Model-Assisted Pattern Search[1]

Christopher Siefert
Department of Computer Science
University of Illinois, Urbana-Champaign

Virginia Torczon
Department of Computer Science
College of William & Mary

Michael Trosset
Department of Mathematics
College of William & Mary

[1]With thanks to the National Science Foundation, ICASE and the National Aeronautics and Space Administration, the William & Mary Endowment Association, the Batten Scholarship Fund, and the Exxon–Mobil Corporation.

Outline:

- the optimization problem
- the MAPS strategy
- an illustrative example
- the MAPS algorithm
- an engineering control problem
- conclusions
- references

MAPS currently solves problems of the form:

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & \ell \leq x \leq u, \end{array}$$

where $f : \mathbf{R}^n \rightarrow \mathbf{R}$, $\ell, x, u \in \mathbf{R}^n$, and $\ell < u$.

If we assume that f is continuously differentiable, then we can guarantee convergence of an easily identified subsequence to a Karush-Kuhn-Tucker (KKT) point. [Lewis/Torczon, SIOPT, 1999; Lewis/Torczon, SIOPT, 2000]

Operating Assumptions:

- Evaluation of the objective f depends on the output of computer simulations that may involve large numbers of state or system variables.

  The objective f is *not* a simple algebraic expression exclusively in terms of the variables x.

- The objective f may be expensive to compute because the simulations upon which its value is based are expensive to compute.

- The objective f may have “interesting” features (numerical and/or analytical).

The MAPS approach is applicable to general nonlinear optimization problems, but problems with the above features were the motivation for the development of MAPS.

The primary goals of the MAPS strategy:

- Use approximations $\hat{f}_k$ of the objective f to accelerate pattern search.
- Use search criteria S_k to encourage a wider exploration of the feasible region.
- Choose both $\hat{f}_k$ and S_k in a way that preserves the theoretical guarantees of pattern search provided by the large body of analysis and takes advantage of our growing computational experience.

References for pattern search:

1991 Virginia Torczon
On the convergence of the multidirectional search algorithm
SIAM Journal on Optimization.

1997 Virginia Torczon
On the convergence of pattern search algorithms
SIAM Journal on Optimization.

1996 Robert Michael Lewis and Virginia Torczon
Rank ordering and positive bases in pattern search algorithms
ICASE Technical Report 96–71.

1998 Robert Michael Lewis, Virginia Torczon, and Michael W. Trosset
Why pattern search works
Optima.

1999 Robert Michael Lewis and Virginia Torczon
Pattern search methods for bound constrained minimization
SIAM Journal on Optimization.

1999 Elizabeth D. Dolan
Pattern Search Behavior in Nonlinear Optimization
Honors Thesis, The College of William & Mary.

2000 Robert Michael Lewis and Virginia Torczon
Pattern search methods for linearly constrained minimization
SIAM Journal on Optimization.

2000 Robert Michael Lewis and Virginia Torczon
A globally convergent augmented Lagrangian pattern search algorithm for optimization with general constraints and simple bounds
Revised for *SIAM Journal on Optimization*.

2000 Elizabeth D. Dolan, Robert Michael Lewis, and Virginia Torczon
On the local convergence of pattern search
Submitted to *SIAM Journal on Optimization*.

2000 Adam P. Gurson
Simplex Search Behavior in Nonlinear Optimization
Honors Thesis, The College of William & Mary.

The basic MAPS strategy:

- Use an approximation $\hat{f}$, created by kriging known values of the objective $f(x)$, to predict minimizers of the objective f.

  The goal is to create an approximation that:

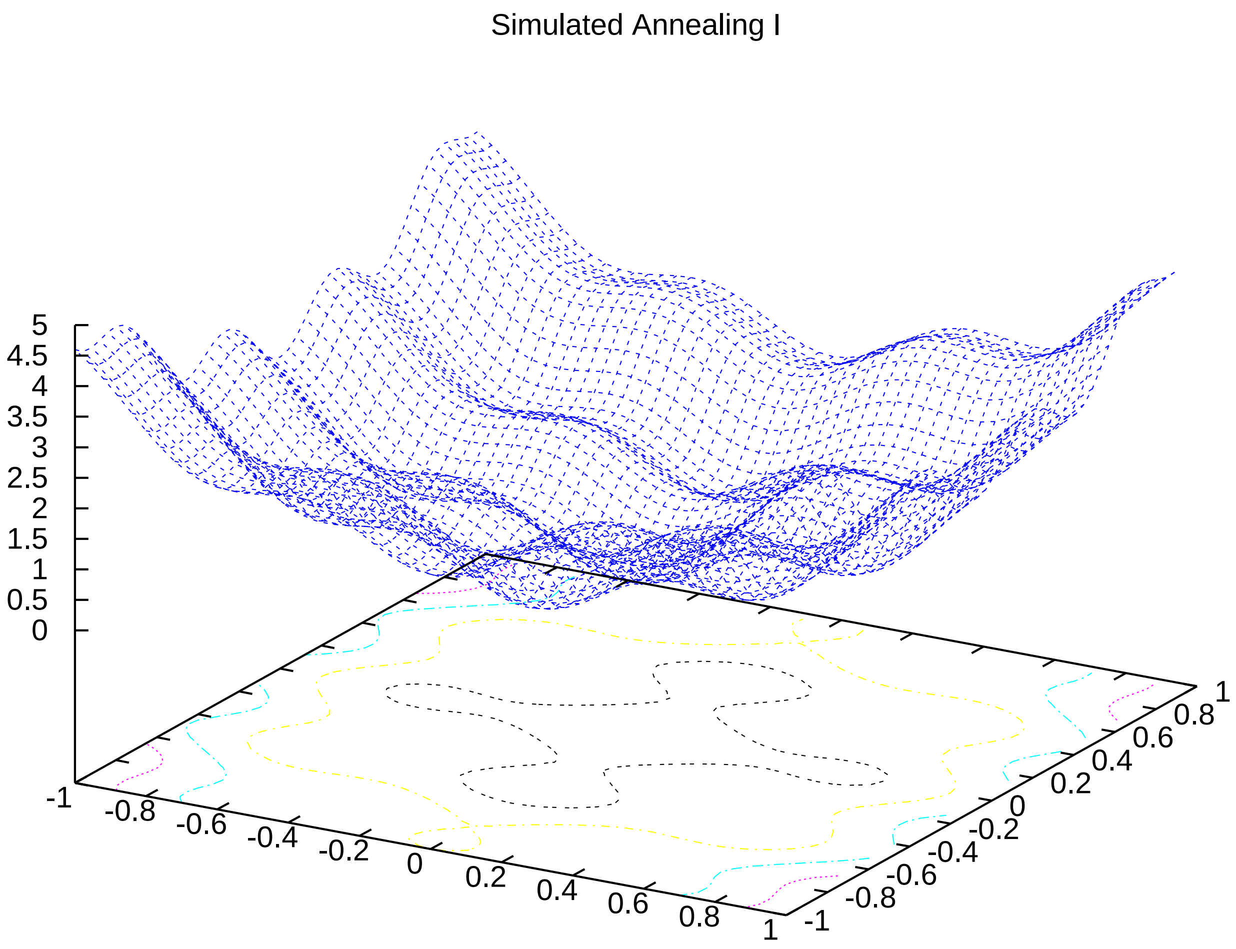
  - makes use of the known values of $f(x)$ since it is assumed that such values are computationally expensive and
  - is inexpensive to evaluate so that finding minimizers of the approximation $\hat{f}$ is computationally tractable.

- Use a search criterion S, created by taking a weighted sum of the approximation $\hat{f}$ and some measure of the quality of the approximation to determine where next to evaluate the objective f.

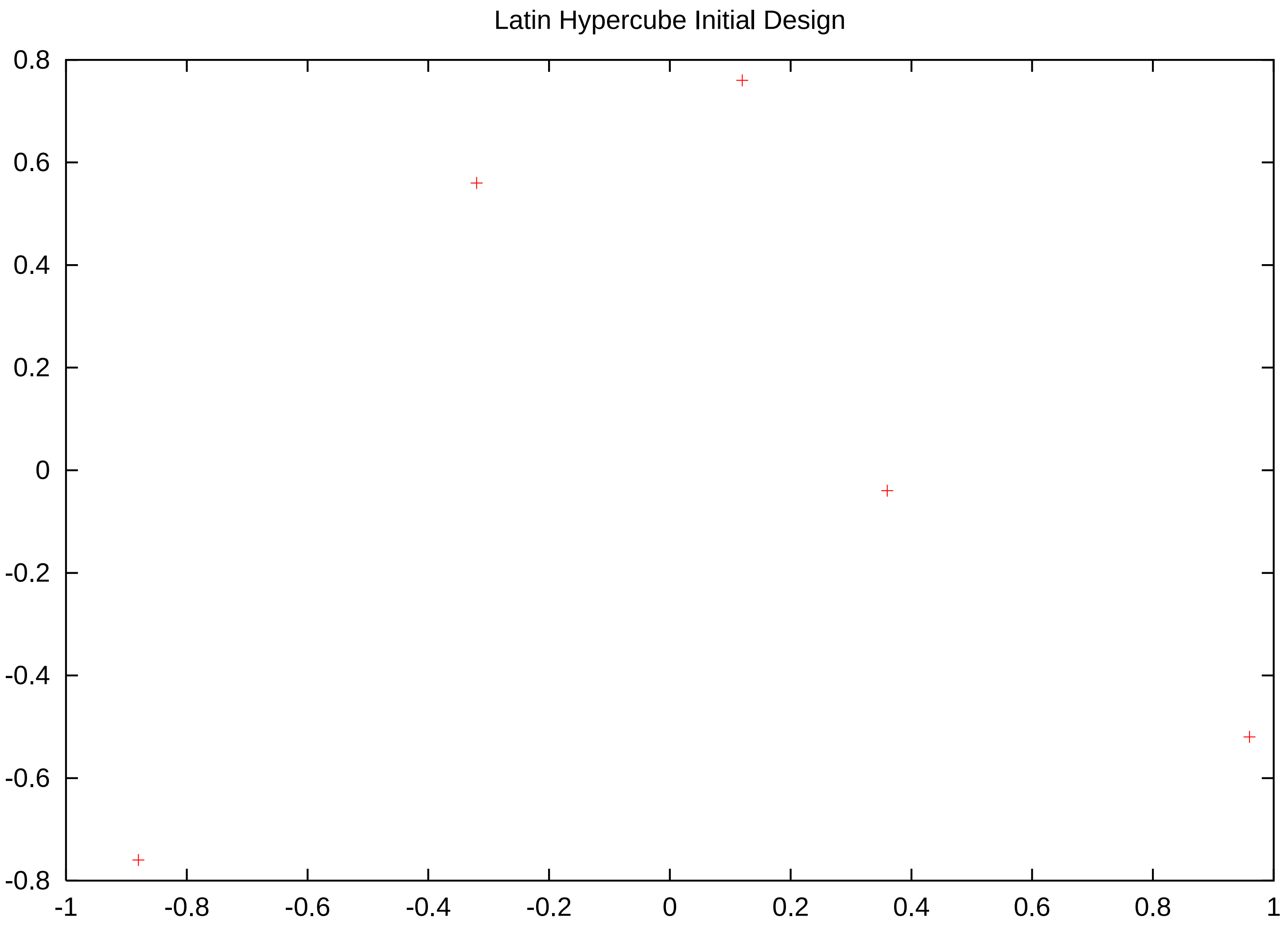
  The goal of the search criterion is to trade-off between:

  - a single-minded pursuit of a (constrained) stationary point and
  - a wider exploration of the feasible region to look for a more promising solution.

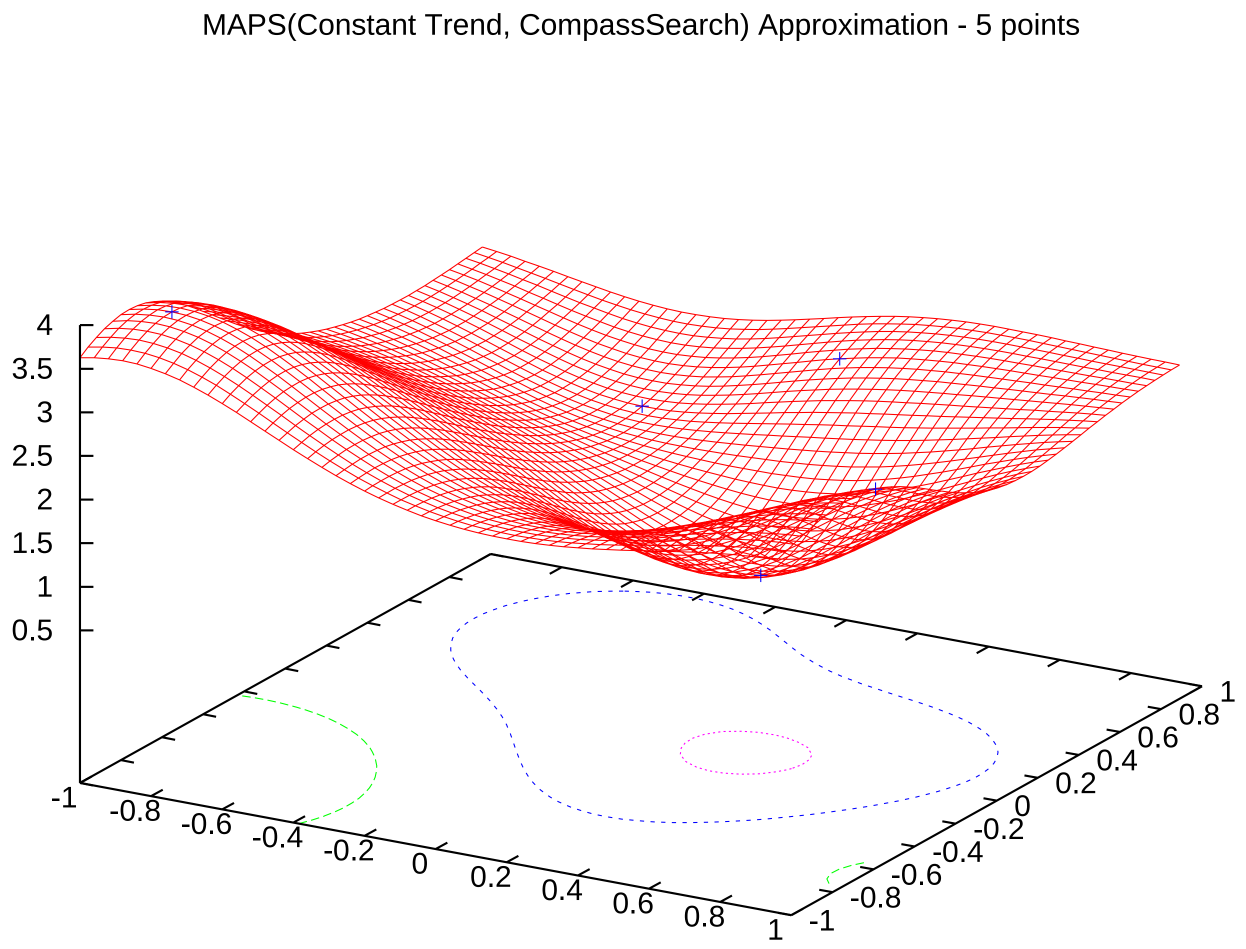
The objective f graphed over the feasible region



