SAND2014-19052C

Exceptional service in the national interest

Sandia National Laboratories

WSTAT

CPAT

Challenges in Varying Fixed Parameters in a Real-World Large-Scale Army Acquisition Problem

Frank M. Muldoon
System Readiness and Sustainment Technologies
Sandia National Laboratories

INFORMS Annual Meeting 2014
November 9th-12th, 2014

U.S. DEPARTMENT OF ENERGY

NNSA National Nuclear Security Administration

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

Ground Combat Fleet Modernization

Sandia National Laboratories

Work Sponsor: Shatiel Edwards, Program Executive Office Ground Combat Systems (PEO GCS)

Organizational Team: PEO GCS, Sandia National Labs, Booz Allen Hamilton, Teledyne Brown

- PEO GCS has a large, diverse fleet of vehicles (Abrams Tanks, Bradleys, Strykers, etc.) that it maintains and must upgrade over the next few decades
- Sandia has developed multiple analytic tools tailored to different aspects of this management challenge, including:
  - **Whole System Trades Analysis Tool (WSTAT)** examines decisions at the individual system level, presenting tradeoffs in design
  - **Capability Portfolio Analysis Tool (CPAT)** examines decisions at the fleet level, attempting to find the best mixture of fixed systems through time
- How do we integrate WSTAT’s capability for system design tradeoffs and CPAT’s ability to modernize the entire fleet over time?

WSTAT Introduction

Whole System Trades Analysis Tool

- WSTAT looks at the design of a single system by examining many potential configurations in an effort to meet multiple competing requirements and objectives (e.g., performance, procurement cost, O&S cost, growth, and risk)
- WSTAT uses a multi-objective genetic algorithm to find the Pareto frontier of design “sweet spots” that balance these multiple competing criteria

Notional

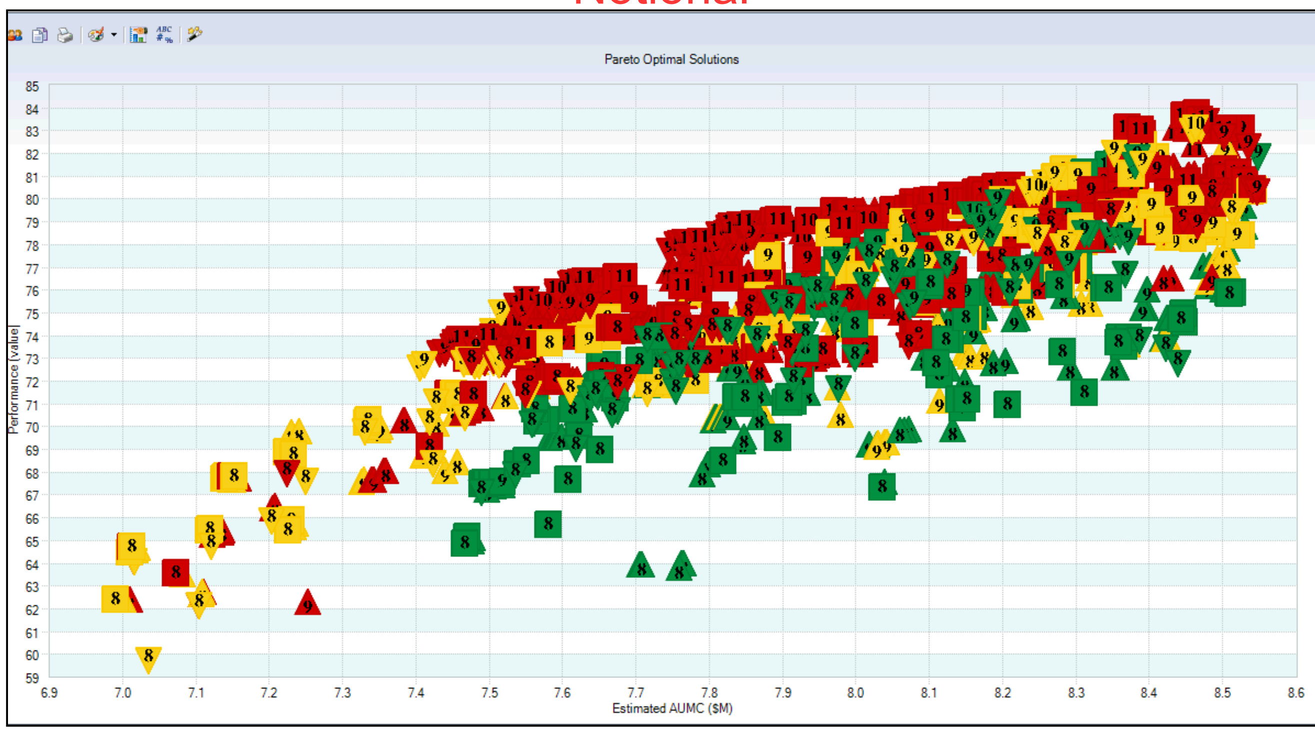
Performance:
Y-Axis

Procurement Cost:
X-Axis

O&S Cost:
20 year cost ($M)

Growth:
△ High
□ Moderate
▽ Low

Risk:
High (red)
Medium (yellow)
Low (green)



