Robust design:
seeking the best of all possible worlds

Susan M. Sanchez
Operations Research Department Seminar
Naval Postgraduate School
February 10, 2000

Overview

What is robust design?

Why is it different for simulation experiments?

How can it improve decision-making?

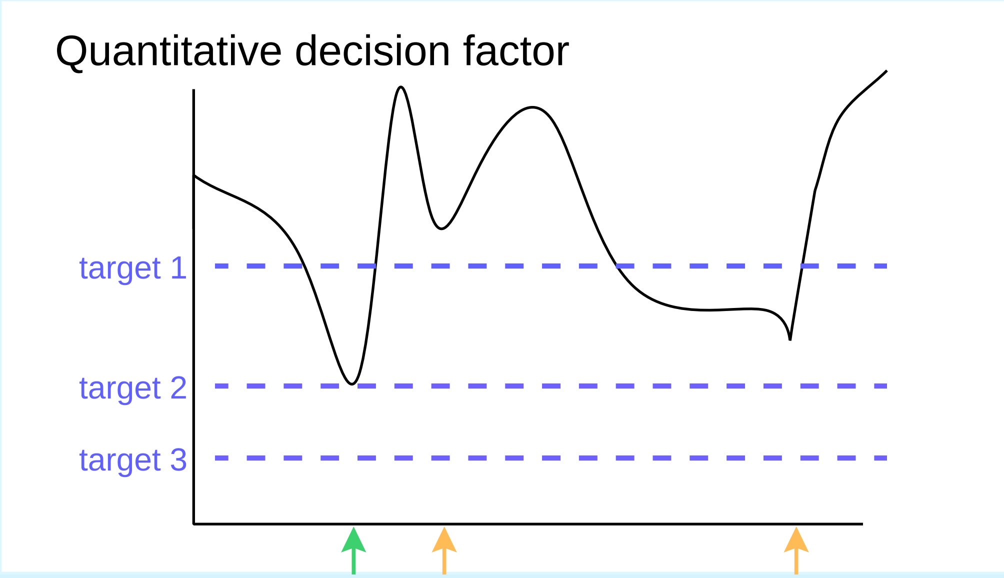
Where can we go from here?

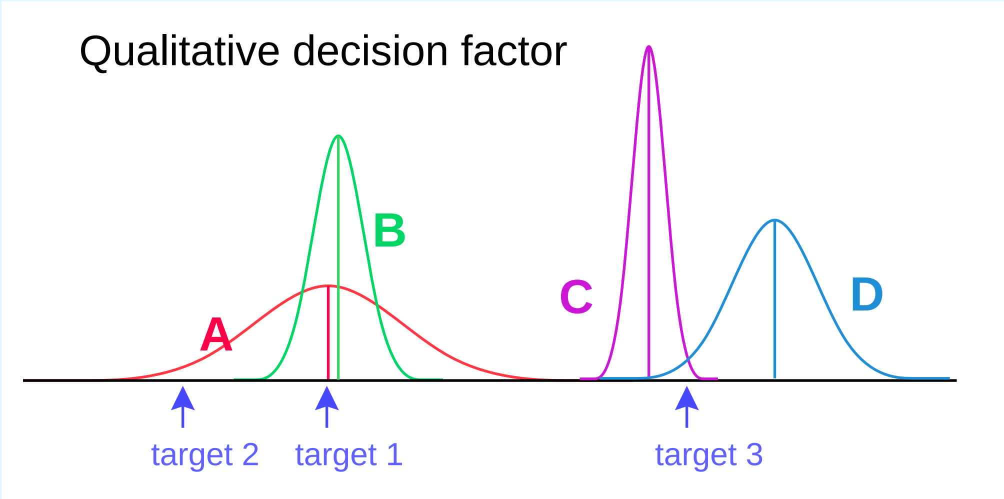
Robust design philosophy

We have a ***model*** of a system
(simulation model, analytic model,
statistical model, physical prototype)
on which we can perform experiments
and collect performance information.

