# Bayesian Dirichlet Cost Rules of Thumb

Samuel Fleischer and Melissa Hooke
Jet Propulsion Laboratory, California Institute of Technology

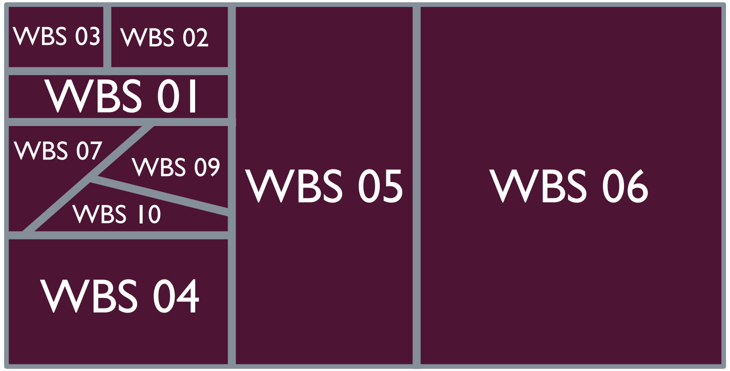
2023 NASA Cost and Schedule Symposium
May 2 - 4, 2023
Presented by Melissa Hooke
Melissa.A.Hooke@jpl.nasa.gov

Jet Propulsion Laboratory
California Institute of Technology

# Cost Rules of Thumb (ROT)

- Cost rules of thumb are critical components of early project formulation, specifically for low CML (Concept Maturity Level)
- Rough costs of spacecraft subsystems must be generated before more detailed estimates can be made
- It is generally thought that there is a “typical” percent allocation of total mission cost to the Level 2 WBS (Work Breakdown Structure) elements

Mission A




Mission B


| | | | |
|---|---|---|---|
| WBS 01 | Project Management | WBS 06 | Flight System / Spacecraft |
| WBS 02 | Systems Engineering | WBS 07 | Mission Operations System (MOS) |
| WBS 03 | Safety and Mission Assurance | WBS 08 | Launch Vehicle / Services |
| WBS 04 | Science / Technology | WBS 09 | Ground Data System (GDS) |
| WBS 05 | Payload(s) | WBS 10 | System Integration, Assembly, Test & Check |

# Traditional ROT Approach

| Mission | Category | Fixed FY WBS 01 Cost % | Fixed FY WBS 02 Cost % | Fixed FY WBS 03 Cost % | Fixed FY WBS 04 Cost % | Fixed FY WBS 05 Cost % | Fixed FY WBS 06 Cost % | Fixed FY WBS 07 Cost % | Fixed FY WBS 09 Cost % | Fixed FY WBS 10 Cost % |
|---|---|---|---|---|---|---|---|---|---|---|
| x | MIDEX | 4.5% | 3.6% | 3.4% | 18.1% | 25.6% | 39.5% | 4.0% | 1.3% | |
| x | MIDEX | 8.5% | 1.0% | 1.5% | 1.7% | 45.3% | 34.1% | 3.9% | 3.9% | |
| x | MIDEX | 5.1% | 3.2% | 1.7% | 8.3% | 34.4% | 30.9% | 1.6% | 12.0% | 2.7% |
| x | MIDEX | 3.6% | 2.1% | 2.2% | 4.7% | 30.9% | 39.8% | 10.9% | 2.1% | 3.7% |
| x | MIDEX | 6.5% | 1.8% | 7.9% | | 40.7% | 35.4% | | 4.1% | 3.5% |
| x | MIDEX | 6.7% | 4.1% | 2.0% | 4.4% | 43.1% | 34.5% | | 2.6% | 2.6% |
| x | MIDEX | 1.7% | | | | 45.0% | 33.5% | 8.5% | 11.3% | |
| x | MIDEX | 6.0% | 10.0% | | 2.2% | 18.7% | 53.1% | 4.9% | 5.1% | |
| Average Percent (ignoring missing data): | | 5.3% | 3.7% | 3.1% | 6.6% | 35.5% | 37.6% | 5.6% | 5.3% | 3.1% |
| Rescaled Rule of Thumb: | | 5.0% | 3.5% | 3.0% | 6.2% | 33.5% | 35.5% | 5.3% | 5.0% | 2.9% |

# Dirichlet ROT Approach

- The ***Dirichlet distribution*** describes the uncertainty of multivariate distributions
- Can be understood as the multivariate extension of the Beta distribution
  - Binomial = model for # successes of *n* trials with probability *p*
  - Beta = model for uncertainty around *p*
  - Multinomial = multivariate extension to binomial distribution, defined by vector of probabilities which sum to 1
- The Dirichlet distribution is the simplest distribution on the ***simplex*** (vectors whose components sum to one)