CPAT Introduction

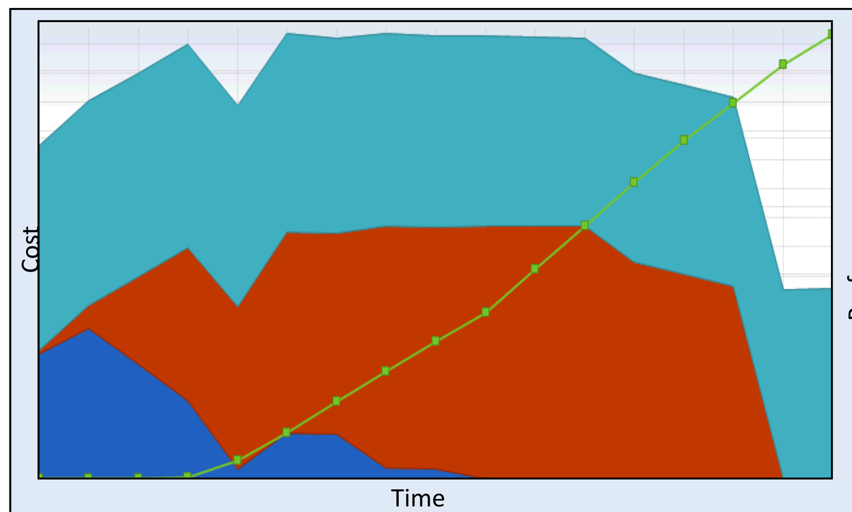
Capability Portfolio Analysis Tool

- CPAT optimizes the mixture of systems within the entire fleet through time (the systems themselves are not modified)
- CPAT uses a single objective multi-stage mixed-integer linear programming (MILP) to perform this optimization
  - Objective: maximize the sum of the performance of all systems in the fleet over the study horizon

Notional

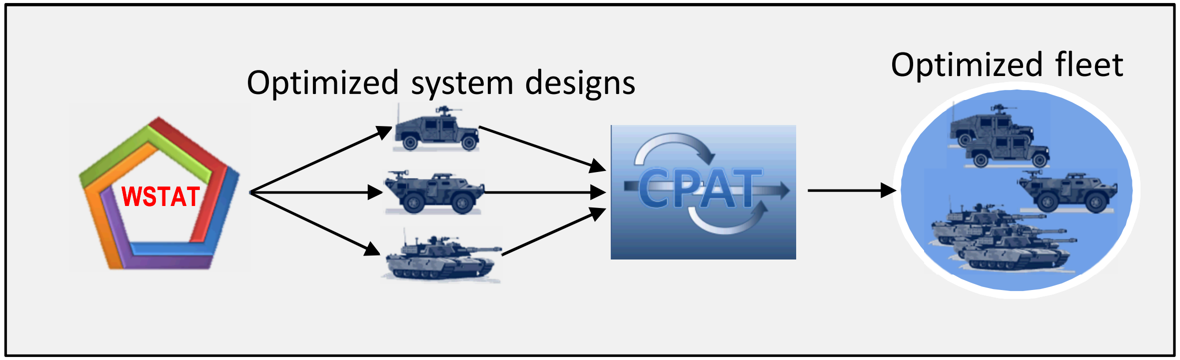
		FY14	FY15	FY16	FY17	FY18	FY19	FY20	FY21	FY22	FY23	FY24	FY25	FY26	FY27	FY28	FY29	FY30
Mission1	Vehicle4	1680	1680	1680	1680	1540	1330	1120	910	700	490	280	140	70				
	Vehicle5					140	350	560	700	840	910	1050	1120	1120	1120	1120	1120	1120
	Vehicle7								70	140	280	350	420	490	560	560	560	560
Mission2	Vehicle20	600	600	600	550	500	450	400	350	300	250	200	150	100	50			
	Vehicle21				50	100	150	200	250	300	350	400	450	500	550	600	600	600
Mission3	Vehicle23	360	360	360	360	345	315	285	255	225	210	180	150	120	90	60	30	
	Vehicle32					15	45	75	105	135	150	150	150	150	150	150	150	150
	Vehicle33											30	60	90	120	150	180	210
Mission4	Vehicle24	1080	1080	1080	1080	990	900	810	720	630	585	495	405	315	225	135	45	
	Vehicle32					90	180	270	360	450	495	495	495	495	495	495	495	495
	Vehicle33											90	180	270	360	450	540	585
Mission5	Vehicle25	1200	1200	1200	1200	1100	1050	950	850	750	700	600	500	400	300	200	100	
	Vehicle32					100	150	250	350	450	500	500	500	500	500	500	500	500
	Vehicle33											100	200	300	400	500	600	700
Mission6	Vehicle26	480	480	480	480	480	440	400	360	320	280	240	200	160	120	80	40	
	Vehicle32						40	80	120	160	200	200	200	200	200	200	200	200
	Vehicle33											40	80	120	160	200	240	280

Notional



Two-Stage Analysis

- The need for both an optimized fleet and optimized systems within that fleet has traditionally been approached in two stages
  - One stage optimizes the individual systems configurations
    - Only selected systems are passed to the next stage
  - The second stage optimizes the mix of **fixed** systems within the fleet



