Getting More Out of Data Competition Results with Pareto Fronts

Christine Anderson-Cook, Los Alamos National Laboratory

candcook@lanl.gov

Lu Lu, University of South Florida

Sarah Burke, The Perduco Group

75 Los Alamos NATIONAL LABORATORY EST. 1943

October 2019

NNSA National Nuclear Security Administration

Managed by Triad National Security, LLC for the U.S. Department of Energy's NNSA

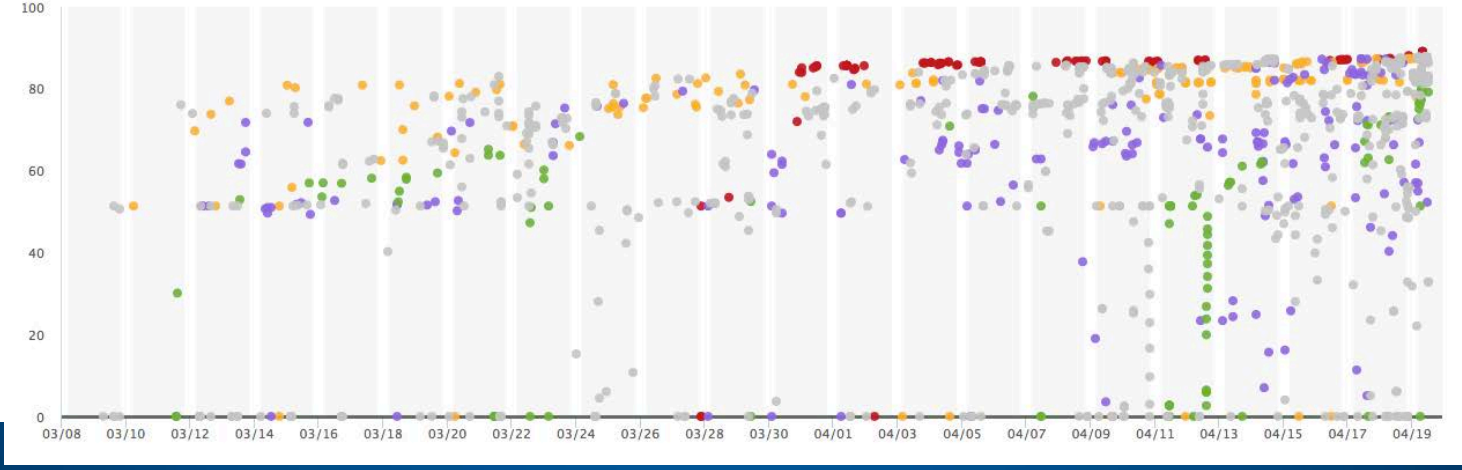
LA-UR-

Outline

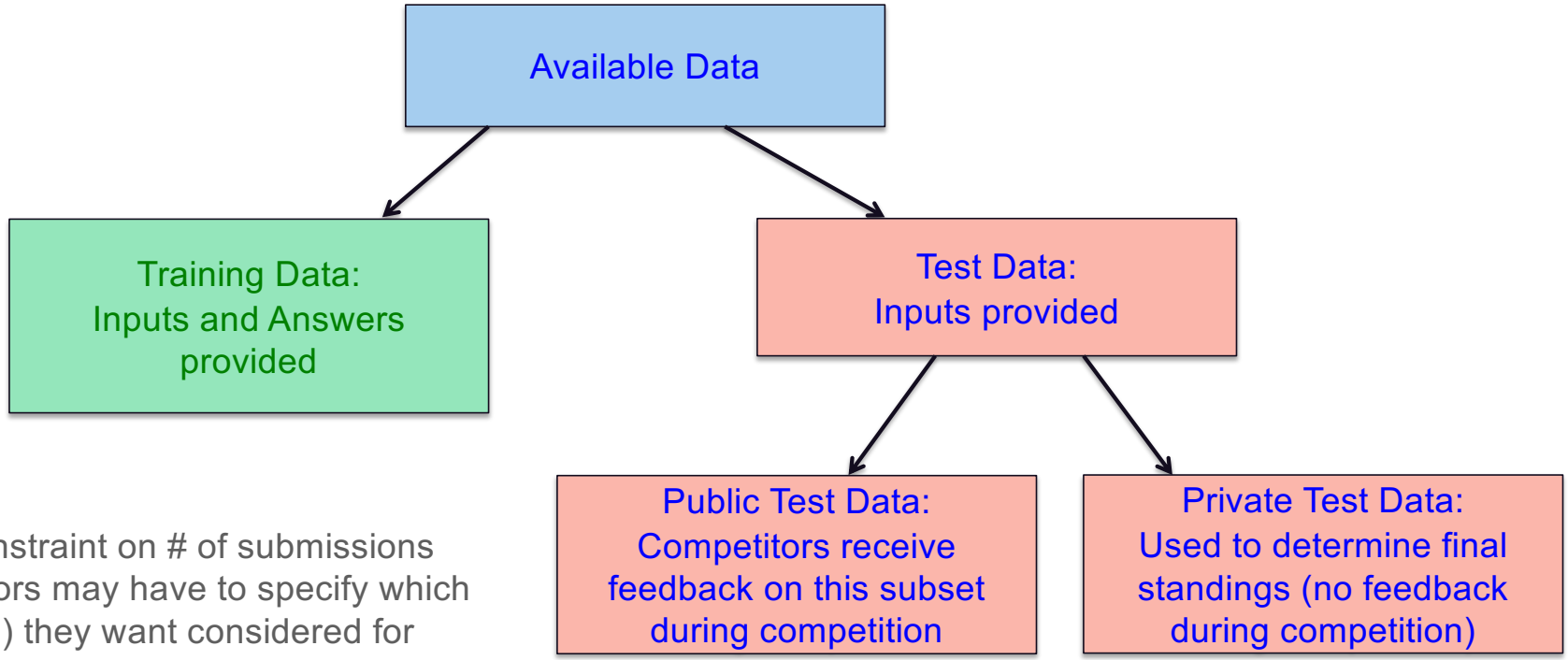
- **Data Competitions – The Basics**
- **Our Urban Radiation Search Competitions**
  - https://www.topcoder.com/lp/detect-radiation
- **Tools for Getting More Out of Results**
  - **Pareto Fronts**
  - **Trade-off Plot**
- **Selecting the Right Solution for Different Scenarios**
  - Design of Experiments
  - Optimizing a Multi-Response System
- **A Few Final Thoughts**