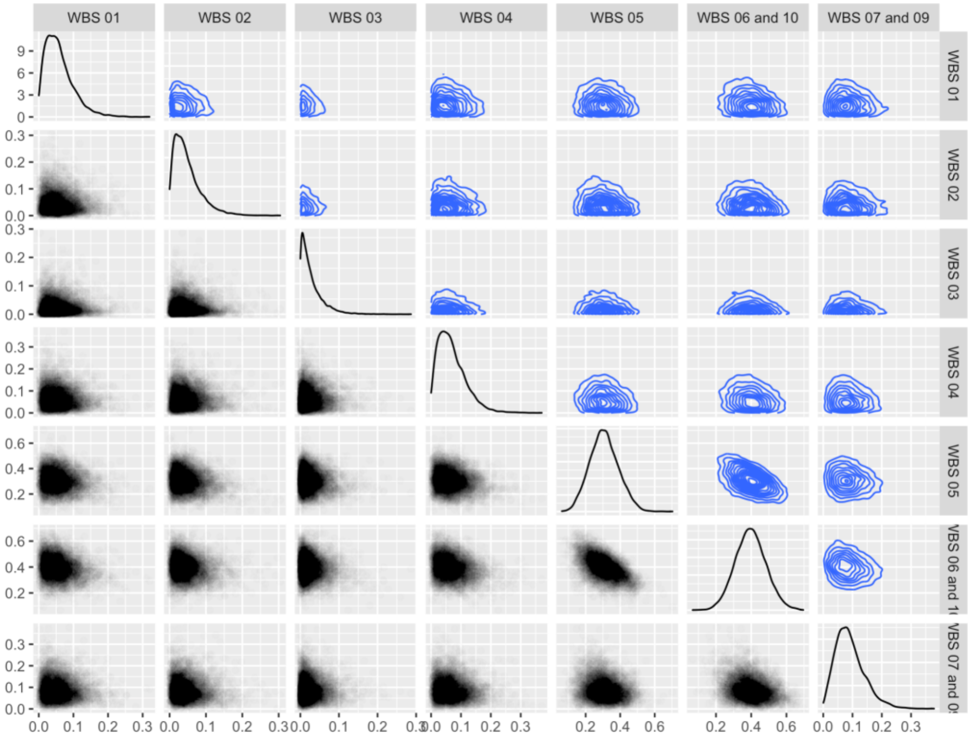



**Figure:** Full uncertainty and correlation between WBS elements of posterior predictions

# Aleatoric and Epistemic Uncertainty

- ***Aleatoric Uncertainty***: uncertainty which arises from inherent randomness
  - *e.g. normal distribution standard deviation*
- ***Epistemic Uncertainty***: uncertainty in the model form or in the uncertainty in the parameters of the model form(s)
  - *e.g. uncertainty in the value of the std. deviation of the normal distribution*
  - *e.g. the uncertainty could be described by a normal distribution or a t distribution*
  - For our model we only consider uncertainty in the parameters
  - The Dirichlet distribution is parameterized by a mean vector $\vec{\alpha}$ and a one-dimensional precision parameter $\phi$

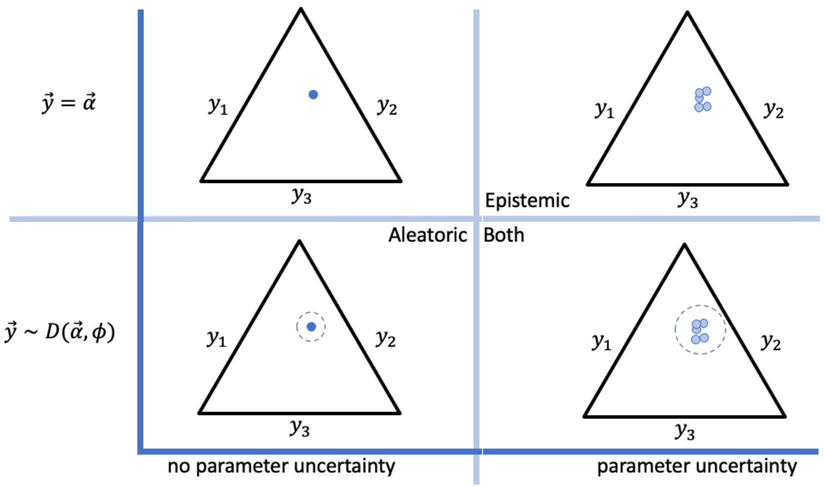



Figure: Schematic of aleatoric and epistemic uncertainty for a Dirichlet model on $\Delta^2$ (the simplex on $\mathbb{R}^3$; the points $[x, y, z]$ such that $x + y + z = 1$ and $x, y, z \geq 0$)

# Mission Classes and Model Form

- Each mission class has a different set of cost categories, i.e. different WBS and mission phases included based on the AO
- We split the data based on mission AO and form separate models for each

| Mission Class | Mission Cost Range | Number of Records | Cost Categories |
|---|---|---|---|
| Flagship | > $2B | 3 | Phase B-D WBS 1-10 except 8 |
| New Frontiers | $1B | 4 | Phase A-D WBS 1-10 except 8 |
| Discovery | $500M | 13 | Phase B-D WBS 1-10 except 8, and Phase E-F WBS 1-4, 7, and 9 |
| MIDEX | $300M | 8 | Phase A-F WBS 1-10 except 8 |

For Flagship, New Frontiers, and MIDEX:

$$\vec{c} \sim \text{Dirichlet}(\vec{\alpha}, \phi)$$

$\vec{c}$ is a nine-dimensional vector on the simplex representing WBS 1, 2, 3, 4, 5, 6, 7, 9, and 10 for the Phases in the table