The decision maker has specified
* performance measure (Y)
* performance 'target value' (τ)
* goal associated with Y, τ
  e.g., the smaller (or bigger) the better
  the closer to target the better

Y is affected by
* decision factors($\boldsymbol{x}$)
  controllable
* internal / external noise factors ($\boldsymbol{w}$)
  uncontrollable or controllable only
  at great expense





Robust design philosophy

Goal is achieved via a loss function ℓ
e.g.
$$\ell\left(y_{\mathbf{x}}\right) = c\left(y_{\mathbf{x}} - \tau\right)^2$$

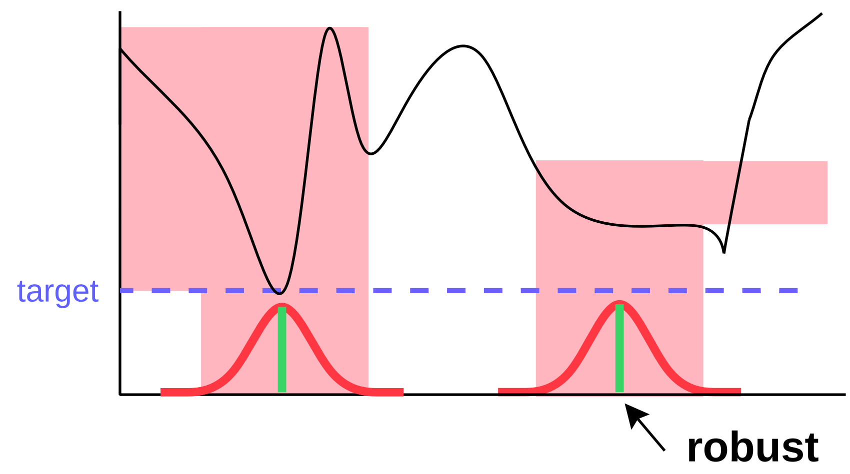
where *c* is a cost-conversion constant

Then, taking expectation over the noise space Ω we find
$$E_{\Omega}\left[\ell\left(y_{\mathbf{x}}\right)\right] = c\left[\sigma_y^2 + \left(\mu_y - \tau\right)^2\right]$$

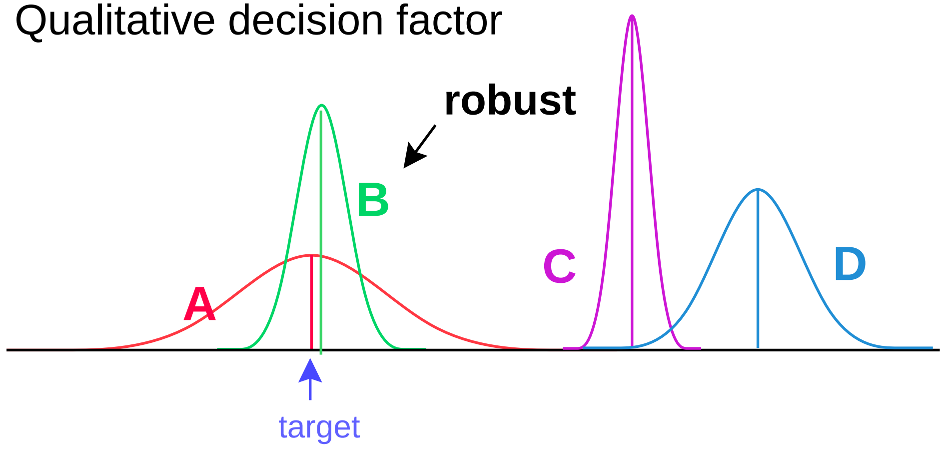
$E_{\Omega}\left[\ell\left(y_{\mathbf{x}}\right)\right]$ can be thought of as the *loss to society* (*long run business loss)* associated with a particular design $\mathbf{x}$

The ***robust design*** is that which ***minimizes expected loss***

Quantitative decision factor



Qualitative decision factor



Robust design benefits

Benefits of robust design include:

* fewer surprises moving from ‘lab’ to implementation
* better communication between analyst and client via expected loss
* ability to evaluate trade-offs between noise reduction costs & performance quality
* facilitates continuous improvement
* better decisions

simultaneously improve performance and decrease costs!

