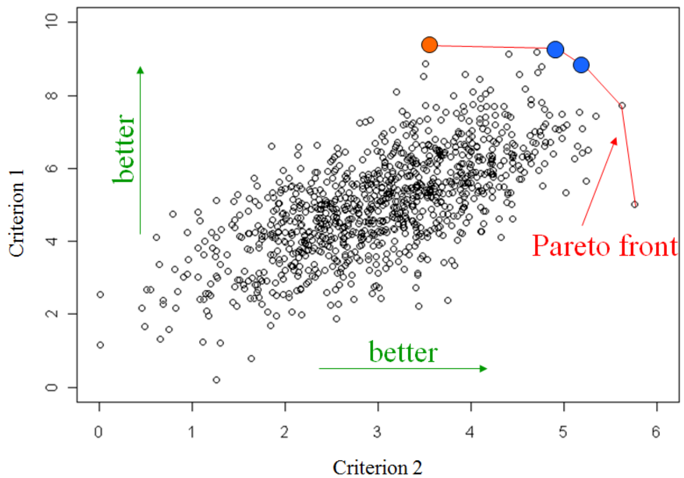
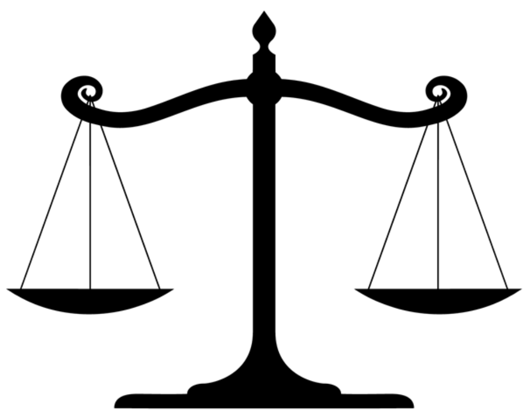
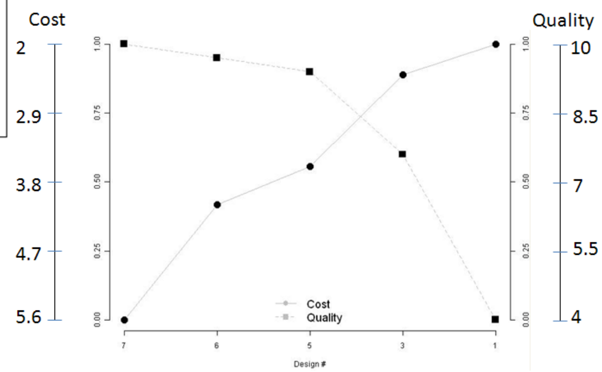
Optimizing in a Resource-Constrained World: Statisticians' Roles in Decision-Making

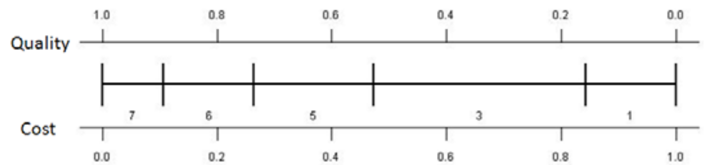


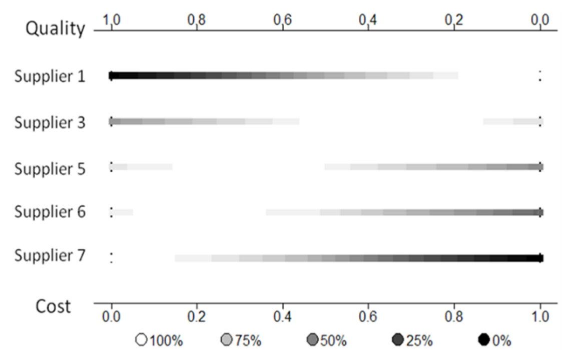
Christine Anderson-Cook
Los Alamos National Laboratory

June 2016









Outline

1. **Motivation**
   - Multiple criteria are everywhere – life is full of trade-offs
   - Statisticians' roles in decision-making
   - Introduction of examples
2. **Define-Measure-Reduce-Combine-Select (DMRCS)**
3. **Pareto Front basics – objectively removing non-contenders**
4. **Deciding among contenders – challenges, strategies, facilitating consensus**
5. **Return to examples to illustrate the process**
6. **Conclusions**

Why Statisticians in Decision-Making?

"Making good decisions relies on forecasting the future, but forecasting is imprecise, …because it forces us to confront how much we don't know." C. Duhigg*

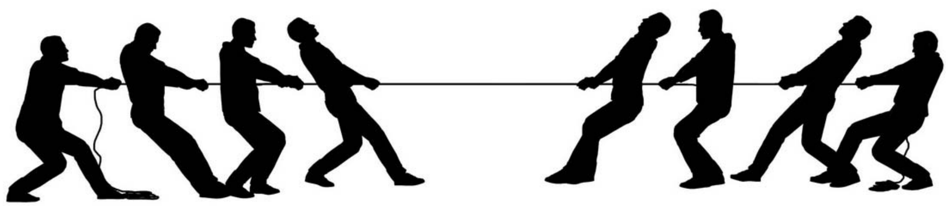
Decisions are all about:

- **Thinking probabilistically about future alternatives (what could happen?, how likely is it?, what do I need now to be prepared for each?) "Envision tomorrow as an array of potential outcomes, all of which have different odds of coming true."***
- **Balancing trade-offs in the presence of uncertainty**
- **Examples: Job hunting, buying a house, choosing a spouse**

* "Smarter Faster Better: The Secrets of Being Productive in Life and Business" (2016)

Christine's Axioms of Decision-Making in a Complex World

1. **There are several aspects to the decision which need to be simultaneously considered and balanced.**
2. **Subject matter expertise is an essential element that needs to be leveraged to make an informed and justifiable decision.**
3. **There are multiple stakeholders who have different priorities based on their position, perspective and experience.**



Potential Pitfalls in Decision-Making

- **Leader dominates process and team members self-suppress alternate ideas**
- **Built in biases**
  - Availability bias – disproportionate influence on readily available information
  - Representative bias – thinking a previous decision (where results of the decision are known) is a close match for current decision
  - Confirmation bias – looking for explanations to match current assessment
  - Planning fallacy – overly optimistic about own role, overly pessimistic about others' roles
- **Group think**
  - Informational cascades (first opinion holds disproportionate influence)
  - Polarization (becoming more entrenched in original views when hearing alternatives)
  - Common knowledge effect (tendency to overemphasize information that all know, and discount individual team member knowledge)

Thoughts on the Decision-Making Process from Christine Fox*

- Decision-makers are hungry for tools to make better decisions.
- It is critical the solutions offered match the actual problem being solved.
- Bigger decisions often involve harder choices with more difficult trade-offs.
- More senior decision-makers make more decisions and have less time to spend on each one.
- Statisticians have tools to participate in the process as facilitators and helpers.

*served as acting US Deputy Secretary of Defense and in the Office of the Secretary of Defense (from her talk at JSM 2015)

Roles for Statisticians

- **Decision support:**
  - The statistician asked to perform analyses or collect data that are used as part of a broader collection of information to guide the decision. Technical expertise is being sought, but an opinion or contribution to the decision is not expected.

- **Provide the decision:**
  - The statistician gathers information to ensure they are answering the right question, and then is expected to provide a final answer with justification.

- **Decision analysis:**
  - The statistician provides decision support as well as is heavily involved in the process, but without the exclusive final say.

Expertise: system is comprised of interconnected process, variation exists in all processes, reducing variation and uncertainty are keys to success

Challenges of Decision-Making

- **Articulating how we make decisions is difficult**
  - Often we have received little or no training in this capacity.
  - Much of our training focuses on univariate objectives or subjectivity of priorities for several objectives is ignored or oversimplified
- **Comparing multiple criteria requires deep understanding of the criteria themselves**
  - For a single objective, whatever the optimum value is that is the best we can do
  - For multiple criteria, we need to be able to quantify how important the loss of moving away from the optimum as we trade-off
- **Shortage of available tools for examining trade-offs**
- **Leading group negotitions toward consensus requires leadership skills and confidence**

Shortage of examples in the literature to show how others have done this

Examples (continued)

1. **Stockpile Maintenance Prioritization – managers need to determine which stockpiles are in most urgent need of investment.**

Index	Reliability	Urgency	Consequence
1	7.97	3.5	8
2	9.98	9.25	3
3	4.3	8.25	8.5
4	7.9	3.75	3
5	9.65	4.25	9
6	4.84	2.75	6.5
7	5.73	4.5	5.5
8	8.84	2.5	3.5
9	7.11	9.25	8

- In a given year, there are limited resources for investment (maintenance and surveillance)
- Several competing aspects to decision (time until anticipated problem with aging, availability of alternate munitions, consequences of a problem)
- Characterizing state of each stockpile needs to be standardized with quantitative metrics
- Different stakeholders have polarized priorities based on different perspectives

Examples (continued)

2. Experiment Design Selection – many natural metrics for judging the goodness of a design

1. Result in good fit of the model to the data
2. Provide good model parameter estimates
3. Provide good prediction throughout the design space.
4. Provide an estimate of "pure" experimental error.