For Discovery:

$$\vec{c}_{\text{phase}} \sim \text{Dirichlet}\left(\vec{\alpha}_{\text{phase}}, \phi_{\text{phase}}\right)$$
$$\vec{c}_{BCD} \sim \text{Dirichlet}\left(\vec{\alpha}_{BCD}, \phi_{BCD}\right)$$
$$\vec{c}_{EF} \sim \text{Dirichlet}\left(\vec{\alpha}_{EF}, \phi_{EF}\right)$$
$$\vec{c} = \left[c_{\text{phase},1}\vec{c}_{BCD} \quad c_{\text{phase},2}\vec{c}_{EF}\right]$$

** We also have priors on the Dirichlet parameters (outside scope of paper & presentation)

# Missing Data Imputation

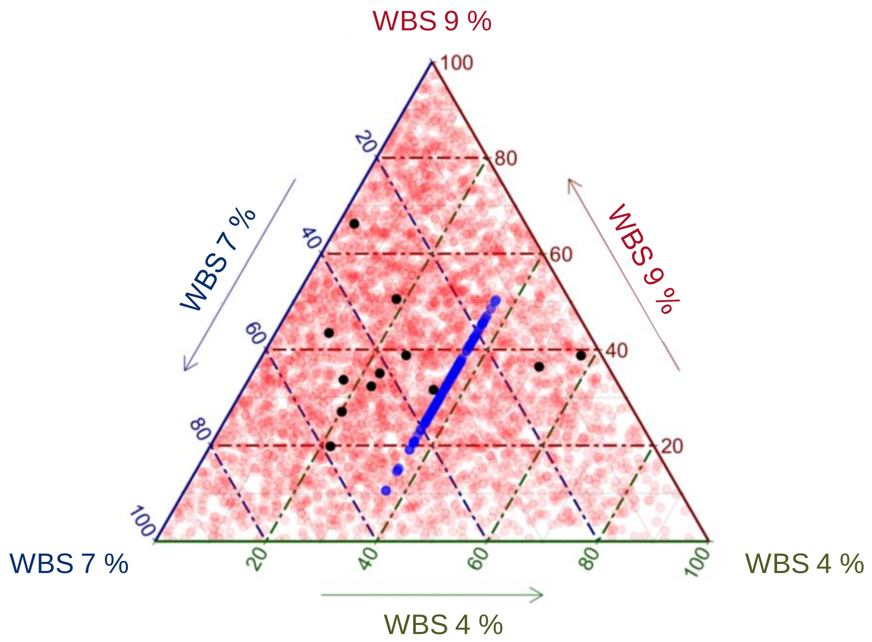
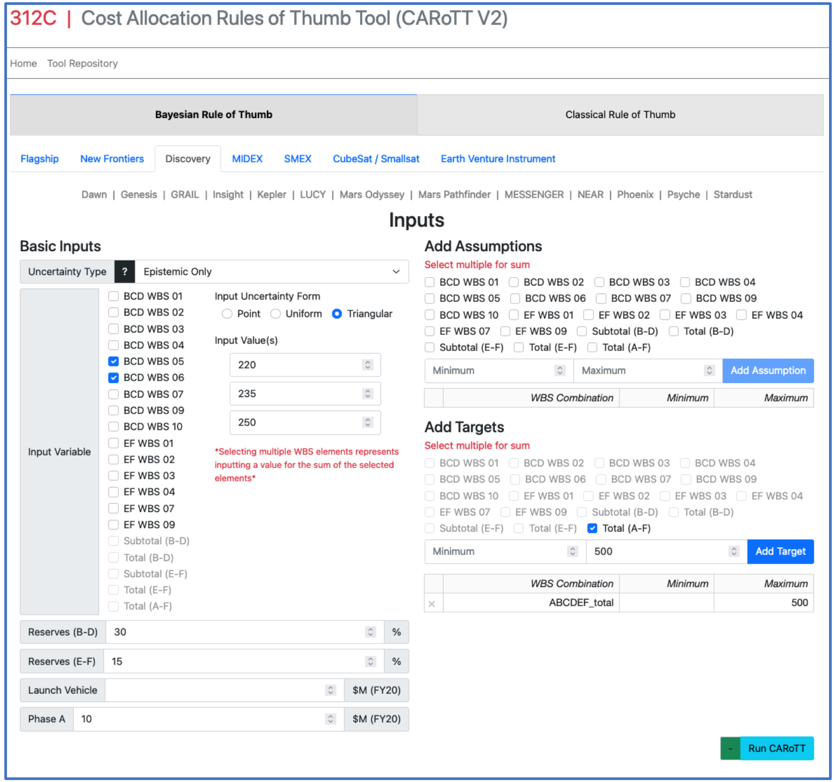



**Figure:** Example of imputation where black points are data from other missions, WBS 4 % is known and blue points show uncertainty on split between WBS 7 & 9, red are posterior predictive for future missions

