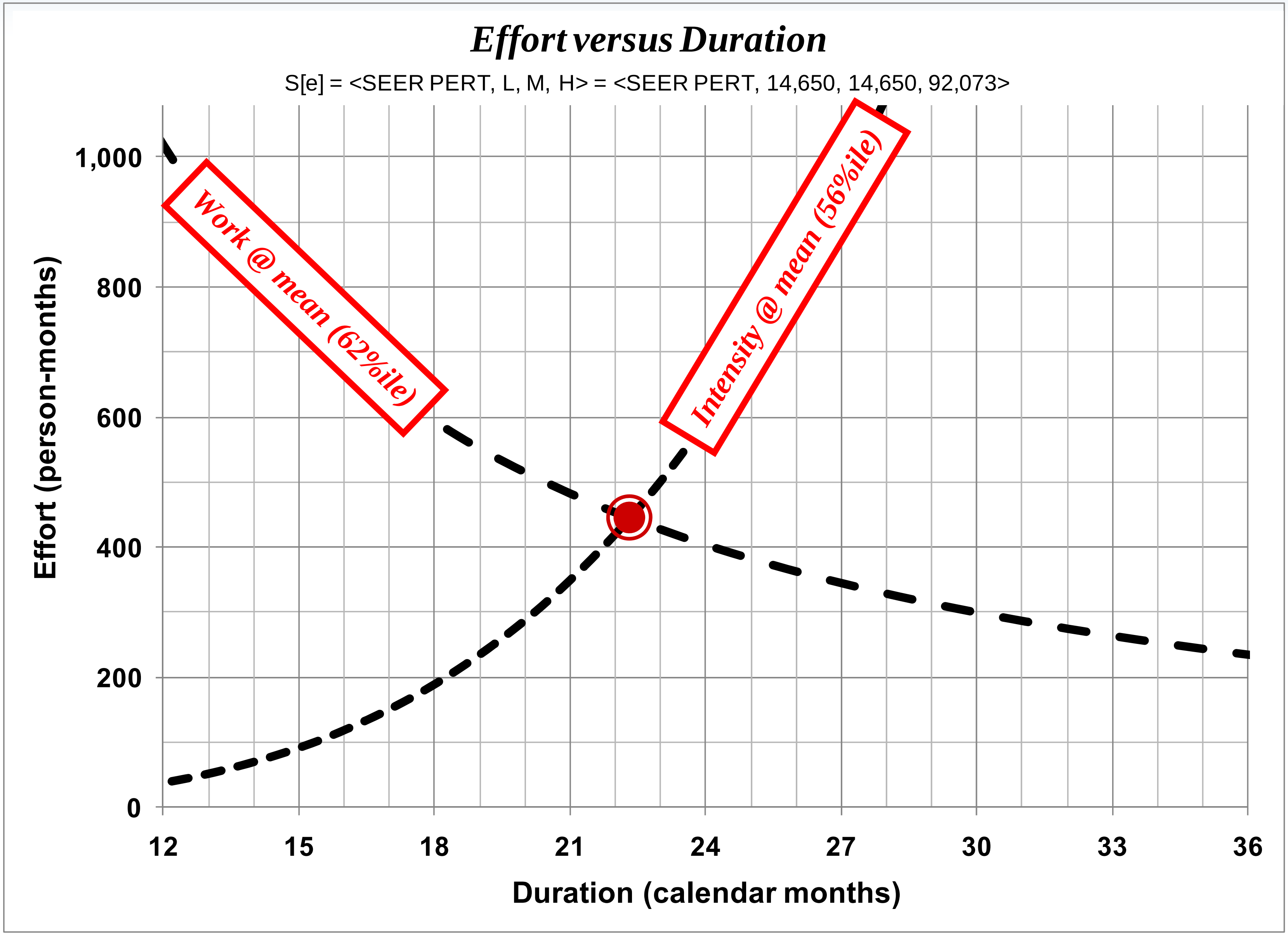
Garvey Question #2

"How likely might the point estimate cost be exceeded for a given schedule?"

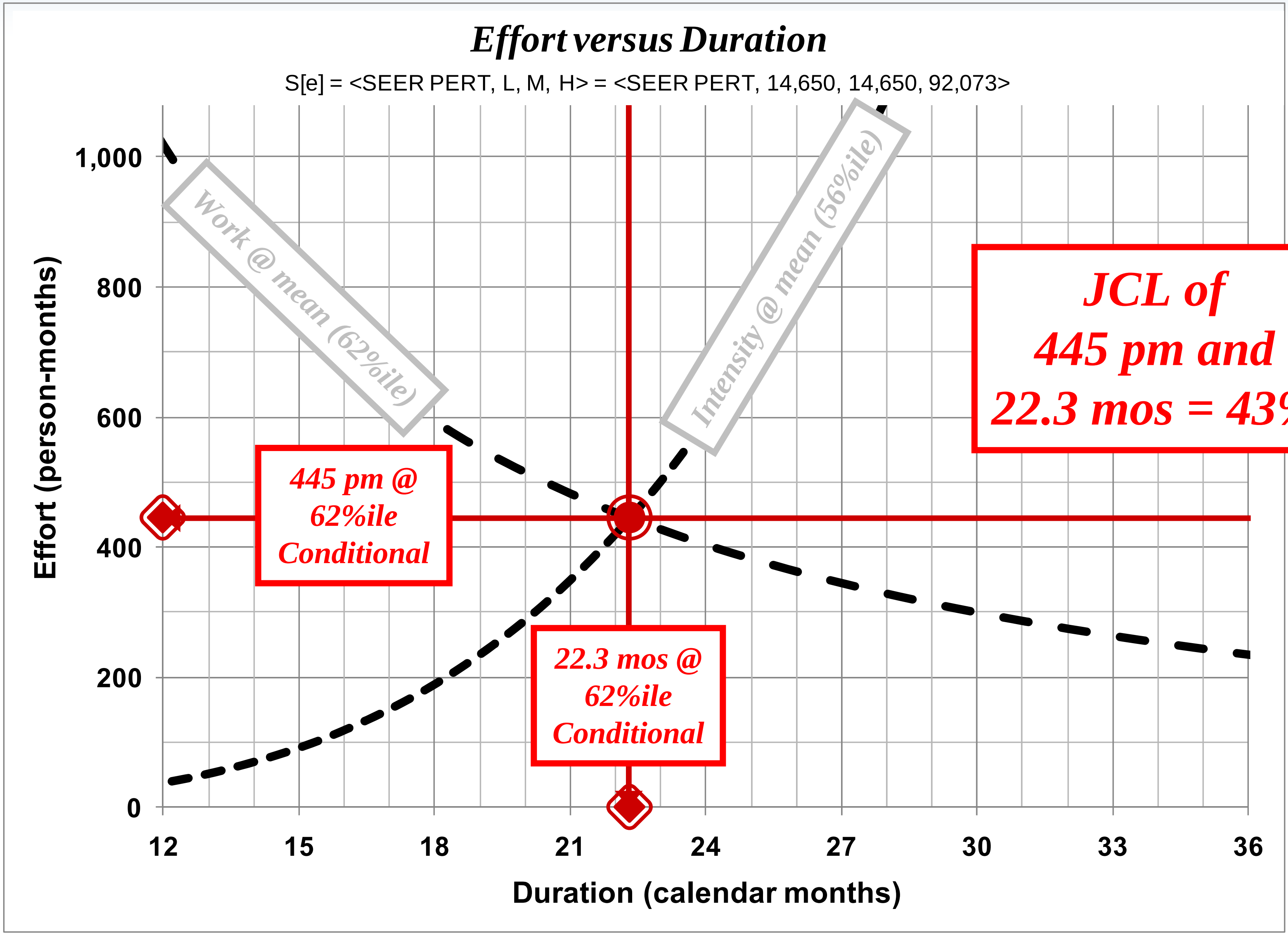
(Garvey, 2000)

TECOLOTE RESEARCH, INC.