Good estimation and prediction for chosen model

5. Give sufficient information to allow for lack of fit test.
6. Provide a check on the homogeneous variance assumption.

Ability to test various aspects of the model

7. Be insensitive (robust) to the presence of outliers in the data.
8. Be robust to to errors in the control of design levels.

Protection if things go wrong

9. Allow models of increasing order to be constructed sequentially.
10. Allow for experiments to be done in blocks.

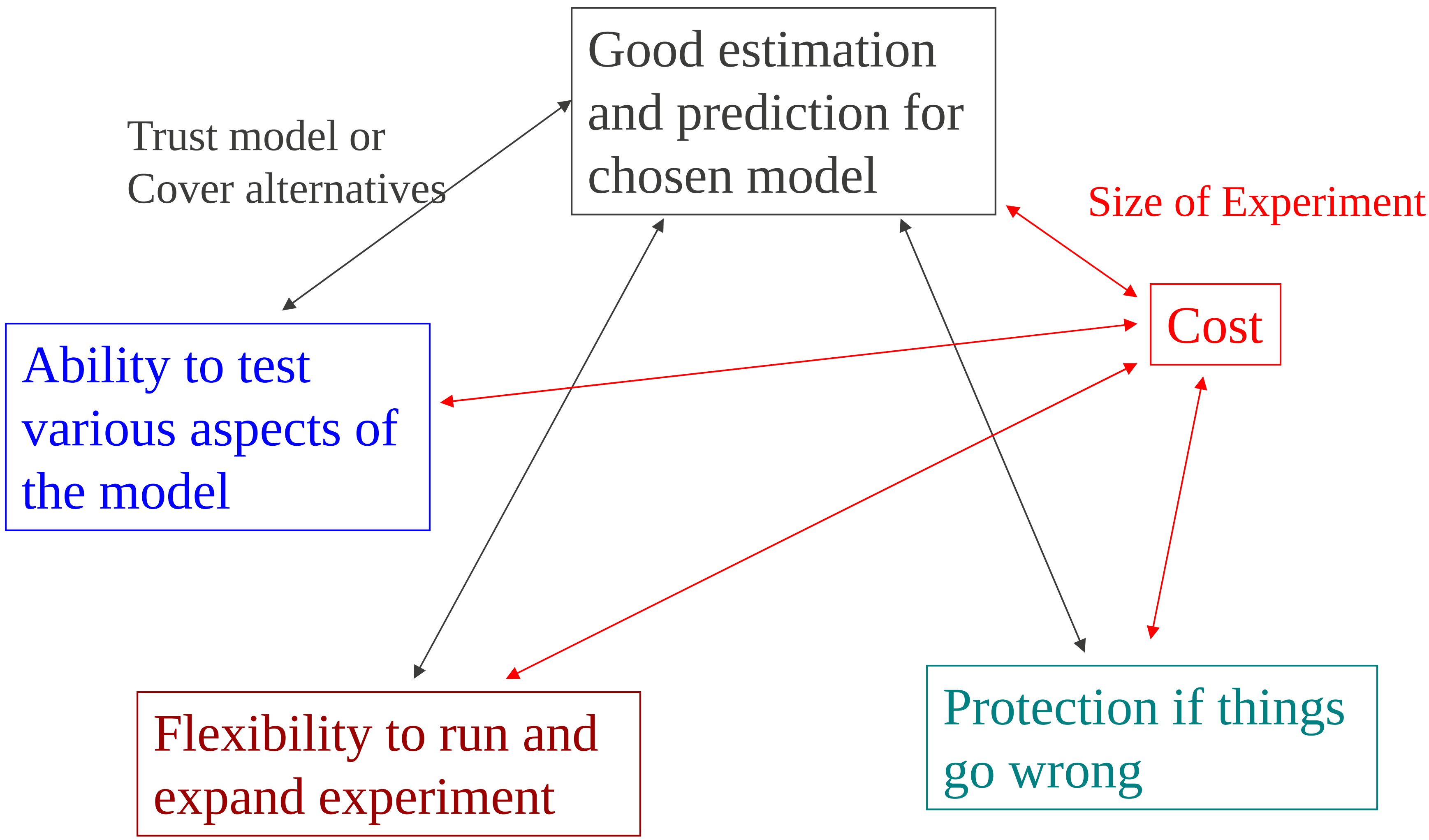
Flexibility to run and expand experiment

Cost

11. Be cost-effective.

Myers, Montgomery, Anderson-Cook RSM (2016) p. 370

Trade-offs


Good estimation and prediction for chosen model
Trust model or
Cover alternatives
Size of Experiment
Cost
Ability to test various aspects of the model
Flexibility to run and expand experiment
Protection if things go wrong

Examples (continued)

3. Multiple Response Optimization – most experiments run collect data on multiple attributes.

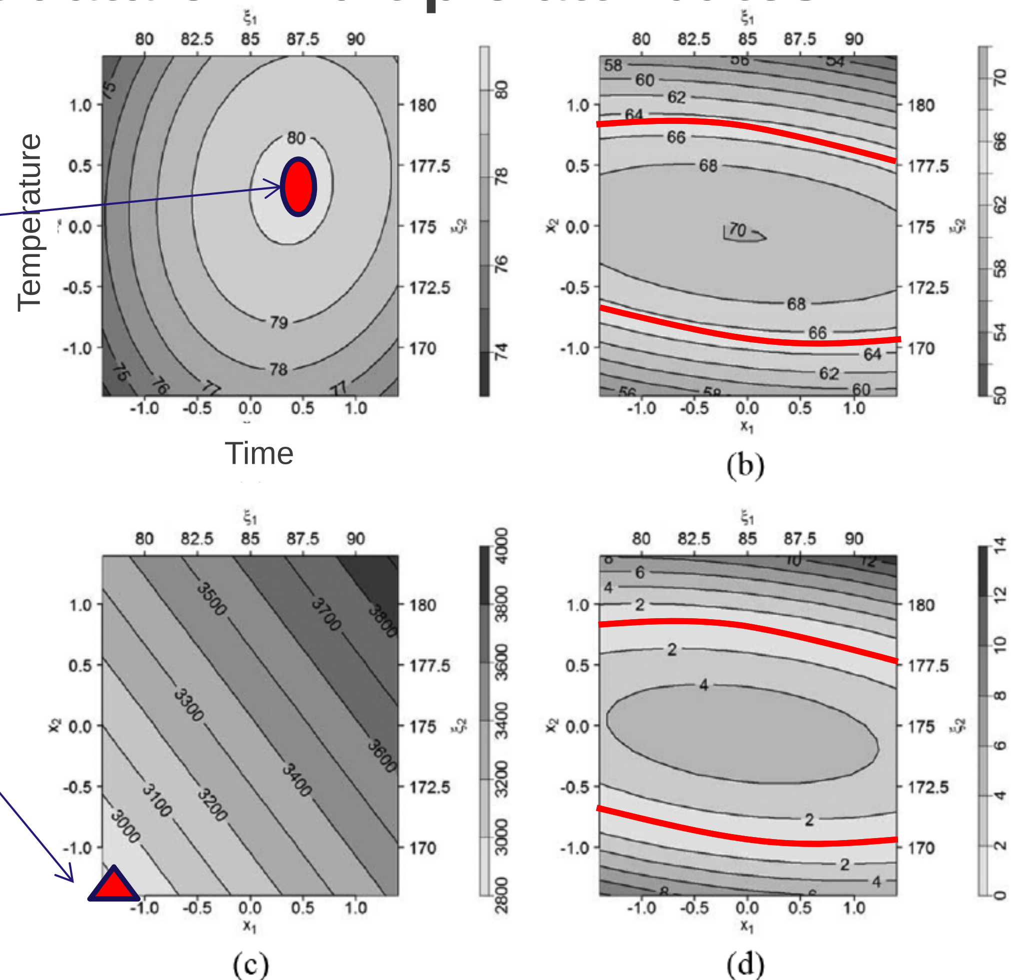
Yield
(maximize)

No universally
best location

Viscosity
(close to 65)

Molecular weight
(minimize)

|Visc – 65|
(minimize)



Response surfaces are estimated – therefore have uncertainty

Outline

1. **Motivation**
   - Multiple criteria are everywhere – life is full of trade-offs
   - Roles in decision-making
   - Introduction of examples
2. **Define-Measure-Reduce-Combine-Select (DMRCS)**
3. **Pareto Front basics – objectively removing non-contenders**
4. **Deciding among contenders – challenges, strategies, facilitating consensus**
5. **Return to examples to illustrate the process**
6. **Conclusions**

Formalized Process for Balancing Multiple Objectives in Decision-Making

- **Buying a new car – what would a formal process for considering alternative look like?**

Define
- What is important for the right car for me?
- What cars are available to choose between?

Measure
- What is the right criterion to measure the important attribute?
- Is high quality data available for the identified criteria?
- Gather relevant information

Reduce
- Are some aspects secondary in importance? Remove or defer?
- Can some cars be eliminated as implausible or non-contenders?

Combine
- How can I evaluate trade-offs between criteria (that are probably on different scales)?
- How much do I value the different criteria?

Select
- Which solution is best given my priorities?
- How does this solution compare to other options?
- Can I defend my choice?