- Some missions have incomplete WBS breakdowns
  - WBS 7 is sometimes bookkept in WBS 9
  - WBS 10 is sometimes bookkept in WBS 6
  - WBS 1, 2, 3 often not split rigorously
- Costs imputed using ***multiple imputations*** with a Dirichlet distribution on elements that are often bookkept together
- After imputation, the larger Dirichlet model across all WBS elements is fitted to imputed datasets and the posteriors are combined to account for imputation uncertainty

# Cost Allocation Rule of Thumb Tool (CARoTT)

python
Dash by plotly
django

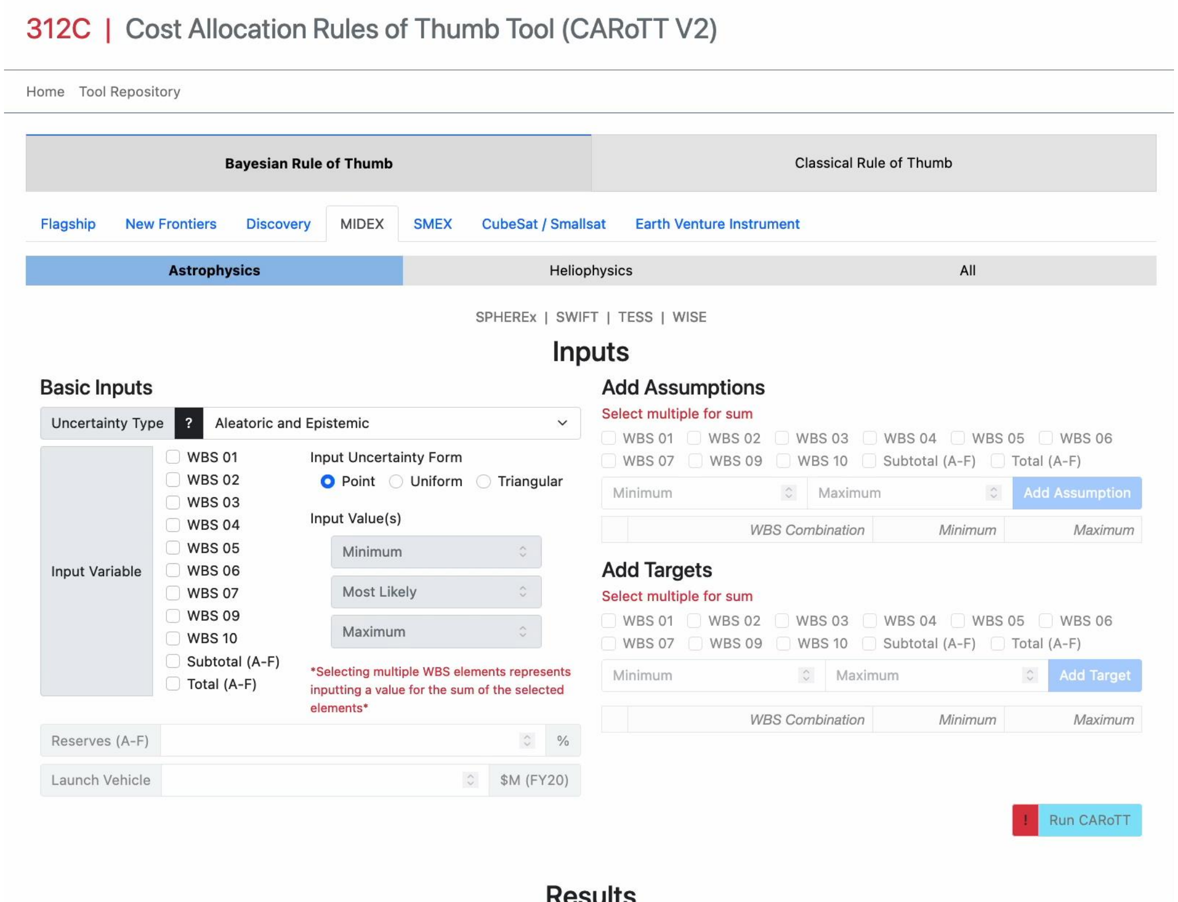



CARoTT Capabilities:

1. Predict unknown individual WBS costs
2. Predict total mission cost
3. Estimate probabilities of meeting cost targets
4. Visualize WBS correlations
5. Extract posterior samples
6. Choose what types of uncertainty to include

