Blind Validation Study of PRICE TruePlanning and SEER-H

Paul Friz and Jordan Klovstad
Systems Analysis and Concepts Directorate
NASA Langley Research Center

With
Benjamin Leser and Benjamin Towle
University of Michigan

And
Serhat Hosder
Missouri University of Science and Technology

Presented to:
NASA Cost and Schedule Symposium
23-25 April 2024

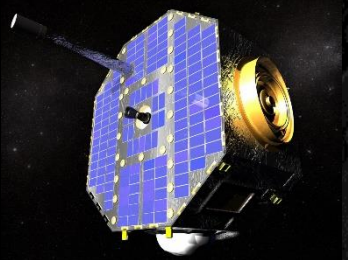
Motivation

- This work was originally presented at AIAA SPACE Conference in 2018 and published in Acta Astronautica in 2020
- Before this study no independent validation of PRICE and SEER was publicly available
- PRICE Systems and Galorath have performed validation studies of PRICE and SEER for NASA missions
  - PRICE is advertised to have an average error of +1% and standard deviation of 13%
  - SEER is advertised to have an average error of -1% and standard deviation of 19%
- The Goal of this Study is to independently assess the accuracy and precision of PRICE TruePlanning – Space Mission Catalog and SEER-H in a blind study

Methods

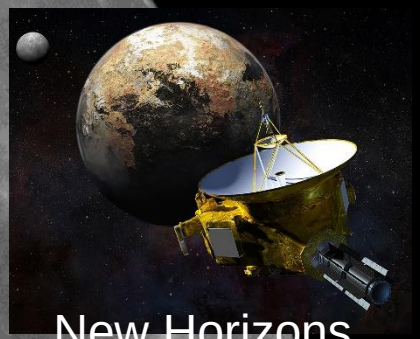
- 12 missions were selected from NASA's ONCE (One NASA Cost Engineering) database
- CADRes (Cost Analysis Data Requirements)
  - CADRes are summaries of technical and cost data for a science mission
  - Based on documents generated at major mission reviews
  - Often have important information missing such as the heritage of a particular component
- Supporting documents
  - Presentations/technical documents from CDR used to simulate estimation environment
- CADRes and supporting documents had all cost information removed before estimators had access to them
- Estimators were two aerospace engineering interns who were trained in the use of PRICE and SEER prior to beginning the study

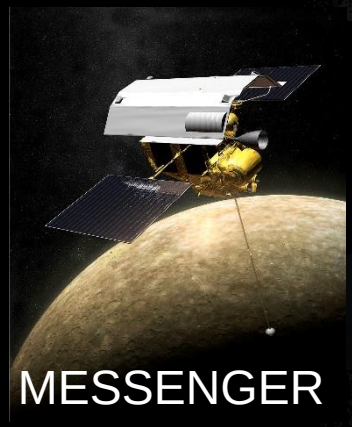
Missions Selected

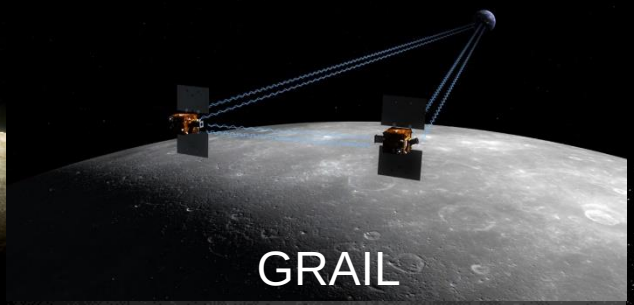

IBEX

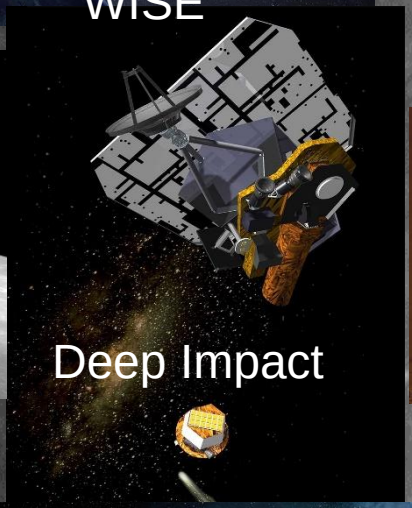

CONTOUR

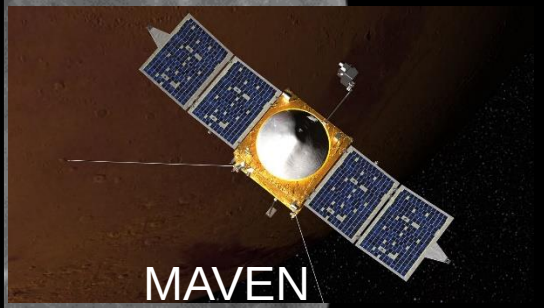

WISE

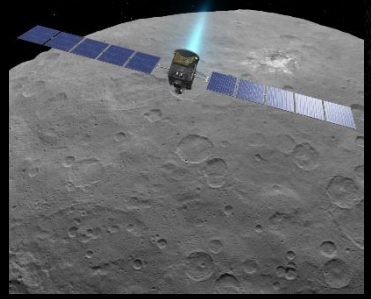

New Horizons

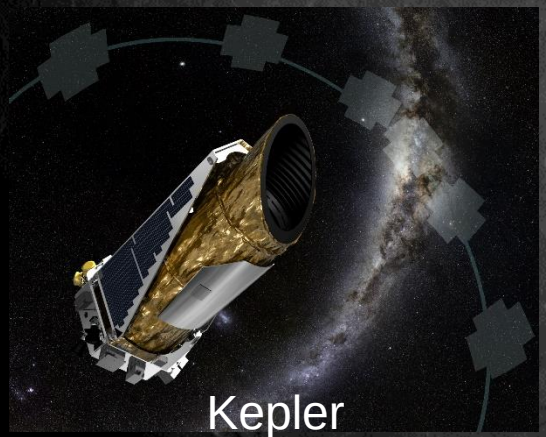

MESSENGER

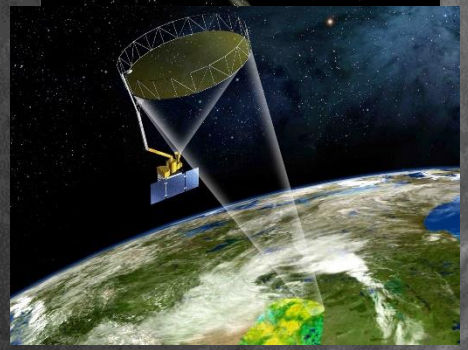

GRAIL


Deep Impact


MAVEN


Dawn


Kepler


SMAP


Juno

What costs are included in this study?