Locate Point Estimate Position

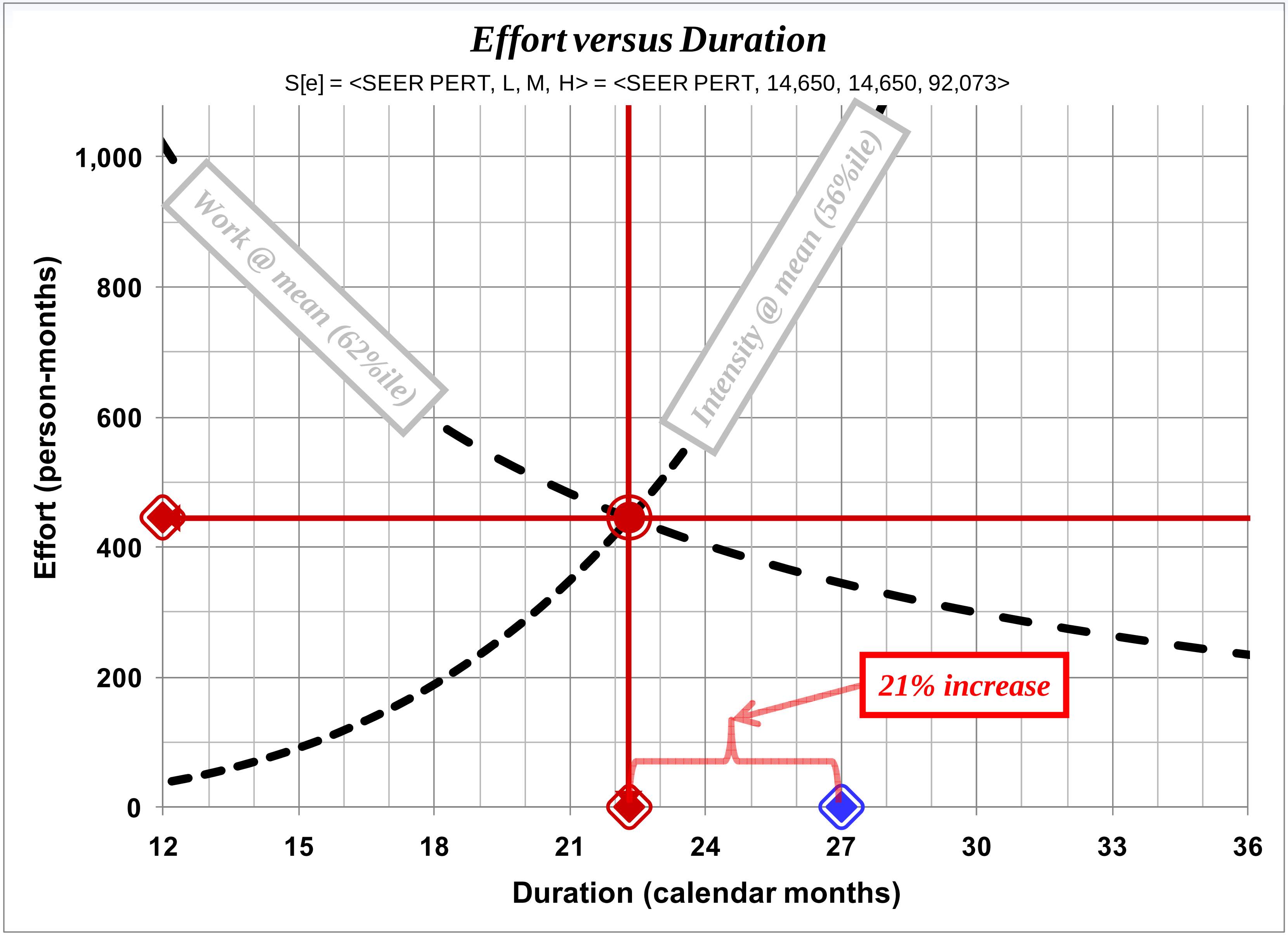


Point Estimate Values



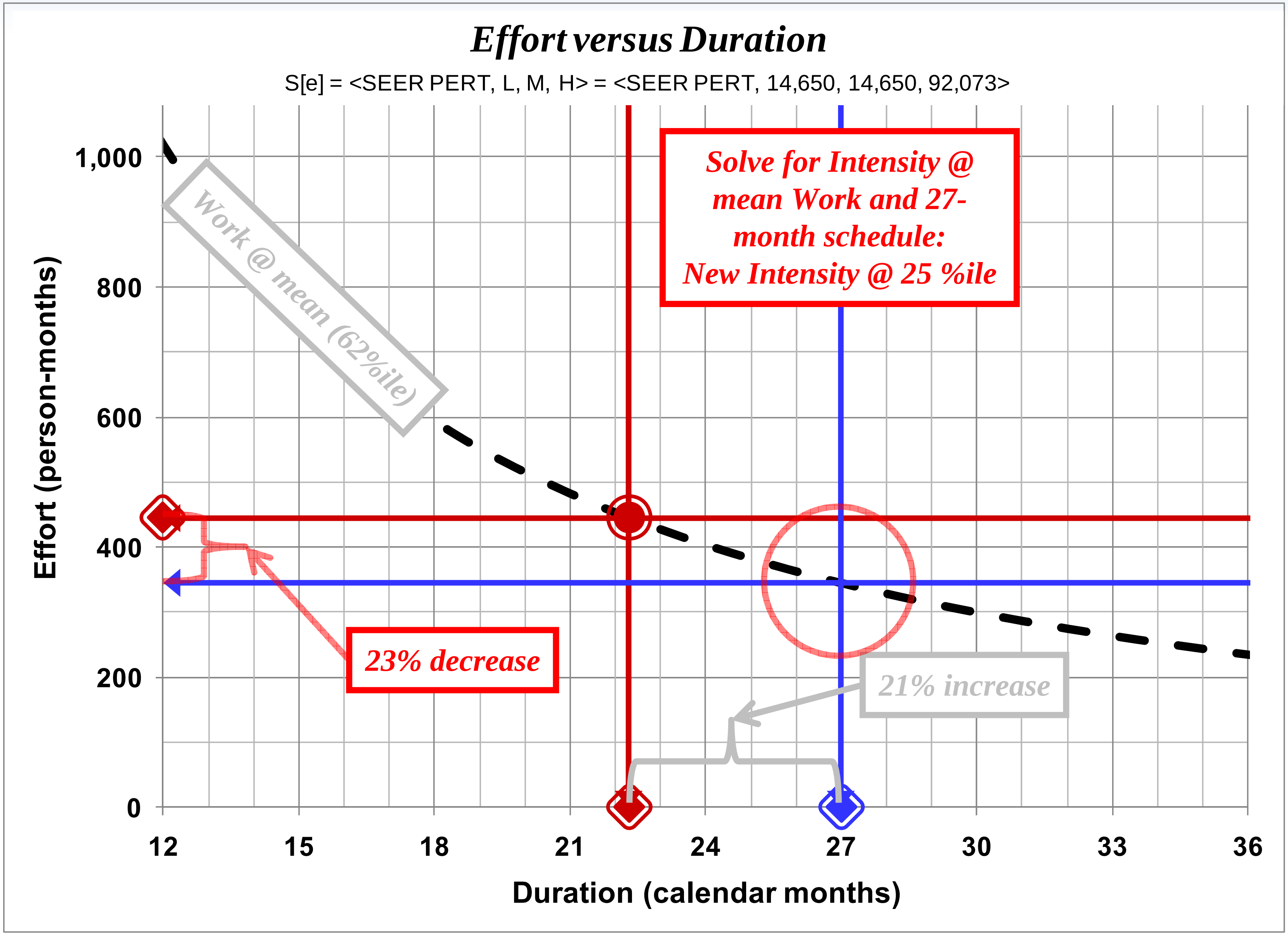
TECOLOTE RESEARCH, INC.