Other benefits in simulation:

* treats variability as a critical component of performance
  not solely a nuisance when estimating $E[Y]$, to be overcome by larger samples
* rapid model evaluation and scenario analysis (more efficient than trial-and-error)
* ability to test whether model performance is highly sensitive distribution ‘parameters’
* ability to test whether component models need more detail, or whether changing a component will materially affect performance

Response surface approach for quantitative decision factors

Approach:

1. **Select performance measure(s)**
   metrics & stat summaries to be used
2. **Specify target value and loss function**
3. **Identify factors, regions of exploration**
   classify as decision, noise or artificial
4. **Plan experiment**
5. **Conduct experiment**
6. **Analyze results**
   collapse data over noise space, construct regression metamodels for μ_y, $\log(\sigma_y)$
7. **Refine metamodel**
   parameters in metamodel should have effects greater than noise threshold
8. **Select and confirm choice of x**
   metamodels/contour plots suggest desirable x, confirm w/ extra runs for previously unexamined designs

Designed experiment, can use different data collection plans for decision factor & noise spaces

plan for **x**: allow estimation of quadratic & interaction terms
e.g., 3-level factorials, central composite designs, Box-Behnken

plan for **w**: many possibilities, less concern for detail
e.g., 2-level fractional factorials, frequency domain oscillation

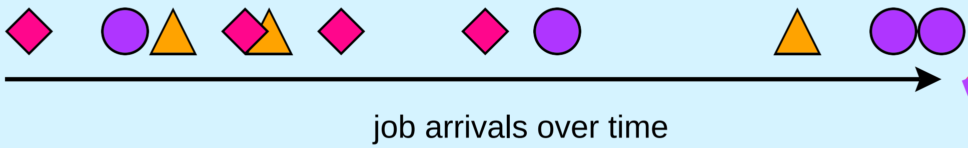
design for simulation-specific ‘artificial factors’ may enhance efficiency
e.g., common / antithetic random numbers

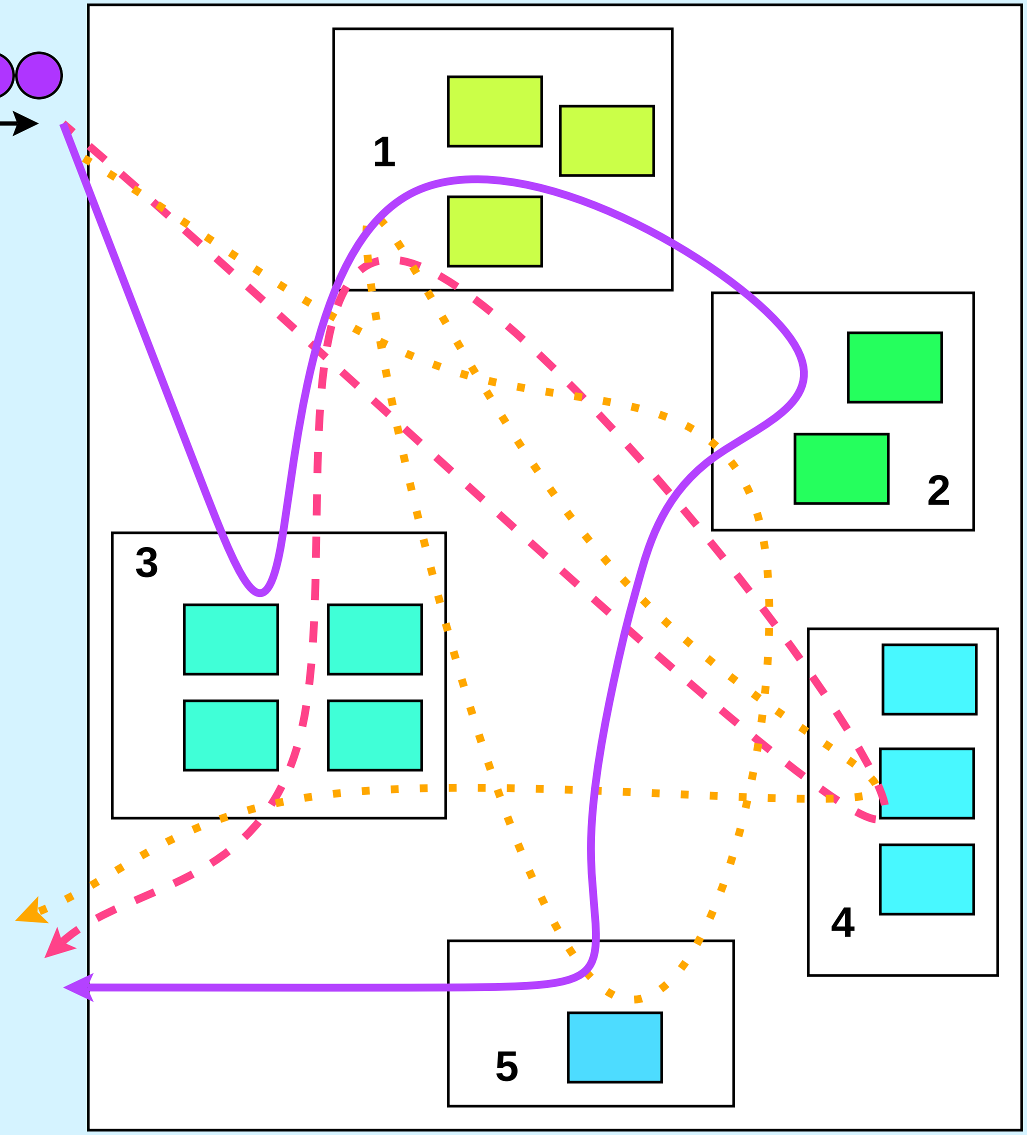
can cross or combine these plans to come up with final plan

