Simulation Calibration with Correlated Knowledge-Gradients

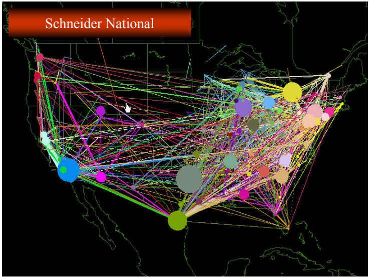
Peter Frazier[1]
Warren Powell[2]
Hugo Simão[2]

[1]Operations Research & Information Engineering, Cornell University
[2]Operations Research & Financial Engineering, Princeton University

Friday March 1, 2013
SIAM CSE, Boston

Simulation Model Calibration at Schneider National

- The logistics company Schneider National uses a large approximate dynamic programming (ADP) model to try "what if" scenarios.
- The model has several input parameters that must be tuned to make its behavior match reality before it can be used.
- The model is tuned by hand once per year on the most recent data. Each tuning effort requires between 1 and 2 weeks.



Incomplete simulation runs provide useful information

- Calibration requires adjusting the inputs to make the ADP model's output match reality.
- Running the simulator to convergence for one set of bonuses takes 3 days, making calibration difficult.
- The model may be run for shorter periods of time, e.g. 12 hours, to obtain output statistically related to the 3 day "gold standard" value.
- Human calibrators use these short runs to great advantage, but their unconventional relationship to the objective function make them difficult to use in existing algorithms.
- We design a new algorithm for incorporating information from these short runs, and for choosing which short runs to do.

Simulation Model Calibration: Notation and Definition

- $\theta_j(x)$ is the model's limiting output for variable j when given input x.
- g_j is our goal for output variable j.
- $f(x)$ is the fitting error with input parameters $x \in \mathscr{X} \subseteq \mathbb{R}^p$,

$$f(x) = \sum_{j=1}^{J} (\theta_j(x) - g_j)^2.$$

- Calibration is the following global optimization problem with expensive noisy measurements:

$$\min_x f(x),$$

We follow a Bayesian Global Optimization approach

- Bayesian Global Optimization (BGO) [Mockus 1989, Jones et al. 1998] is a general approach for global optimization of functions that are expensive or time-consuming to evaluate.
- We follow a BGO approach:
  - We begin with a Gaussian process prior distribution on the unknown function, which is generally a Gaussian process.
  - The parameters of the prior were chosen using data from past calibrations and conversations with the calibration expert at Schneider.
  - We combine the function evaluations observed so far with the prior to obtain a posterior.
  - Then, we use the posterior to choose the next point to evaluate.
- Most BGO algorithms assume our observations are unbiased estimates of our objective $f(x)$. Here, we design an algorithm for observations with different statistical properties.