Benefits of Competitions

1. **Crowdsourcing** – participants from many broad areas, bringing lots of expertise to generate solutions
2. **Competitive fervor** – a deadline for the finish line and desire to win spur accelerated improvements
3. **Targeted goal and data** – fair comparisons between solutions are straightforward and can (should) be tailored to the specific problem
4. **Highlighting interesting problems** – possible recruiting tool



Modern Data Competitions – the Fundamentals



Other rules:

- Some constraint on # of submissions
- Competitors may have to specify which solution(s) they want considered for final standings
- Sometimes options for teams to combine/collaborate

Competitors don't know which runs are in each of the test data sets

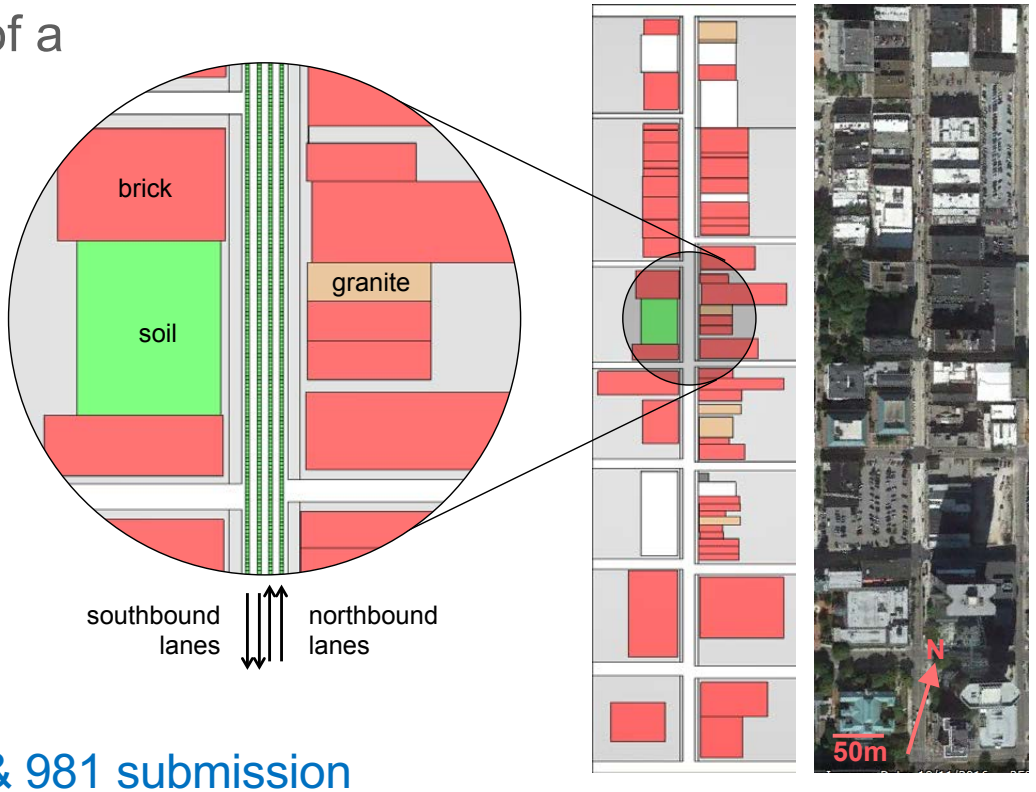
Leaderboard scoring: pre-specified scalar that combines all desired aspects of solution

Outline

- **Data Competitions – The Basics**
- **Our Urban Radiation Search Competitions**
  - https://www.topcoder.com/lp/detect-radiation
- **Tools for Getting More Out of Results**
  - **Pareto Fronts**
  - **Trade-off Plot**
- **Selecting the Right Solution for Different Scenarios**
  - Design of Experiments
  - Optimizing a Multi-Response System
- **A Few Final Thoughts**

The Urban Radiological Search Competitions

- ORNL designed multiple versions of a 0.5 mile street model with characteristics similar to a street in Knoxville.
- 56 buildings
  - –48 brick, 7 granite, 1 concrete
- Side streets, sidewalks, 6 parking areas.
- Ability to vary levels of K, U, Th.



2 competitions:

Government (2018) – 16 participants & 981 submission

TopCoder (2019) – 71 participants & 1614 submissions

Competition format: *datacompetitions.lbl.gov*

Competitors are provided with

- A training set of **list mode data** for ~10k runs.
- A test set of ~16k runs: 43% **public**, 57% **private**.
- Energy spectra for each source type.

For each run in the test set, competitors must

- **Detect** whether there is an extraneous source.
- **Identify** the type of source.
- **Locate** when the detector is closest to it.

Time since last photon (μs)	Photon energy (keV)
1020	88.72
91	179.65
9453	446.41
820	942.51
4295	182.96
1313	262.20
2858	354.80
2687	1295.18
1392	1459.02
…	…

1. HEU: Highly enriched uranium
2. WGPu: Weapons grade plutonium
3. ^{131}I: Iodine, a medical isotope
4. ^{60}Co: Cobalt, an industrial isotope
5. ^{99m}Tc: Technetium, a medical isotope
6. A combination of HEU and ^{99m}Tc