Establish Updated Schedule Constraint



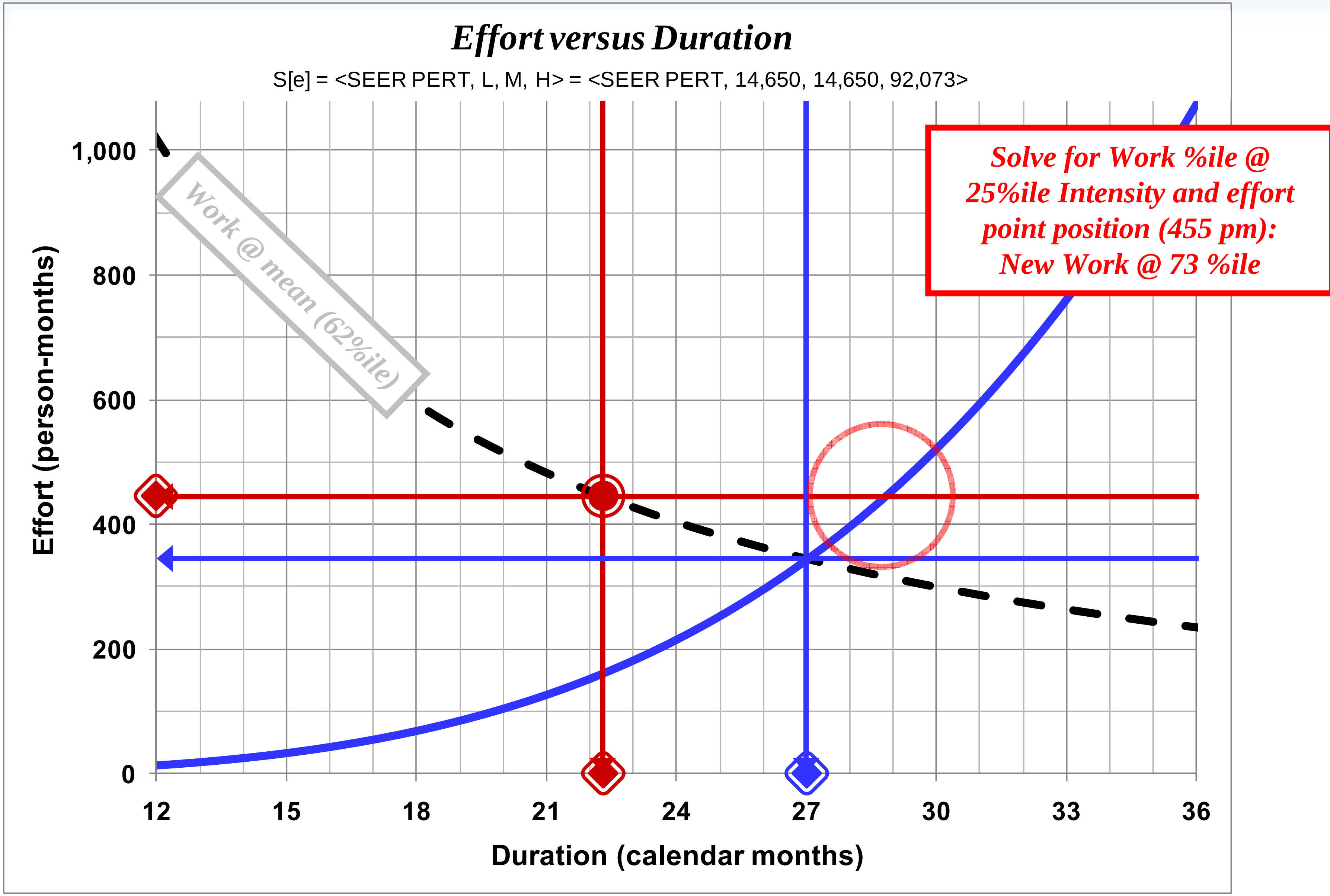
TECOLOTE RESEARCH, INC.

Locate Position for Intensity Update



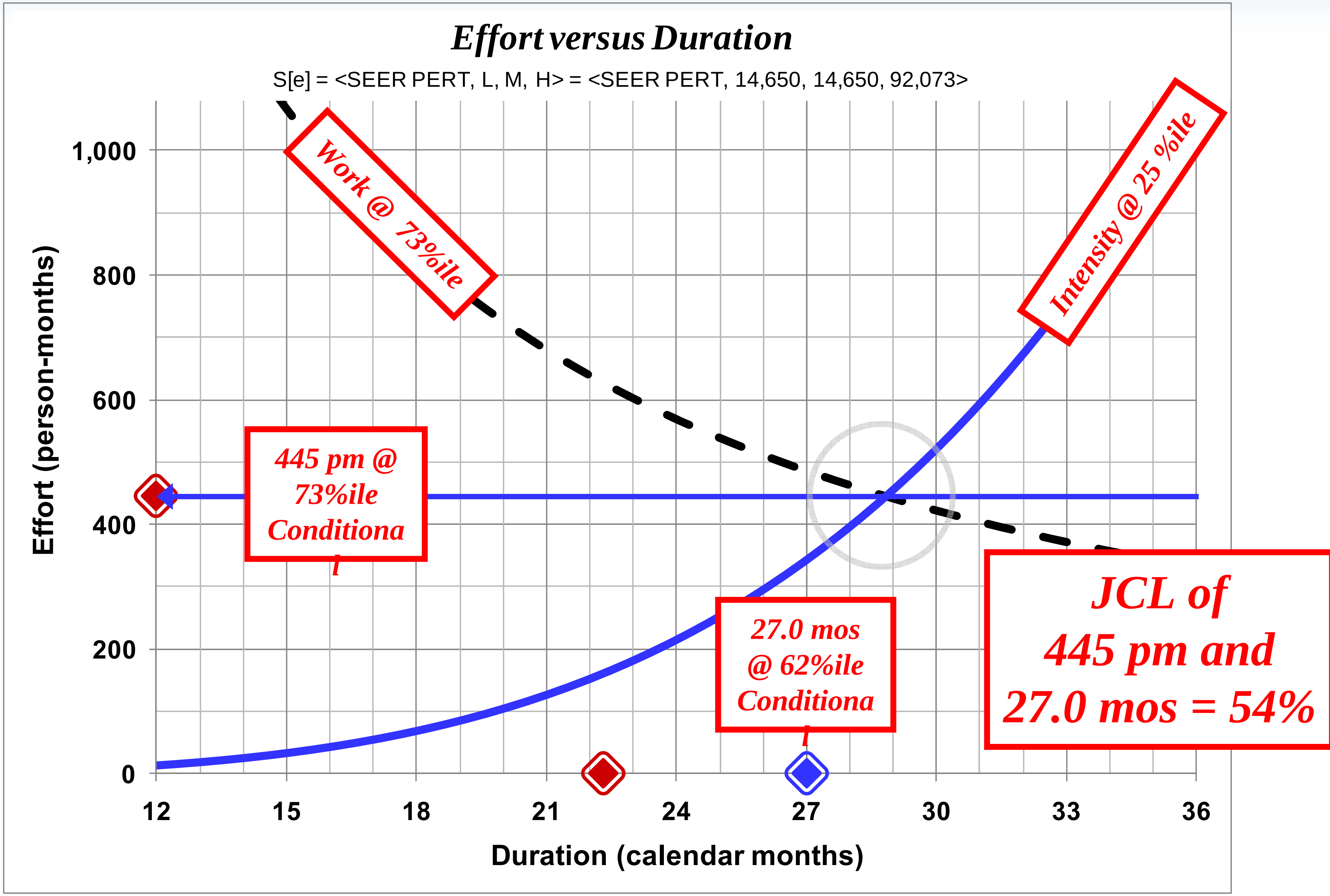
TECOLOTE RESEARCH, INC.

Update Intensity Location



TECOLOTE RESEARCH, INC.

Solution



TECOLOTE RESEARCH, INC.