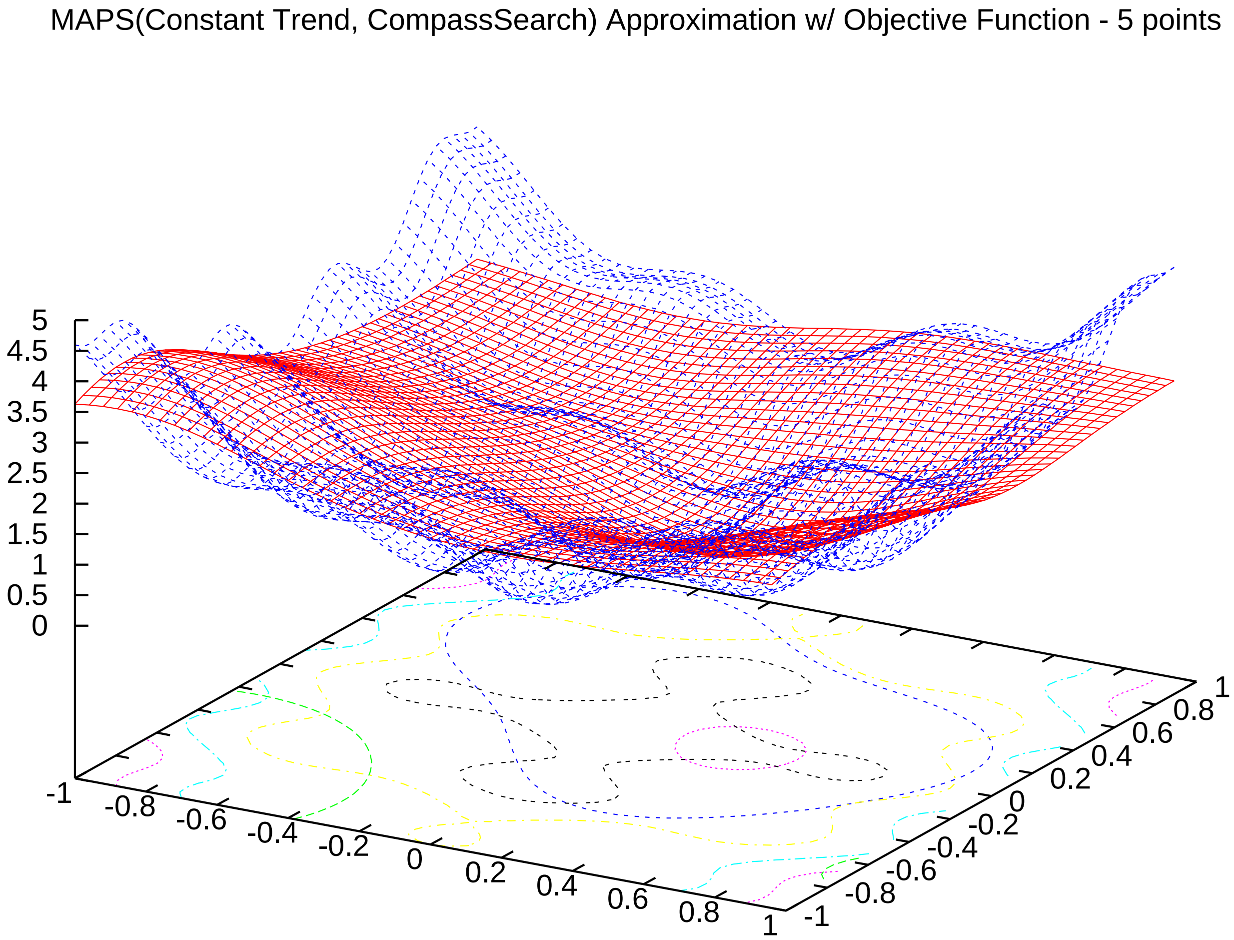
The initial design sites $x^1, \ldots, x^5$, selected from the feasible region



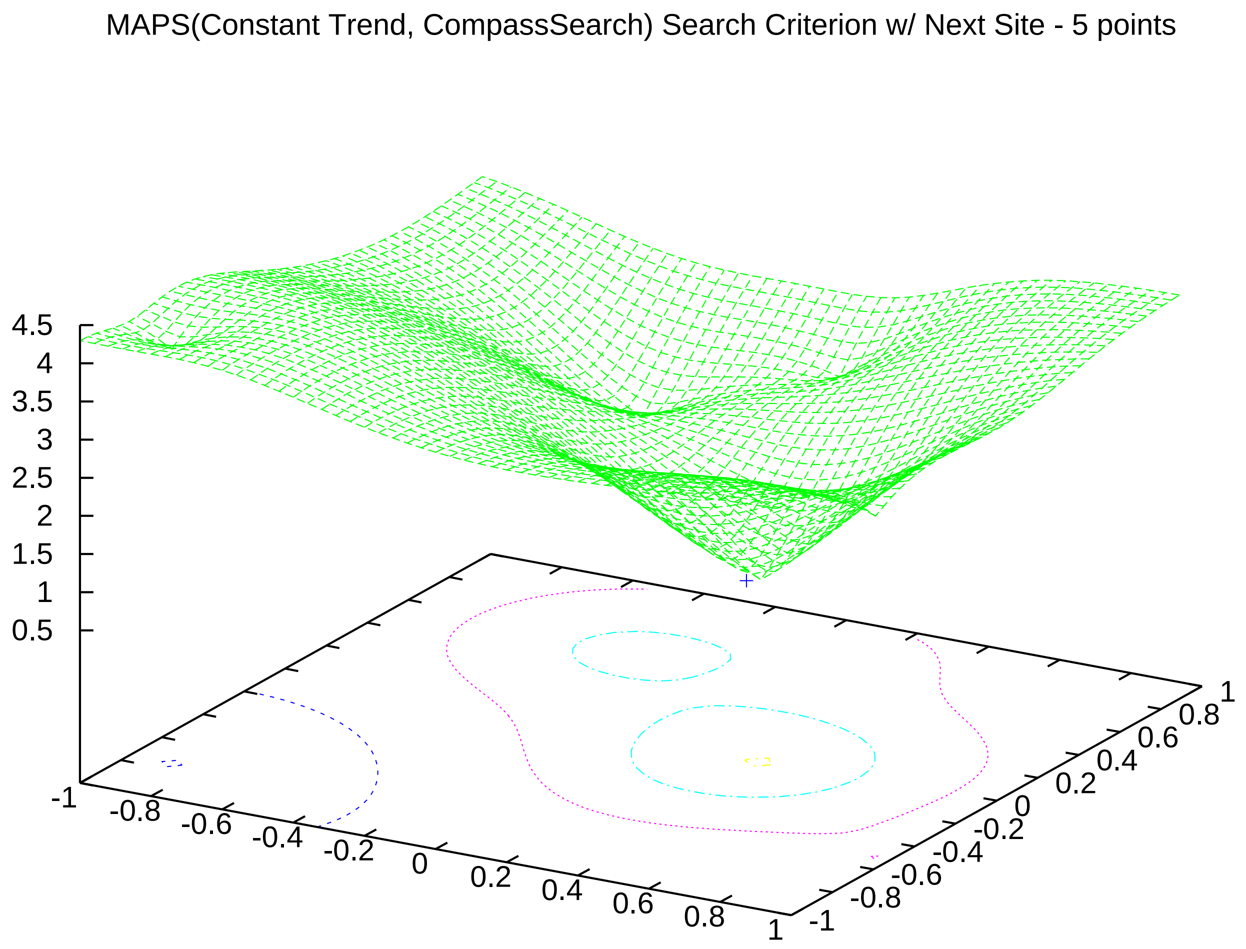
The initial approximation $\hat{f}_0$ graphed over the feasible region