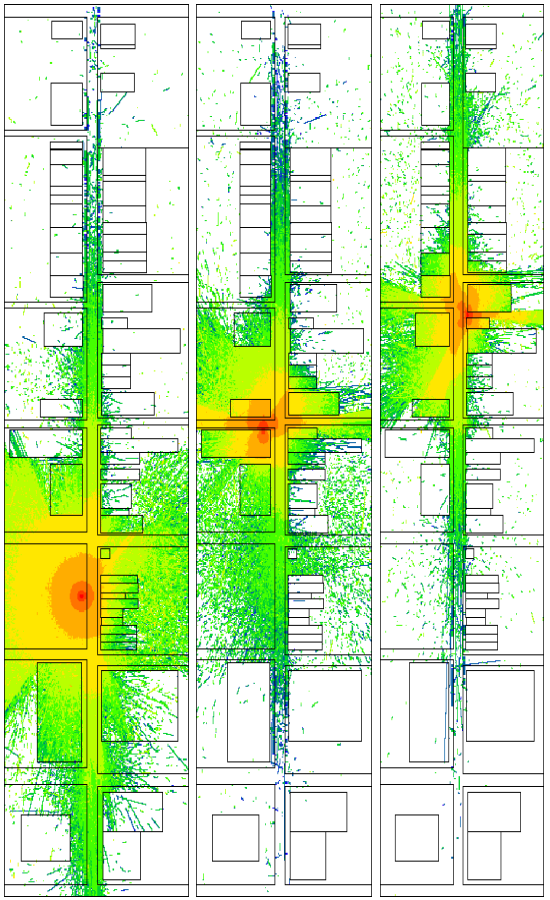
Movie

Factors considered in designing the competition

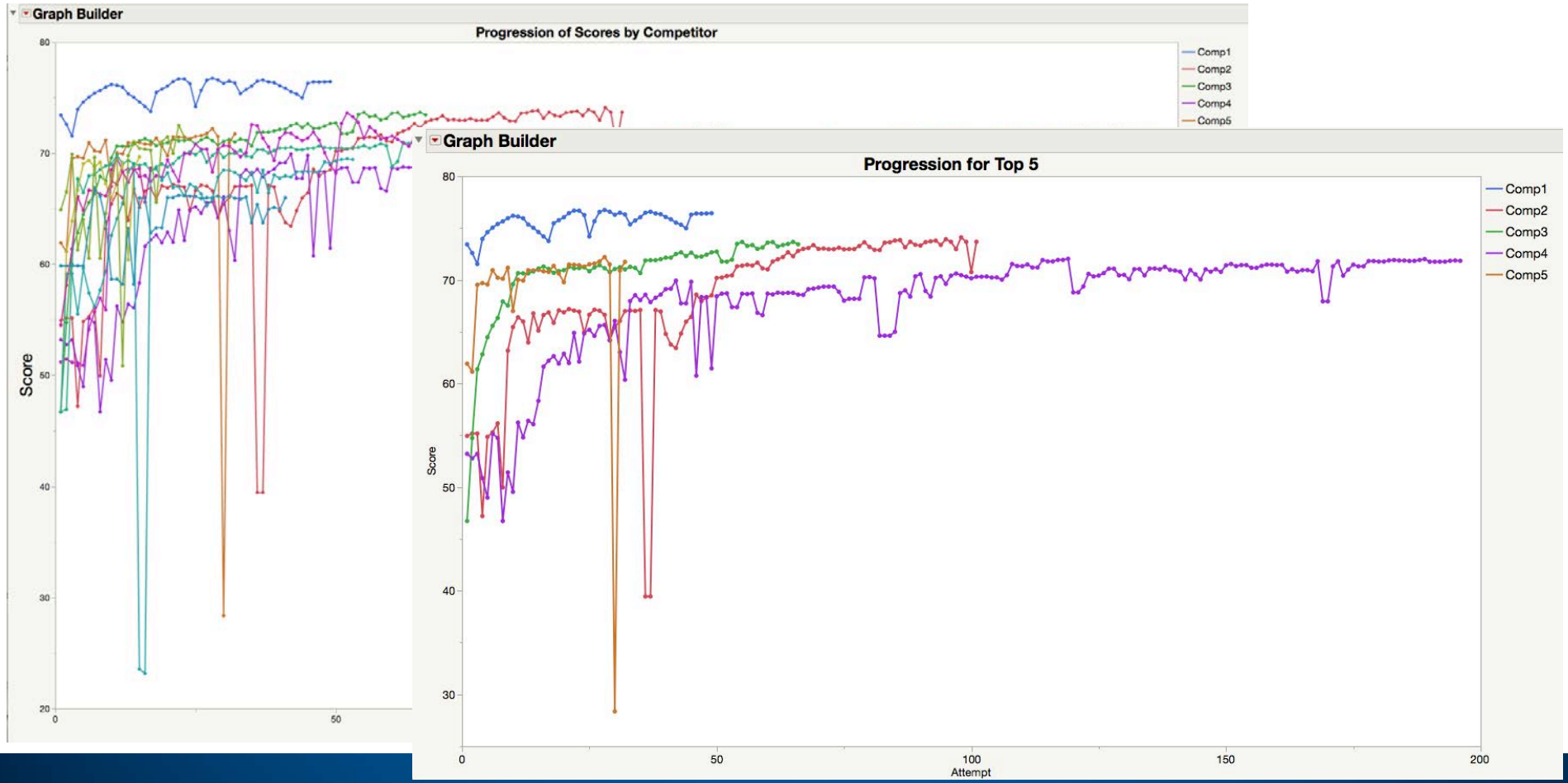
A. **Background:** 8 background models, each with 82 parameters

B. Source
- **6 types**
- **2 shielding settings** (On/Off)
- **strength**
- **15 source locations**

C. Other
- **speed** of detector
- **proximity to source** (lane for detector)
- length of path, starting location



Results


Graph Builder
Progression of Scores by Competitor
Comp1
Comp2
Comp3
Comp4
Comp5
Score
80
70
60
50
40
30
20
0
50
Graph Builder
Progression for Top 5
Comp1
Comp2
Comp3
Comp4
Comp5
Score
80
70
60
50
40
30
0
50
100
150
200
Attempt

The Results

- The Leaderboard

Member	TC_Pub	TC_Priv
pfr	90.26719	76.4289
p_kuzmin	86.67468	73.6693
gardn999	87.45759	73.42756
rayvanve	87.88758	71.83694
cyril.v	86.82957	71.7367
wleite	87.41078	71.38199
smg478	85.62803	71.33488
cannab	84.57155	69.66864
pasda	85.69037	69.42756
ZFTurbo	78.64456	68.63696

Top 10 received prize money

But there is so much more information there!