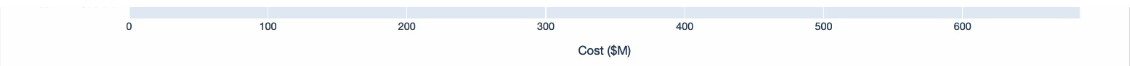
# CARoTT Demo

1. Predict unknown individual WBS costs from one or more known WBS costs or a known total cost target



# CARoTT Demo

## 2. Predict total cost from one or more known WBS element costs




### Posterior Predictive Distribution Summary

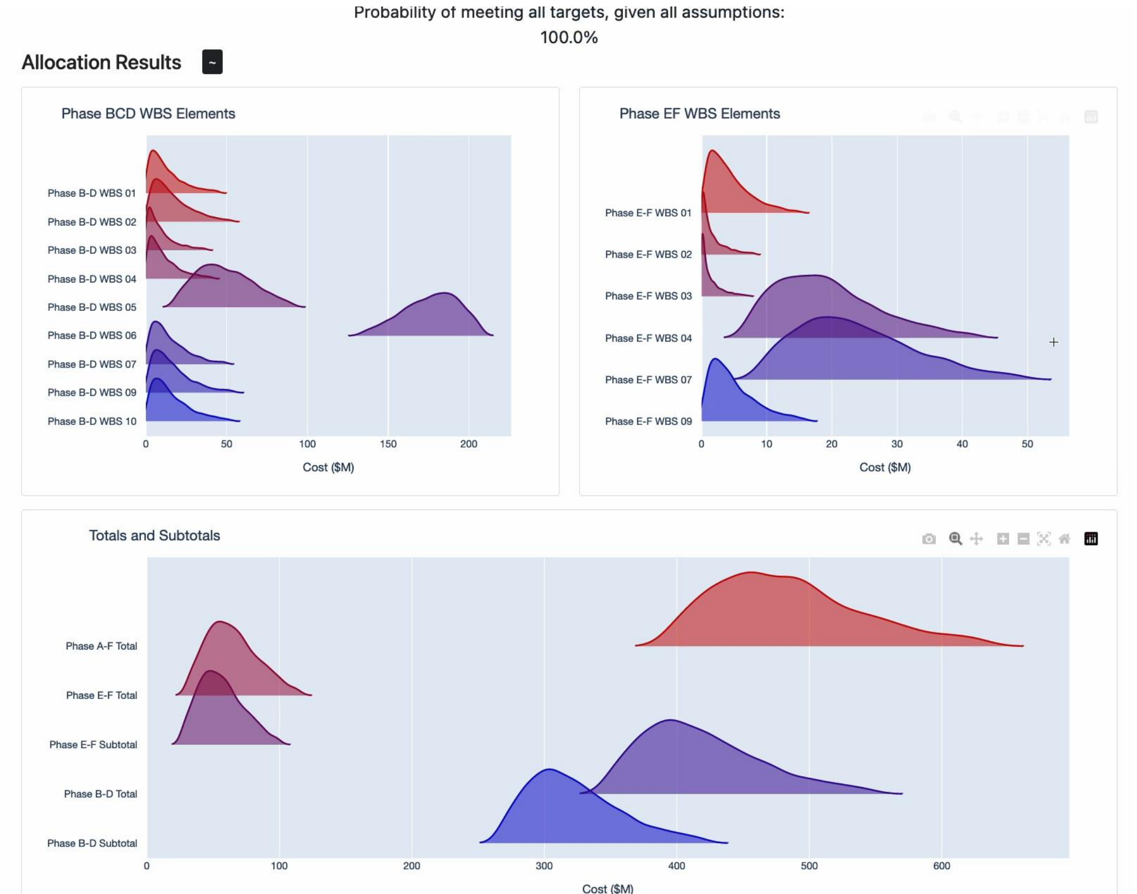
Export

| Line Item | 2.5th Percentile | 30th Percentile | 50th Percentile | 70th Percentile | 97.5th Percentile |
|---|---|---|---|---|---|
| Phase B-D WBS 01 | 0.478 | 5.366 | 9.418 | 15.611 | 45.607 |
| Phase B-D WBS 02 | 1.089 | 7.372 | 12.363 | 19.376 | 53.691 |
| Phase B-D WBS 03 | 0.109 | 2.773 | 5.93 | 10.682 | 36.002 |
| Phase B-D WBS 04 | 0.355 | 4.266 | 7.95 | 13.6 | 42.637 |
| Phase B-D WBS 05 | 17.533 | 37.387 | 47.377 | 58.225 | 91.374 |
| Phase B-D WBS 06 | 133.26 | 166.769 | 177.49 | 187.685 | 207.717 |
| Phase B-D WBS 07 | 0.769 | 6.471 | 10.986 | 17.631 | 47.862 |
| Launch Vehicle | 0 | 0 | 0 | 0 | 0 |
| Phase B-D WBS 09 | 1.197 | 7.758 | 12.843 | 19.747 | 52.081 |
| Phase B-D WBS 10 | 0.995 | 7.202 | 12.475 | 19.387 | 53.902 |
| Phase B-D Subtotal | 265.023 | 297.927 | 314.971 | 336.201 | 419.145 |
| Phase B-D Reserves | 79.507 | 89.378 | 94.491 | 100.86 | 125.744 |
| Phase B-D Total | 344.529 | 387.305 | 409.462 | 437.062 | 544.889 |
| Phase E-F WBS 01 | 0.237 | 2.081 | 3.501 | 5.441 | 15.373 |
| Phase E-F WBS 02 | 0.002 | 0.343 | 0.925 | 1.911 | 8.42 |
| Phase E-F WBS 03 | 0 | 0.19 | 0.604 | 1.454 | 7.333 |
| Phase E-F WBS 04 | 6.136 | 13.821 | 17.947 | 22.862 | 41.979 |
| Phase E-F WBS 07 | 8.148 | 17.225 | 22.055 | 27.802 | 49.22 |
| Phase E-F WBS 09 | 0.38 | 2.392 | 3.894 | 6.186 | 16.138 |
| Phase E-F Subtotal | 25.705 | 44.132 | 53.458 | 64.228 | 103.12 |
| Phase E-F Reserves | 3.856 | 6.62 | 8.019 | 9.634 | 15.468 |
| Phase E-F Total | 29.561 | 50.751 | 61.477 | 73.862 | 118.589 |
| Phase A | 0 | 0 | 0 | 0 | 0 |
| Phase A-F Total | 388.345 | 445.209 | 473.205 | 507.28 | 628.634 |