- PRICE and SEER estimated the WBS elements highlighted in *blue* below, i.e. WBS 1, 2, 3, 5, 6, 10
- NASA Work Breakdown Structure (WBS)
  - *1 Project Management*
  - *2 Systems Engineering*
  - *3 Safety & Mission Assurance*
  - 4 Science/Technology
  - *5 Payload*
  - *6 Spacecraft*
  - 7 Mission Operations
  - 8 Launch Vehicles/Services
  - 9 Ground Systems
  - *10 Systems Integration & Testing*
  - 11 Education and Public Outreach

Tool Comparison: Primary Cost Drivers

PRICE

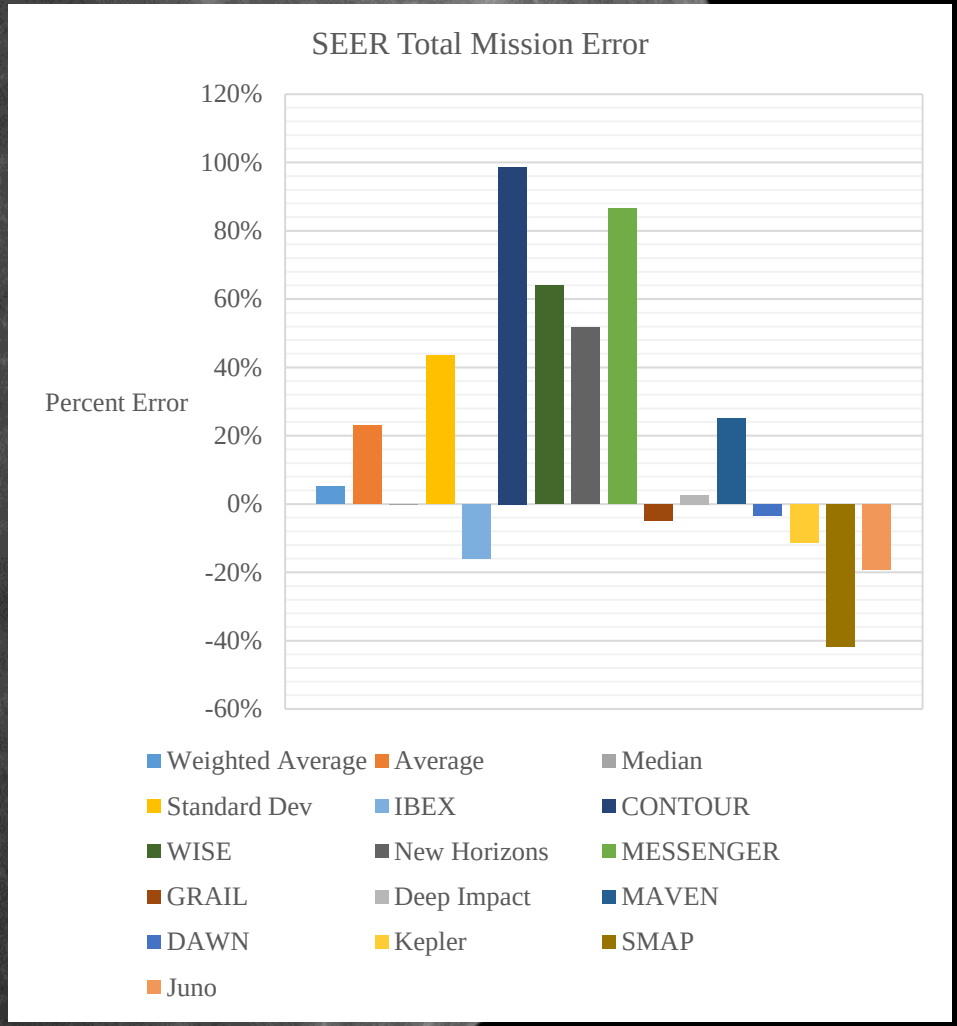
- Mass is primary scaling factor for all components
- All component types
  - Function
  - Heritage
  - Material
  - Complexity
  - Etc.
- “Calculators” can be used to guide input values

SEER

- Structures/Mechanisms
  - Mass, materials, complexity of fit/form, heritage
- Electronics
  - Number/function of boards, number of components/IO pins, clock speed, heritage, FPGAs/custom chips
- Optics
  - Element type, size, quantity, heritage
- Sensors
  - Sensor type, pixel size/quantity

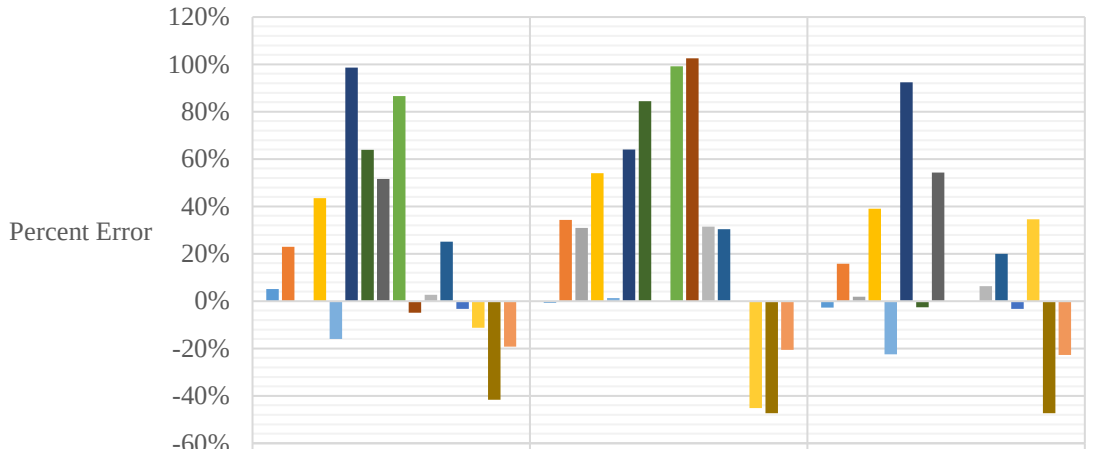
SEER's Error in Estimating Total Mission Costs

- Average error weighted by cost of missions: 5%
  - Meaning small systematic error
- Average error: 23%
- Median error: -0.3%
- Standard Deviation: 43%
- SEER is equally likely to overestimate or underestimate cost
- SEER tends to over estimate small missions and under estimate large missions