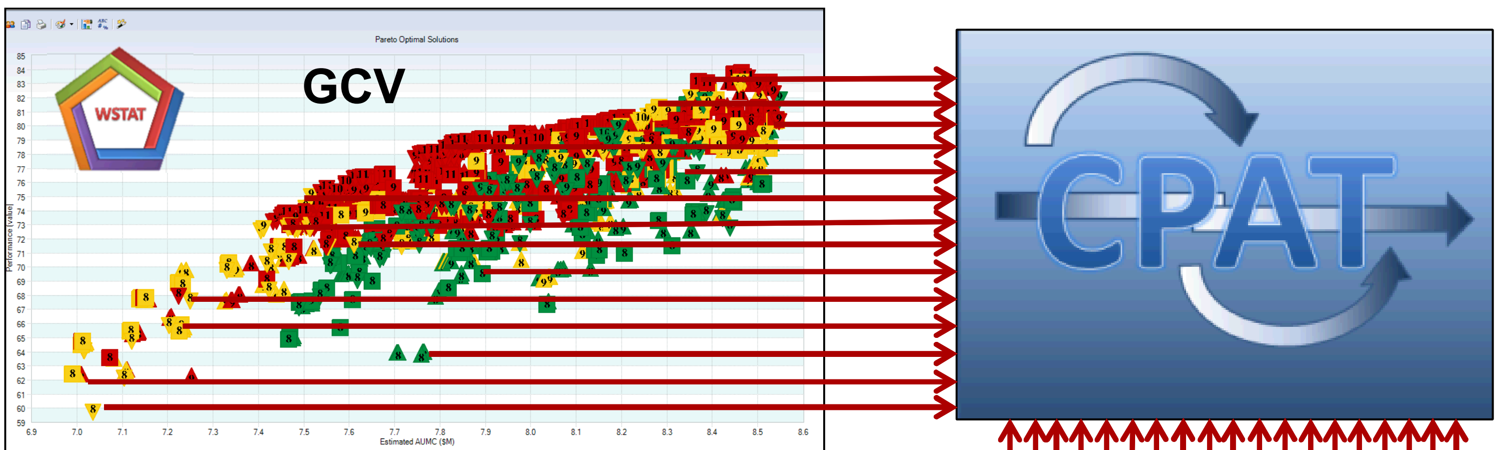
- Currently, analysis stages do not communicate
  - WSTAT does not know how configurations will be incorporated into the fleet
  - CPAT does not know about all possible system configurations

Combined Analysis Process

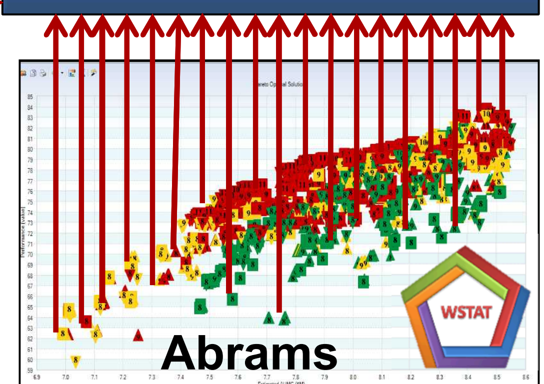
Sandia National Laboratories

- Individual systems are still optimized, but instead of selecting one system to pass to the next stage, all individual Pareto optimal solutions are passed
- Thousands of new system configurations for each system type are created in CPAT and the optimization chooses the best configuration for the fleet

Notional



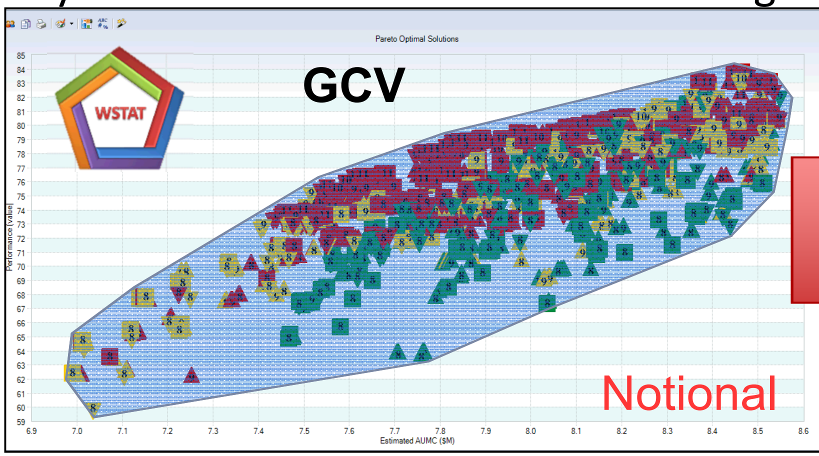
- Attempt to include each Pareto point as a unique system in CPAT
- This approach is not feasible. Each new system configuration in CPAT requires additional variables and constraints and the problem size becomes too big to solve