Garvey Question #2: Mathematical Model

$$T_{point} \equiv 22.3 \quad p_{T_{point}} \equiv 52\% \quad T_1 \equiv 27.0 \quad p_{T_1} \equiv 52\% \quad E_{point} \equiv 445 \quad p_{E_{Point}} \equiv 52\%$$

$$\left[I_1 = \left(\frac{\boldsymbol{F}_{\boldsymbol{\Psi}}^{-1}\left(p_{T_1}\right)}{T_1^{\gamma\alpha_E+\alpha_T}} \right)^{1/\alpha_E} = \left(\frac{\boldsymbol{F}_{\boldsymbol{\Psi}}^{-1}\left(52\%\right)}{22.3^{3.96(0.57+0.76)}} \right)^{1/0.57} = 7.321961\text{E}-04 \right]_{<*>}$$

$$\left[\boldsymbol{E}_{I_1} = \left(I_1^{\alpha_T} \boldsymbol{\Psi}^{\gamma} \right)^{1/(\gamma\alpha_E+\alpha_T)} = \left(I_1^{0.76} \boldsymbol{\Psi}^{3.96} \right)^{1/(3.96(0.57+0.76))} \right]_{<*>}$$

$$\therefore \boldsymbol{P}\left(\boldsymbol{E}_{I_1} > E_{point}\right) = 1 - \boldsymbol{F}_{\boldsymbol{E}_{I_1}}\left(E_{point}\right) = 1 - \boldsymbol{F}_{\boldsymbol{E}_{I_1}}\left(445\right) = 1 - 73\% = 27\%$$

Note: <*> indicates Aerospace 2004: Military Ground Operational

Note: $\boldsymbol{F}_{\boldsymbol{X}}\left(x\right) \equiv$ CDF of random variable $\boldsymbol{X}$ at value x

Note: $\boldsymbol{F}_{\boldsymbol{X}}^{-1}\left(p\right) \equiv$ inverse CDF of random variable $\boldsymbol{X}$ at probability p

TECOLOTE RESEARCH, INC.

Garvey Question #2: Example Answer

"How likely might the point estimate cost be exceeded for a given schedule?"

Original Position: 100% − 62% = 38% → JCL = 43%

Answer: 100% − 73% = 27% → JCL = 54%

TECOLOTE RESEARCH, INC.

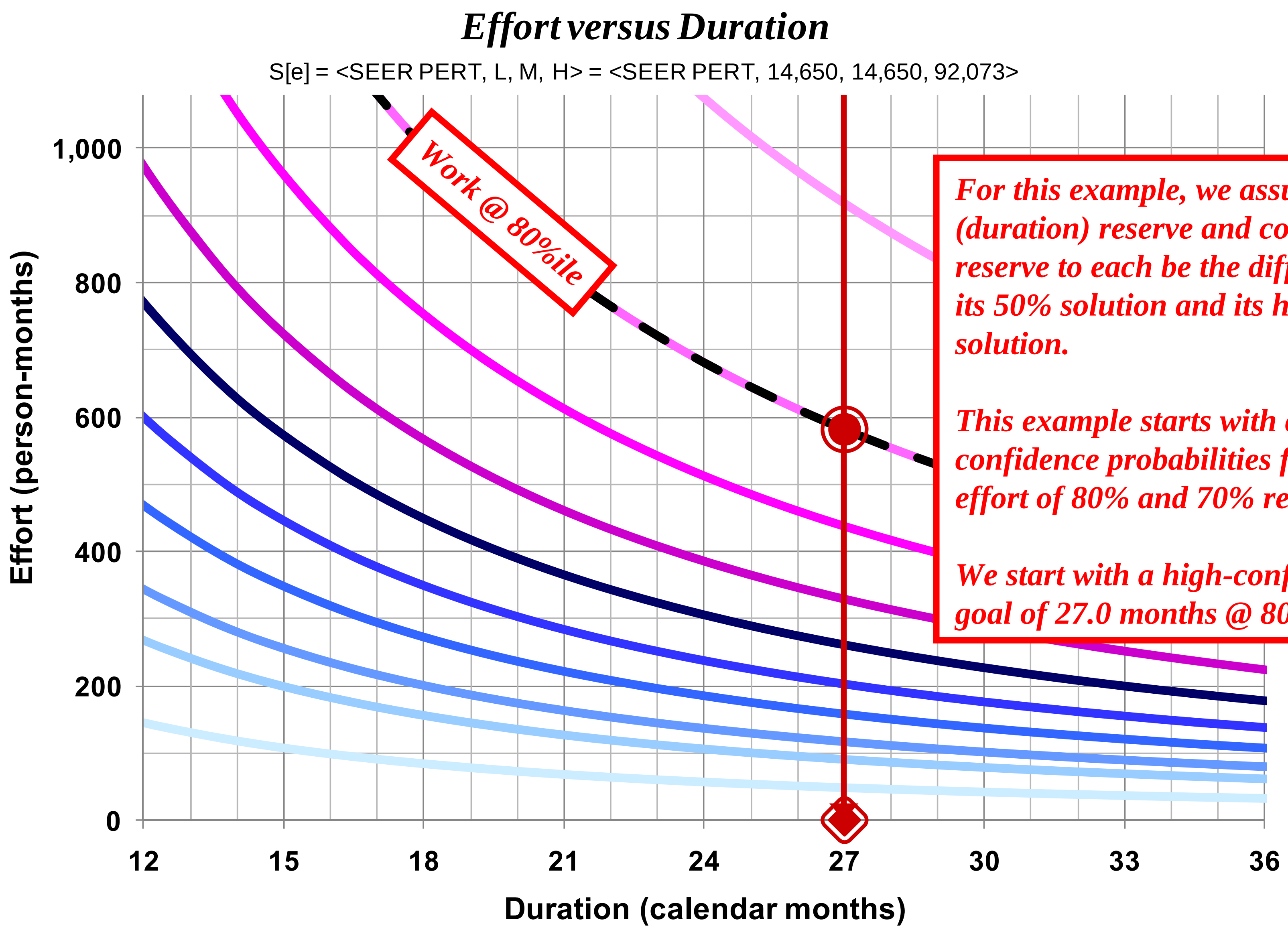
Garvey Question #3

"How are cost reserve recommendations affected by schedule risk?"

(Garvey, 2000)

TECOLOTE RESEARCH, INC.