Roles in DMRCS

Subject Matter Experts		Statisticians
Identify key aspects for right choice, clarify solution space	**Define**	Encourage big picture thinking to be inclusive, discourage premature oversimplification
Metrics for characterizing aspects of interest, gather data	**Measure**	Translating qualitative measures into quantitative metrics, data pedigree, uncertainty/variation in measurements
Select primary/secondary criteria, eliminate implausible solutions	**Reduce**	Elimination of non-contenders
Interpret range of choice, prioritize importance of criteria	**Combine**	Measures more comparable, visualization of alternatives
Final selection, description of why choice is best	**Select**	Tools for comparing top solutions, quantifying relative performance

Revisiting Potential Decision-Making Pitfalls

- **Leader dominates process and team members self-suppress alternate ideas**

Process encourages participation and divergent thinking – less leader centric

- **Built in biases**
  - Availability bias, Representative bias, Confirmation bias, Planning fallacy

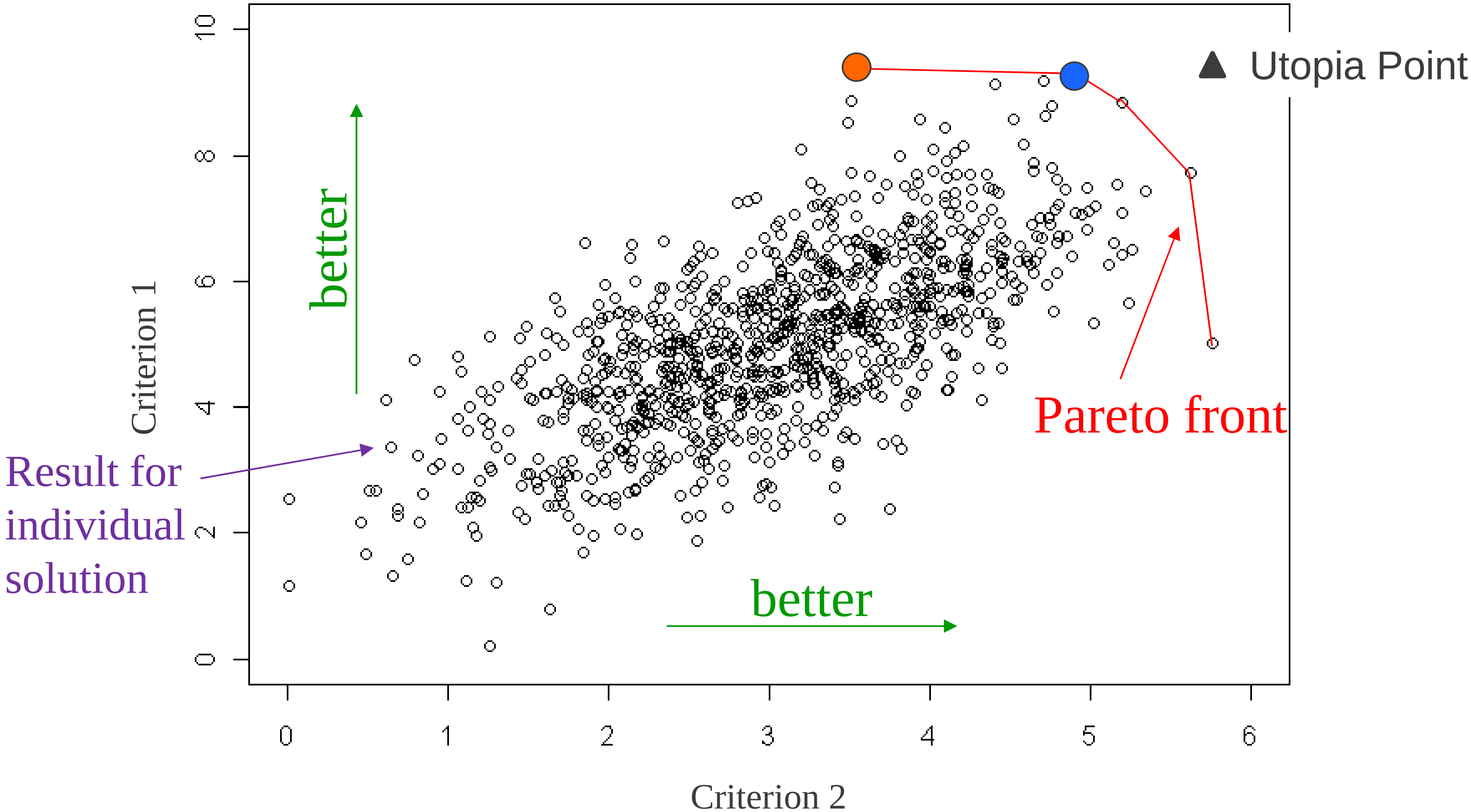
Having data on multiple criteria encouraged better perspective, more realism

- **Group think**
  - Informational cascades — Early parts of process encourage different priorities
  - Polarization — Later stages remind of need for compromise
  - Common knowledge effect — Multiple criteria formally integrated for discussion and evaluation

Outline

1. **Motivation**
   - Multiple criteria are everywhere – life is full of trade-offs
   - Roles in decision-making
   - Introduction of examples
2. **Define-Measure-Reduce-Combine-Select (DMRCS)**
3. **Pareto Front basics – objectively removing non-contenders**
4. **Deciding among contenders – challenges, strategies, facilitating consensus**
5. **Return to examples to illustrate the process**
6. **Conclusions**

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

Constructing a Pareto Front

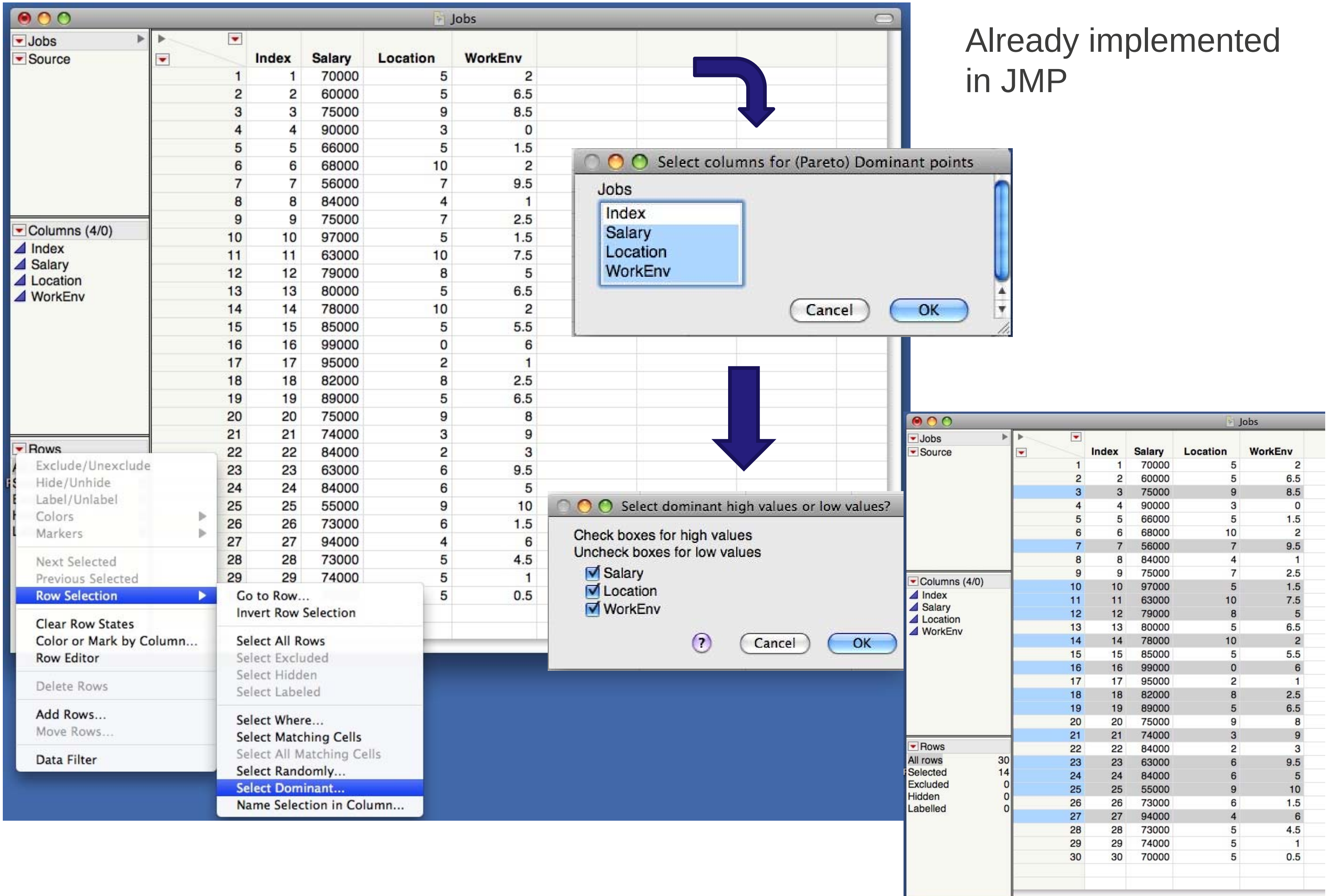
1. From an Enumerated List

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

2. When searching a solution space

- Two distinguishing aspects of search algorithm for constructing PF
  a) How do define when algorithm has found an improvement and should update
  b) Updating PF for each constructed solution