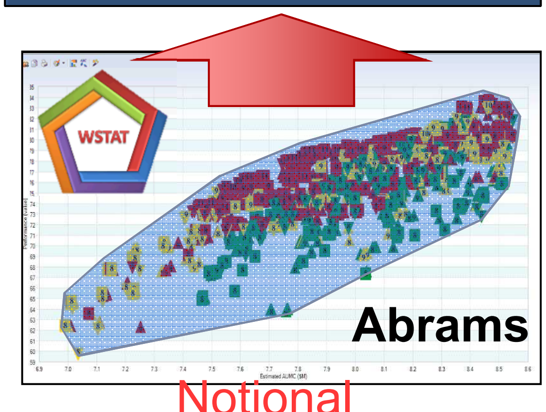


Notional

Combined Analysis Process

- Individual systems are still optimized, but instead of selecting one system to pass to the next stage, all Pareto optimal solutions are passed
- To do this, the convex hull of the five dimensional Pareto region is formed
- For each individual system, CPAT is allowed to choose any set of parameters, performance, procurement cost, O&S cost, growth, and risk, within that system's convex hull of the Pareto region

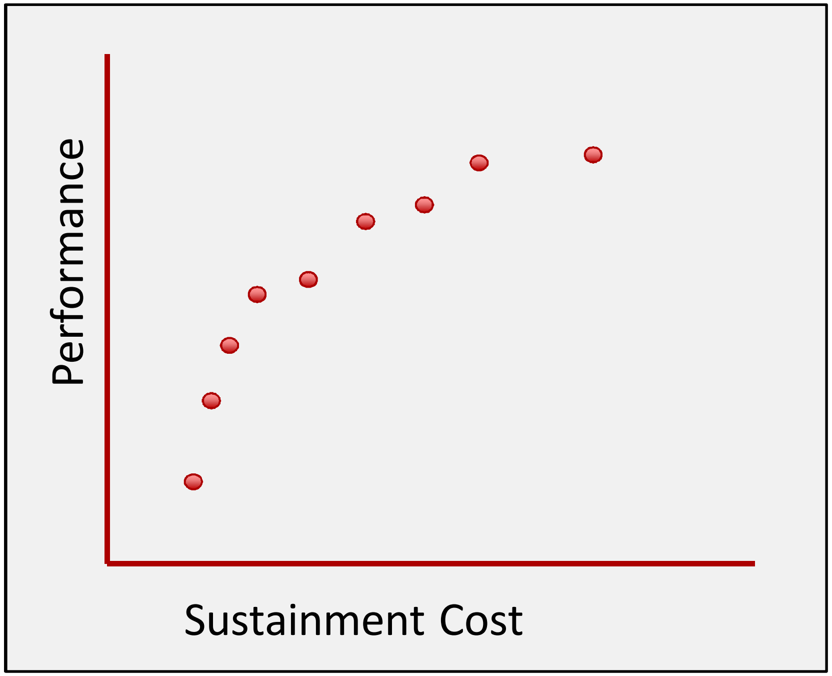




- Fixed parameters for each system, such as cost and performance, are now variable, but restricted to that system's Pareto frontier
- CPAT chooses individual system configurations in order to "holistically" optimize the entire fleet

Example Fleet Optimization Problem

- Consider the following Pareto frontier for system *s* in the two dimensions of performance and sustainment cost
- Also, consider the following simplified fleet modernization problem with the goal of optimizing performance limited by only budget constraints

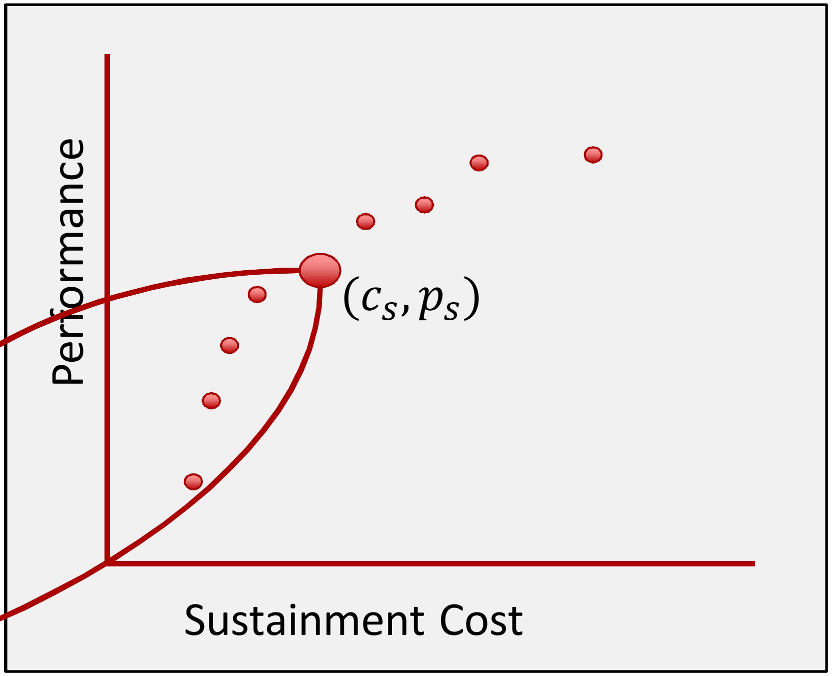


$$\text{Max} \sum_{s,t} p_s * i_{s,t}$$
$$\text{s.t.} \quad \sum_{s} c_s * i_{s,t} \leq b_t \quad \forall\, t$$
$$i_{s,t}\ integer \ \forall\ s,t$$

- p_s is the performance of system *s*
- c_s is the sustainment cost of system *s*
- $i_{s,t}$ is the number of systems *s* in the fleet at time period *t*
- b_t is the budget for time period *t*