SEER Results – Payload and Spacecraft

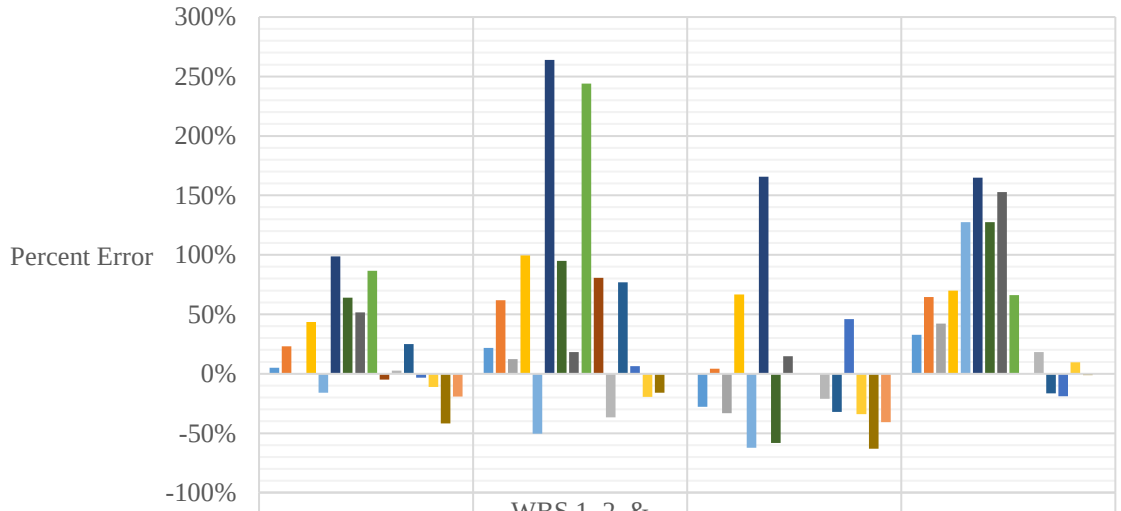
- Large variance in payload estimates likely due to lack of technical details provided in CADRe data
- WBS 5 Payload
  - Cost weighted average error-1%
  - Average error: 34%
  - Median error: 31%
  - Standard Deviation: 54%
- WBS 6 Spacecraft
  - Average error (weighted by cost): -3%
  - Average error: 16%
  - Median error: 2%
  - Standard Deviation: 39%



	Mission Total	WBS 5 - Payload Total	WBS 6 - Spacecraft Bus Total
Weighted Average	5%	-1%	-3%
Average	23%	34%	16%
Median	-0.3%	31%	2%
Standard Dev	43%	54%	39%
IBEX	-16%	1%	-22%
CONTOUR	99%	64%	92%
WISE	64%	84%	-3%
New Horizons	52%		54%
MESSENGER	87%	99%	
GRAIL	-5%	103%	
Deep Impact	3%	32%	6%
MAVEN	25%	30%	20%
DAWN	-3%		-3%
Kepler	-11%	-45%	35%
SMAP	-42%	-47%	-47%
Juno	-19%	-21%	-23%

SEER Results – System Level Costs

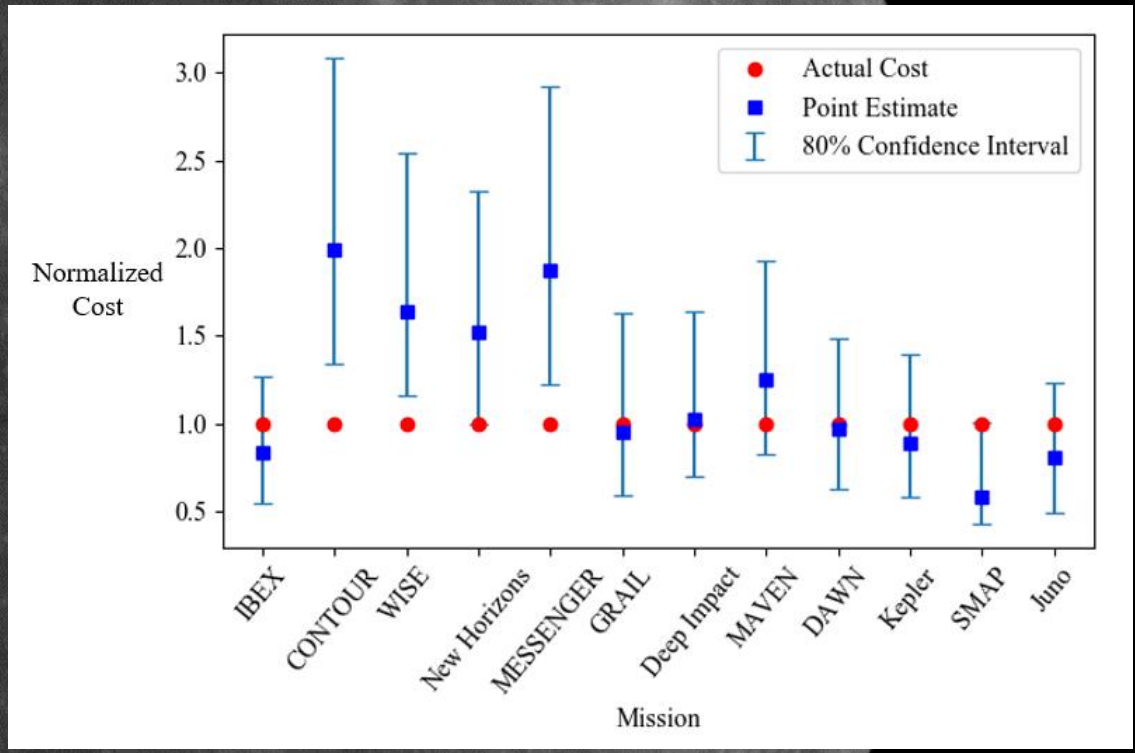
- SEER overestimates mission level systems costs
  - Cost weighted average error: 22%
  - Average error: 62%
  - Median error: 12%
  - Standard Deviation: 99%
- SEER underestimates spacecraft systems costs
  - Cost weighted average error: -28%
  - Average error: 4%
  - Median error: -33%
  - Standard Deviation: 67%
- WBS 10 IAT
  - Cost weighted average error: 33%
  - Average error: 64%
  - Median error: 42%
  - Standard Deviation: 70%
- Definitions:
  - PM (Project Management)
  - SE (Systems Engineering)
  - SMA (Safety and Mission Assurance)
  - IAT (Integration Assembly and Test)



	Mission Total	WBS 1, 2, & 3 - PM, SE, SMA	Spacecraft PM, SE, IAT	WBS 10 IAT
Weighted Average	5%	22%	-28%	33%
Average	23%	62%	4%	64%
Median	-0.3%	12%	-33%	42%
Standard Dev	43%	99%	67%	70%
IBEX	-16%	-50%	-62%	128%
CONTOUR	99%	264%	166%	165%
WISE	64%	95%	-58%	127%
New Horizons	52%	18%	15%	153%
MESSENGER	87%	244%		66%
GRAIL	-5%	81%		
Deep Impact	3%	-37%	-21%	18%
MAVEN	25%	77%	-32%	-16%
DAWN	-3%	6%	46%	-19%
Kepler	-11%	-19%	-34%	10%
SMAP	-42%	-16%	-63%	-1%
Juno	-19%	1%	-41%	

SEER Uncertainty Quantification

- User inputs optimistic, most likely, and pessimistic estimates for all inputs
  - Most of inputs are automatically filled by SEER
- 9 out of 12 (75%) of the missions fell in the 80% confidence interval
  - SEER's uncertainty capabilities performed as expected