### Posterior Exploration +

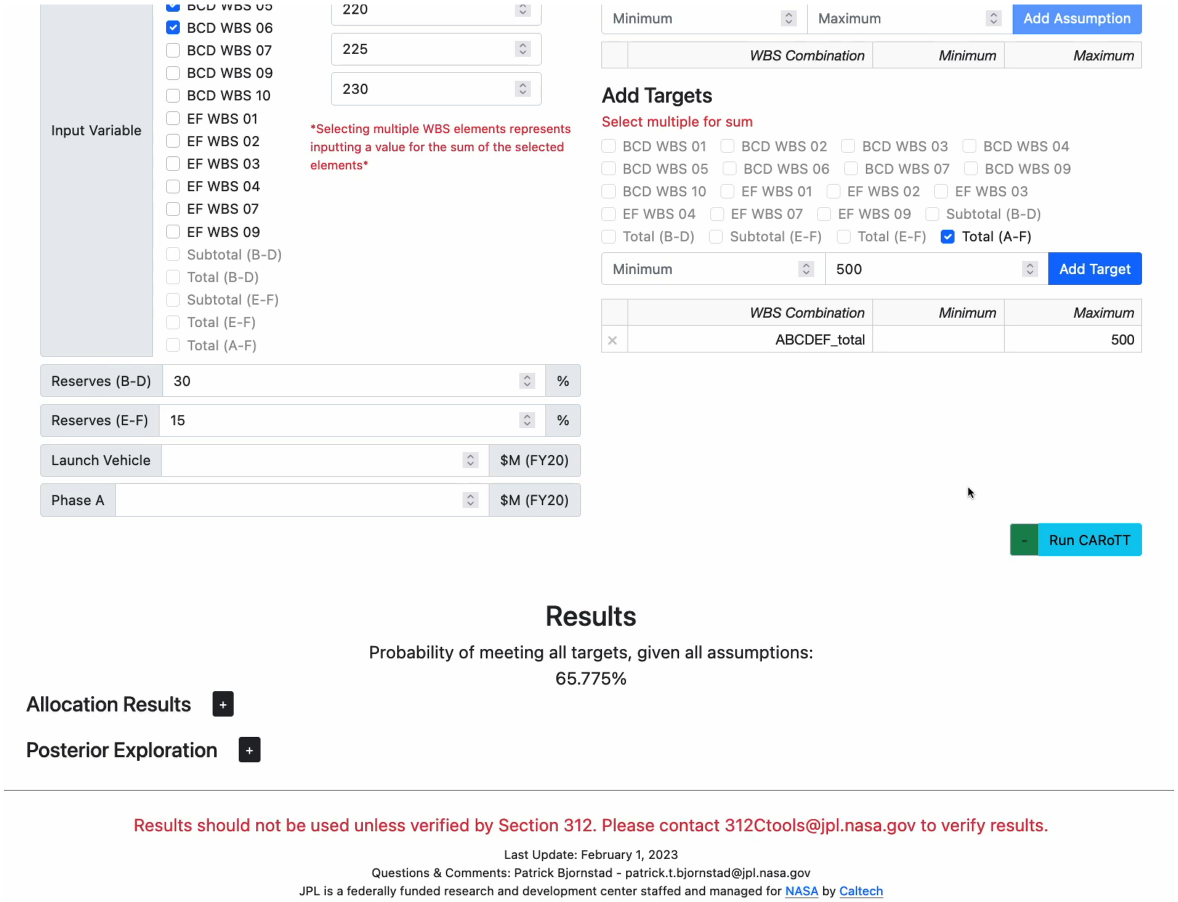
# CARoTT Demo

3. Estimate probabilities of simultaneously meeting one or more cost targets



# CARoTT Demo

4. Visualize marginal or two-dimensional joint predictive densities



# CARoTT Demo

5. Extract a large sample of the full posterior predictive distribution

Probability of meeting all targets, given all assumptions:
65.825%

**Allocation Results** +

**Posterior Exploration** -

**X Variable(s) for Scatter**

(•) Native ( ) Log

Select multiple for sum

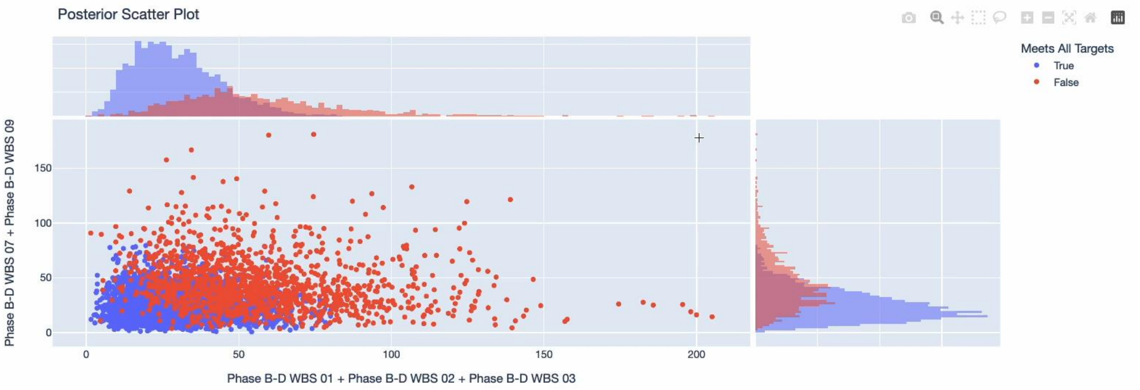
- [x] BCD WBS 01 [x] BCD WBS 02 [x] BCD WBS 03 [ ] BCD WBS 04
- [ ] BCD WBS 05 [ ] BCD WBS 06 [ ] BCD WBS 07 [ ] BCD WBS 09
- [ ] BCD WBS 10 [ ] EF WBS 01 [ ] EF WBS 02 [ ] EF WBS 03 [ ] EF WBS 04
- [ ] EF WBS 07 [ ] EF WBS 09 [ ] Subtotal (B-D) [ ] Total (B-D)
- [ ] Subtotal (E-F) [ ] Total (E-F) [ ] Total (A-F)

**Y Variable(s) for Scatter**

(•) Native ( ) Log

Select multiple for sum

- [ ] BCD WBS 01 [ ] BCD WBS 02 [ ] BCD WBS 03 [ ] BCD WBS 04