Already implemented in JMP



Outline

1. **Motivation**
   - Multiple criteria are everywhere – life is full of trade-offs
   - Roles in decision-making
   - Introduction of examples
2. **Define-Measure-Reduce-Combine-Select (DMRCS)**
3. **Pareto Front basics – objectively removing non-contenders**
4. **Deciding among contenders – challenges, strategies, facilitating consensus**
5. **Return to examples to illustrate the process**
6. **Conclusions**

The Subjective Phase

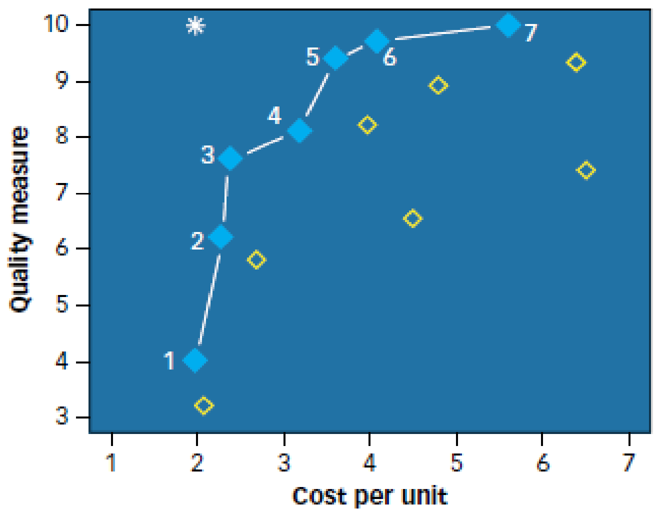
- **Complicating aspects:**
  - Comparing characteristics on different scales – what is the equivalent of one unit of cost on the quality scale?
  - Different stakeholders have different priorities
  - We often have not used a formal process for decision-making in the past (“when I see the right solution, I will know it”)
  - We are often tempted to bring in new criteria to help validate our gut-feeling favorite

Goal: Provide some graphical and quantitative tools that can help keep the discussion focused on defensible solutions and build common understanding

Different tools needed for different situations

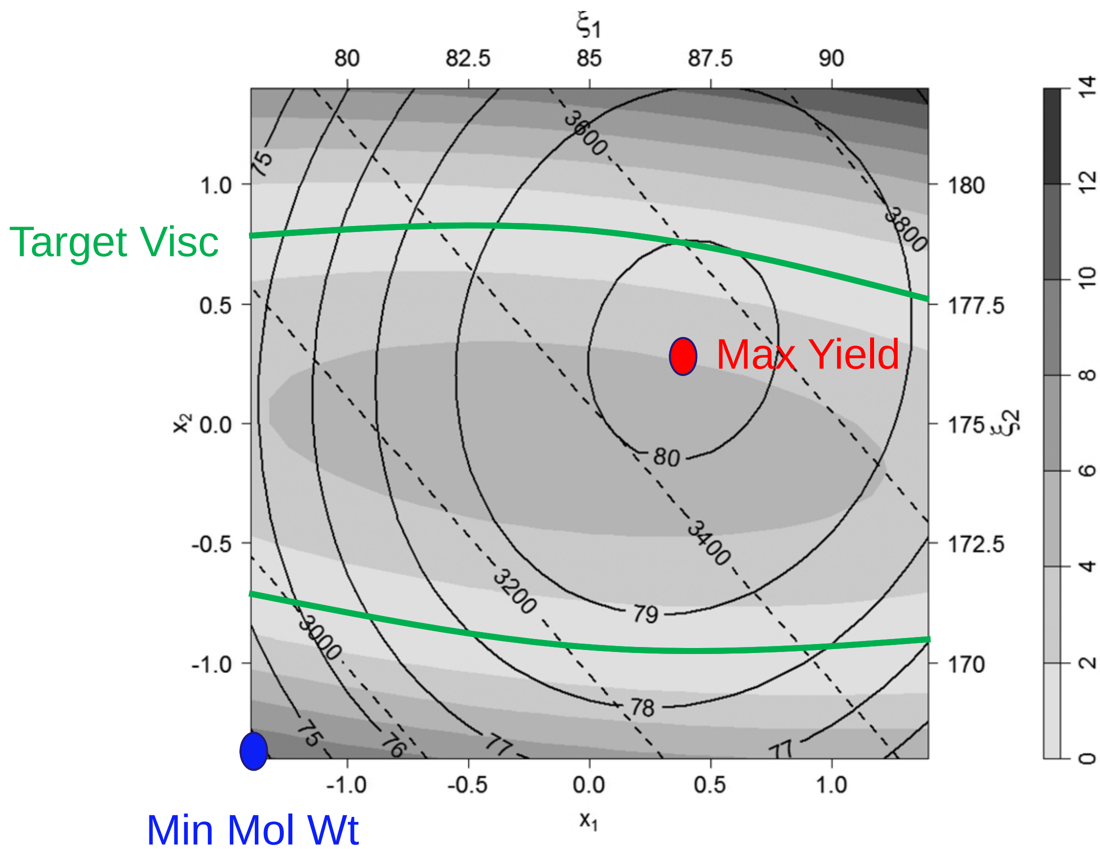
Possible tools: 1. Raw Data

In low dimensions (small # of choices & easy to visualize) – consider the solutions directly