The simulation environment: what's different?

	Taguchi's robust design	Robust design for simulation
setting	physical product prototypes	simulation model of product, process or system
x, w	few factors / few potential designs	many factors / many potential designs
model	$E[Y]$ is a linear function of $\mathbf{x}$, main effects only (ANOVA)	$E[Y]$ highly nonlinear, interactions expected (response surface methods)
noise structure	i.i.d. Normal, no time dependence	input: random number seeds output: correlated data streams
variability	completely removed once $\mathbf{w}$ specified; unequal variance results from (unfit) interactions between $\mathbf{x}$ and $\mathbf{w}$	inherent variability remains, may also be function of $\mathbf{x}$
unit of analysis	single data point	one run or batch (appropriate run/batch sizes?)
data collection	"Orthogonal arrays" or 2-level factorials	simulation-specific factors in data collection, innovative approaches possible
data cost	data expensive -> few data pts	either data or analyst may be the most expensive

Example: job shop





Three types of jobs (exponential interarrivals)

Differences by job type allowed for
* processing times (gamma),
* machine routings (fixed), and
* batch setup times (fixed)

FIFO queues form if all machines at a station are busy

Performance: time in system (τ = 10 hrs)

Loss: scaled squared error

$$\ell_s\left(y_{\mathbf{x}}\right) = \ell\left(y_{\mathbf{x}}\right)/c = \left(y_{\mathbf{x}} - \tau\right)^2$$

Job shop experiment

Data collection

Parameters:

M_1, M_2, M_3, M_4 : # machines at stations 1-4

B: common batch size

2^{5-1} resolution V design
augmented with 2 center pts (18 pts)

Noise factors:

PJ_1, PJ_2 : proportion jobs of types 1, 2

2^2 factorial (4 pts)

Crossed design: 18 $\times$ 4 = 72 configurations

Output: 18 pairs $\left(\overline{Y}_{i\cdot},\ s^2_{i\cdot}\right)$

Metamodels

Initial (full): 5 main effects, 10 interactions, 1 quadratic

$R^2 > .999$ for both μ, $ln(\sigma^2)$ metamodels

Final: (after augmenting with 6 configs)

μ metamodel:
3 main effects, 2 interactions, 2 quadratic
$R^2 = .969$, (adjusted = .955)
all included terms have $p < .01$,
no excluded term has $p < .05$

$ln(\sigma^2)$ metamodel:
2 main effects, 1 interaction, 1 quadratic
$R^2 = .979$ (adjusted=.974),
all included terms have $p < .01$,
no excluded term has $p < .05$

Selected systems

Selection criterion	Configuration(s) B	M_1	M_2	M_3	M_4	# machines stations 1-4	Scaled Loss (90% CI), validation
Low loss	**2**	**5**	**3**	**4**	**5**	17	**5.33** (4.43, 5.87), 5.66
Low mean	**1**	**5**	**3-4**	**4-6**	**5**	17-20	**8.03** ***57% worse!*** (7.50, 8.64), 7.64-7.67
Low variance	**3**	**5**	**3-4**	**4-6**	**3**	15-18	**6.84** ***28% worse!*** (6.03, 7.73), 6.99-7.16
	3	**5**	**3-4**	**4-6**	**4**	16-19	**6.34** (5.63, 7.13), 6.76-6.93
	3	**5**	**3-4**	**4-6**	**5**	17-20	**5.90** (5.22, 6.62), 6.65-6.95
Low loss from first 18 configurations	**1**	**5**	**3**	**6**	**5**	19	**7.61** ***43% worse!***