Summary of SEER Results

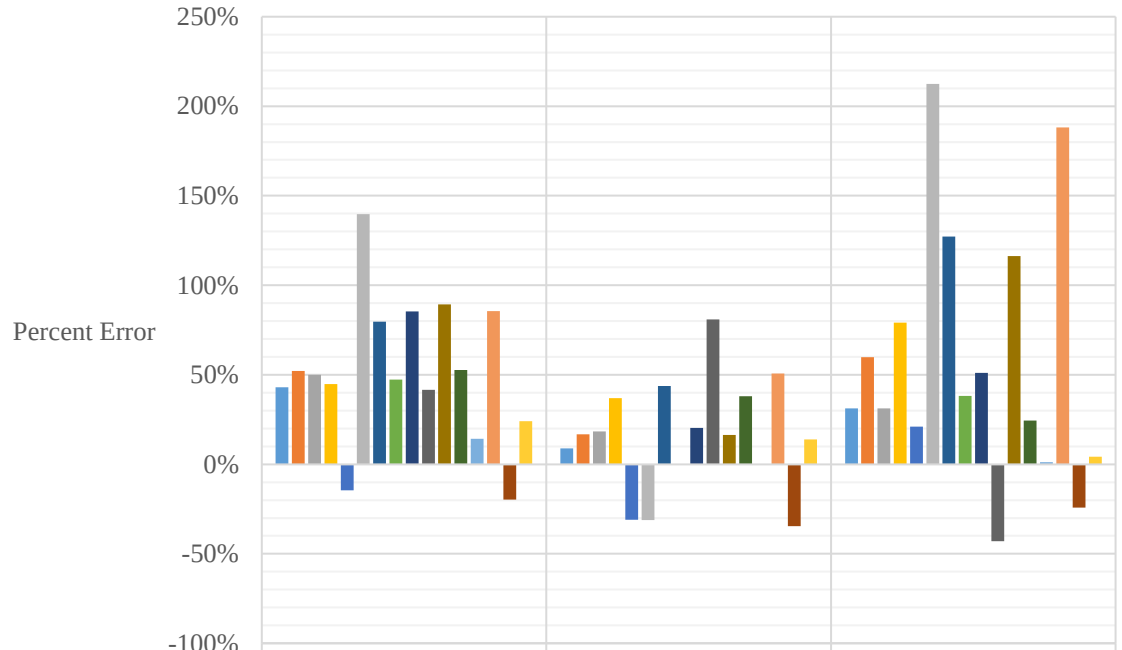
- Median error of -0.3%:
  - SEER just as likely to overestimate costs as underestimate them
- Average mission error of 23%:
  - When SEER overestimates the error is greater than when it underestimates
- Cost weighted average error: 5%
  - Meaning a low systematic error in the model as a whole
- Standard Deviation of 43%:
  - Point estimates have high variance
- SEER's uncertainty quantification capabilities performed as expected
  - 9 out of 12 (75%) of the missions fell in the 80% confidence interval

PRICE's Error in Estimating Total Mission Costs

- Average error (weighted by cost): 43%
- Average error: 52%
- Median error: 50%
- Standard Deviation: 45%
- All missions except for two were overestimated
- Large systematic error, but standard deviation not much larger than SEER
  - PRICE has similar precision to SEER

PRICE Results - Payload and Spacecraft

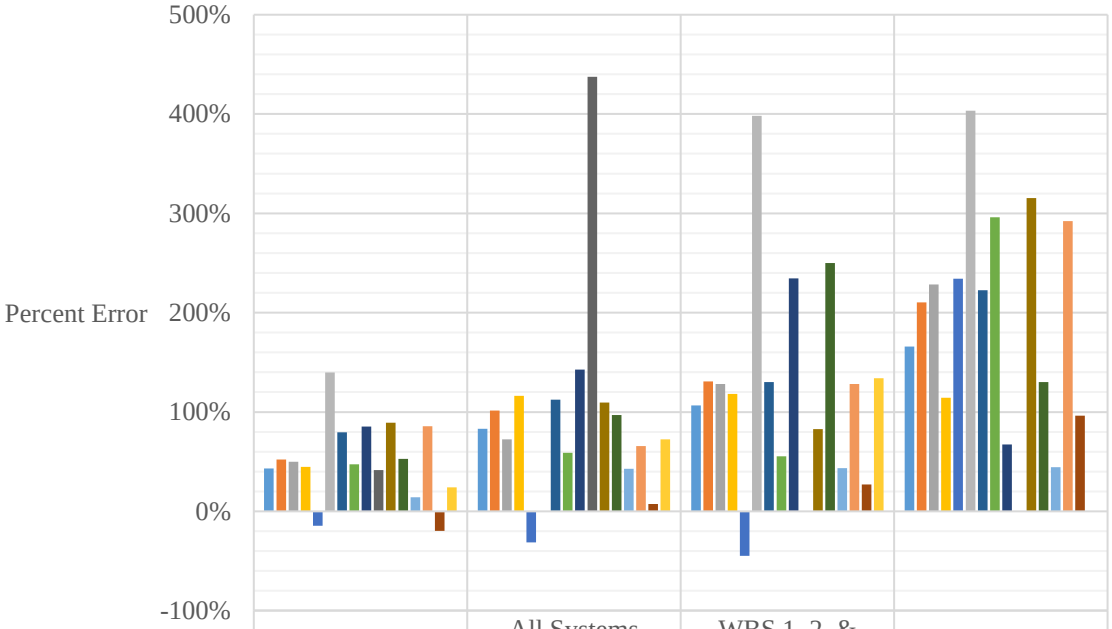
- PRICE's error and standard deviation for Payloads is much smaller than it is for spacecraft
- WBS 5 Payload
  - Average error (weighted by cost): 9%
  - Average error: 17%
  - Median error: 18%
  - Standard Deviation: 37%
- WBS 6 Spacecraft
  - Average error (weighted by cost): 31%
  - Average error: 60%
  - Median error: 31%
  - Standard Deviation: 79%



	Mission Total	Payload Subsystems Total	Spacecraft Bus Subsystem Total
Weighted Average	43%	9%	31%
Average	52%	17%	60%
Median	50%	18%	31%
Standard Dev	45%	37%	79%
IBEX	-15%	-31%	21%
CONTOUR	140%	-31%	213%
WISE	80%	44%	127%
New Horizons	47%		38%
MESSENGER	85%	20%	51%
GRAIL	42%	81%	-43%
Deep Impact	89%	16%	116%
MAVEN	53%	38%	24%
DAWN	14%		1%
Kepler	86%	51%	188%
SMAP	-20%	-35%	-24%
Juno	24%	14%	4%