Cost vs Quality

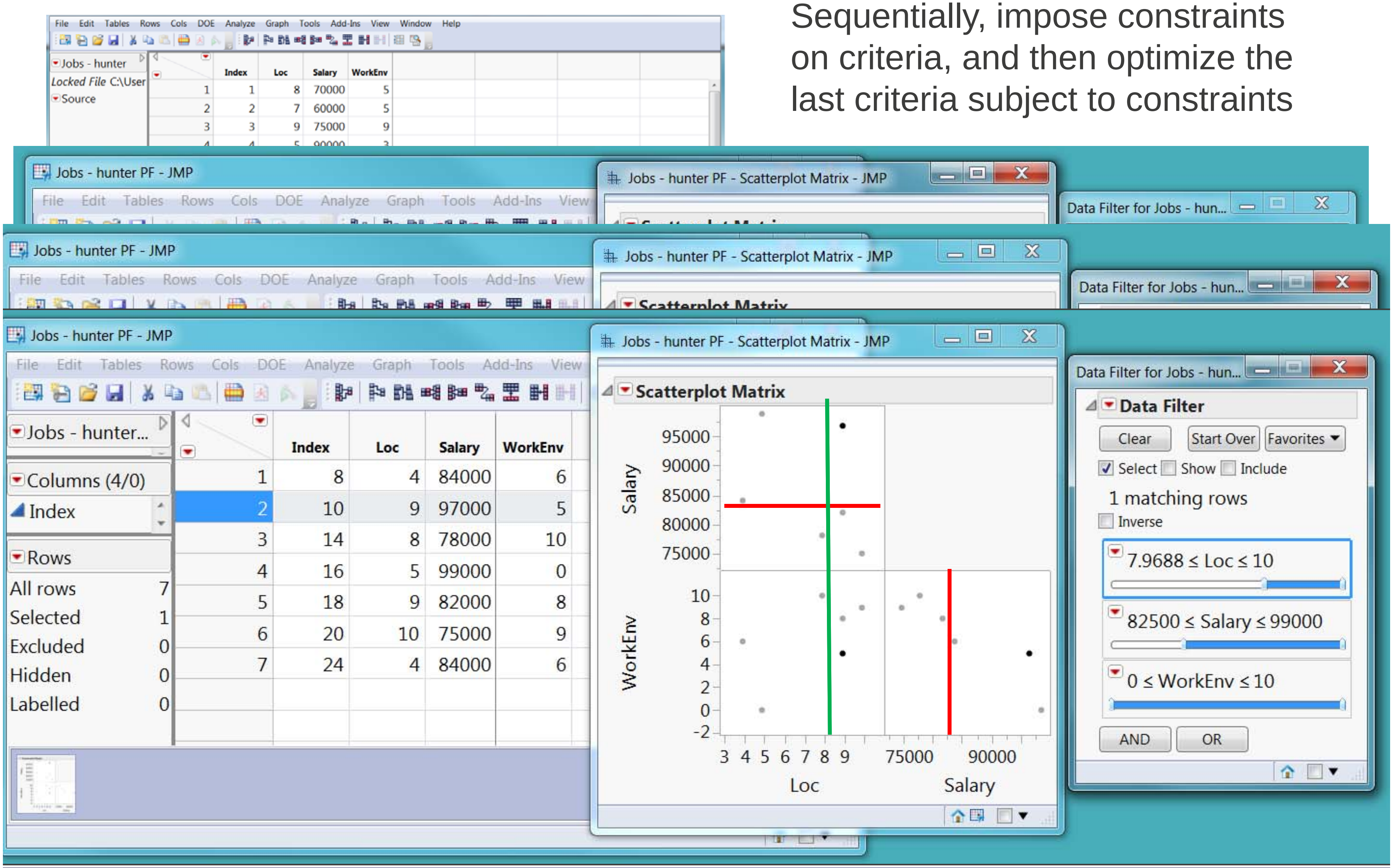
Yield / Viscosity / Molecular Weight



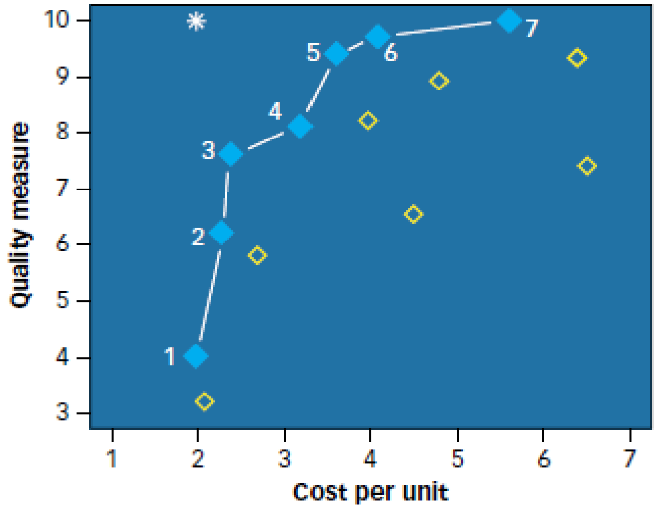
Real limitations on how this scales – often keeps it at the gut-feeling level

2. Conditional Optimization (Thresholds)

Sequentially, impose constraints on criteria, and then optimize the last criteria subject to constraints



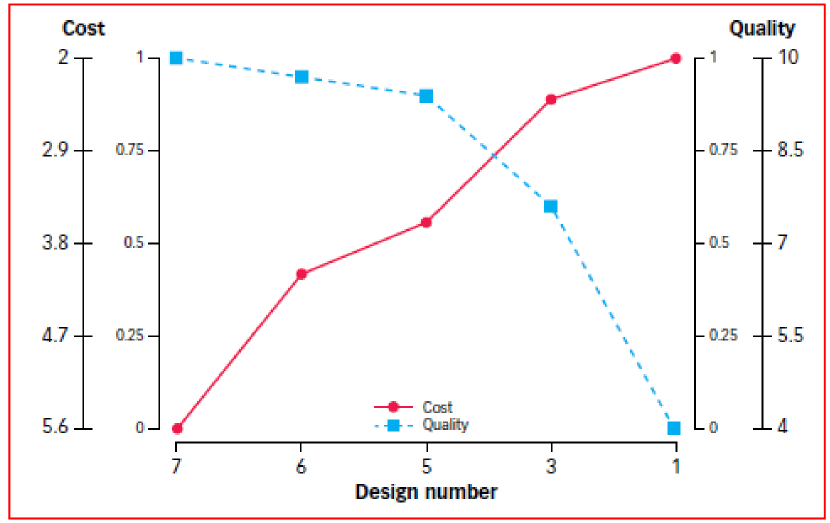
Desirability Based Graphical Tools



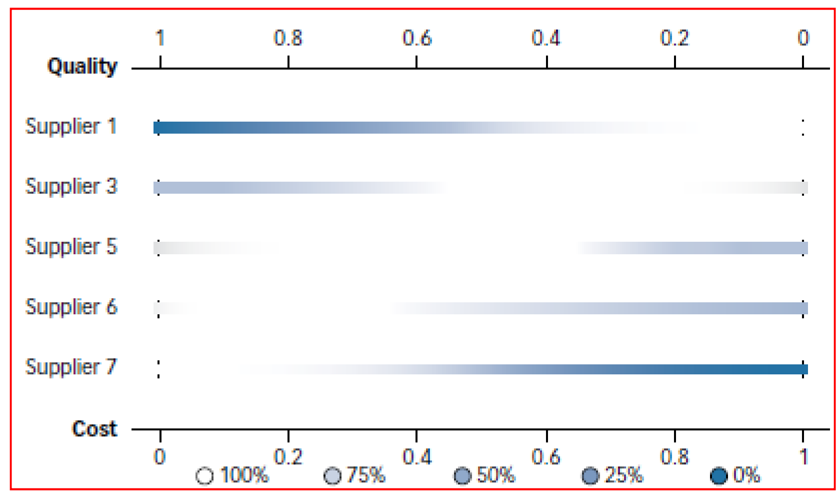
$$DF_{add} = w_{\text{cost}} d_{\text{cost}} + w_{\text{qual}} d_{\text{qual}}$$

$$DF_{mult} = (d_{\text{cost}})^{w_{\text{cost}}} \bullet (d_{\text{qual}})^{w_{\text{qual}}}$$

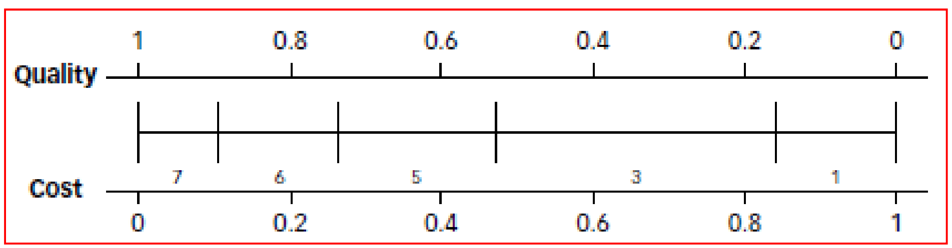
Trade-off Plot



Synthesized Efficiency Plot



Mixture Plot



Subjective Elements:

1. Scaling of DF scores, 2. DF form, 3. Weights

Outline

1. **Motivation**
   - Multiple criteria are everywhere – life is full of trade-offs
   - Roles in decision-making
   - Introduction of examples
2. **Define-Measure-Reduce-Combine-Select (DMRCS)**
3. **Pareto Front basics – objectively removing non-contenders**
4. **Deciding among contenders – challenges, strategies, facilitating consensus**
5. **Return to examples to illustrate the process**
6. **Conclusions**

Example 1. Stockpile Maintenance Prioritization –determine which stockpiles are in most urgent need of investment.

Most challenging part was facilitating competing priorities

Having process helped!

1. Brainstorm for possible metrics
   1. Number of rounds required in inventory
   2. Time required to spool up production

2. Group metrics into categories –redundant or too similar, missing metrics
   Potential categories emerge:

3. Make definition of each metric precise enough to assign scores

4. Populate scores for each of the stockpiles

5. Translate score in table to a consistent scale

6. Calculate group scores

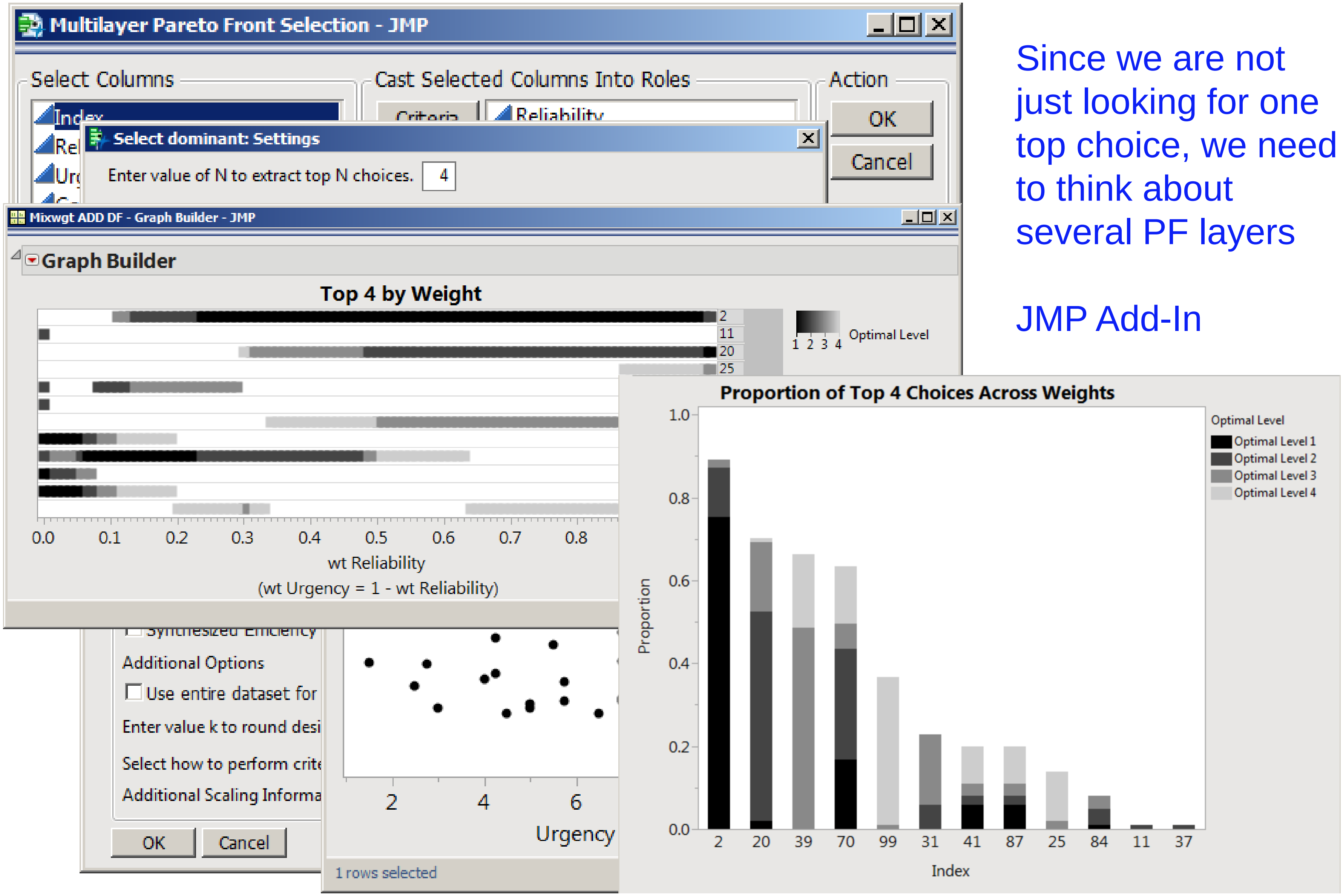
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1		ric #1	ric #2	ric #3	up 1 Score	ric #4	ric #5	up 2 Score	ric #6	ric #7	ric #8	ric #9	up 3 Score	ric #10	ric #11	Metric #12	Group 4 Score	
2	Stockpile #1																	
3	Stockpile #2																	
4	Stockpile #3																	
5	Stockpile #4																	
6	Stockpile #5																	
7	Stockpile #6																	
8	Stockpile #7																	