Specifying the cost coefficient c allows direct computation of total cost of machine + loss:

config x_1 = (2,4,3,4,5) has scaled loss 5.33, while x_2 = (2,5,3,4,5) has scaled loss 6.01

Let k_2 = machine cost at station 2, then x_1 preferred to x_2 if $k_2 / c > .68$

Phase II: Variance attribution
interesting questions we can answer

Approach:

System Evaluation

1. **Select noise factors**
   decision factor deviation from nominal is considered noise
2. **Plan the experiment**
   one plan, 2 or 3 levels per factor
3. **Conduct experiment**
4. **Assess system capability**
   estimate overall mean, variance

Noise Factor Assessment

5. **Construct metamodels**
   include inherent system variability
6. **Assess sensitivity**
   which noise sources propagate?

System Optimization

7. **Evaluate current configuration**
   which noise sources propagate?
8. **Evaluate alternatives**
   how do we improve the system?

Model complexity assessment:
Setup time was modeled as fixed for given job/machine combos, but it's random. Does this matter?

Data needs assessment:
Mean interarrival time, mean processing times are estimates. Does this matter?

Assessing environmental changes:
Marketing dept. thinks volatility for Product 1 demand will rise. What's the impact?

Assessing decision factor changes:
How much would performance suffer if we removed one machine from station 3?

Effective use of limited budget: best option?
(a) Refurbish station 1 machines for $16K each, reducing processing times by 10%,
(b) Replace station 1 machines for $35K each, reducing processing times by 30%
(c) Install new fixtures for $3K each, reducing setup times by 20%

Variance attribution details

Idea: a little work up front will save time, effort later on

we can answer questions using the metamodel results, without rerunning the simulation model unnecessarily!

The j^{th} run in the sampling plan yields output $\overline{Y}_j,\ s_j$ (after suitable truncation)

Fit metamodels

$$\hat{\mu} = \hat{\beta}_0 + \sum_{i=1}^{w} \hat{\beta}_i W_i$$

changes in mean

$$\hat{\sigma} = \hat{\gamma}_0 + \sum_{i=1}^{w} \hat{\gamma}_i W_i$$

changes in inherent variance

Estimate the system variance

$$Var(Y) \approx \underbrace{\hat{\gamma}_0^2}_{\textit{inherent}} + \underbrace{\sum_{i=1}^{w} \left(\hat{\beta}_i^2 + \hat{\gamma}_i^2\right) Var(W_i)}_{\textit{transmitted}}$$

Assess transmitted variances as

- % of total variance
- % of non-inherent variance

system may amplify or dampen noise factor variance

Job shop experiment...

Factors:

PJ_1, PJ_2: proportion jobs of type 1, 2

M_3: # machines at station 3

IA: interarrival time mean

α_{t1}: processing time multiplier at station 1

α_s: setup time multiplier

Sampling plan:

nearly-saturated factorial (8 runs)

Factor	Trans. Var	% of total	% non-inherent
Inherent	4.811	97.3	
M_3	.013	.3	9.6
PJ_1	.004	.1	3.3
PJ_2	---	---	---
α_{t1}	.033	.7	24.8
IA	.019	.4	14.7
α_s	.063	1.3	47.6
Total	4.943	100	100

Job shop results...

Model complexity assessment: Setup time was modeled as fixed for given job/machine combos, but it's random. Does this matter?

Data needs assessment: Mean interarrival time, mean processing times are estimates. Does this matter?

PROBABLY NOT -- changes in these factors have a relatively small impact on total system variance, though setups have most

Factor	% of total
Inherent	97.3
M_3	.3
PJ_1	.1
PJ_2	--
α_{t1}	.7
IA	.4
α_s	1.3
Total	100

Assessing environment: Marketing thinks the volatility for Product 1 demand will rise. What's the impact?

ALMOST NONE: doubling the demand variability doubles the transmitted variance, but increase is only 0.1% of total

Still more job shop results...

Assessing decision factors:
How much would performance suffer if we removed one machine from station 3?