19 criteria:

When there is no source:

- Correctly identify those situations

When there is a source:

- **Detect**
- **Identify**
- **Locate**

} x 6 sources

Score is a single particular weighted average of these 19 criteria

For different scenarios, we may want to pick algorithms / solution with different strength

Data to Work with

Raw data (1479 valid submissions)

	A	B	C	D	Q	R	S	T	U	V
1	Sub Name	S1 #Iden	S1 #Detect	S1 #Missed	S6 #Iden	S6 #Detect	S6 #Missed	NoS Correct	NoS Wrong	Score
2	C13-1	0	0	1062	0	0	1100	2801	0	46.726166
3	C13-2	178	312	572	208	554	338	2729	72	47.072412
4	C13-3	175	299	588	203	543	354	2741	60	47.235201
5	C13-4									204389
6	C13-5									377422
7	C13-6									775294
8	C13-7									476067
9	C13-8									341563
10	C13-9									341563
11	C13-10									414647
12	C13-11									256971
13	C13-12									007021

Processed summary

	A	B	G	H	M	N	O	P	Q
1	Sub Name	S1 prop D	S6 prop D	S1 prop I	S6 prop I	NoS prop Co	Ave D	Ave I	Score
2	C13-1	0	0	0	0	1	0	0	46.726166
3	C13-2	0.4613936	0.6927273	0.1676083	0.1890909	0.9742949	0.416671	0.1286901	47.072412
4	C13-3	0.4463277	0.6781818	0.1647834	0.1845455	0.9785791	0.4026808	0.1240386	47.235201
5	C13-4	0.1177024	0.2681818	0.0433145	0.1590909	0.9714388	0.1892071	0.0969336	46.204389
6	C13-5	0.413371	0.6745455	0.1214689	0.3072727	0.9796501	0.4138998	0.1657279	47.377422
7	C13-6	0.413371	0.6745455	0.1214689	0.3072727	0.9796501	0.4138998	0.1657279	45.775294
8	C13-7	0.413371	0.6745455	0.1214689	0.3072727	0.9796501	0.4138998	0.1657279	47.476067
9	C13-8	0.220339	0.4518182	0.0762712	0.2063636	0.9703677	0.2833847	0.1513386	45.341563
10	C13-9	0.220339	0.4518182	0.0762712	0.2063636	0.9703677	0.2833847	0.1513386	45.341563
11	C13-10	0.220339	0.4518182	0.0762712	0.2063636	0.9703677	0.2833847	0.1513386	52.414647
12	C13-11	0.220339	0.4518182	0.0762712	0.2063636	0.9703677	0.2833847	0.1513386	54.256971
13	C13-12	0.220339	0.4518182	0.0762712	0.2063636	0.9703677	0.2833847	0.1513386	55.007021

Possible Criteria to Consider for Choosing Between Algorithms

- Correct Detection Proportion (Source 1-6)
- Correct Identification Proportion (Source 1-6)
- Correct Proportion for No Source or False Positive proportion
- Average Correct Detection Proportion (averaged over 6 sources)
- Average Correct Identification Proportion (averaged over 6 sources)
- Score from Leaderboard

minimize

maximize

	A	B	G	H	I	J	K	L	M
1	Sub Name	S1 prop D	S6 prop D	S1 prop I	S6 prop I	NoS prop Co	Ave D	Ave I	Score