7. Plot the data to look for leading candidates – (Pareto front tools)

$DF_{additive} = \sum w_i SM_i$ where $\sum w_i = 1$

$DF_{multiplicative} = \prod (SM_i)^{w_i}$ where $\sum w_i = 1$

TopN-PFS (Burke et al) – software to help with



Since we are not just looking for one top choice, we need to think about several PF layers

JMP Add-In

2. Experiment Design Selection – many natural metrics for judging the goodness of a design

1. Result in good fit of the model to the data
2. Provide good model parameter estimates
3. Provide good prediction throughout the design space.
4. Provide an estimate of "pure" experimental error.

Good estimation and prediction for chosen model

5. Give sufficient information to allow for lack of fit test.
6. Provide a check on the homogeneous variance assumption.

Ability to test various aspects of the model

7. Be insensitive (robust) to the presence of outliers in the data.
8. Be robust to to errors in the control of design levels.

Protection if things go wrong

9. Allow models of increasing order to be constructed sequentially.
10. Allow for experiments to be done in blocks.

Flexibility to run and expand experiment

11. Be cost-effective.

Cost

Myers, Montgomery, Anderson-Cook RSM (2016) p. 370

Robust Parameter Design Experiment Example

- **Process with 4 control (A-D) and 3 noise (E-G) factors**

$$\begin{aligned} y = {} & \beta_0 + \beta_A A + \beta_B B + \beta_C C + \beta_D D + \beta_{AB} AB \\ & + \beta_{AC} AC + \beta_{AD} AD + \beta_{BC} BC + \beta_{BD} BD + \beta_{CD} CD \\ & + \beta_E E + \beta_F F + \beta_G G + \beta_{AE} AE + \beta_{AF} AF + \beta_{AG} AG \\ & + \beta_{BE} BE + \beta_{BF} BF + \beta_{BG} BG + \beta_{CE} CE + \beta_{CF} CF \\ & + \beta_{CG} CG + \beta_{DE} DE + \beta_{DF} DF + \beta_{DG} DG + \varepsilon. \end{aligned}$$

Terms involving control factors

Terms involving noise factors

- **Goal: create a design that (1) estimates the mean model based on the control factors well, and (2) estimates the control x noise factor terms well**
- **Levels of each of 7 factors: (-1,0,+1)**
- **Available budget: 27-30 runs**

Lu, Anderson-Cook, Robinson (2011)

Translate Goal into Criteria

Cost: Total Experiment size: 27, 28, 29, 30

Precision of estimation for mean and variance models:

$$X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \Rightarrow \mathbf{M} = |\mathbf{X}'\mathbf{X}| = \begin{pmatrix} \mathbf{M}_{11} & \mathbf{M}_{12} \\ \mathbf{M}_{21} & \mathbf{M}_{22} \end{pmatrix}$$

D_s-optimality allows focus on subsets of terms in model

$$(\mathbf{M}^{11})^{-1} = \mathbf{M}_{11} - \mathbf{M}_{12}\mathbf{M}_{22}^{-1}\mathbf{M}_{21} \Rightarrow \mathrm{D_S\text{-}mean\text{-}eff} = |(\mathbf{M}^{11})^{-1}|^{1/11}$$

$$(\mathbf{M}^{22})^{-1} = \mathbf{M}_{22} - \mathbf{M}_{21}\mathbf{M}_{11}^{-1}\mathbf{M}_{12} \Rightarrow \mathrm{D_S\text{-}var\text{-}eff} = |(\mathbf{M}^{22})^{-1}|^{1/15}$$

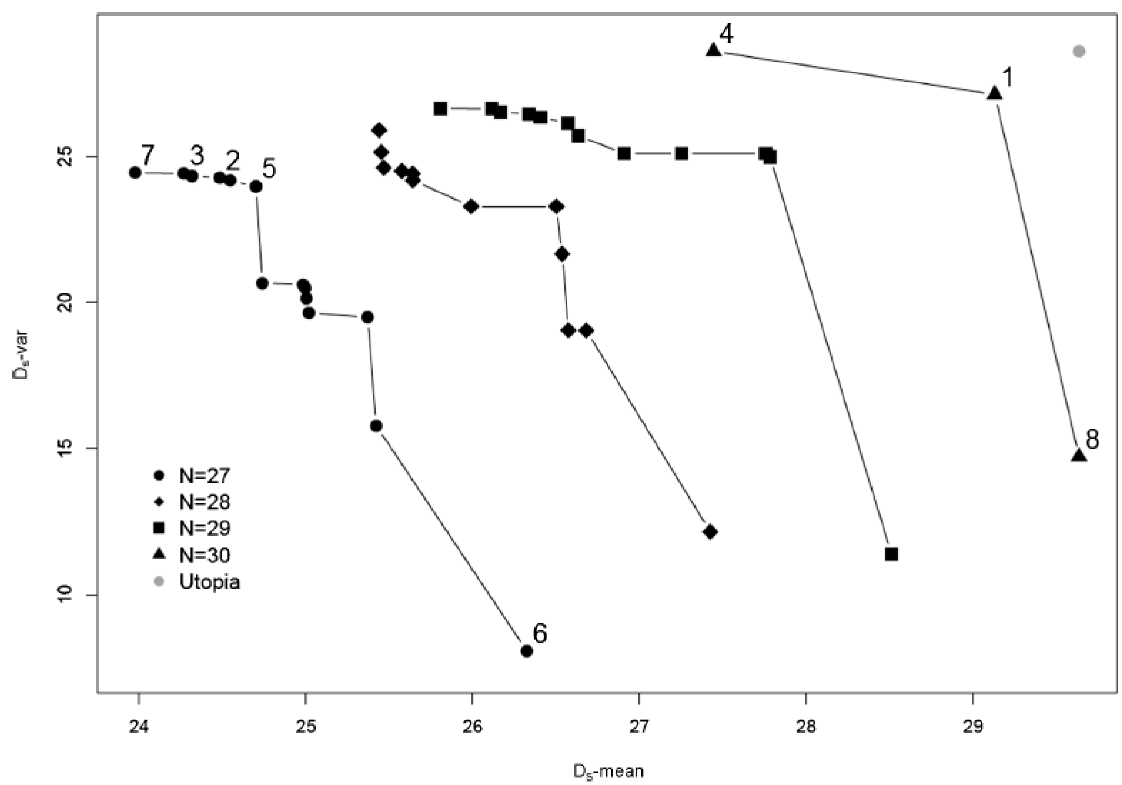
Example – Pareto front for N, D_s-mean, D_s-var