Example Fleet Optimization Problem

- Consider the following Pareto frontier for system *s* in the two dimensions of performance and sustainment cost
- Also, consider the following simplified fleet modernization problem with the goal of optimizing performance limited by only budget constraints



$$\text{Max } \sum_{s,t} p_s * i_{s,t}$$
$$\text{s.t. } \sum_{s} c_s * i_{s,t} \le b_t \quad \forall\, t$$
$$i_{s,t} \; integer \; \forall \; s,t$$

Typically, only one system configuration is selected and given to the fleet optimization problem

- p_s is the performance of system *s*
- c_s is the sustainment cost of system *s*
- $i_{s,t}$ is the number of systems *s* in the fleet at time period *t*
- b_t is the budget for time period *t*

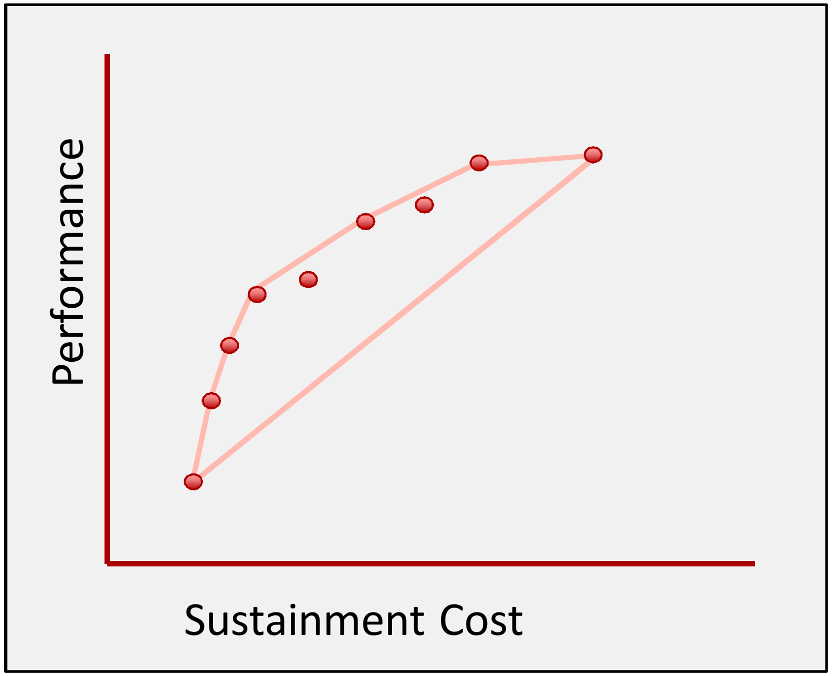
Non-Linear Optimization Problem

- The constraints of the convex hull of the Pareto region for each system are represented using matrices as

$$\boldsymbol{A}_s\boldsymbol{p}_s + \boldsymbol{B}_s\boldsymbol{c}_s \le \boldsymbol{d_s} \;\; \forall\, s$$

- The parameters p_s and c_s for each system *s* are now variable
- The non-linear fleet optimization problem is

$$\begin{aligned} &\text{Max } \textstyle\sum_{s,t} p_s * i_{s,t} \\ &\text{s.t.} \quad \textstyle\sum_s c_s * i_{s,t} \le b_t \;\; \forall\, t \\ &\qquad \boldsymbol{A}_s\boldsymbol{p}_s + \boldsymbol{B}_s\boldsymbol{c}_s \le \boldsymbol{d_s} \;\; \forall\, s \\ &\qquad i_{s,t} \; integer \;\; \forall \;\; s,t \end{aligned}$$



- This fleet optimization problem has *bilinear* terms $p_s * i_{s,t}$ and $c_s * i_{s,t}$ and is difficult to solve
- One solution is to linearize these non-linear terms

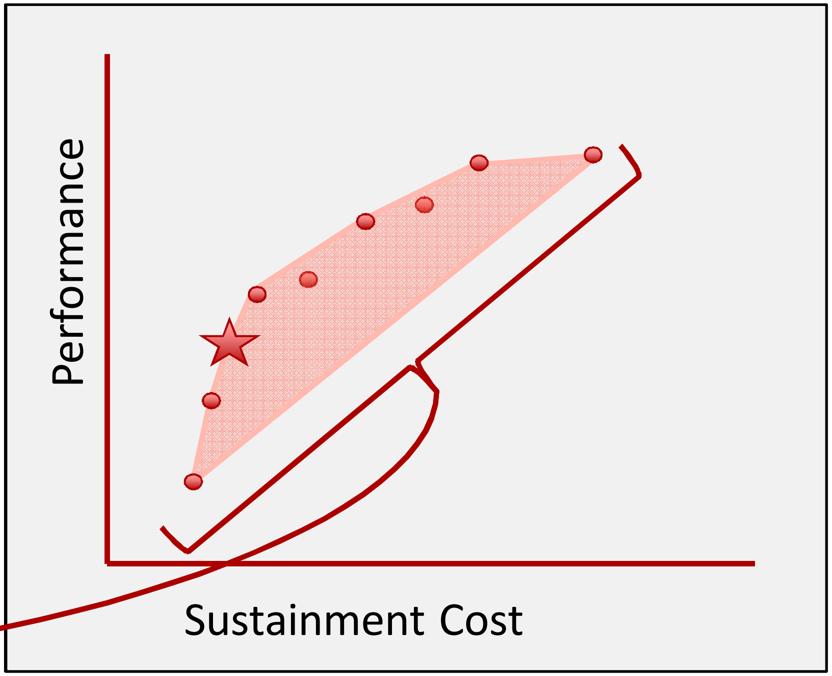
Non-Linear Optimization Problem

- The constraints of the convex hull of the Pareto region for each system are represented using matrices as

$$\boldsymbol{A}_s\boldsymbol{p}_s + \boldsymbol{B}_s\boldsymbol{c}_s \le \boldsymbol{d}_s \ \ \forall\, s$$