x 6

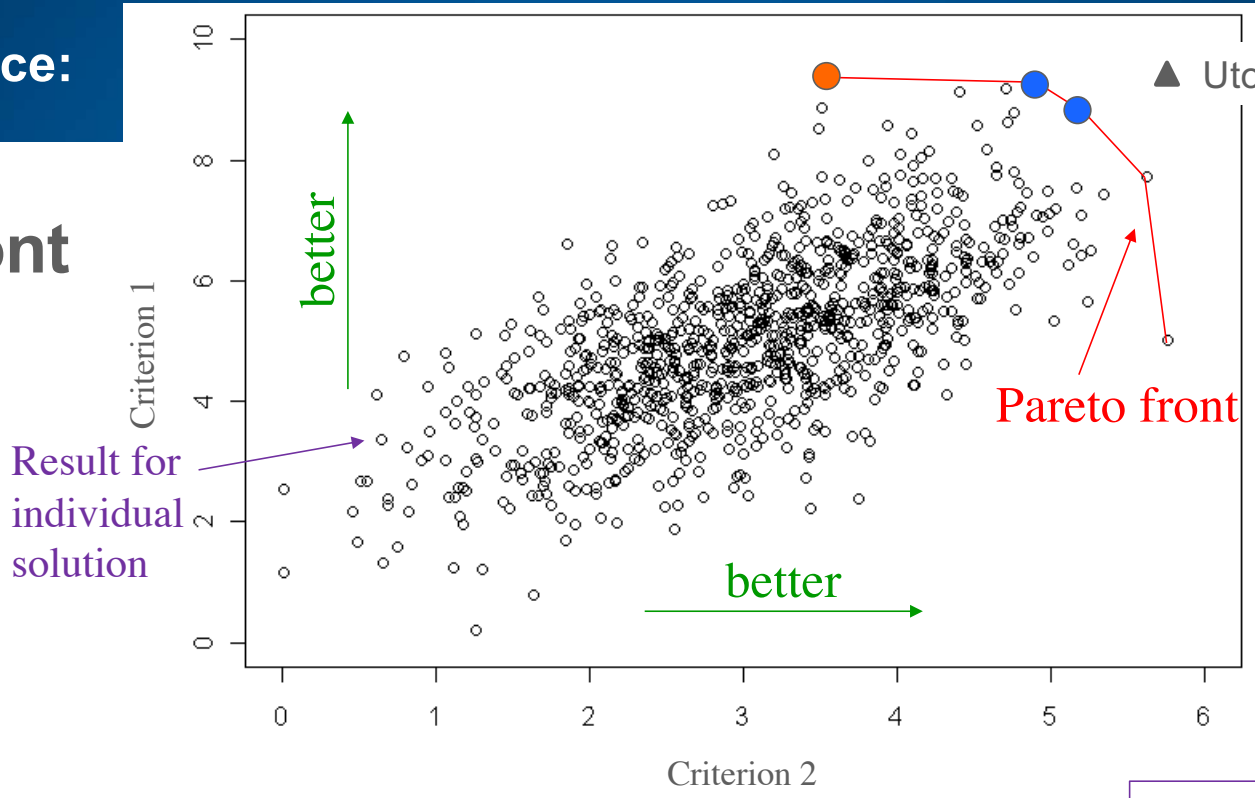
x 6

Outline

- **Data Competitions – The Basics**
- **Our Urban Radiation Search Competitions**
  - https://www.topcoder.com/lp/detect-radiation
- **Tools for Getting More Out of Results**
  - **Pareto Fronts**
  - **Trade-off Plot**
- **Selecting the Right Solution for Different Scenarios**
  - Design of Experiments
  - Optimizing a Multi-Response System
- **A Few Final Thoughts**

Help for Reduce:

Pareto Front



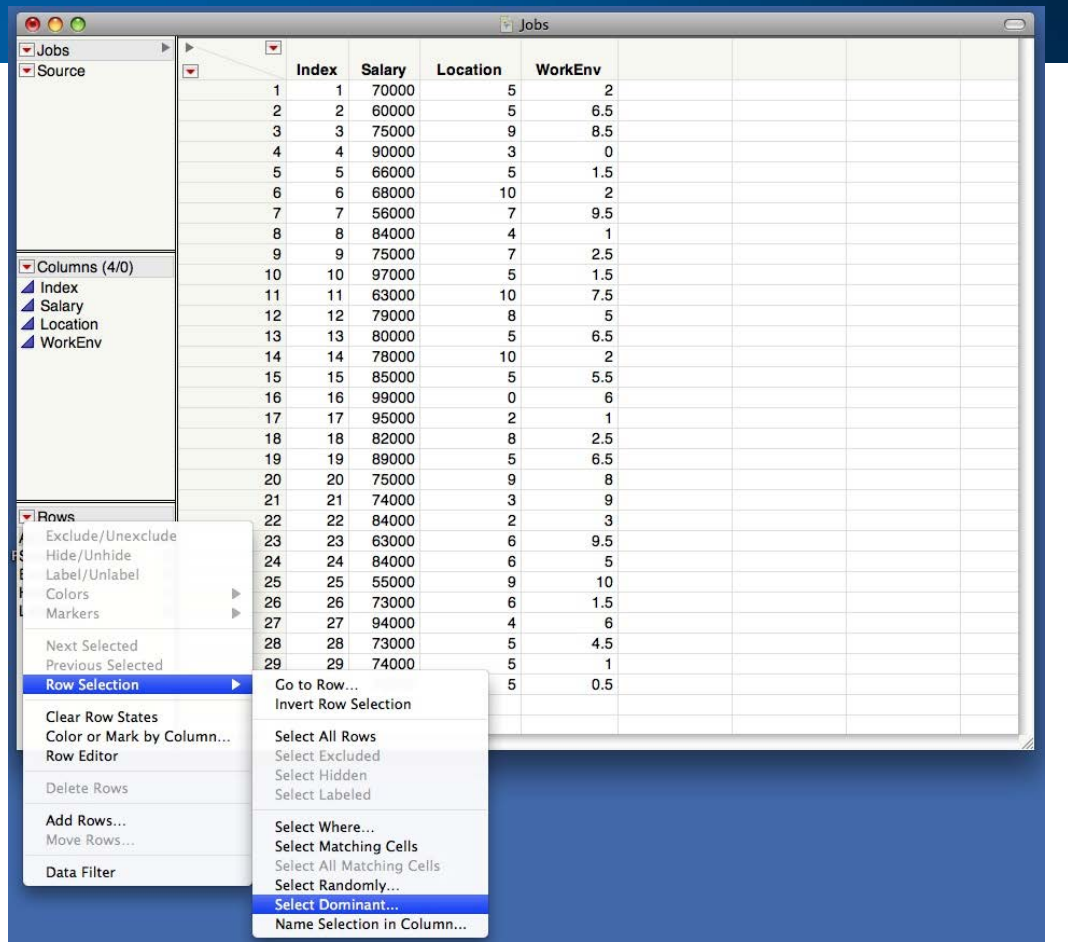
- Choices not on Pareto front are not rational.
- The Pareto front represents the objective set of best solutions for any combination of weighting, scaling and DF choices (allowing non-contenders to be discarded)
- Collection of choices bounds ranges for criterion – gives basis for comparison of alternatives

Pareto front = set of all non-dominated solutions
A solution *i dominates* another solution *j* if all of the criterion value of *i* are at least as good as *j*, and at least one is better