Locate Appropriate Work Curve



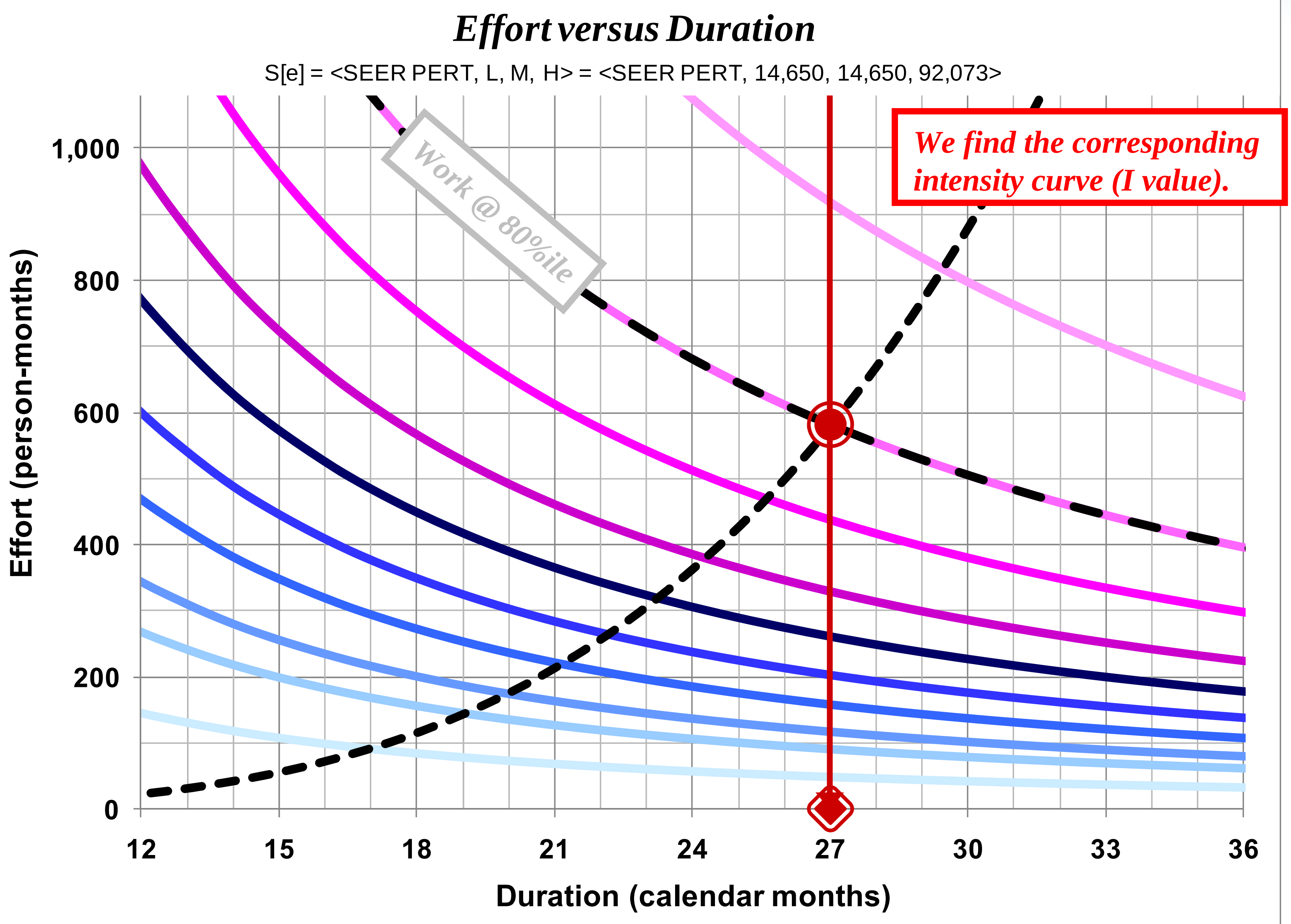
For this example, we assume schedule (duration) reserve and cost (effort) reserve to each be the difference between its 50% solution and its high-confidence solution.

This example starts with desired high-confidence probabilities for duration and effort of 80% and 70% respectively.

We start with a high-confidence duration goal of 27.0 months @ 80%.

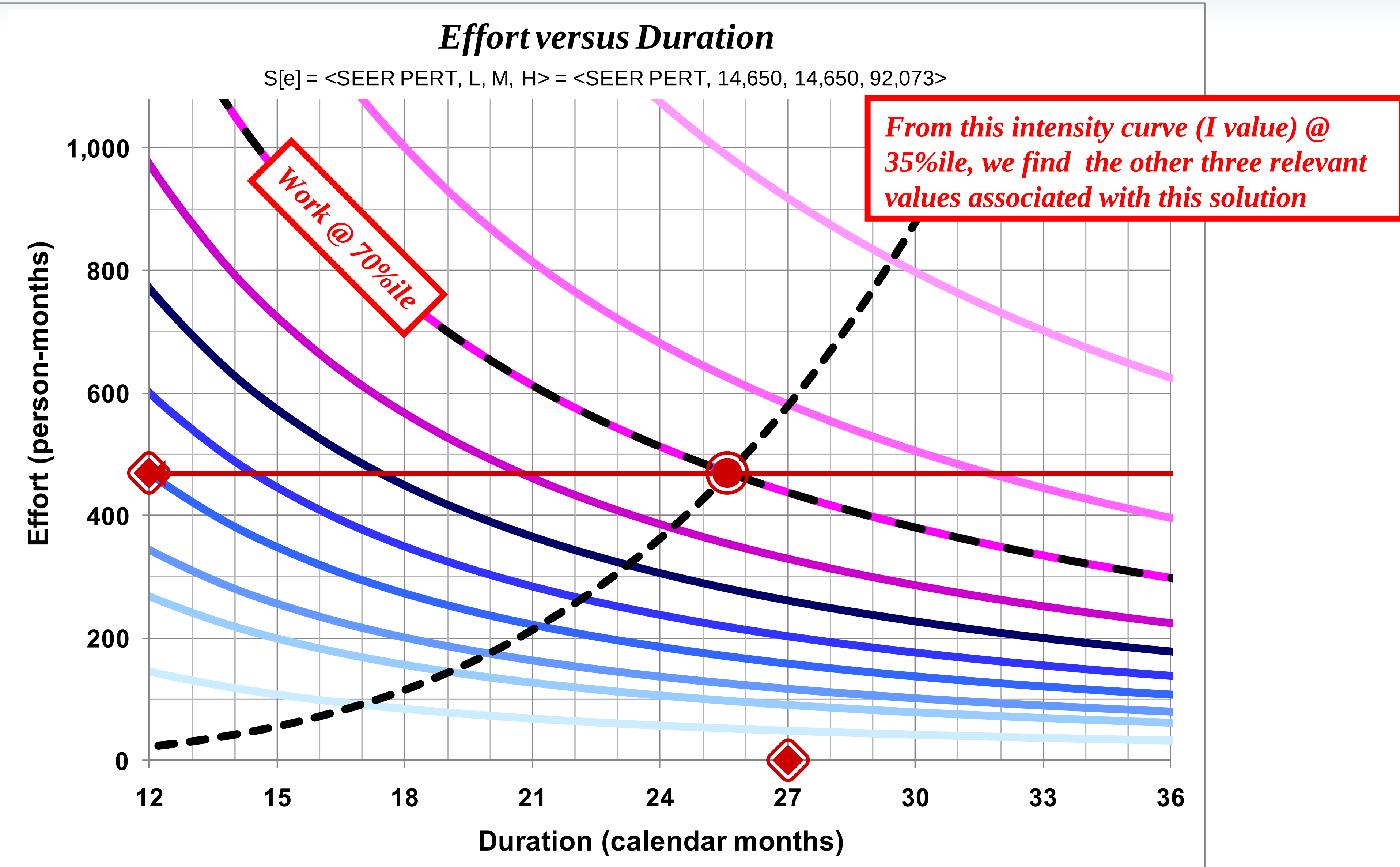
TECOLOTE RESEARCH, INC.