- The parameters p_s and c_s for each system *s* are now variable
- The non-linear fleet optimization problem is

$$\begin{aligned} \text{Max } & \textstyle\sum_{s,t} p_s * i_{s,t} \\ \text{s.t. } & \textstyle\sum_s c_s * i_{s,t} \le b_t \quad \forall\, t \\ & \boldsymbol{A}_s\boldsymbol{p}_s + \boldsymbol{B}_s\boldsymbol{c}_s \le \boldsymbol{d}_s \ \ \forall\, s \\ & i_{s,t} \ integer \ \ \forall \ \ s,t \end{aligned}$$



The whole Pareto frontier is included and the optimization can choose any point in the convex hull

- This fleet optimization problem has *bilinear* terms $p_s * i_{s,t}$ and $c_s * i_{s,t}$ and is difficult to solve
- One solution is to linearize these non-linear terms

Linearizing Non-Linear Terms

- Various ways exist to linearize the bilinear terms $p_s * i_{s,t}$ and $c_s * i_{s,t}$
- One such way using unary expansions (Gupte, et al. 2013) of the integer variables $i_{s,t}$ and replacing $c_s * i_{s,t}$ with $w_{s,t}$ for each *s* and *t* requires the following constraints

$$cCont_{s,t,b} \leq M * cBinary_{s,t,b} \qquad \forall\, s,t,b$$
$$cCont_{s,t,b} \leq c_s \qquad \forall\, s,t,b$$
$$cCont_{s,t,b} \geq c_s \ - M * \left(1 - cBinary_{s,t,b}\right) \quad \forall\, s,t,b$$
$$\sum_b cBinary_{s,t,b} \leq 1 \qquad \forall\, s,t$$
$$i_{s,t} = \sum_b b * cBinary_{s,t,b} \qquad \forall\, s,t$$
$$w_{s,t} = \sum_b b * cCont_{s,t,b} \qquad \forall\, s,t$$

- This method requires $s * t * U$ additional binary and continuous variables which can be computational difficult
- Techniques to relax these binary variables to be continuous with the addition of $s * t * \log_2(U)$ binary variables and constraints (Adams and Henry 2012) did not help

Linearizing Non-Linear Terms

- The technique we used is a binary expansion (Gupte, et al. 2013) of the integer variables $i_{s,t}$
- Replace $c_s * i_{s,t}$ with $w_{s,t}$ for all *s* and *t* and add the following set of constraints

$$cCont_{s,t,b} \leq M * cBinary_{s,t,b} \qquad \forall\, s,t,b$$

$$cCont_{s,t,b} \leq c_s \qquad \forall\, s,t,b$$

$$cCont_{s,t,b} \geq c_s - M * (1 - cBinary_{s,t,b}) \qquad \forall\, s,t,b$$

$$i_{s,t} = \sum_b 2^b * cBinary_{s,t,b} \qquad \forall\, s,t$$

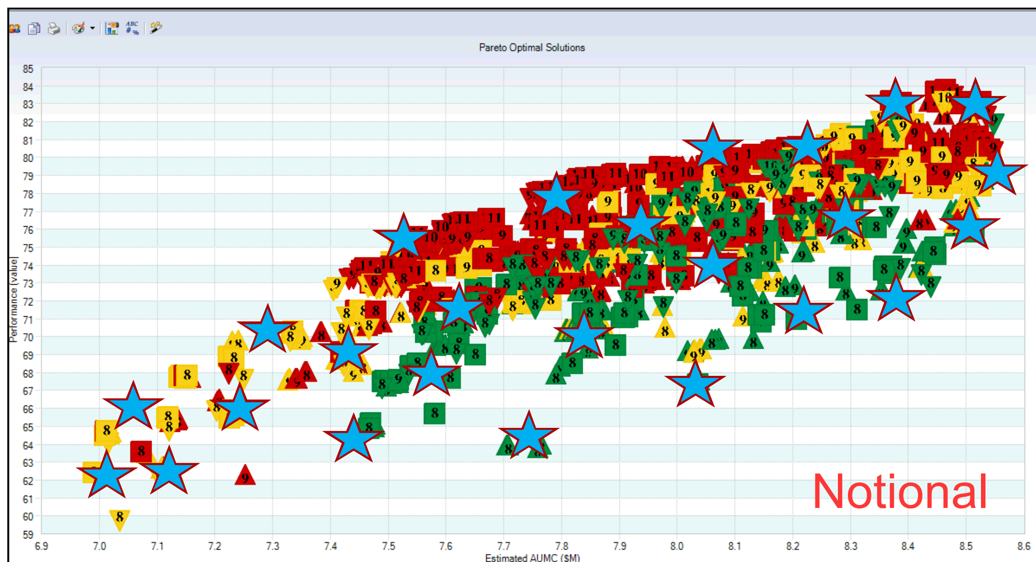
$$w_{s,t} = \sum_b 2^b * cCont_{s,t,b} \qquad \forall\, s,t$$