- [ ] BCD WBS 05 [ ] BCD WBS 06 [x] BCD WBS 07 [x] BCD WBS 09
- [ ] BCD WBS 10 [ ] EF WBS 01 [ ] EF WBS 02 [ ] EF WBS 03 [ ] EF WBS 04
- [ ] EF WBS 07 [ ] EF WBS 09 [ ] Subtotal (B-D) [ ] Total (B-D)
- [ ] Subtotal (E-F) [ ] Total (E-F) [ ] Total (A-F)




**Posterior Predictive Distribution**

Export

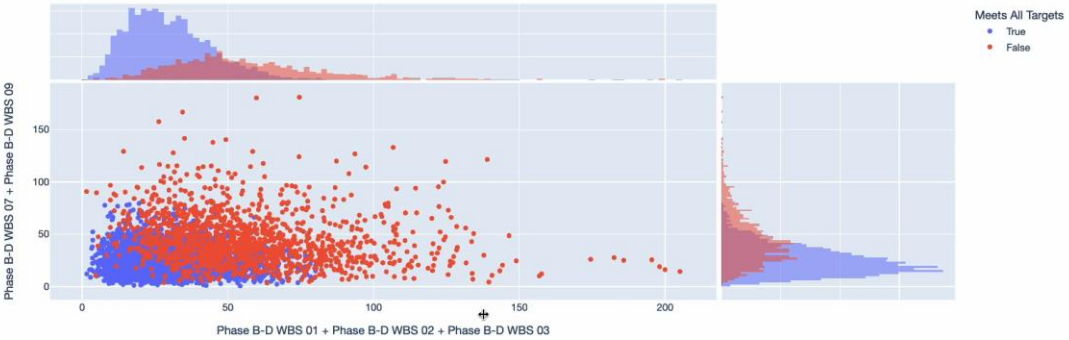
| BCD WBS 01 | BCD WBS 02 | BCD WBS 03 | BCD WBS 04 | BCD WBS 05 | BCD WBS 06 | BCD WBS 07 | LV | BCD WBS 09 | BCD WBS 10 | BD Subtotal | BD Reserves | BD Total | EF WBS 01 | EF WBS 02 | EF WBS 03 | EF WBS 04 | EF WBS 07 | EF WBS 09 | Subto |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 29.906 | 6.152 | 39.92 | 2.051 | 76.906 | 147.465 | 29.708 | 0 | 22.865 | 5.22 | 360.192 | 108.058 | 468.25 | 5.257 | 3.356 | 0.582 | 41.383 | 23.779 | 4.114 | 78.4 |

# CARoTT Demo

6. Choose what types of uncertainty they want to include (aleatoric only, epistemic only, aleatoric and epistemic, neither aleatoric nor epistemic)

☐ Subtotal (E-F) ☐ Total (E-F) ☐ Total (A-F) ☐ Subtotal (E-F) ☐ Total (E-F) ☐ Total (A-F)