Illustration of a BGO algorithm using unbiased observations

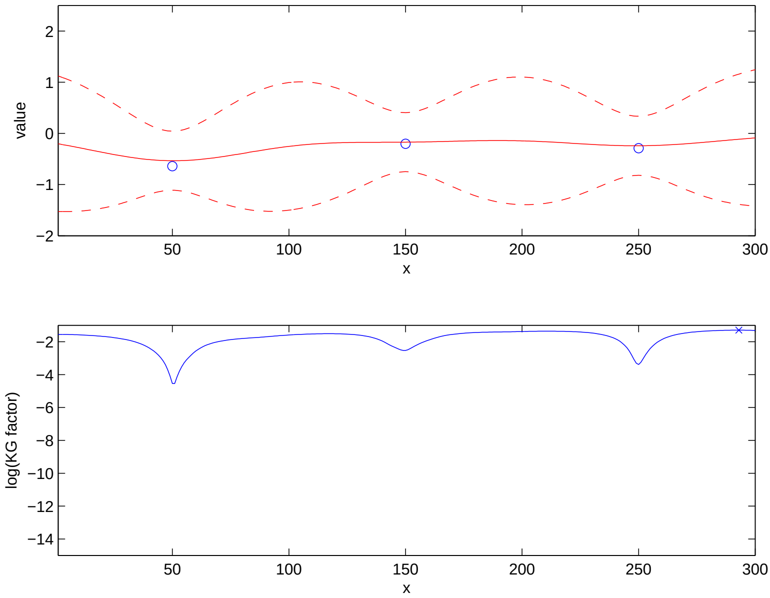


Illustration of a BGO algorithm using unbiased observations

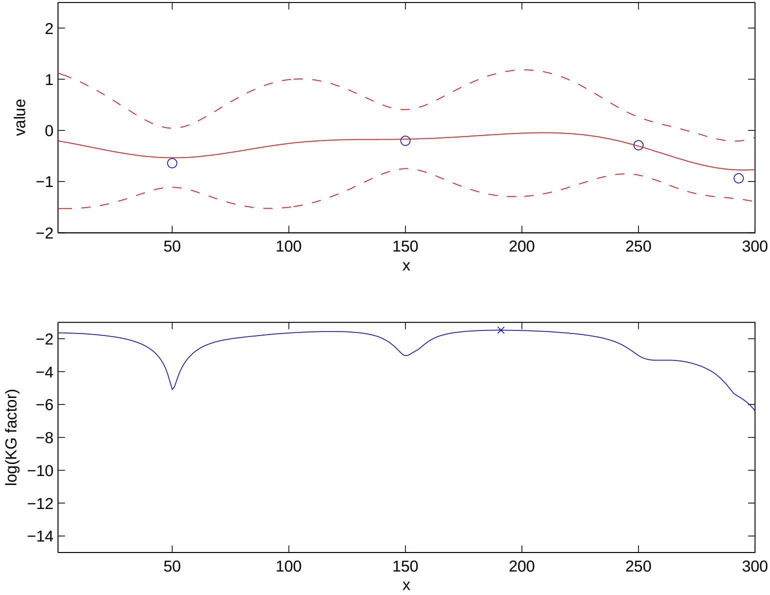


Illustration of a BGO algorithm using unbiased observations

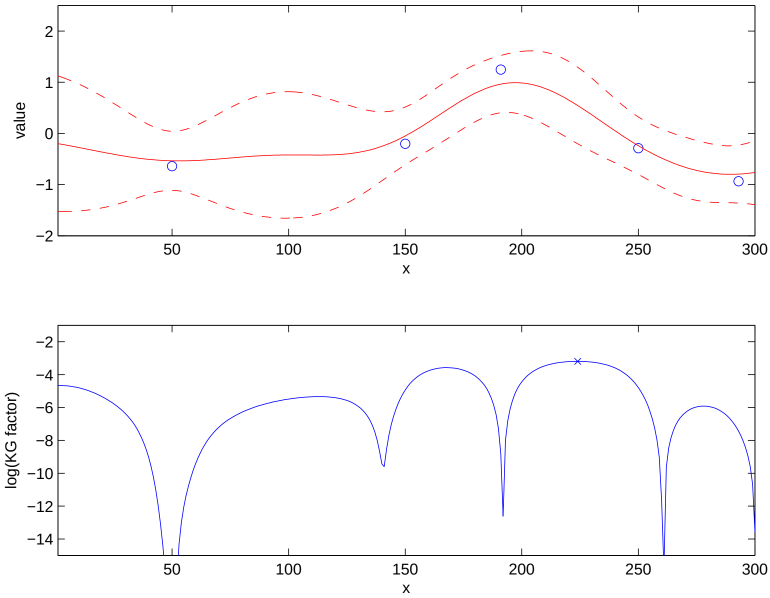


Illustration of a BGO algorithm using unbiased observations

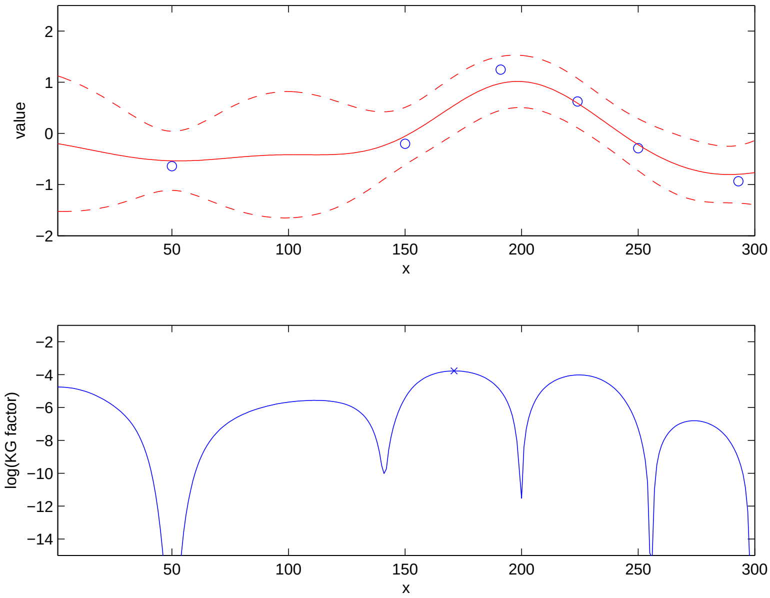


Illustration of a BGO algorithm using unbiased observations

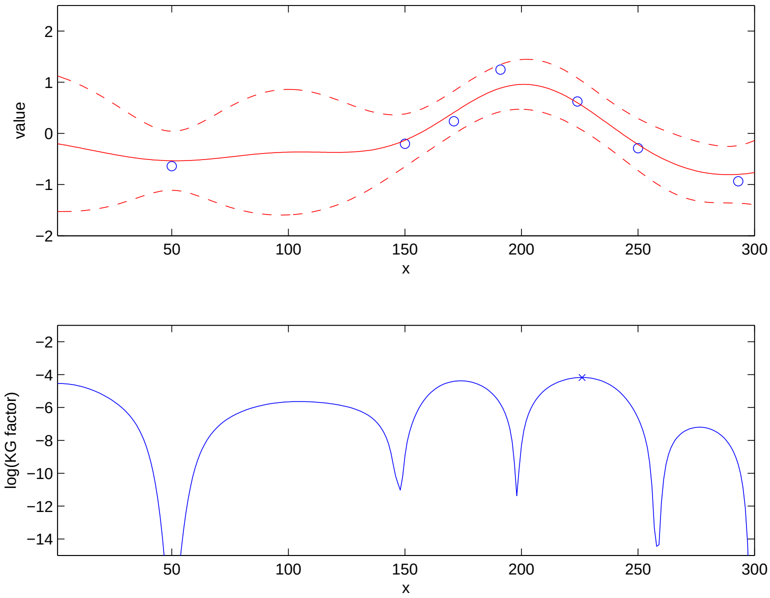


Illustration of a BGO algorithm using unbiased observations