PRICE Results – System Level Costs

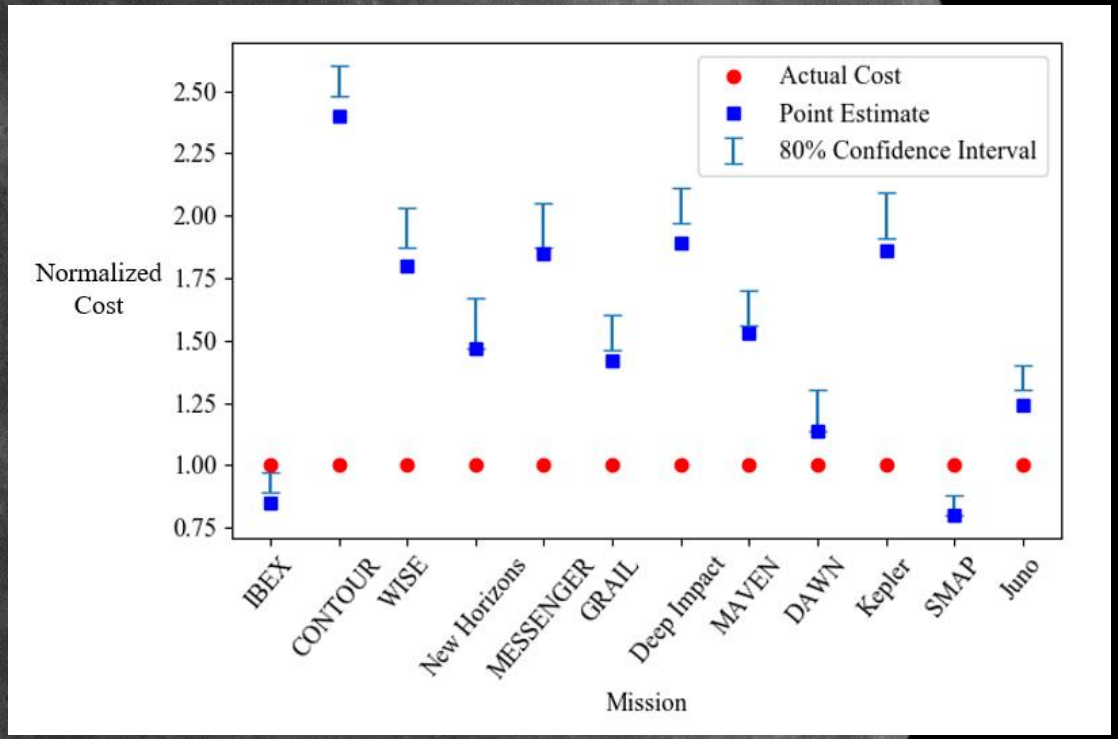
- Extremely large errors in predicting system level costs
- WBS 1, 2, & 3:
  - Average error (weighted by cost): 106%
  - Average error: 131%
  - Median error: 128%
  - Standard Deviation: 118%
- WBS 10 IAT:
  - Average error (weighted by cost): 166%
  - Average error: 210%
  - Median error: 228%
  - Standard Deviation: 114%
- System level errors are likely driven by subsystem level errors.



	Mission Total	All Systems Costs (WBS 1, 2, 3, & 10)	WBS 1 ,2, & 3 - PM, SE, SMA	WBS 10 IAT
Weighted Average	43%	83%	106%	166%
Average	52%	101%	131%	210%
Median	50%	72%	128%	228%
Standard Dev	45%	116%	118%	114%
IBEX	-15%	-31%	-45%	234%
CONTOUR	140%		398%	403%
WISE	80%	113%	130%	223%
New Horizons	47%	59%	56%	296%
MESSENGER	85%	143%	235%	67%
GRAIL	42%	437%		
Deep Impact	89%	110%	83%	315%
MAVEN	53%	97%	250%	130%
DAWN	14%	43%	43%	44%
Kepler	86%	66%	128%	292%
SMAP	-20%	7%	27%	96%
Juno	24%	72%	134%	

PRICE Uncertainty Quantification

- User inputs optimistic, most likely, and pessimistic estimates for some inputs
  - No uncertainty inputs are auto filled in PRICE
- None of the missions fell in the 80% confidence interval
- Point estimates typically fell at the 0-10% confidence level, i.e. outside the 80% confidence interval