- $s * t * \log_2(U)$ auxiliary binary variables and continuous variables are required for an exact representation of the product term which was less computationally difficult
- A similar set of constraints and variables are used to replace the non-linear terms $p_s * i_{s,t}$

Solving CPAT Optimization

- CPAT uses CPLEX 12.6 under the hood to solve the MILP
  - The model was written in OPL
- Sample test problem included
  - ~ 70 systems
  - ~ 30 time periods
  - ~ 70,000 constraints
  - ~ 20,000 variables

- We found the Pareto frontier for 3 systems and selected 25 non-dominated points
  - ~ 5,000 constraints were add for these 3 systems
  - ~ 2,000 variables were add for these 3 systems
- Solution times under an hour



Numerical Instability Issues

- CPAT now contains more information about the Pareto frontiers and is represented by a MILP solved using CPLEX 12.6
- After implementing this code we noticed that solving the same problem resulted in different results
  - Parallel optimization was utilized so we expected variation in the branch-and-bound tree, but did not expect optimal solutions to differ more than the CPLEX tolerance
  - With parallel optimization disabled, changes to parameter settings also resulted in different optimal solutions
- Enabling the option “kappastats” in CPLEX revealed the nature of the problem
  - Kappastats evaluates the condition number of the optimal bases during the solution of an MILP model
    - Stable Bases (condition number less than 1e7)
    - Suspicious Bases (condition number between 1e7 and 1e10)
    - Unstable Bases (condition number between 1e10 and 1e14)
    - Ill-Posed Bases (condition number greater than 1e14)

Numerical Instability Issues

- Our test problem had a high percentage of bases matrices with unstable condition numbers indicating severe numerical instability issues

Stable	0.0%
Suspicious	16.8%
Unstable	82.5%
Ill-Posed	0.7%

- These issues resulted from the vast number of Big M constraints added

$$cCont_{s,t,b} \leq M * cBinary_{s,t,b} \qquad \forall\, s,t,b$$
$$cCont_{s,t,b} \geq c_s - M * \left(1 - cBinary_{s,t,b}\right) \ \forall\, s,t,b$$

- Costs in the Army acquisition problem ran into the tens of millions so *M* was in that range which was orders of magnitude higher than other coefficients in the problem

Stable	90.3%
Suspicious	9.3%
Unstable	0.4%
Ill-Posed	0.0%

- Rescaling the costs resulted in a majority of bases matrices having stable condition numbers

Memory Issues

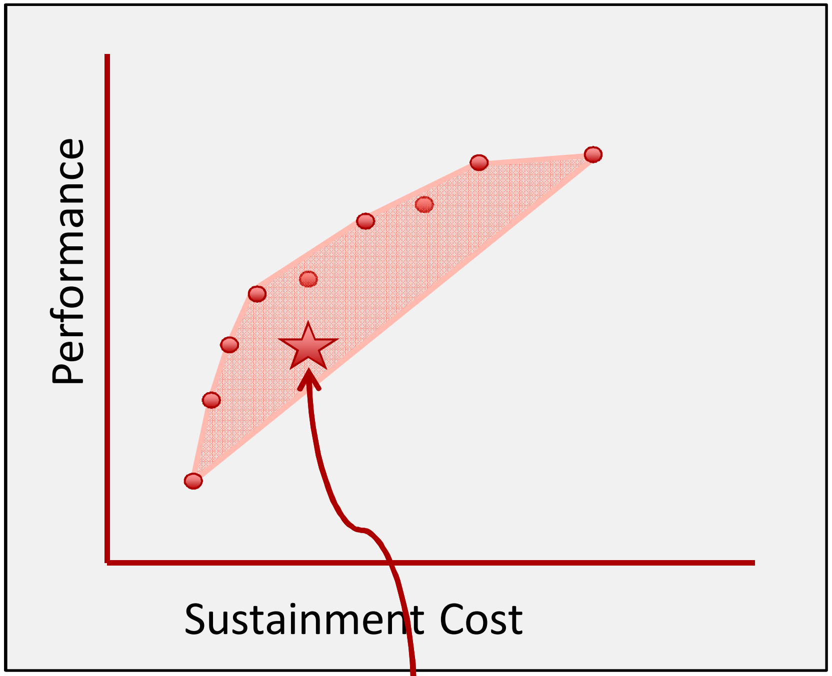
- Further testing revealed issues with the constraint matrix representing the convex hull of the Pareto frontier

$$\boldsymbol{A}_s \boldsymbol{p}_s + \boldsymbol{B}_s \boldsymbol{c}_s \leq \boldsymbol{d}_s \ \forall s$$

- The constraint matrices $\boldsymbol{A}_s$ and $\boldsymbol{B}_s$ were almost fully dense
  - There were also a substantial number of coefficients that were close to zero (i.e., within 1e-4)
- The number of constraints from the convex hull of the Pareto frontier added to the problem is limited by memory.
- Having more than 10,000 convex hull constraints resulted in an out-of-memory error in CPLEX
- The limitation on constraints also limits the number of Pareto points that could be included in the optimization
  - This limits the richness of solutions provided by the Pareto frontier