$$(3^7)^{27} + \cdots + (3^7)^{30}$$

$$= 10^{100} \text{ possible designs}$$

44 contenders
on Pareto front

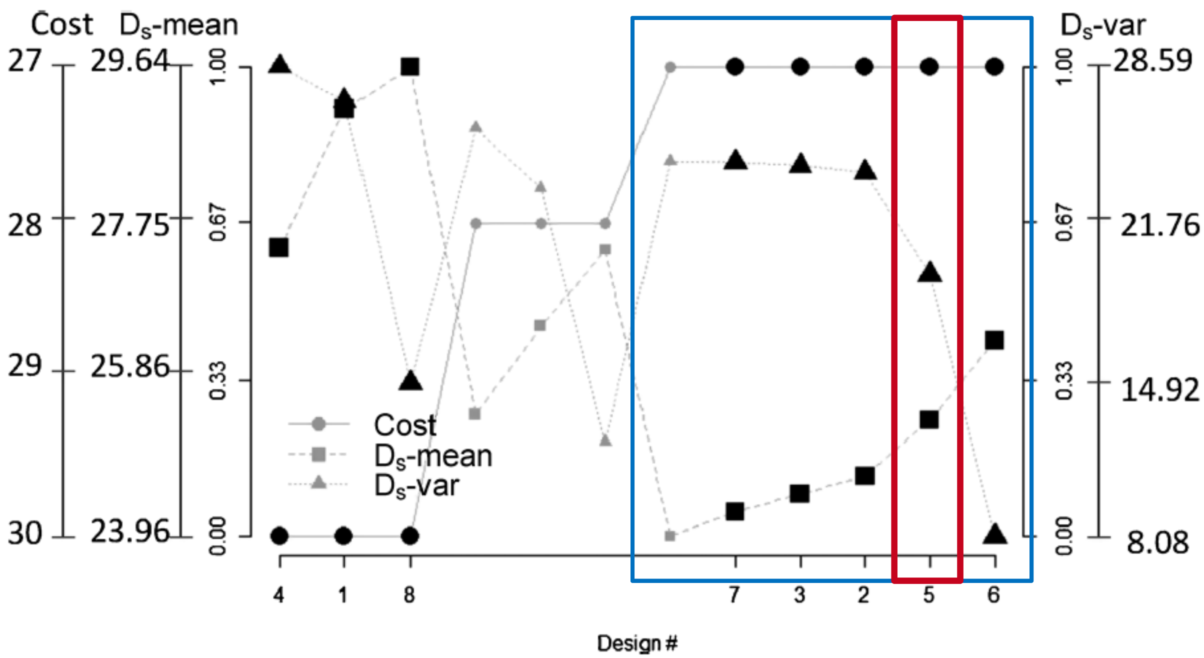
PAPE (Pareto Aggregating Point Exchange) Algorithm



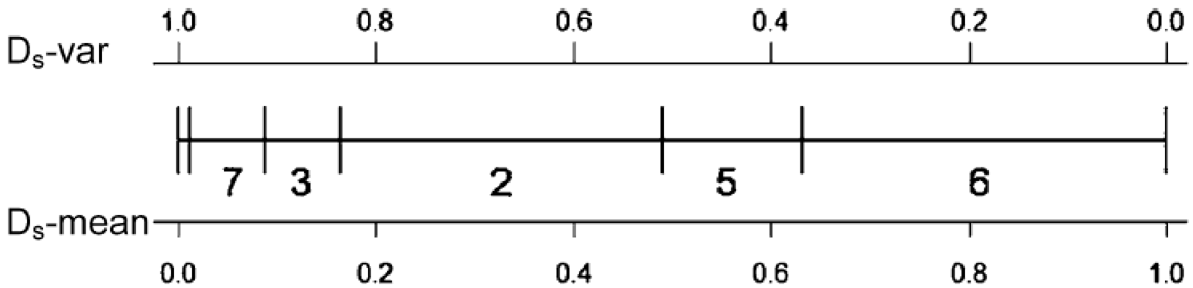
Still work to be done:
Ultimately only run one design – but now know choices and can find what is best for needs of experiment

Lu, Anderson-Cook, Robinson (2011)

Understanding the Choices



$$Desirability = w_{Cost}C_{Cost} + w_{Dm}C_{Dm} + w_{Dv}C_{Dv}$$



10^{100} designs

44 designs

12 designs

1 design

Other Examples for Design Construction

1. **Screening Experiment**:
   - D-optimality [maximize $|X'X|$]
   - Good estimation of pure error [maximize df_{PE}]
   - Good estimation of lack of fit [maximize $tr(R'R)/(m-p)$]

2. **Split Plot Design**
   - Good estimation of terms when WP to SP variance ratio is unknown [max D(0.1), max D(10)]
   - Size of experiment [min N]
   - Number of Whole Plots [min #WP]

3. **Sequential Experimentation for Estimating System Reliability**
   - Good prediction of
     - system reliability [min uncertainty interval]
     - mechanical sub-system reliability [min uncertainty interval]
     - electrical sub-system reliability [min uncertainty interval]

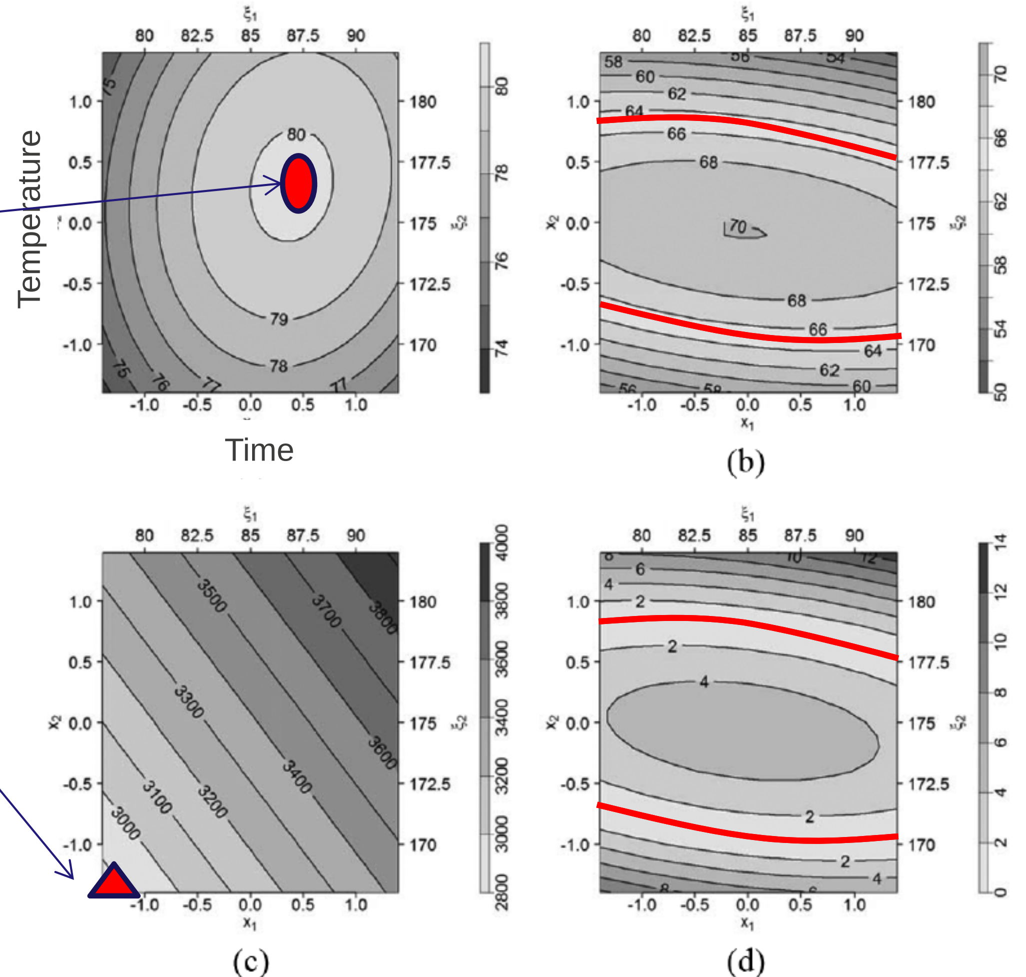
4. **Supersaturated Designs**
   - Avoid correlation between terms in model [min $E(s^2)$]
   - Reduce bias on terms in primary model [min $tr(AA')$]

3. Multiple Response Optimization – most experiments run collect data on multiple attributes.

Yield
(maximize)

No universally
best location

Molecular weight
(minimize)



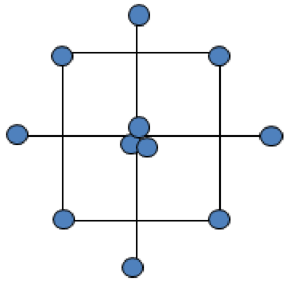
Viscosity
(close to 65)

|Visc – 65|
(minimize)