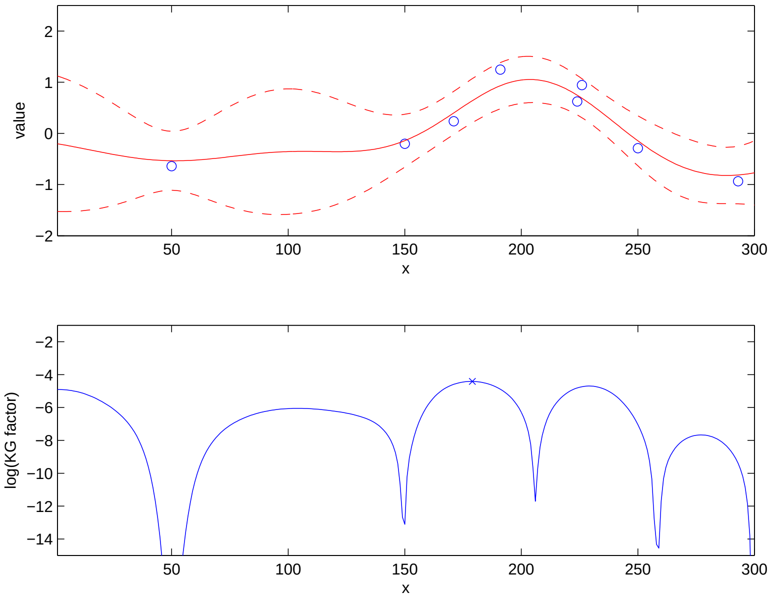
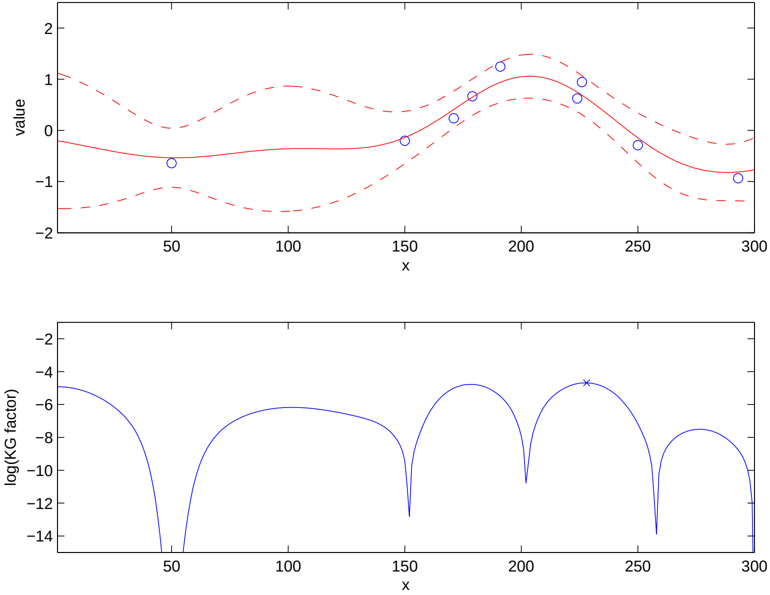


Illustration of a BGO algorithm using unbiased observations



Knowledge-Gradient Policy

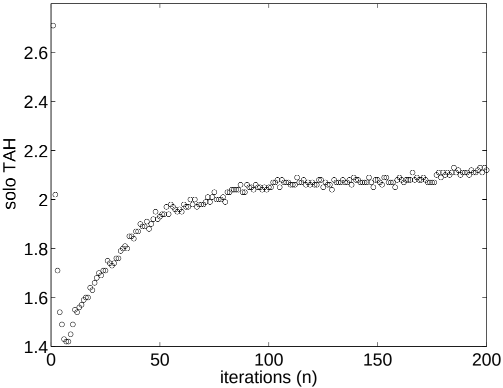
The **knowledge-gradient (KG) policy** is defined to be the policy that chooses its next measurement x_m to maximize the KG factor,

$$\nu^{KG}(x) = \left[\min_{x'} \mu_m(x')\right] - \mathbb{E}_m\left[\min_{x'} \mu_{m+1}(x') \mid x_m = x\right].$$

- $\mu_m(x') := \mathbb{E}_n[f(x')]$ is the expected loss at x' given what we know at time m.
- $\min_{x'} \mu_m(x')$ is the best we can do given what we know at m.
- $\min_{x'} \mu_{m+1}(x')$ is the best we will be able to do given what we know at m and what we learn from our measurement x_m.
- The KG factor is similar to expected improvement [Jones et al. 1998], and is the expected value of sampling information [Howard 1966].

ADP Output

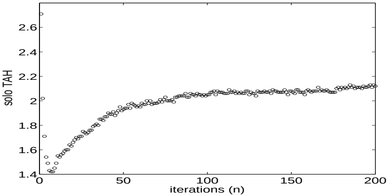
Typical ADP Output at **one** choice of the parameter vector x.
The plot shows sampled time-at-home for one particular driver type over 200 ADP training iterations.