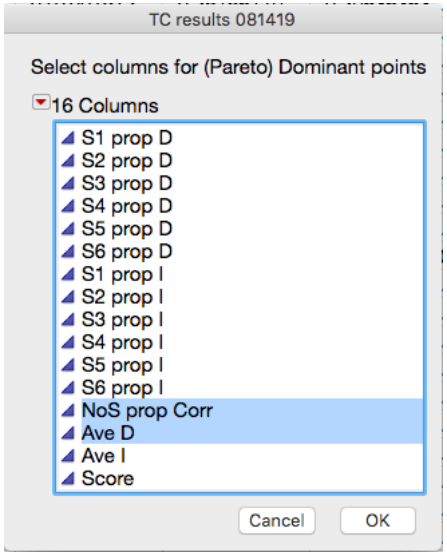
Constructing a Pareto Front

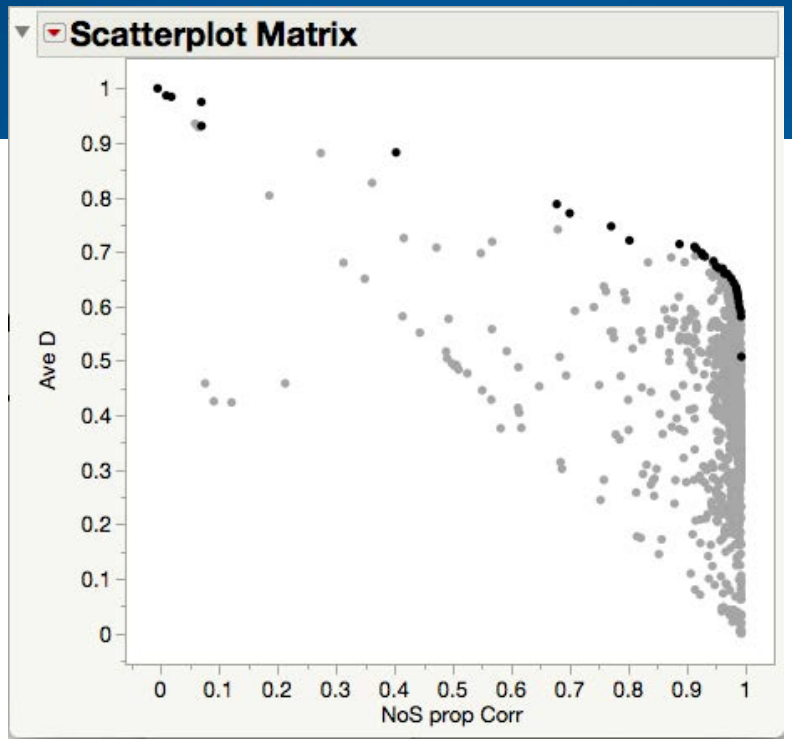
From an Enumerated List

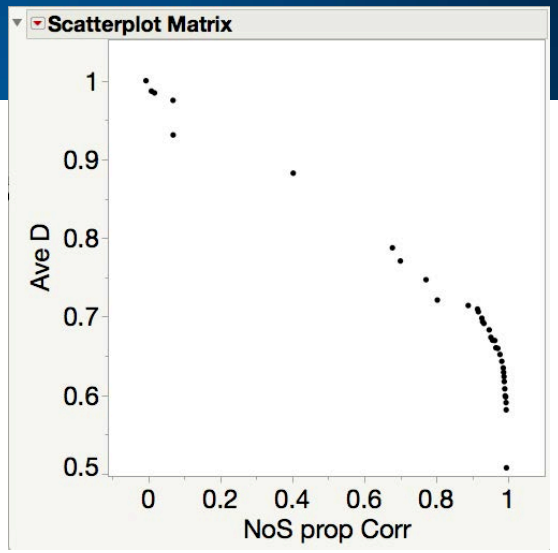
- Straightforward for optimizing (minimize, maximize, hit target) for any number of criteria
- No practical computational constraints

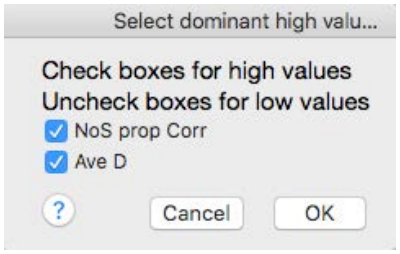


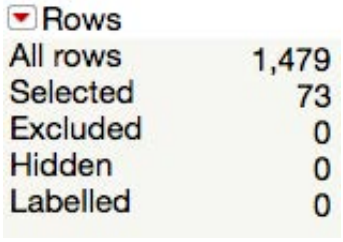
Sample Solution

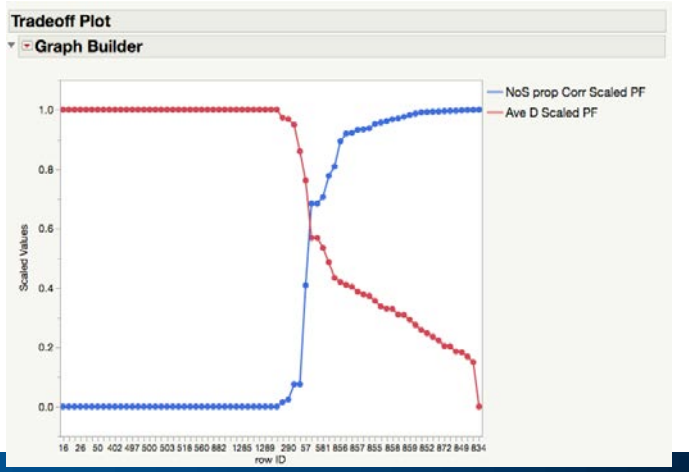












More Criteria … more challenges with Visualization

Tradeoff Plot Results

Variables	Best Value	Worst Value (PF)	Worst Value (All)
S1 prop D	1	0.39077	0
S1 prop I	1	0	0
FalsePositive	0	1	1

Trade Off Plot Application

Select columns for Pareto dominant points.

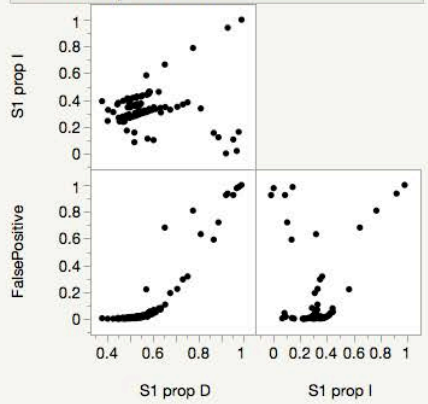
S6 prop D
S1 prop I
S2 prop I
S3 prop I
S4 prop I
S5 prop I
S6 prop I
NoS prop Corr
FalsePositive
Ave D
Ave I
Score

Maximize: S1 prop D, S1 prop I, *optional numeric*

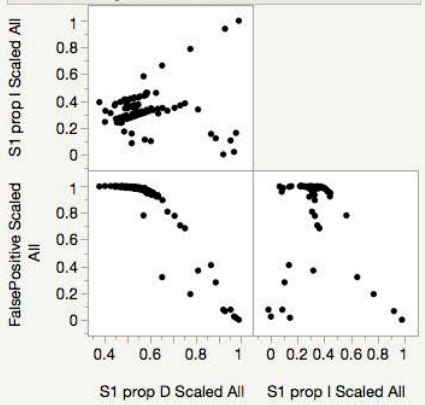
Minimize: FalsePositive, *optional numeric*