Response surfaces are estimated – therefore have uncertainty

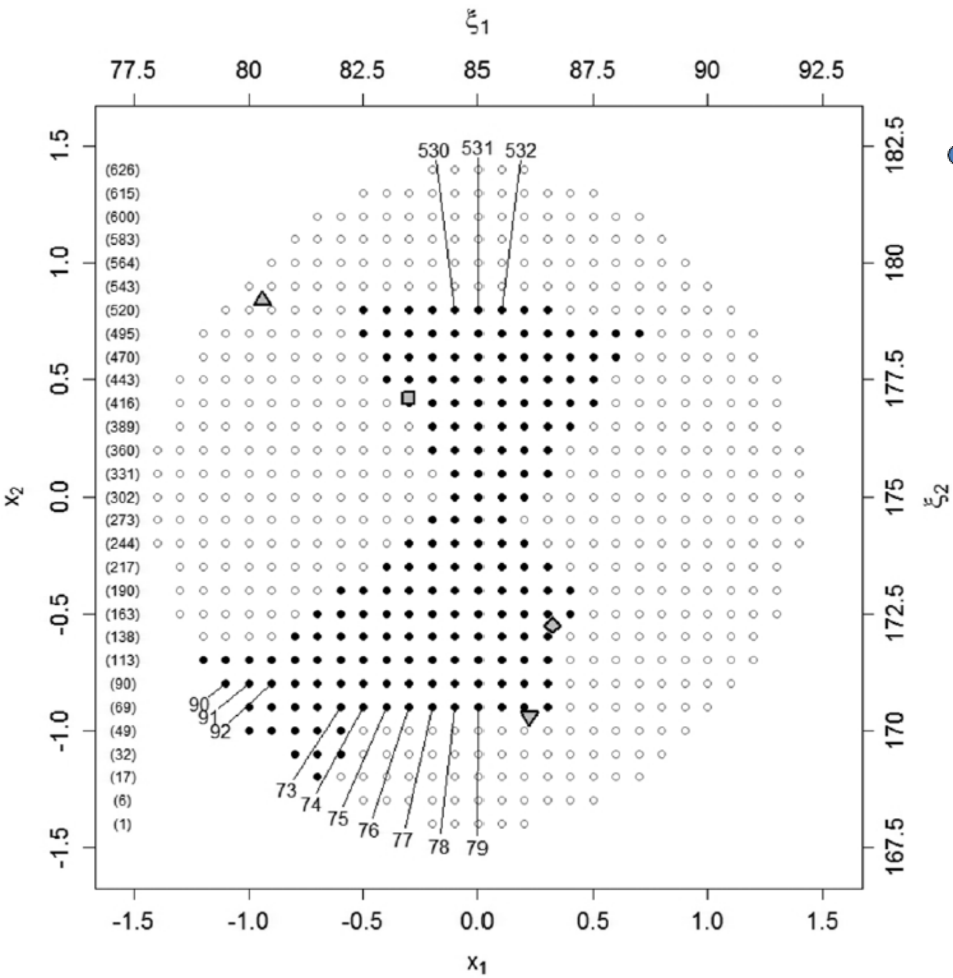
Experiment: 2 factor Central Composite Design with 3 responses

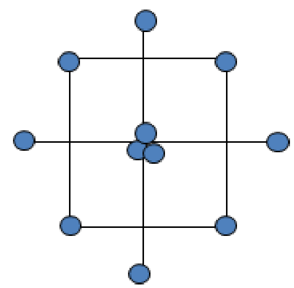
Natural variables		Coded variables		Responses		
ξ_1 (time)	ξ_2 (temperature)	x_1	x_2	y_1 (yield)	y_2 (viscosity)	y_3 (molecular weight)
80	170	−1	−1	76.5	62	2,940
80	180	−1	1	77.0	60	3,470
90	170	1	−1	78.0	66	3,680
90	180	1	1	79.5	59	3,890
85	175	0	0	79.9	72	3,480
85	175	0	0	80.3	69	3,200
85	175	0	0	80.0	68	3,410
85	175	0	0	79.7	70	3,290
85	175	0	0	79.8	71	3,500
92.07	175	1.414	0	78.4	68	3,360
77.93	175	−1.414	0	75.6	71	3,020
85	182.07	0	1.414	78.5	58	3,630
85	167.93	0	−1.414	77.0	57	3,150

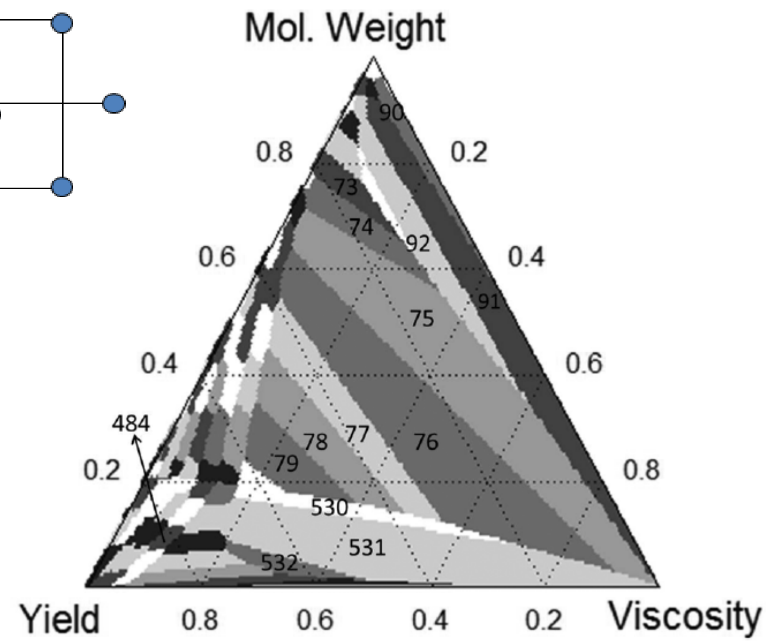


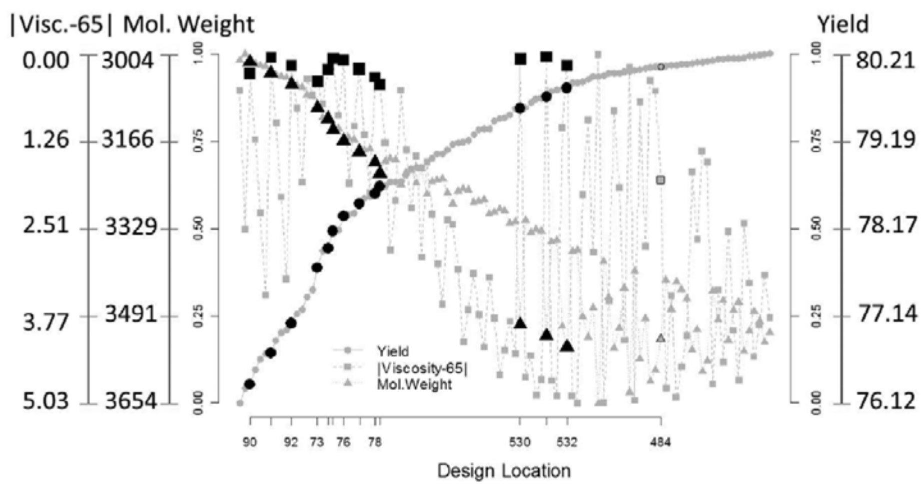
Goals: maximize target (close to 65) minimize

Optimizing – Finding the Pareto Front









$$Desirability = (C_{Yield})^{w_{Yield}} (C_{MW})^{w_{MW}} (C_{Visc})^{w_{Visc}}$$

Chapman, Lu, Anderson-Cook (2014a & b)

Outline

1. **Motivation**
   - Multiple criteria are everywhere – life is full of trade-offs
   - Roles in decision-making
   - Introduction of examples
2. **Define-Measure-Reduce-Combine-Select (DMRCS)**
3. **Pareto Front basics – objectively removing non-contenders**
4. **Deciding among contenders – challenges, strategies, facilitating consensus**
5. **Return to examples to illustrate the process**
6. **Conclusions**

How has this changed me?

- **Clearer sense of what is objective and what is subjective:**
  - More open-minded about alternative solutions
  - Learning more from my colleagues
  - Fairer in my assessments (eg. As journal reviewer)
    - “This is wrong” versus “I would have done it a different way”
- **More thorough in my own decision-making in daily life**
- **More charitable about different opinions, and finding it easier to survive the political season ☺**

Conclusions

- Challenging and realistic decisions have multiple competing objectives
- A structured decision-making process can lead to better results
  - Define – making a decision based on the right criteria, avoiding potential biases
  - Measure – using good quality quantitative data for comparisons between alternatives
  - Reduce – elimination of non-contenders
  - Combine – thinking about how comparable alternatives are, visualizing of alternatives
  - Select – making the right choice and being able to quantifying performance
- A flexible set of tools for different stages of the process will help adapt the process for different decisions
- Scientists, engineers and statisticians can benefit from early and frequent exposure to the decision-making process (presenting alternatives, fair pros and cons, defending a decision)

References

1. 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.
2. Chapman, J., Lu, L., Anderson-Cook, C.M. (2014a) "Process Optimization for Multiple Responses Utilizing the Pareto Front Approach" *Quality Engineering* 26 253–268.
3. Chapman, J.L., Lu, L., Anderson-Cook, C.M. (2014b) "Incorporating Response Variability and Estimation Uncertainty into Pareto Front Optimization" *Computers & Industrial Engineering* 76 253-267
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., Anderson-Cook, C.M., Robinson, T.J. (2012) "A Case Study to Demonstrate Pareto Frontiers for Selecting a Best Response Surface Design with Simultaneously Optimizing Multiple Criteria" *Applied Stochastic Models in Business and Industry* 28 206-221.
6. Lu, L., Chapman, J., Anderson-Cook, C.M. (2013) "A Case Study on Selecting A Best Allocation of New Data for Improving the Estimation Precision of System and Sub-System Reliability using Pareto Fronts" *Technometrics* 55 473-487.
7. Lu, L., Chapman, J.L., Anderson-Cook, C.M. (2016) "Multiple Response Optimization for Higher Dimensions in Factors and Responses" Los Alamos National Laboratory Technical Report LAUR 16-20568
8. Myers, R.H., Montgomery, D.C., Anderson-Cook, C.M. (2016) *Response Surface Methodology: Product and Process Optimization Using Designed Experiments.* New York: Wiley.