Reconciling ADP Output and BGO Formulation

- The classic BGO formulation assumes that an observation at x has distribution $\mathrm{Normal}(f(x), \lambda(x))$.
- If we run our ADP model to convergence (say, to 200 iterations), then this assumption is met...
- ... but running to convergence at a single x takes 3 days.
- If our x seems bad after a few iterations, we should stop early.
- Human calibrators use early stopping to their advantage.

Statistical Model of ADP Output



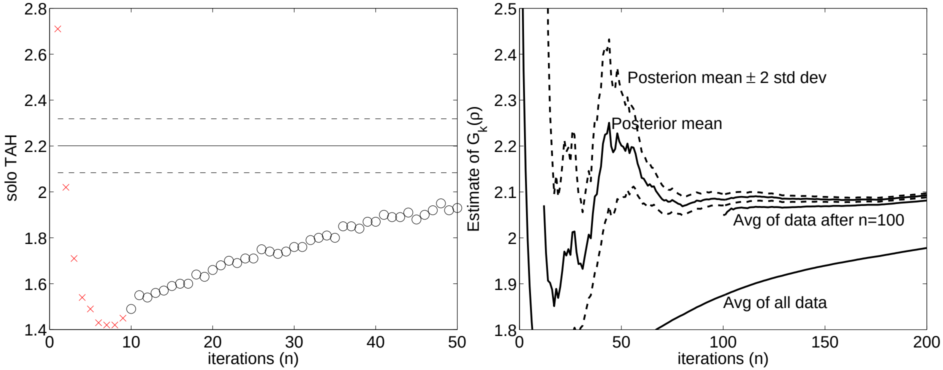
We model the ADP output as

$$Y_j^n(x) = B_j(x) + [\theta_j(x) - B_j(x)]\,[1 - \exp(-nR_j(x))] + \varepsilon^n, \quad n > n_0.$$

- $Y_j^n(x)$ are direct observations from the ADP model;
- $\theta_j(x)$ is the limiting value to which this output converges;
- $R_j(x)$ is the rate at which the output converges to its limiting value;
- $n_0 = 10$ allows us to ignore erratic initial output;
- $B_j(x)$ is, roughly speaking, the output at the first iteration;
- ε^n is an independent unbiased normal random variable.

Working with Non-stationary Output

Using the model, we may obtain an estimate of $\theta_j(x)$ from observations $Y_j^n(x)$, $n = 1, \ldots, m$, for m less than 200.



KG method

- Recall that the KG factor is given by

$$\nu^{KG}(x) = \min_{x'} \mu_m(x') - \mathbb{E}_m \left[\min_{x'} \mu_{m+1}(x') \mid x_m = x \right],$$

  and that the KG policy measures the x with the largest KG factor.
- This KG factor is well-defined even when the observations are non-stationary.
- To compute the KG factor, we use the predictive distribution of $(\mu_{n+1}(x'))_{x' \in \mathscr{X}}$ given that we measure at x.

Computing the KG policy (Approximately)

- We have $\mu_{m+1}(x) = \mathbb{E}_{m+1}\left[\sum_j(\theta_j(x) - g_j)^2\right]$, which is a function of the time $m+1$ posterior mean and variance of $\theta_j(x)$.
- We calculate the predictive distributions for the time $m+1$ posterior mean and variance of $\theta_j(x)$ from the statistical model.
- We then calculate $\mathbb{E}_m[\mu_{m+1}(x)] = \mu_m(x)$ and $\tilde{\sigma}^m(x, x_m) = \sqrt{Var_m[\mu_{m+1}(x) \mid x_m]}$.
- $\max_{x\in\mathscr{X}} \mu_{m+1}(x) \approx \max_{x\in\mathscr{X}'} \mu_m(x) + \tilde{\sigma}^m(x, x_m)Z$ where $\mathscr{X}' \subset \mathscr{X}$ is a finite subset and Z is a one-dimensional standard normal random variable.
- Then, the KG factor is approximated by the expectation of a piecewise linear function of a standard normal random variable. This expectation can be computed analytically.