Posterior Scatter Plot




**Posterior Predictive Distribution**

Export

| BCD WBS 01 | BCD WBS 02 | BCD WBS 03 | BCD WBS 04 | BCD WBS 05 | BCD WBS 06 | BCD WBS 07 | LV | BCD WBS 09 | BCD WBS 10 | BD Subtotal | BD Reserves | BD Total | EF WBS 01 | EF WBS 02 | EF WBS 03 | EF WBS 04 | EF WBS 07 | EF WBS 09 | Subto |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 29.906 | 6.152 | 39.92 | 2.051 | 76.906 | 147.465 | 29.708 | 0 | 22.865 | 5.22 | 360.192 | 108.058 | 468.25 | 5.257 | 3.356 | 0.582 | 41.383 | 23.779 | 4.114 | 78.4 |
| 39.125 | 15.155 | 7.774 | 1.449 | 30.466 | 195.808 | 24.155 | 0 | 13.573 | 9.793 | 337.3 | 101.19 | 438.49 | 16.185 | 0.52 | 2.272 | 26.723 | 16.312 | 18.897 | 80. |
| 25.756 | 3.124 | 0.353 | 2.536 | 67.992 | 158.567 | 25.67 | 0 | 2.162 | 25.295 | 311.454 | 93.436 | 404.89 | 0.942 | 0.186 | 0.049 | 7.872 | 13.419 | 11.125 | 33.5 |
| 13.857 | 3.319 | 14.43 | 4.828 | 46.48 | 180.022 | 12.275 | 0 | 2.958 | 25.675 | 303.842 | 91.153 | 394.995 | 18.377 | 15.6 | 0 | 32.496 | 18.114 | 12.295 | 96.8 |
| 19.469 | 5.885 | 7.275 | 12.159 | 90.228 | 134.548 | 11.993 | 0 | 3.66 | 8.038 | 293.255 | 87.976 | 381.231 | 0.93 | 0.075 | 0.035 | 9.506 | 15.123 | 3.279 | 28.9 |
| 1.642 | 22.59 | 8.875 | 6.788 | 30.234 | 191.02 | 14.39 | 0 | 7.249 | 11.742 | 294.53 | 88.359 | 382.889 | 2.006 | 0.045 | 0.542 | 20.854 | 29.848 | 2.569 | 55.8 |
| 1.963 | 8.521 | 0.104 | 4.278 | 58.558 | 168.181 | 6.838 | 0 | 3.772 | 9.534 | 261.749 | 78.525 | 340.274 | 4.475 | 3.683 | 2.471 | 12.835 | 26.475 | 1.406 | 51.3 |
| 0.453 | 7.554 | 0.229 | 1.416 | 47.995 | 181.313 | 4.215 | 0 | 20.556 | 3.74 | 267.472 | 80.242 | 347.713 | 8.585 | 0.208 | 6.851 | 11.907 | 24.196 | 7.068 | 58.8 |
| 20.471 | 9.967 | 10.505 | 14.258 | 44.2 | 181.406 | 25.332 | 0 | 20.112 | 11.364 | 337.615 | 101.285 | 438.9 | 6.284 | 0.996 | 0.067 | 30.82 | 34.575 | 4.304 | 77.0 |
| 16.971 | 7.513 | 4.68 | 0.623 | 64.032 | 159.544 | 3.737 | 0 | 23.397 | 6.753 | 287.25 | 86.175 | 373.425 | 4.121 | 2.161 | 0.194 | 28.164 | 14.054 | 7.902 | 56.5 |

« ‹ 1 / 400 › »

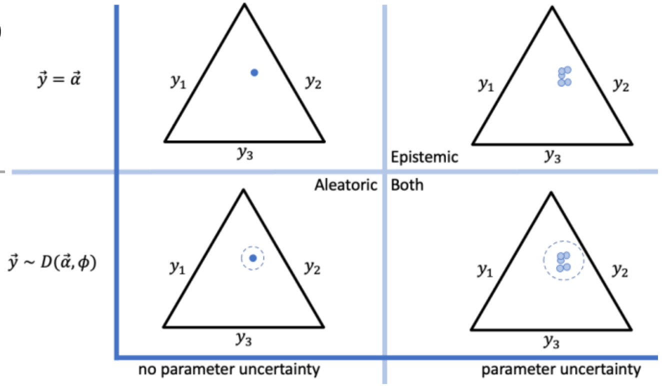
# Using Aleatoric vs. Epistemic mode?

## Neither

- Traditional ROT validation
- Want to show example allocation for a given mission class and cost cap

## Epistemic Only

- Estimate of WBS means based on data
- Model what future cost models may predict




## Aleatoric Only

- Assuming we have the right mean, how much do projects vary around that mean?
- Probably not very useful in practice

## Aleatoric & Epistemic

- Quantify as much uncertainty as possible to assess probability of achieving a certain cost
- Correlation between WBS elements

# Conclusion & Future Work

## Conclusion

- We can and should provide more robust cost estimates with an understanding of uncertainty
- This is important at all stages of project formulation, including very early project formulation
- Back-of-the-envelope calculations which use average ratios between WBS elements may provide engineers a rough sense of whether a concept is feasible or not, but comes with no insight into the risk of overrunning cost
- Here we have shown we can ingest high level cost allocation data with missing or incomplete information and generate probabilistic predictions of allocations

## Future Work

- Future work should continue to improve upon the selection of an informative prior
- Other applications of the Dirichlet model: mass budgets, workforce allocation models for center management

Jet Propulsion Laboratory
California Institute of Technology

jpl.nasa.gov

Questions? Contact Melissa Hooke (Melissa.A.Hooke@jpl.nasa.gov)