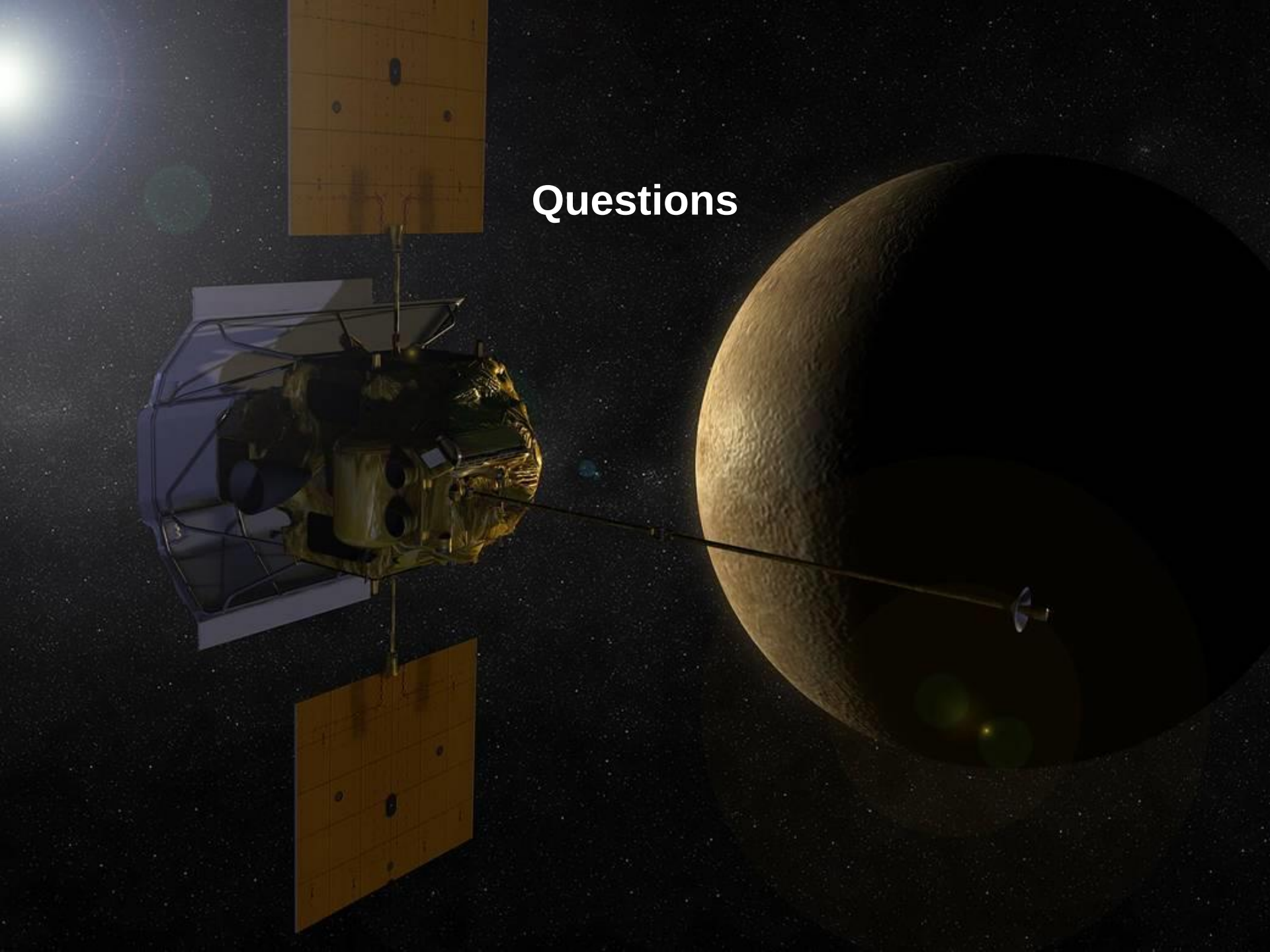
Discussion

- Mass Estimating
  - Common Assumptions
    - Optimistic: Current best estimate
    - Most Likely: Current best estimate + contingency
    - Pessimistic: Most Likely + 30%
  - Mass estimates from CDR (Critical Design Review) were used
    - 30% estimate was likely excessive and drove up the SEER estimate
- Using CDR rather than launch data added excess uncertainty
- Prototyping Assumptions
  - Standard assumption of 1.3 prototypes for components without documented number of prototypes
- Subject Matter Experts/Missing Documentation
  - Missions in this study took place over the past 20 years
  - It was not possible to ask clarifying questions about the hardware
  - This results in large uncertainty in inputs to PRICE and SEER
- Experience of Estimators
  - Estimators had formal training from PRICE Systems LLC in the use of PRICE
  - Estimators had informal training in SEER and completed several training exercises prior to beginning the study
  - Estimators had access to experienced estimators for questions

Conclusions

- SEER's uncertainty quantification capabilities performed as expected
  - 9 out of 12 (75%) of the missions fell in the 80% confidence interval

- Both SEER and PRICE had large errors and standard deviations
  - External factors which may have affected the results include:
    - Mass margin assumptions
    - Conflating uncertainty in design with uncertainty in the models
    - Prototyping assumptions
    - Inability to ask clarifying questions to subject matter experts
    - Experience of the estimators


Questions

Backup

Contact information

- Paul D. Friz PhD
  - Paul.D.Friz@nasa.gov
  - 757.864.2050

Outline

- Motivation
- Methods
- Tool Comparison (PRICE vs. SEER)
- SEER Results
- PRICE Results
- Discussion
- Conclusions

Introduction

- What is parametric costing?
  - Using a set of cost estimating relationships to predict cost when only a few key pieces of data are known
  - Cost estimating relationships are derived from historical data.
- Parametric costing tools are…
  - Commonly used at NASA and in industry to estimate life cycle costs of future space missions
  - Allow users to quickly estimate the cost of a mission concept before detailed designs have been completed
  - Frequently used by NASA to evaluate spacecraft and instrument proposals
- Two parametric costing tools commonly used at NASA are…
  - PRICE TruePlanning (Space Mission Catalog) by PRICE Systems LLC
  - SEER-H (with EOS and IC plugins) by Galorath Inc

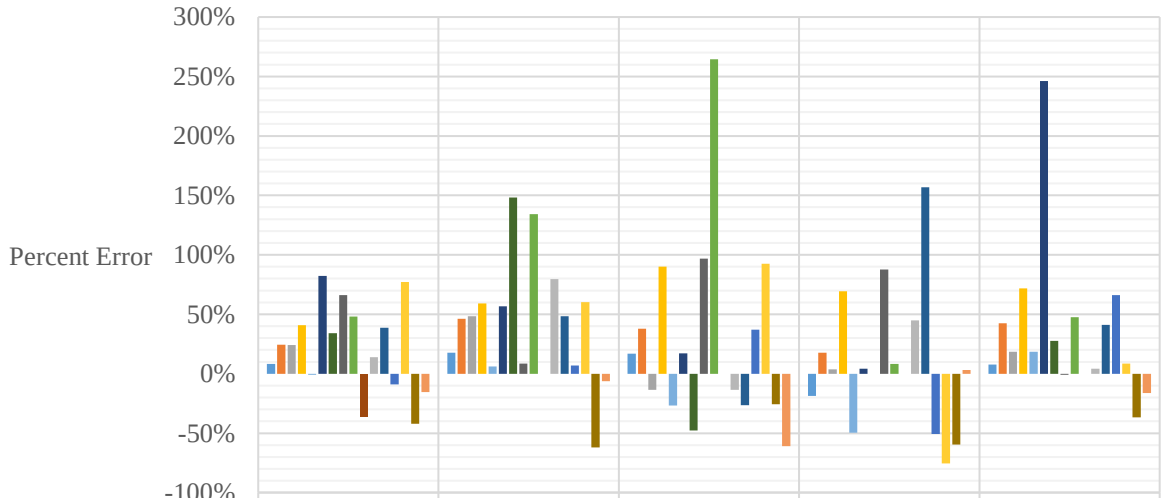
What Do PRICE and SEER Estimate?

- PRICE and SEER were used to estimate the following items.
- NASA Work Breakdown Structure (WBS)
  - 1 Project Management
  - 2 Systems Engineering
  - 3 Safety & Mission Assurance
  - 4 Science/Technology
  - 5 Payload
  - 6 Spacecraft
  - 7 Mission Operations
  - 8 Launch Vehicles/Services
  - 9 Ground Systems
  - 10 Systems Integration & Testing
  - 11 Education and Public Outreach

- WBS 6: Spacecraft, is broken down further
  - Project Management
  - Systems Engineering
  - Integration, Assembly, and Test
  - Structures
  - Thermal
  - Propulsion
  - Guidance, Navigation, and Control (GN&C)
  - Communication
  - Electrical Power
  - Harness
  - Command and Data Handling (C&DH)

SEER Results – Spacecraft Subsystems

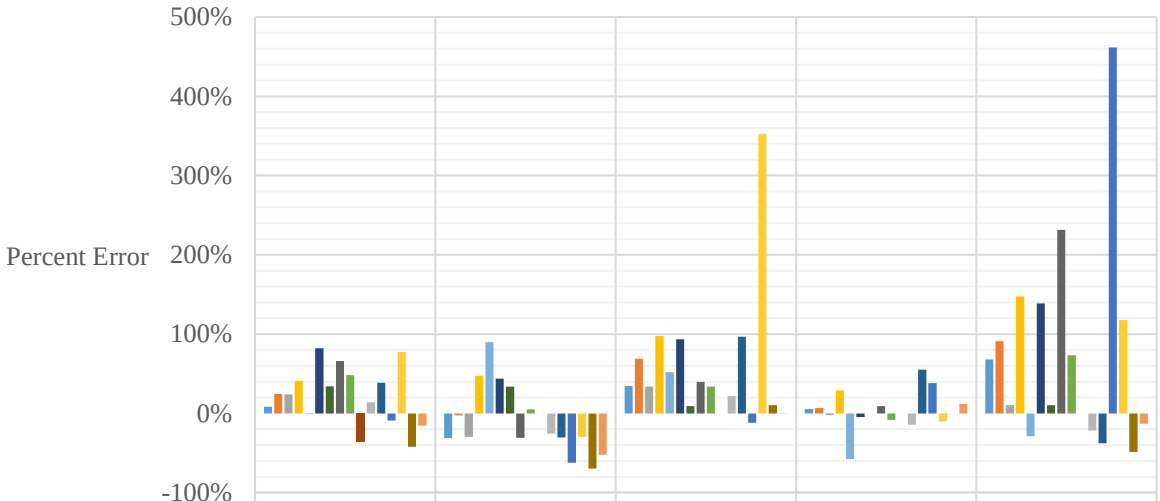
- Large errors in estimating individual subsystems tend to average each other out.