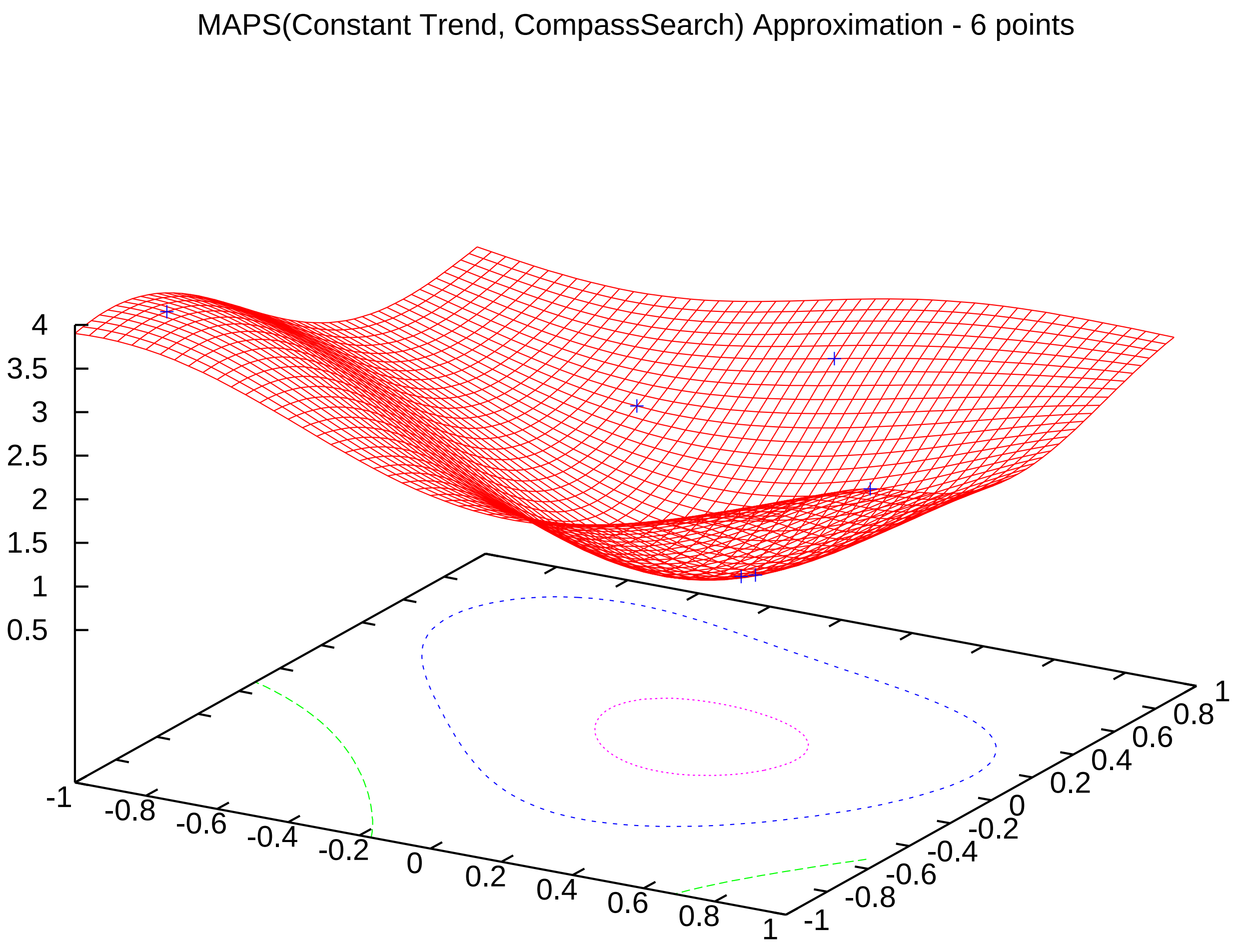
The objective f and the initial approximation $\hat{f}_0$