Locate Corresponding Intensity Curve



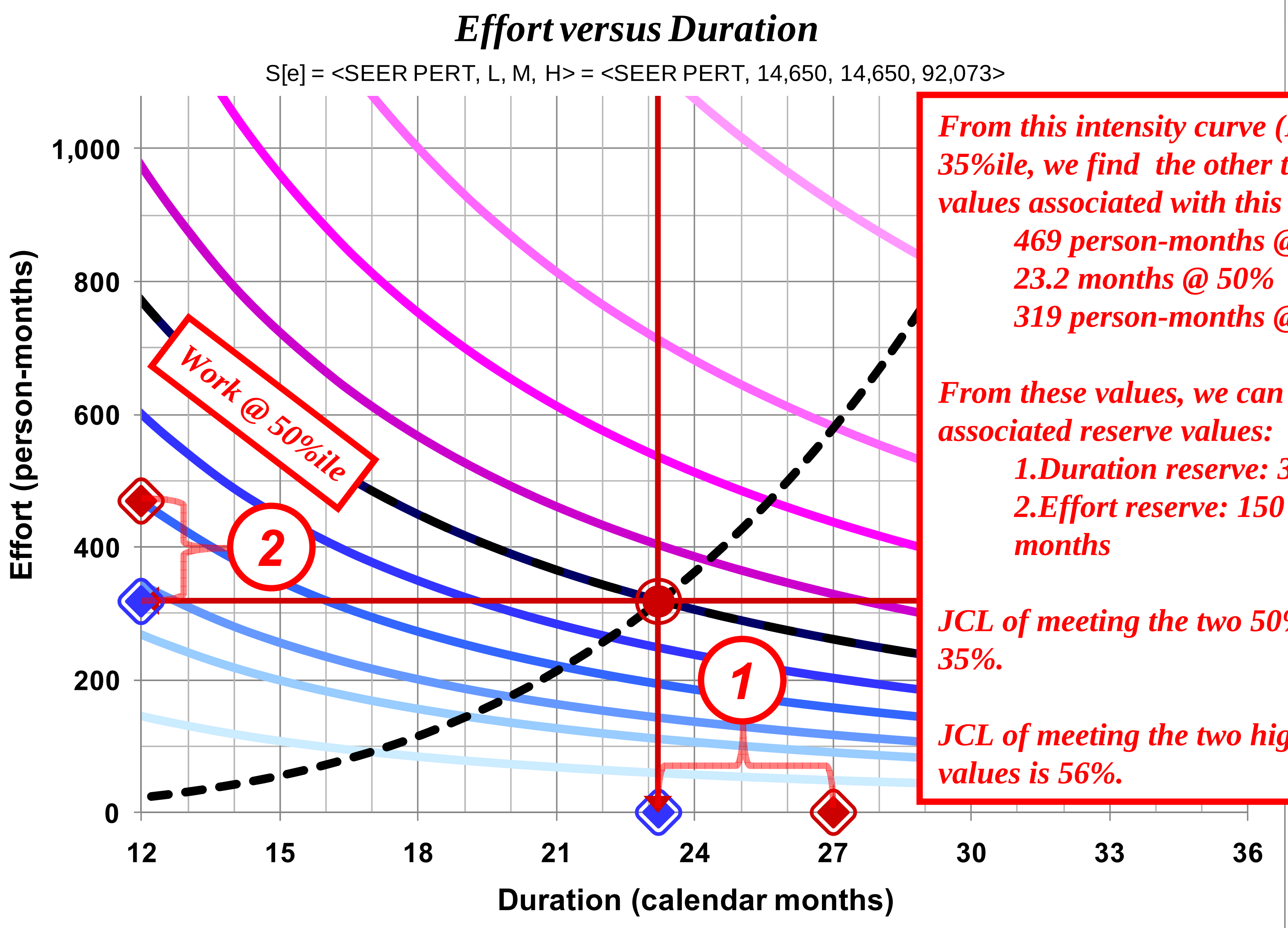
TECOLOTE RESEARCH, INC.

Locate and Project Constrained Effort Solution



TECOLOTE RESEARCH, INC.

Locate and Project Constrained 50% Confidence Solutions



From this intensity curve (I value) @ 35%ile, we find the other three relevant values associated with this solution:

469 person-months @ 70%
23.2 months @ 50%
319 person-months @ 50%

From these values, we can find the associated reserve values:

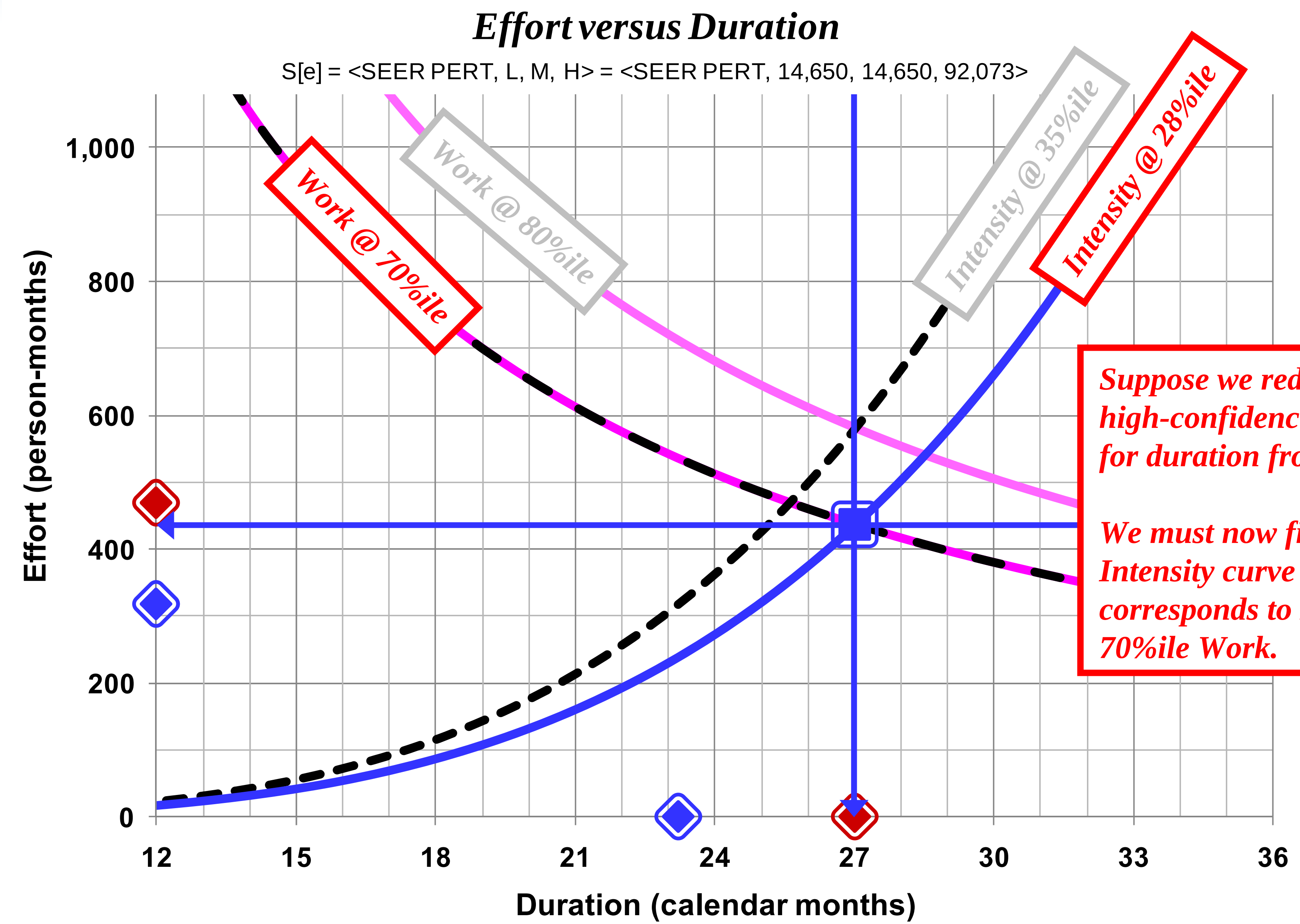
1.Duration reserve: 3.8 months
2.Effort reserve: 150 person-months

JCL of meeting the two 50% values is 35%.

JCL of meeting the two high-confidence values is 56%.

TECOLOTE RESEARCH, INC.