Non-Pareto Point Issue

- The fleet optimization problem is allowed to choose any point in the convex hull of the Pareto frontier
- The optimization may select a point that is no where near a Pareto point
- The Pareto frontier for each system is a representation of the true Pareto frontier
- If the point looks Pareto optimal it is possible to go back to the WSTAT optimization and try to fill in some more Pareto points



The optimization may pick a solution that is not a Pareto point

- If the point select is not on the Pareto frontier, it is possible to go back to the WSTAT for the system and search for a configuration that matches the performance and cost

Non-Pareto Point Issue

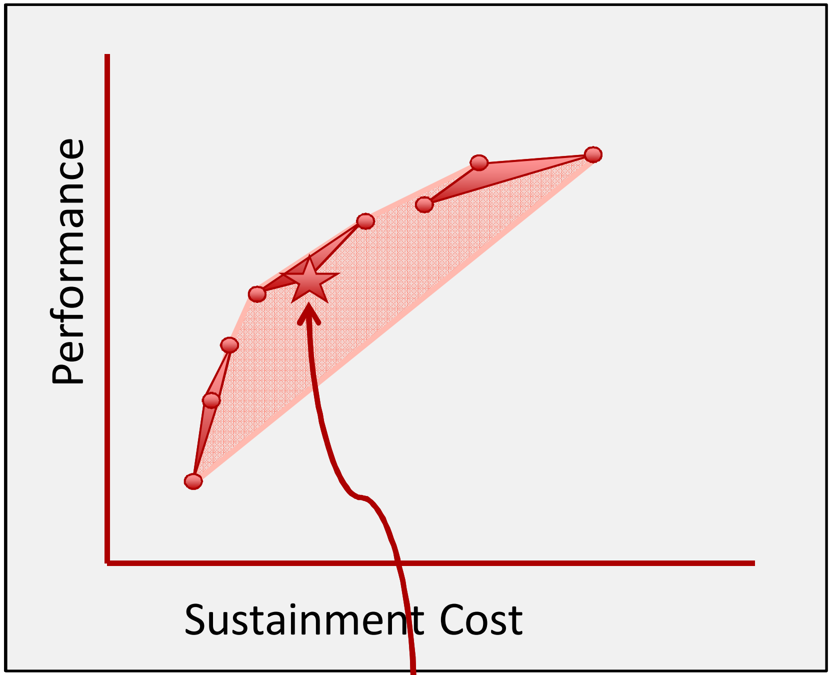
- In order to prevent the optimization from choosing a non-Pareto point we could take the union of disjoint polytopes and model using disjunctive programming (Balas 1979)
- Each region is made of a set of constraints $\boldsymbol{A}_i\boldsymbol{x} \leq \boldsymbol{b}_i$
- We multiply a binary variable λ_i by each region and linearize

$$\boldsymbol{A}_i\boldsymbol{w}_i \leq \boldsymbol{b}_i\lambda_i \ \forall i$$

$$\sum_i \lambda_i = 1$$

$$\sum_i w_i = x$$

$$\lambda_i \ binary$$

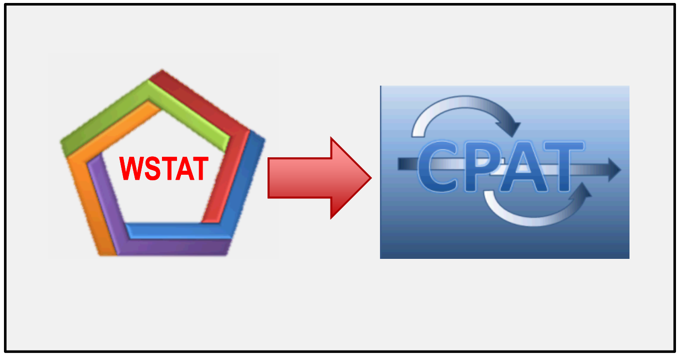


The optimization is forced to pick a point in one of the disjoint polytopes

- The disjoint polytopes may need to be refined until the optimization selects a corner point

Conclusions

- It is possible to combine separate optimization approaches in order to provide a "holistic" fleet optimization problem
- Combining CPAT and WSTAT answers new questions that simultaneously combine elements of system and fleet design
- Numerical issues must be address for consistent results due to numerous big M constraints added to the formulation
- Memory issues can also be prohibitive on the size of the Pareto region incorporated into the fleet optimization problem
- Need to determine the best way to handle system configurations that are not Pareto optimal solutions



References

Thank you!

- Adams, W.P. and Henry, S.M., "Base-2 Expansions for Linearizing Products of Functions of Discrete Variables," *Operations Research*, Vol. 60, No. 6, pp 1477-1490, 2012.
- Balas, E., "Disjunctive Programming," *Annals of Discrete Mathematics*, Vol. 5, pp 3-51, 1979.
- Gupte, A., Ahmed, S., Cheon, M.S., and Dey, S.S., "Solving Mixed Integer Bilinear Problems Using MILP Formulations," *SIAM Journal on Optimization*, Vol. 23, No. 2, pp 721-744, 2013.