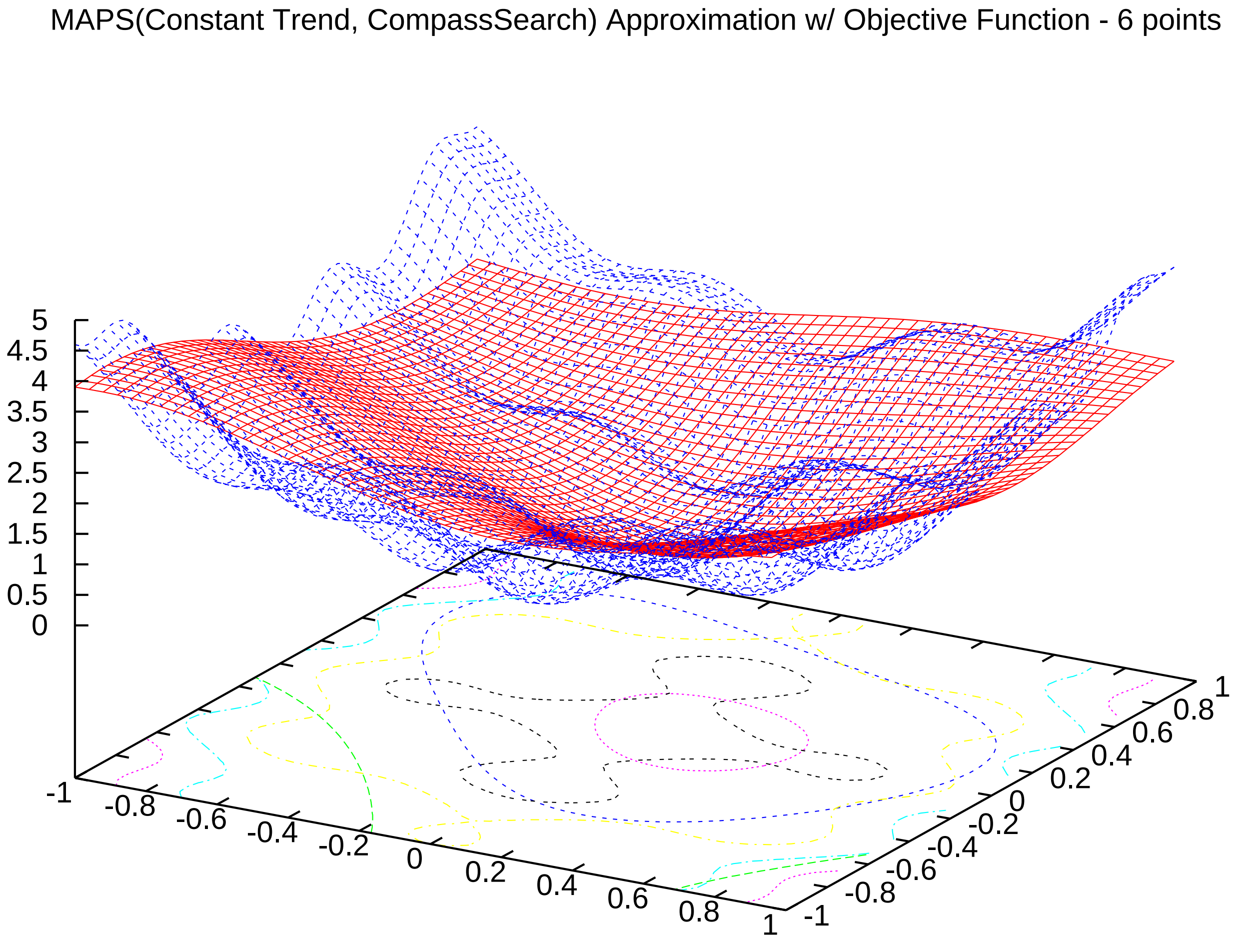
The initial search criterion S_0 graphed over the feasible region