	Bus Subsystems Total	Structures	Thermal	Propulsion	GN&C
Weighted Average	8%	18%	17%	-19%	8%
Average	24%	46%	38%	18%	43%
Median	24%	48%	-14%	4%	19%
Standard Dev	41%	59%	90%	69%	72%
IBEX	-1%	6%	-27%	-50%	19%
CONTOUR	82%	57%	17%	4%	246%
WISE	34%	148%	-48%		28%
New Horizons	66%	8%	97%	88%	-1%
MESSENGER	48%	134%	264%	8%	48%
GRAIL	-36%				
Deep Impact	14%	80%	-14%	45%	4%
MAVEN	39%	48%	-27%	157%	41%
DAWN	-9%	7%	37%	-51%	66%
Kepler	77%	60%	93%	-75%	9%
SMAP	-42%	-62%	-26%	-60%	-37%
Juno	-15%	-6%	-61%	3%	-16%

SEER Results – Spacecraft Subsystems

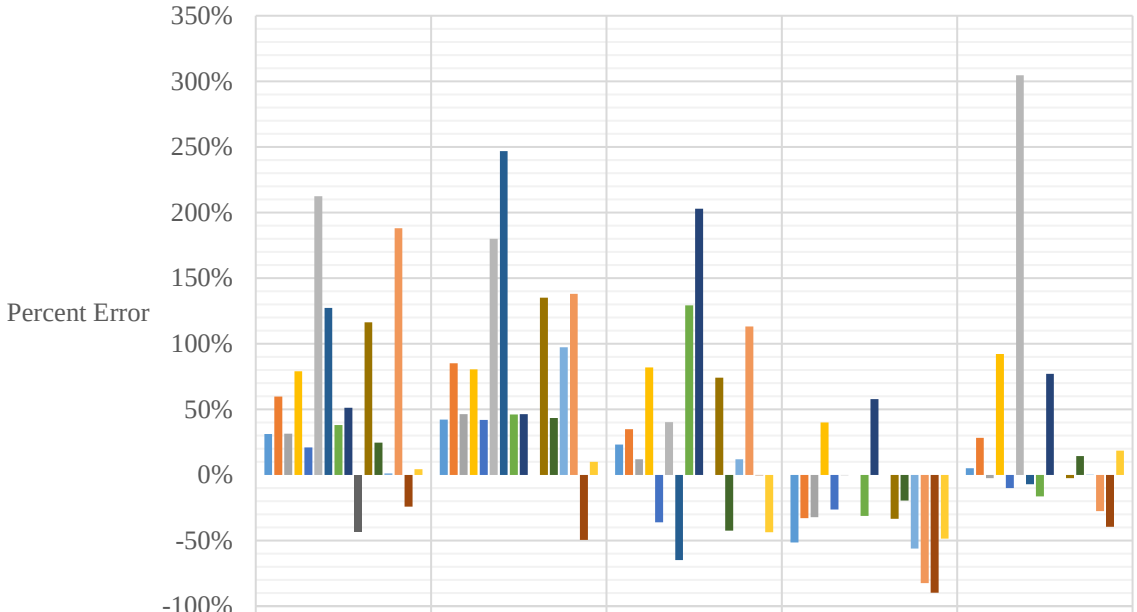
- Extremely large errors in C&DH and Power subsystems due to lack of knowledge of heritage.



	Bus Subsystems Total	Comm	Electrical Power	Harness	C&DH
Weighted Average	8%	-31%	35%	6%	68%
Average	24%	-2%	69%	7%	91%
Median	24%	-30%	34%	-2%	10%
Standard Dev	41%	47%	97%	29%	148%
IBEX	-1%	90%	52%	-57%	-29%
CONTOUR	82%	44%	93%	-5%	139%
WISE	34%	34%	9%		10%
New Horizons	66%	-31%	40%	9%	232%
MESSENGER	48%	5%	34%	-8%	73%
GRAIL	-36%				
Deep Impact	14%	-26%	21%	-14%	-22%
MAVEN	39%	-30%	97%	55%	-38%
DAWN	-9%	-62%	-12%	38%	462%
Kepler	77%	-30%	352%	-10%	118%
SMAP	-42%	-70%	10%	0%	-48%
Juno	-15%	-52%	0%	12%	-13%

PRICE Results – Spacecraft Subsystems

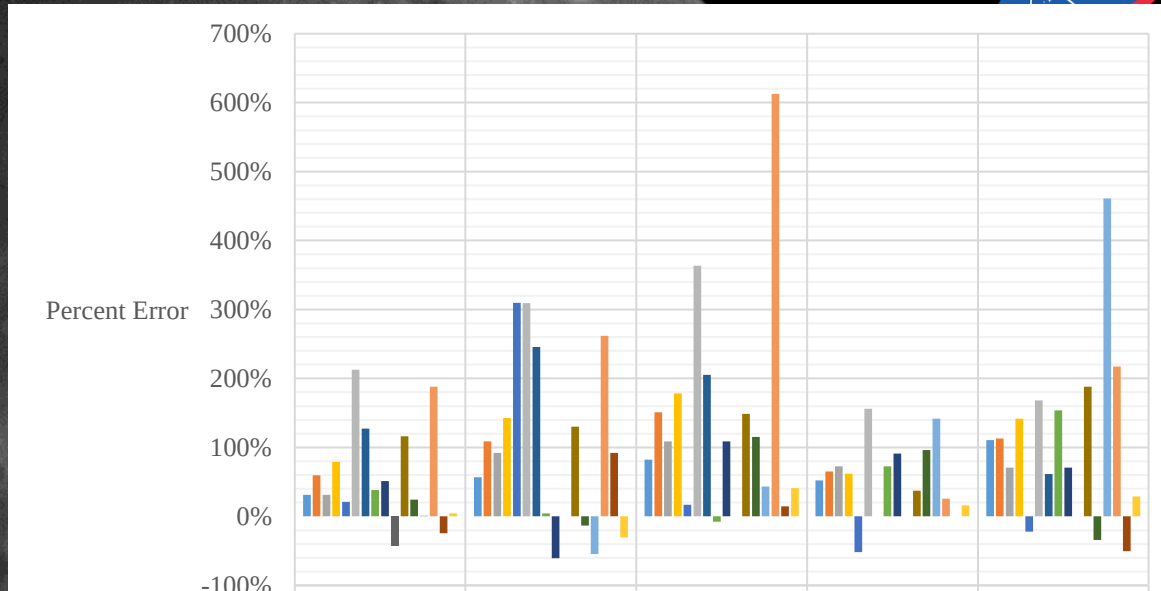
- Large errors in predicting costs of individual subsystems.