Update Intensity Curve Location

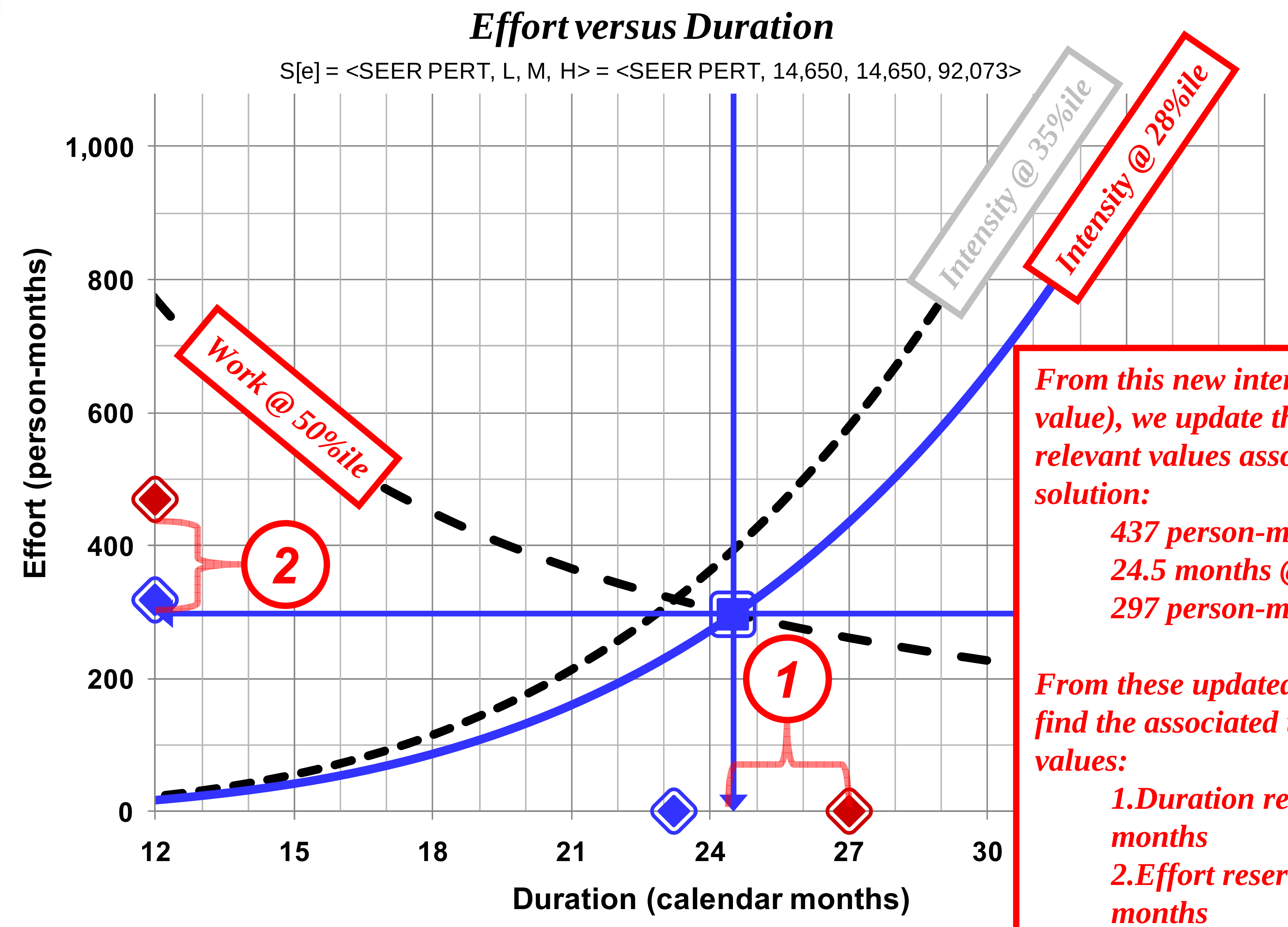


Suppose we reduce the desired high-confidence probability for duration from 80% to 70%.

We must now find a new Intensity curve (I value) that corresponds to 27.0 months @ 70%ile Work.

TECOLOTE RESEARCH, INC.

Project Updated Effort and Duration Positions



From this new intensity curve (I value), we update the other three relevant values associated with this solution:

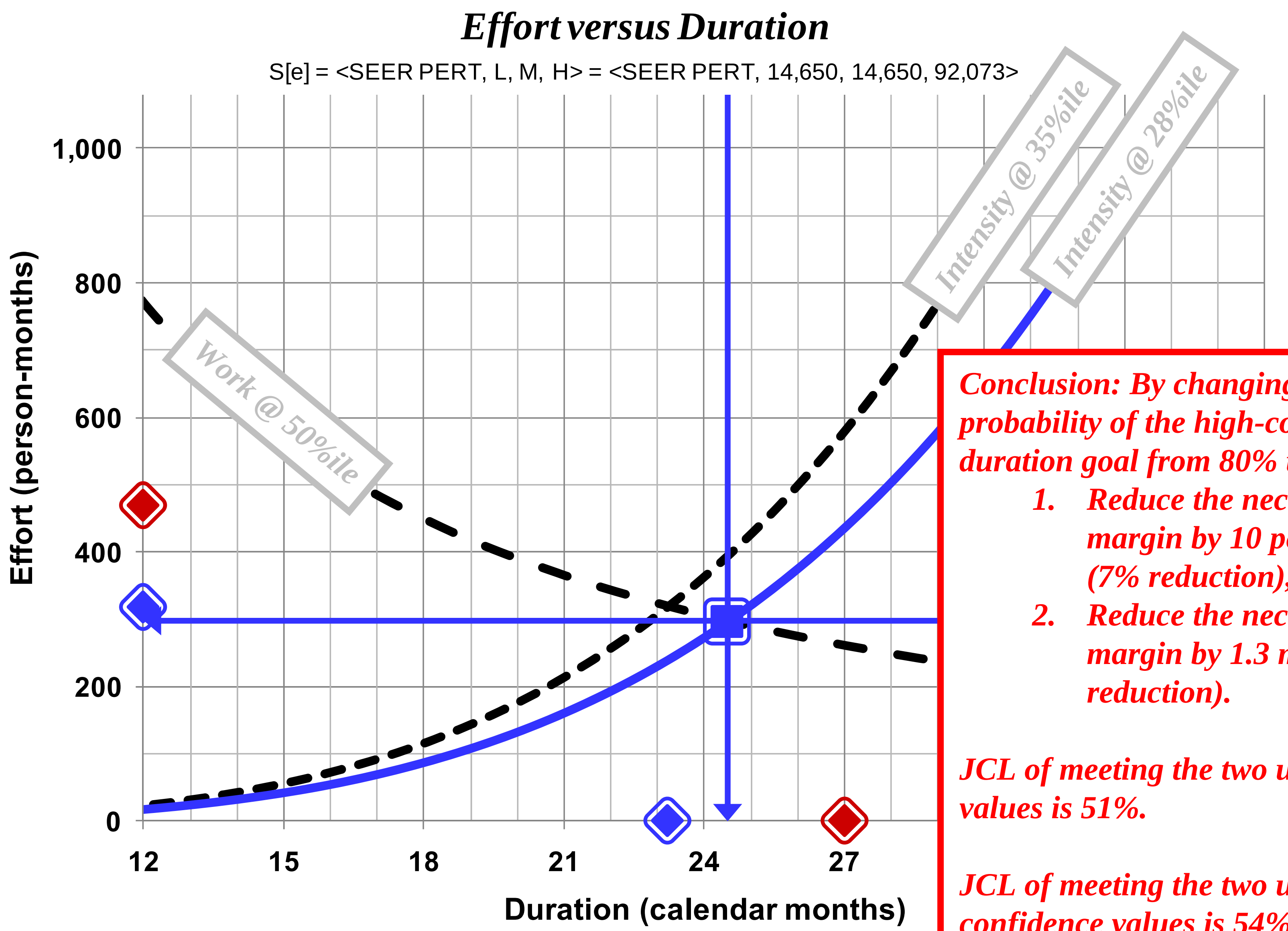
- *437 person-months @ 70%*
- *24.5 months @ 50%*
- *297 person-months @ 50%*

From these updated values, we can find the associated updated reserve values:

1. *Duration reserve: 2.5 months*
2. *Effort reserve: 140 person-months*

TECOLOTE RESEARCH, INC.

Solution



Conclusion: By changing the desired probability of the high-confidence duration goal from 80% to 70% we:

1. *Reduce the necessary effort margin by 10 person-months (7% reduction),*
2. *Reduce the necessary duration margin by 1.3 months (34% reduction).*

JCL of meeting the two updated 50% values is 51%.

JCL of meeting the two updated high-confidence values is 54%.

TECOLOTE RESEARCH, INC.