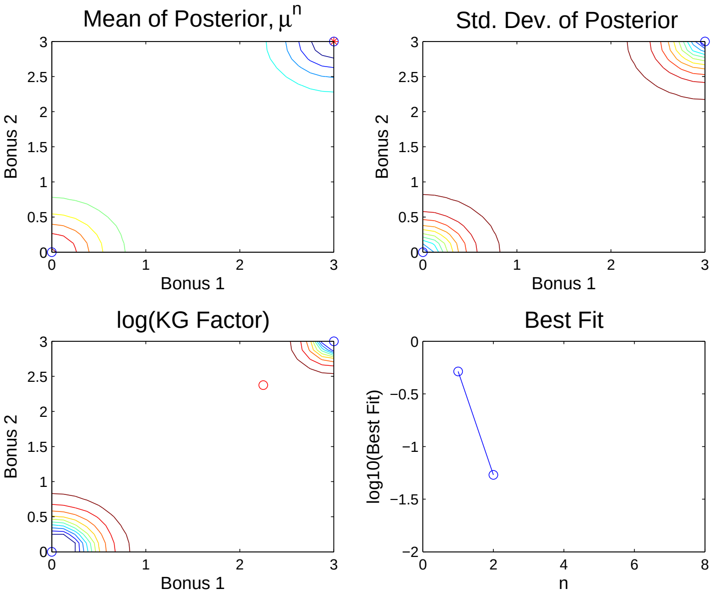
The Knowledge-Gradient Policy is a One-Step Approximation to the Bayes-Optimal Policy

- If just one measurement remains, then the knowledge-gradient policy is optimal (in the average case under our prior).
- If more than one measurement remains, it is not optimal:
  - We view KG as a one-step approximation to optimal.
  - The optimal policy is the solution to a (very difficult) dynamic program.
  - In the calibration problem presented here, I have no idea what the solution to this dynamic program is.
  - In other BGO problems, the solution to the dynamic program is known. In some of these problems, KG is close to optimal. In others, it is far.

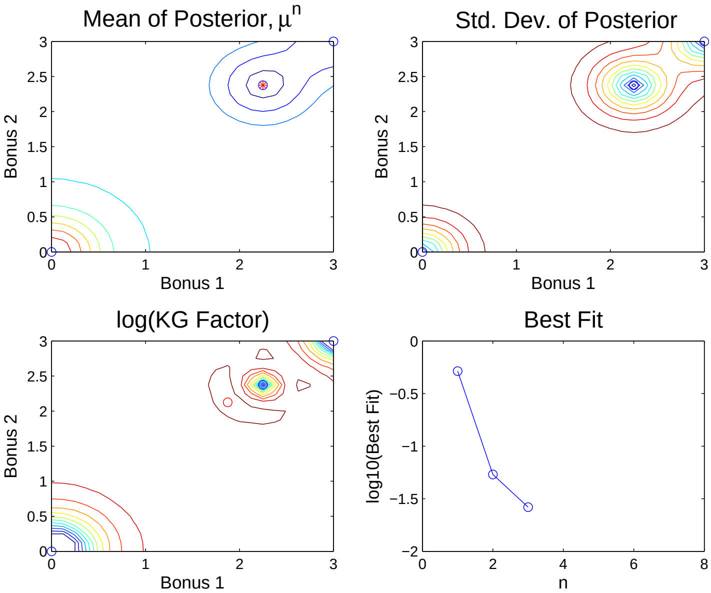
Time-at-Home

- The most critical input parameters are the time-at-home (TAH) bonuses.
- The optimization model awards a bonus to itself each time it brings a truck driver home. The amount awarded depends on the type of driver.
- The most critical driver types are solo company drivers, and solo independent contractors.
- Current company practice gets solo company drivers home 2 times per month, and independent contractors 1.7 times per month, on average.
- If we tune these so that the average number of time at home events are correct for these two driver types, then the other outputs also tend to match.