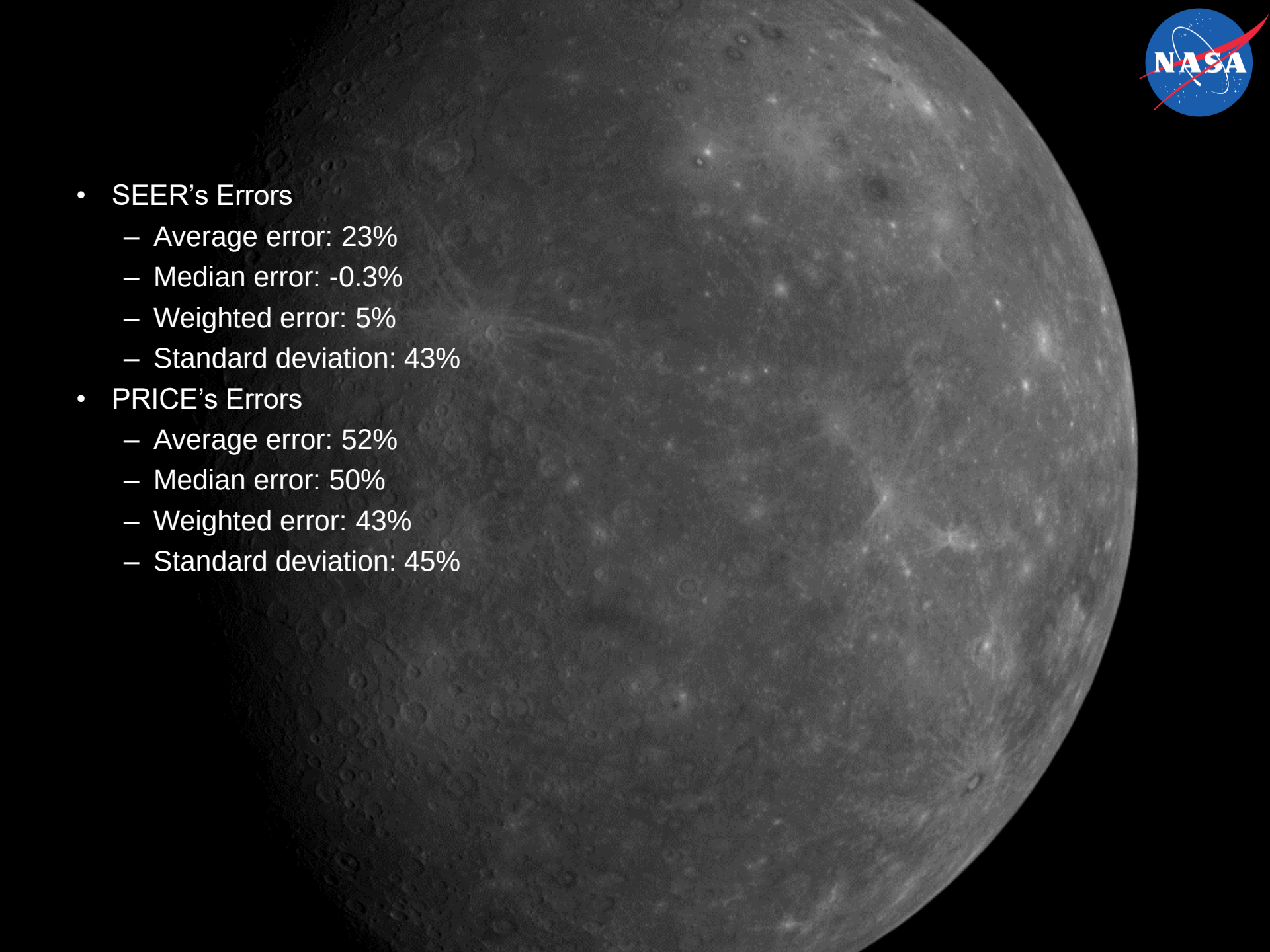
	Spacecraft Bus Subsystem Total	Structures	Thermal	Propulsion	GN&C
Weighted Average	31%	42%	23%	-51%	5%
Average	60%	85%	35%	-33%	28%
Median	31%	46%	12%	-32%	-2%
Standard Dev	79%	80%	82%	40%	92%
IBEX	21%	42%	-36%	-26%	-10%
CONTOUR	213%	180%	40%	-0.2%	305%
WISE	127%	247%	-65%		-7%
New Horizons	38%	46%	129%	-31%	-16%
MESSENGER	51%	46%	203%	58%	77%
GRAIL	-43%				
Deep Impact	116%	135%	74%	-34%	-2%
MAVEN	24%	43%	-43%	-20%	14%
DAWN	1%	97%	12%	-56%	0.1%
Kepler	188%	138%	113%	-82%	-28%
SMAP	-24%	-50%	-0.1%	-90%	-40%
Juno	4%	10%	-44%	-49%	19%

PRICE Results – Spacecraft Subsystems

- Largest standard deviations are in subsystems which are primarily electronics.
  - C&DH
  - Power
  - Communications
  - GN&C



	Spacecraft Bus Subsystem Total	Comm	Electrical Power	Harness	C&DH
Weighted Average	31%	57%	82%	52%	111%
Average	60%	109%	151%	65%	113%
Median	31%	92%	109%	73%	71%
Standard Dev	79%	143%	178%	62%	142%
IBEX	21%	310%	17%	-52%	-22%
CONTOUR	213%	309%	363%	156%	168%
WISE	127%	245%	205%		62%
New Horizons	38%	4%	-8%	73%	154%
MESSENGER	51%	-61%	109%	91%	71%
GRAIL	-43%				
Deep Impact	116%	130%	149%	37%	188%
MAVEN	24%	-13%	115%	96%	-34%
DAWN	1%	-55%	43%	142%	461%
Kepler	188%	262%	613%	26%	217%
SMAP	-24%	92%	15%		-51%
Juno	4%	-31%	41%	16%	29%



- SEER’s Errors
  - Average error: 23%
  - Median error: -0.3%
  - Weighted error: 5%
  - Standard deviation: 43%
- PRICE’s Errors
  - Average error: 52%
  - Median error: 50%
  - Weighted error: 43%
  - Standard deviation: 45%