Garvey Question #3 Mathematical Model

$$T_{1\_High\_Confidence} = T_{2\_High\_Confidence} \equiv 27.0$$

$$p_{T_{1\_High\_Confidence}} \equiv 80\% \quad p_{T_{2\_High\_Confidence}} \equiv 70\%$$

$$p_{E_{1\_High\_Confidence}} = p_{E_{2\_High\_Confidence}} \equiv 70\%$$

$$\left[I_1 \equiv \left(\frac{F_{\Psi}^{-1}\left(p_{T_{1\_High\_Confidence}}\right)}{T_{1\_High\_Confidence}^{\gamma\alpha_E + \alpha_t}} \right)^{1/\alpha_E} = \left(\frac{F_{\Psi}^{-1}(80\%)}{27.0^{3.96(0.57+0.76)}} \right)^{1/0.57} = 1.235254\text{E}-03 \right]_{<*>}$$

Note: <*> indicates Aerospace 2004: Military Ground Operational

Note: $F_X(x) \equiv$ CDF of random variable X at value x

Note: $F_X^{-1}(p) \equiv$ inverse CDF of random variable X at probability p

Garvey Question #3 Mathematical Model

$$\left[\boldsymbol{E}_{I_1} \equiv \left(I_1^{\alpha_T} \boldsymbol{\Psi}^{\gamma} \right)^{1/(\gamma\alpha_E+\alpha_T)} = \left(\left(1.235254\text{E}-03\right)^{0.76} \boldsymbol{\Psi}^{3.96} \right)^{1/(3.96(0.57+0.76))} \right]_{<*>}$$

$$E_{1\_High\_Confidence} \equiv \boldsymbol{F}_{\boldsymbol{E}_{I_1}}^{-1}\left(p_{E_{1\_High\_Confidence}} \right) = \boldsymbol{F}_{\boldsymbol{E}_{I_1}}^{-1}\left(70\%\right) = 469$$

$$E_{1\_50\%} \equiv \boldsymbol{F}_{\boldsymbol{E}_{I_1}}^{-1}\left(50\%\right) = 319$$

$$\left[\boldsymbol{T}_{I_1} = \left(\frac{1}{I_1^{\alpha_E}} \boldsymbol{\Psi} \right)^{1/(\gamma\alpha_E+\alpha_T)} = \left(\frac{1}{\left(1.235254\text{E}-03\right)^{0.57}} \boldsymbol{\Psi} \right)^{1/(3.96(0.57+0.76))} \right]_{<*>}$$

$$T_{1\_50\%} \equiv \boldsymbol{F}_{\boldsymbol{T}_{I_1}}^{-1}\left(p_{T_{1\_High\_Confidence}} \right) = \boldsymbol{F}_{\boldsymbol{T}_{I_1}}^{-1}\left(50\%\right) = 23.2$$

$$\Delta_{T_1} = T_{1\_High\_Confidence} - T_{1\_50\%} = 27.0 - 23.2 = 3.8$$

$$\Delta_{E_1} = E_{1\_High\_Confidence} - E_{1\_50\%} = 469 - 319 = 150$$

Garvey Question #3 Mathematical Model

$$\left[I_2 = \left(\frac{\boldsymbol{F}_{\boldsymbol{\Psi}}^{-1}\left(p_{t_{2\_High\_Confidence}}\right)}{T_{2\_High\_Confidence}^{\gamma\alpha_E+\alpha_T}} \right)^{1/\alpha_E} = \left(\frac{\boldsymbol{F}_{\boldsymbol{\Psi}}^{-1}(70\%)}{36.0^{3.96(0.57+0.76)}} \right)^{1/0.57} = 9.287725\text{E}-04 \right]_{<*>}$$

$$\left[\boldsymbol{E}_{I_2} \equiv \left(I_2^{\alpha_T} \boldsymbol{\Psi}^{\gamma} \right)^{1/(\gamma\alpha_E+\alpha_T)} = \left((9.287725\text{E}-04)^{0.76} \boldsymbol{\Psi}^{3.96} \right)^{1/(3.96(0.57+0.76))} \right]_{<*>}$$

$$E_{2\_High\_Confidence} \equiv \boldsymbol{F}_{\boldsymbol{E}_{I_2}}^{-1}\left(p_{E_{2\_High\_Confidence}}\right) = \boldsymbol{F}_{\boldsymbol{E}_{I_2}}^{-1}(70\%) = 437$$

$$E_{2\_50\%} \equiv \boldsymbol{F}_{\boldsymbol{E}_{I_2}}^{-1}(50\%) = 297$$

TECOLOTE RESEARCH, INC.

Garvey Question #3 Mathematical Model

$$\left[\boldsymbol{T}_{\boldsymbol{M_2}} = \left(\frac{1}{I_2^{\alpha_E}} \boldsymbol{\Psi} \right)^{1/(\gamma\alpha_E + \alpha_T)} = \left(\frac{1}{(9.287725\text{E}-04)^{0.57}} \boldsymbol{\Psi} \right)^{1/(3.96(0.57+0.76))} \right]_{<*>}$$

$$T_{2\_50\%} \equiv \boldsymbol{F}^{-1}_{\boldsymbol{T}_{\boldsymbol{M_2}}} \left(p_{T_{2\_High\_Confidence}} \right) = \boldsymbol{F}^{-1}_{\boldsymbol{T}_{\boldsymbol{M_2}}} (50\%) = 24.5$$

$$\Delta_{T_2} = T_{2\_High\_Confidence} - T_{2\_50\%} = 27.0 - 24.5 = 2.5$$

$$\Delta_{E_2} = E_{2\_High\_Confidence} - E_{2\_50\%} = 437 - 297 = 140$$

$$\Delta_T = 2.5 - 3.8 = -1.3 \quad \Delta_{T\%} = \frac{2.5 - 3.8}{3.8} = -34\%$$

$$\Delta_E = 140 - 150 = -10 \quad \Delta_{E\%} = \frac{140 - 150}{150} = -7\%$$

Garvey Question #3: Example Answer

"How are cost reserve recommendations affected by schedule risk?"

Answer: Reducing schedule high-confidence position from 80% to 70% reduces schedule reserve by 34% and cost reserve by 7%

TECOLOTE RESEARCH, INC.

References

Aerospace 2004. Gayek, Jonathan E. , Long, Lutrell G., Bell, Kim D., et. al., "Software Cost and Productivity Model", Aerospace Report No. ATR-2004(8311)-1, Aerospace Corporation, El Segundo, CA.

Bernstein, Peter L. 1996. *Against the Gods: The Remarkable Story of Risk*. New York City : John Wiley & Sons, Inc., 1996. ISBN 0-471-29563-9 (paper).

Garvey, Paul 2000. *Probability Methods for Cost Uncertainty Analysis: A Systems Engineering Perspective*, Chapman-Hall/CRC-Press, Boca Raton, London, New York, NY.

Ross, Michael A. 2008a. "Next Generation Software Estimating Framework: 25 Years and Thousands of Projects Later. [ed.] Stephen A. Book and Edward White III. *Journal of Cost Analysis and Parametrics.* s.l. : Society of Cost Estimating and Analysis - International Society of Parametric Analysts, Fall 2008. Vol. 1, 2, pp. 7-30. ISSN 1941-658X.

Ross, Michael A. 2008b. "Uncertainty to Probability for Cost with Duration Estimating Relationships (CDERs)." *Proceedings, 2008 SSCAG/SCAF/EACE Meeting*, Noordwijk, The Netherlands.

Ross, Michael A. 2011b. "Joint Confidence Level of a Parametric Software Cost and Schedule Estimate: Method and Example." *Proceedings, 2011 ISPA/SCEA Joint Conference & Training Workshop*, Albuquerque, NM.

Valerdi, Ricardo 2009. Gaffney, John & Ross, Michael A., "Approaches to Calculating Systems Engineering Schedule in Parametric Cost Models." *Proceedings, 2009 Conference on Systems Engineering Research*, Loughborough University, Leicestershire, UK.