Scaling:
Full Dataset
Pareto Front
Variable to sort on:
FalsePositive

Scatterplot Matrix (Original Values)

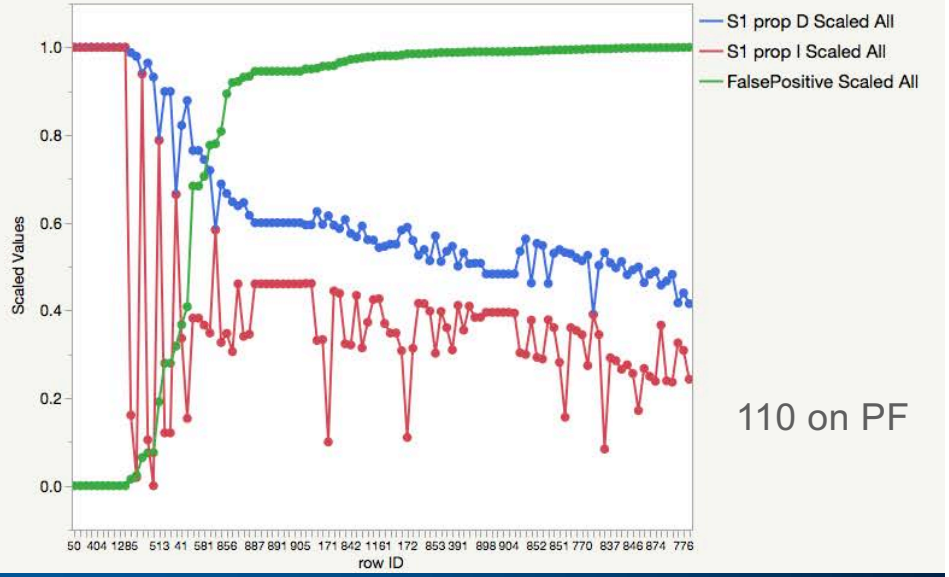


Scatterplot Matrix



Tradeoff Plot

Graph Builder



110 on PF

Even More Criteria ….

Tradeoff Plot Results

Variables	Best Value	Worst Value (PF)	Worst Value (All)
S1 prop D	1	0.37288	0
S1 prop I	1	0	0
NoS prop Corr	1	0	0
Score	76.744	3.00275	0

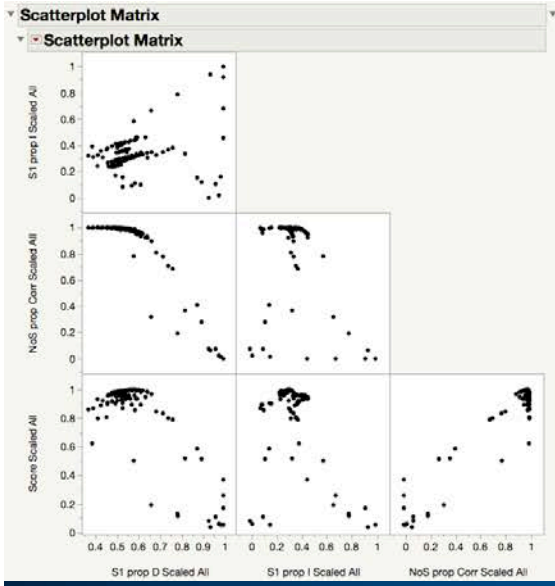
Trade Off Plot Application

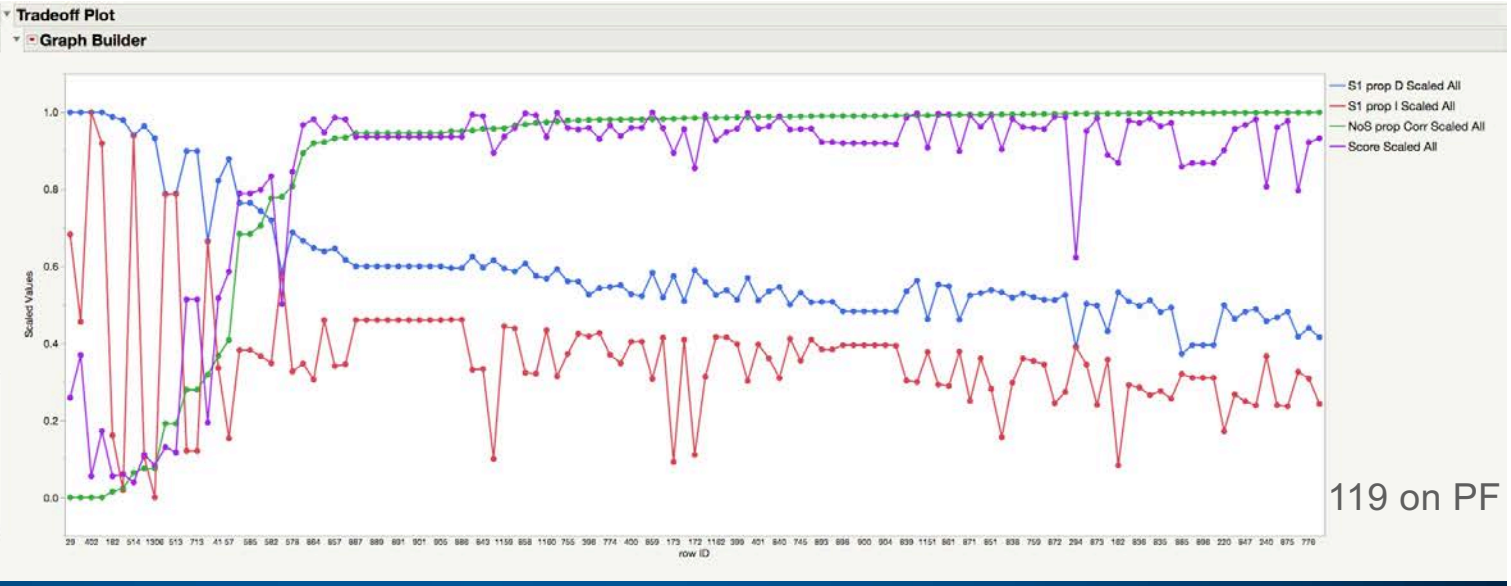
Select columns for Pareto dominant points.