Needed: cost conversion constant, e.g., \$300/unit for a 1-day (10 hr) dev from target

-> c = \$3 per hr^2 per unit

Current system:
mean = 11.075
var = 4.943
E[loss] = $3(1.075^2 + 4.943)$ = \$18.29/unit

New system:
mean = 11.188
var = 4.930
E[loss] = $3(1.188^2 + 4.930)$ = \$19.02/unit

Cost? \$0.73 per unit, or about \$29 per day

Effective use of limited budget:
which of these is the best option?
(a) Refurbish, reduce proc times 10% (\$48K)
(b) Replace, reduce proc times 30% (\$105K)
(c) New fixtures, reducing setups 20% (\$57K)

$\text{Loss}_{\text{init}} = 3[(11.075 - \tau)^2 + 4.923] = \$18.29\text{/unit}$

$\text{Loss}_{(a)} = 3[(10.840 - \tau)^2 + 4.937] = \$17.21\text{/unit}$

$\text{Loss}_{(b)} = 3[(10.531 - \tau)^2 + 4.926] = \$15.63\text{/unit}$

$\text{Loss}_{(c)} = 3[(10.071 - \tau)^2 + 4.903] = \$14.72\text{/unit}$

Over a 5-year time horizon (41,667 units),
(a) net loss of \$3K
(b) net savings of \$5.8K
(c) net savings of \$91.8K

Example: home mortgage portfolio forecast model

Setting:

18-month project with top 10 financial institution to set up forecasting system for portfolio of over 90,000 loans, making use of

* loan characteristics
* market conditions

Semi-markov model with three components:

* state transition (multinomial logit)
* loss or no-loss on defaults (logit)
* proportion lost (regression)

When finished, 146 fitted coefficients

1 hour / run on campus Sparc
9 hrs / run on client machine
up to 4 days / run on campus IBM

Tolerance analysis for portfolio segment:
fixed rate, uninsured S Cal new loans
with initial balance < $150,000

1. Assess impact of 61 fitted coefficients for several performance measures
   nearly saturated 2-level factorial, 64 runs
   coefficient levels at fitted +/- std. dev.

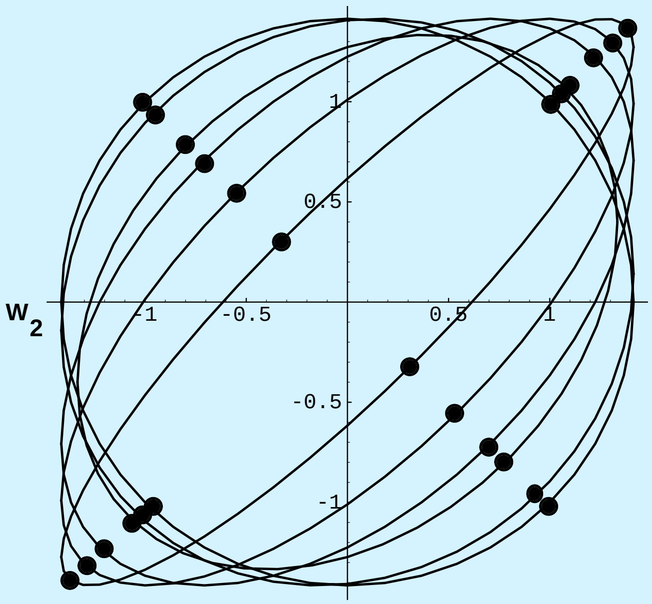
Transmitted Variance%

Coeff. group	**pct default**	**pct loss**	**time on books**
Transitions from			
Current	86	55	66
Delinq 30-89 days	0	0	0
Delinq 90+ days	14	5	32
Loss	0	40	1
Severity	0	0	0
Total	100	100	100

Portfolio forecast model results

2. Assess relative importance of certain loan and market characteristics

needed sampling plans, analysis appropriate for correlated noise factors



correlated

Average Portfolio Performance

	Transmitted Variance%		
Factor group	**pct default**	**pct loss**	**time on books**
Unemploymt Mkt apprec	69.6	82.5	56.4
Conv. rate	0.2	0.3	3.6
Loan/Value Loan amount	23.2	11.8	26.7
Interest rate	6.6	5.0	5.0
Term	0.4	0.4	8.3
Total	100.0	100.0	100.0

Where can we go from here?

Applications:

* Selection / sensitivity for particular system
* Rapid model / system assessment

Extensions:

* Robust selection for qualitative systems
  P(correct selection), P(good selection)
  fully sequential procedures
* Rare event simulation
  procedures that prefer sampling contenders, sample path generation

Efficient procedures

* Data expensive: *fully sequential procedures, exploit artificial factors such as batch overlap, common / antithetic sampling, control variates*
* Analyst expensive: *automated sampling plan generation,within-run noise factor plans (frequency domain oscillation)*

Structural issues:

* Impact of various loss functions
  guidelines for specific problem classes
* Correlated decision factors
* Other metamodel structures
  splines, radial basis functions,...