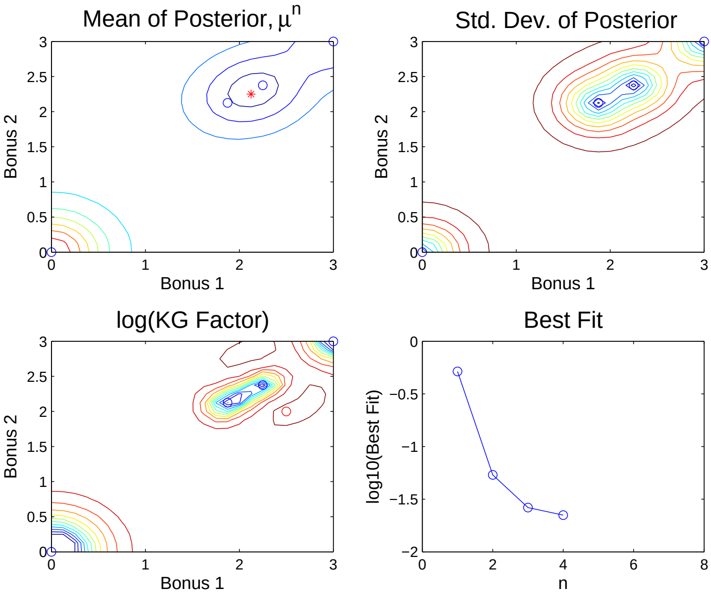
Simulation Model Calibration Results



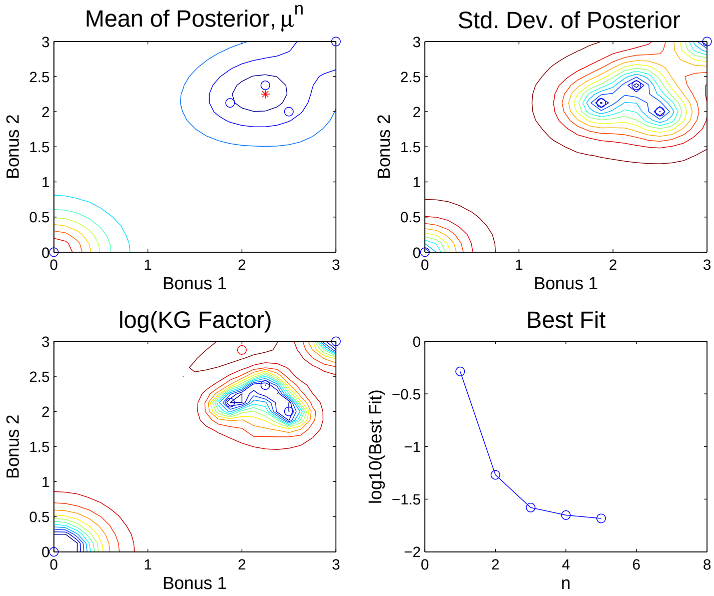
Simulation Model Calibration Results



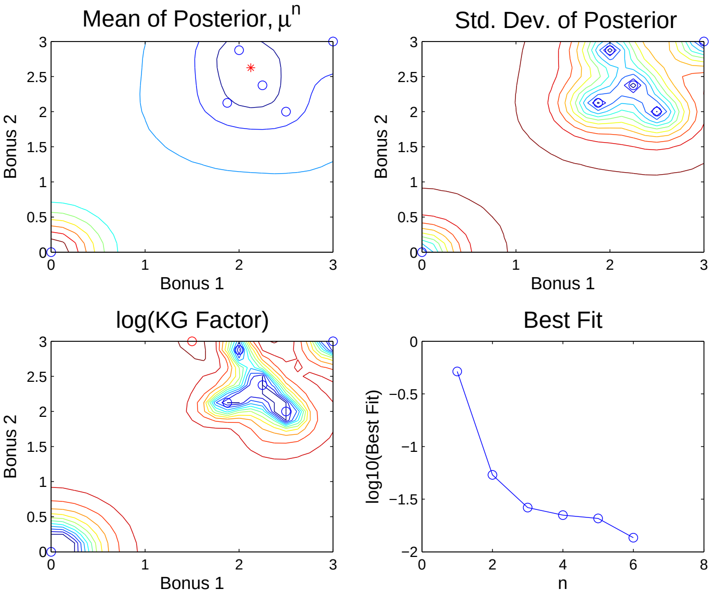
Simulation Model Calibration Results



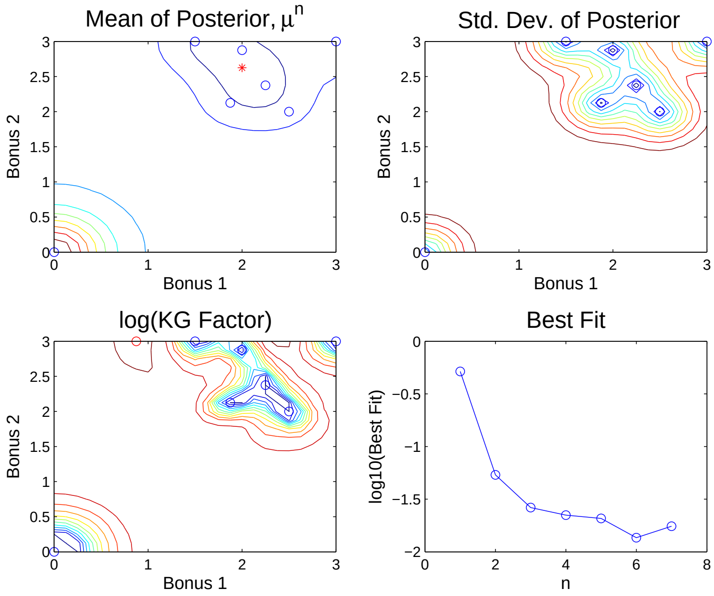
Simulation Model Calibration Results



Simulation Model Calibration Results



Simulation Model Calibration Results



Simulation Model Calibration Results

- The KG method calibrates the model in approximately 3 days, compared to $7 - 14$ days when tuned by hand.
- The calibration is automatic, freeing the human calibrator to do other work.
- The KG method calibrates as accurately or better than does by-hand calibration.
- In this approach, we thought about a problem of multi-information source optimization in a broader decision-theoretic framework. This framework can be applied to other multi-information source optimization problems.

References

Frazier, P., Powell, W., and Simão, H. (2009).
Simulation model calibration with correlated knowledge-gradients.
In *Winter Simulation Conference Proceedings, 2009*. Winter Simulation Conference.

Howard, R. (1966).
Information Value Theory.
Systems Science and Cybernetics, IEEE Transactions on, 2(1):22–26.

Jones, D., Schonlau, M., and Welch, W. (1998).
Efficient Global Optimization of Expensive Black-Box Functions.
Journal of Global Optimization, 13(4):455–492.

Mockus, J. (1989).
Bayesian approach to global optimization: theory and applications.
Kluwer Academic, Dordrecht.

Thank You

Any questions?

Model Parameters

- Input parameters to the model include:
  - time-at-home bonuses.
  - “pacing” parameters describing how fast and far drivers drive per day.
  - gas prices
  - ...
- Output parameters from the model include:
  - billed miles
  - driver utilization
  - average number of trips home per driver per 4 weeks.
  - proportion of drivers without time at home over 4 weeks.
  - ...
- Some of these inputs are known (e.g., gas prices), but some are unknown (e.g. time-at-home bonuses).
- Goal: adjust the inputs to make the optimal solution found by the model match current practice.