S5 prop D
S6 prop D
S1 prop I
S2 prop I
S3 prop I
S4 prop I
S5 prop I
S6 prop I
NoS prop Corr
Ave D
Ave I
Score

Maximize: S1 prop D, S1 prop I, NoS prop Corr, Score, *optional numeric*

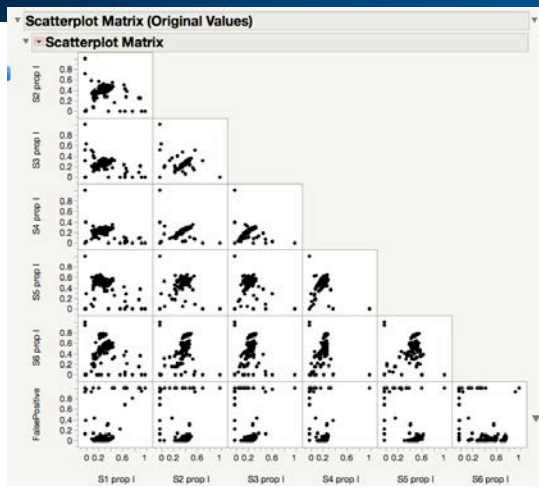
Minimize: *optional numeric*

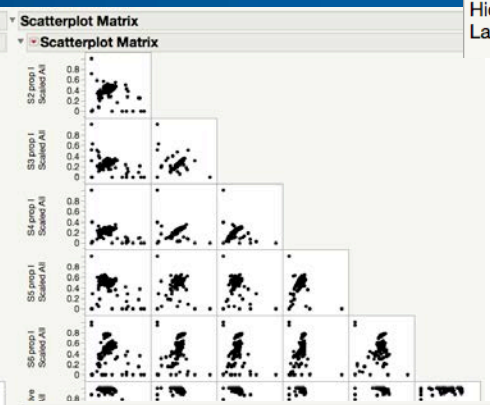


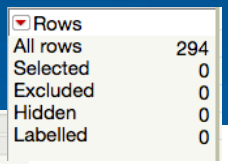


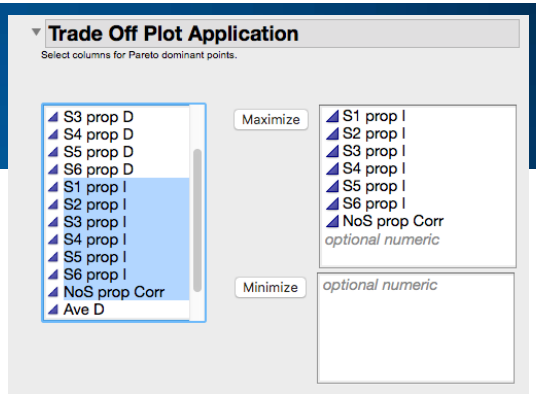
119 on PF

All About Identification









294 on PF

Conclusions

- Data competitions can be a cost-effective way of getting access to diverse solutions that solve challenging problems in different ways.
  - Presenting the right data to the competitors and asking the right questions are key to getting useable results
- Pareto Fronts are a powerful tool to reduce the number of solutions that need to be considered, by eliminating only irrational choices
- Trade-off plots provide an dimension-robust approach to see how prioritizing different criteria impacts the performance of other criteria
- Since many choices of a final decision involve getting consensus from a team, visualizing the options can be helpful.
- The Add-in for creating Trade-off plots is available

References

1. Anderson-Cook, C.M., Lu, L., Myers, K., Quinlan, K., Pawley, N. (2019) “Improved Learning from Data Competitions through Strategic Generation of Informative Data Sets” **Quality Engineering** 31(4) 564-580.
2. Anderson-Cook, C.M., Myers, K., Lu, L., Fugate, M.L., Quinlan, K., Pawley, N. (2019) “Data Competitions: Getting More from a Strategic Design and Post-Competition Analysis” **Statistical Analysis and Data Mining** 12 271-289**.**
3. Anderson-Cook, C.M., Lu, L. (2015) “Much-Needed Structure: A New 5-Step Decision-Making Process Helps You Evaluate, Balance Competing Objectives” **Quality Progress** 48(10) 42-50.
4. Lu, L., Anderson-Cook, C.M., Robinson, T.J. (2011) “Optimization of Designed Experiments Based on Multiple Criteria Utilizing a Pareto Frontier” **Technometrics** 53 353-365.
5. Lu, L., Chapman, J.L., Anderson-Cook, C.M. (2017) “Multiple Response Optimization for Higher Dimensions in Factors and Responses” **Quality and Reliability Engineering International** 33 727-744.

Abstract

Data competitions have attracted considerable attention among the world's community of data and analytics scientists, as well as discipline-specific subject matter experts. Their broad involvement provides a model of crowdsourcing for business and government to solve tough high-impact problems in a cost-effective way. Typically winners are determined through a leaderboard formula that needs to be static throughout the competition, with fixed rewards and penalties for patterns of correct and incorrect responses for different aspects of the solution. However, for different uses of the solution, these aspects might be more or less important. By using the existing capability for constructing flexible high-dimensional Pareto fronts in JMP, it is possible to explore and identify various solutions with their strengths and weaknesses. Pareto fronts allow the user to identify all of the objectively superior solutions across all possible weightings of the different elements of the solution, and discard non-competitive solutions. The approach to using multiple Pareto fronts to highlight different "best" solutions will be demonstrated through a recently completed data competition focused on detecting, identifying and locating radioactive sources in an urban environment (https://www.topcoder.com/lp/detect-radiation).