Other

* Explore robust design in conjunction with optimization models
  relation to stochastic programming?
* Local use of global metamodels
  when should you refit a global model with new local data?
  when should you narrow model scope?
* Robustness with feedback
  can results be obtained when ‘‘noise" is really enemy’s strategic decision?

Summary

Robust design is a great way to analyze models of complex systems because

* **It is flexible:** can be applied to models which are analytic, statistical, terminating or non-terminating simulations, or physical models

* **It is efficient:** it can indicate when model components have sufficient detail, sampling plans can be chosen to keep either data requirements or analysts' time and effort low

* **Solutions are realistic *by design*:** suggested system configurations have shown they'll behave well over a broad range of adverse conditions

* **It facilitates continuous improvement:** it clearly indicates important determinants of system variance, guides efforts for system 'optimization' and improvement by conveying hidden costs to decision-makers

Related Papers

Ramberg, J. S., S. M. Sanchez, P. J. Sanchez, and L. J. Hollick (1991), ‘‘Designing Simulation Experiments: Taguchi Methods and Response Surface Metamodels,’’ *Proceedings of the 1991 Winter Simulation Conference*, 167--176.

Ramberg, J. S., J. J. Pignatiello, Jr. and S. M. Sanchez (1992), ‘‘A Critique and Enhancement of the Taguchi Method.’’ *ASQC Quality Congress Transactions*, May, 491--498.

Schruben, L. W., S. M. Sanchez, P. J. Sanchez and V. A. Czitrom (1992), ‘‘Variance Reallocation in Taguchi’s Robust Design Framework,’’ *Proceedings of the 1992 Winter Simulation Conference*, 548--556.

Sanchez, S. M., J. S. Ramberg, J. Fiero, and J. J. Pignatiello, Jr. (1993), ‘‘Quality by Design.’’ Chapter 10 in *Concurrent Engineering: Automation, Tools, and Techniques*, ed. A. Kusiak. John Wiley and Sons: NY, 235--286.

Sanchez, S. M. (1994). ‘‘A Robust Design Tutorial.’’ *Proceedings of the 1994 Winter Simulation Conference*, 106--113.

Sanchez, S. M. (1994). ‘‘Experiment Designs for System Assessment and Improvement when Noise Factors are Correlated.’’ *Proceedings of the 1994 Winter Simulation Conference*, 290--296.

Sanchez, S. M., L. D. Smith and E. C. Lawrence (1996). ‘‘Sensitivity and Scenario Analysis for Simulation Metamodels.’’ *Proceedings of the 1996 Winter Simulation Conference*, 1440--1447.

Sanchez, S. M., P. J. Sanchez, J. S. Ramberg and F. Moeeni (1996). ‘‘Effective Engineering Design Through Simulation.’’ *International Transactions on Operational Research*, Vol. 3, No. 2, 169--185.

Smith, L. D., S. M. Sanchez and E. Lawrence (1996). ‘‘A Comprehensive Model for Managing Credit Risk and Forecasting Losses on Home Mortgage Portfolios.’’ *Decision Sciences*, Vol. 27, No. 2, 291--317.

Lawrence, E. C., L. D. Smith and S. M. Sanchez (1996), ‘‘Managing Credit Risk in a Changing Economy,’’ *Bankers Magazine*, January/February issue, 58--63.

Moeeni, F., S. M. Sanchez and A. J. Vakharia (1997). ‘‘A Robust Design Methodology for Kanban System Design.’’ *International Journal of Production Research*, Vol. 35, No. 10, 2821--2838.

Sanchez, S. M., P. J. Sanchez, and J. S. Ramberg (1998). ‘‘A Simulation Framework for Robust System Design.’’ Chapter 12 in *Concurrent Design of Products, Manufacturing Processes and Systems*, ed. B. Wang. Gordon and Breach, NY: 279--314.

Sanchez, S. M. and P. Konana (2000). ‘‘Efficient Data Allocation for Frequency Domain Experiments.’’ *Operations Research Letters* 26(2), 81-89.

Sanchez, S. M., L. D. Smith and E. C. Lawrence (2000). ‘‘Tolerance Design Revisited: Assessing the Impact of Correlated Noise Factors." Working paper, University of Missouri -- St. Louis (under review).