The first update of the approximation $\hat{f}_1$

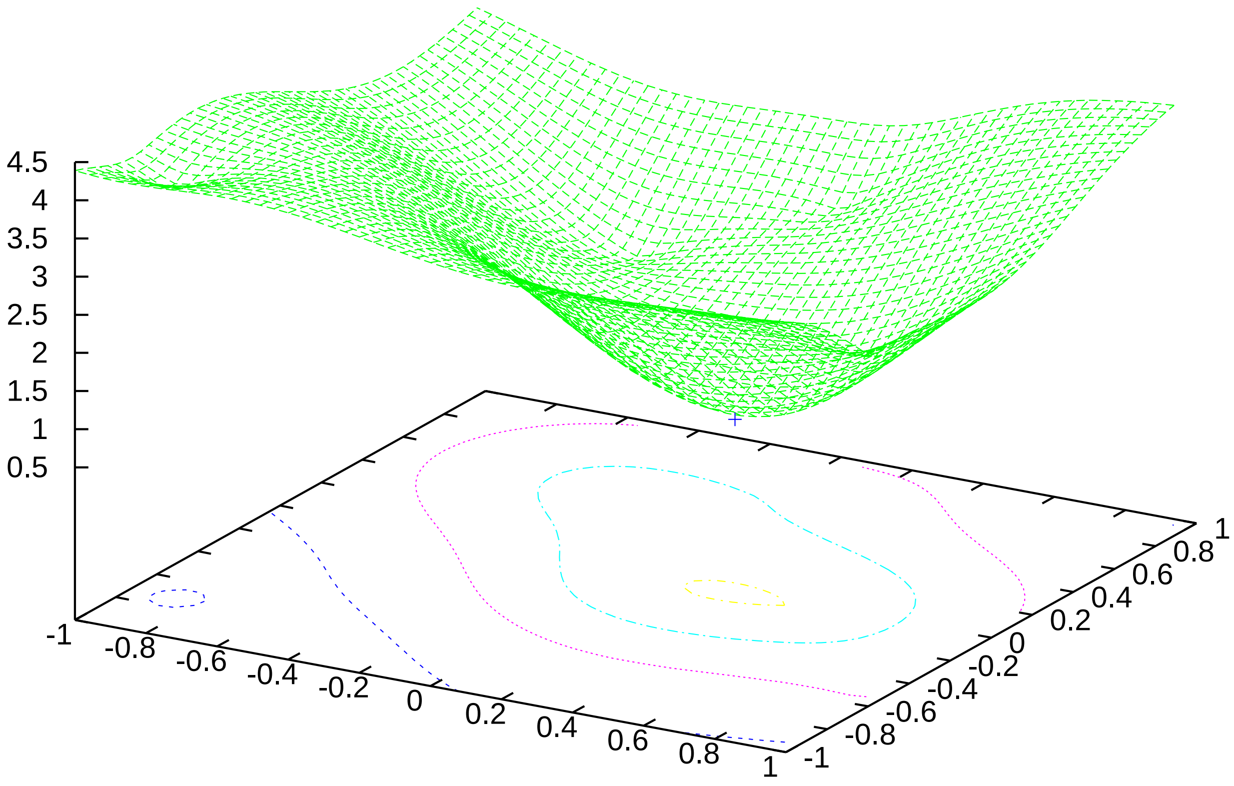


The objective f and the first update of the approximation $\hat{f}_1$

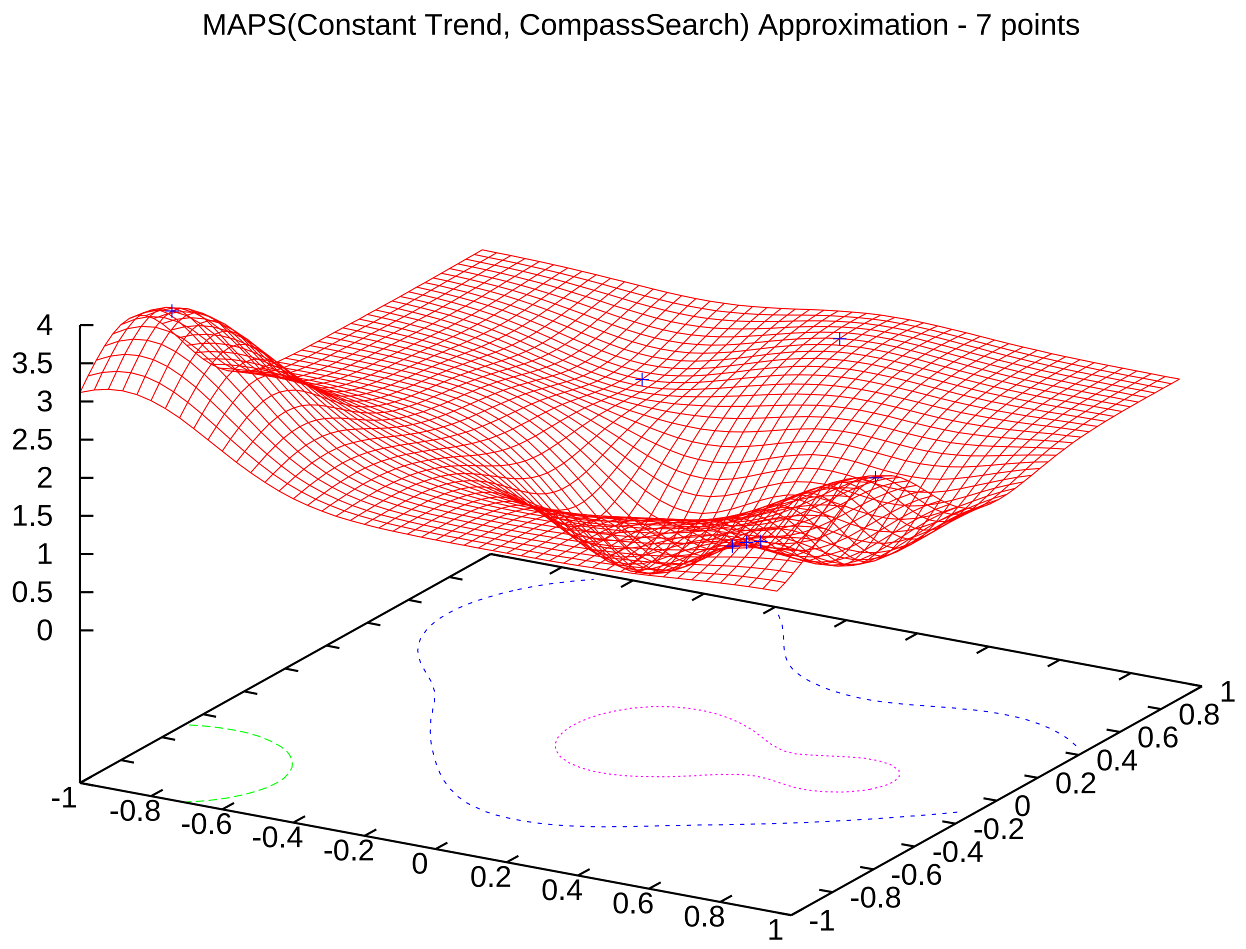


The first update of the search criterion S_1

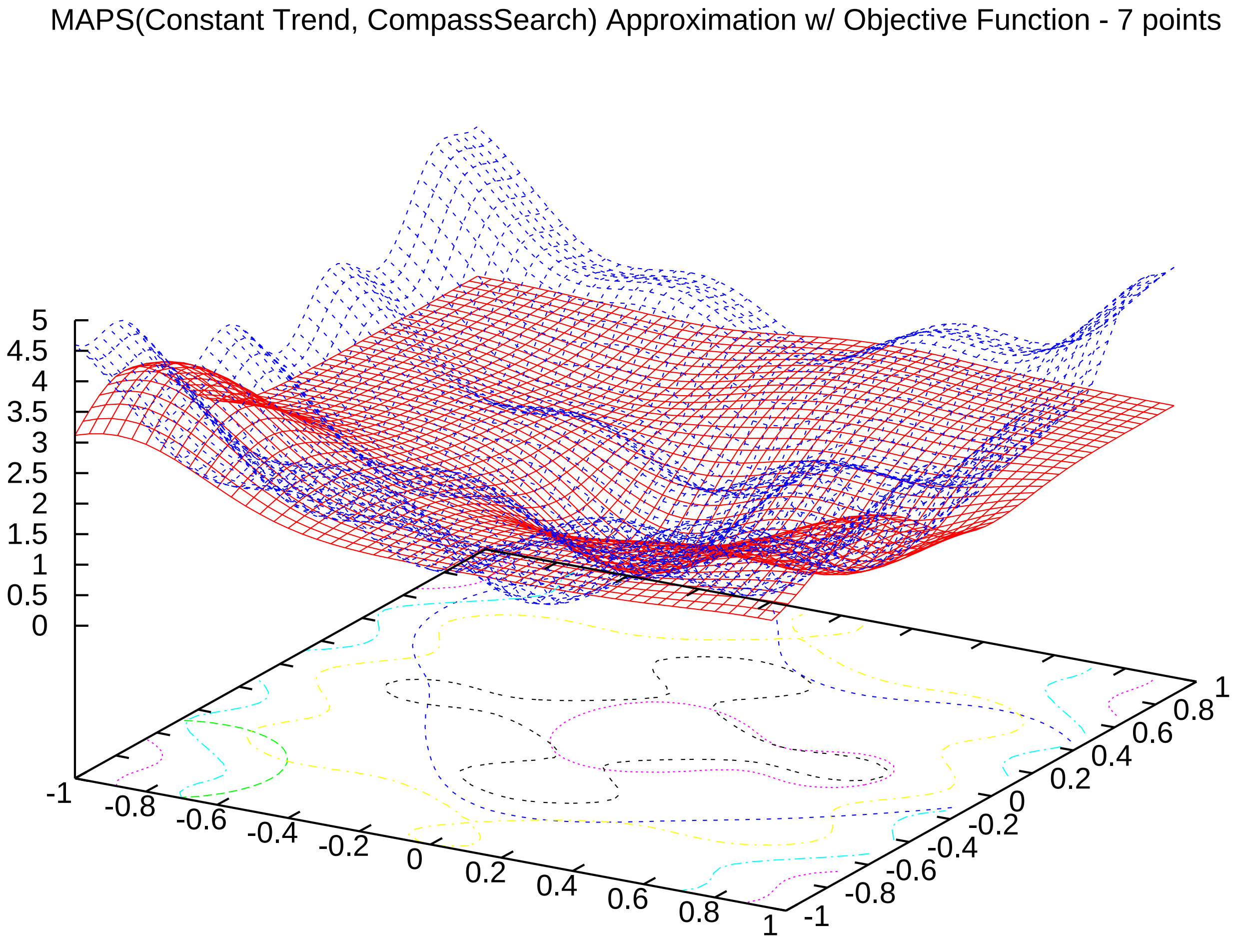
MAPS(Constant Trend, CompassSearch) Search Criterion w/ Next Site - 6 points



The second update of the approximation $\hat{f}_2$



The objective f and the second update of the approximation $\hat{f}_2$



MAPS: Model-Assisted Pattern Search:

1. Set $k = 0$.
   Specify the family of functions from which $\{\hat{f}_k\}$ will be selected.
   Specify an initial grid G_0 that conforms to the feasible region $\ell \leq x \leq u$.

2. Perform an initial computer experiment:

   (a) Select initial design sites $x^1, \ldots, x^N \subset G_0$.
   (b) Compute $f(x^1), \ldots, f(x^N)$.
   (c) Construct $\hat{f}_0$ by kriging.
   (d) Specify S_0.

3. Set $x_0 = \text{argmin}\ \{f(x^1), \ldots, f(x^N)\}$.
   Set $\text{Eval}_0 = \{x^1, \ldots, x^N\}$.

4. Terminate if any of the stopping criteria are satisfied. Otherwise:

   (a) Apply an optimization algorithm to S_k to obtain $x^t \in G_k \setminus \text{Eval}_k$.
   (b) Compute $f(x^t)$. Set $\text{Eval}_k = \text{Eval}_k \cup \{x^t\}$.
   (c) If $f(x^t) < f(x_k)$, then $x_k = x^t$.
       Else if $\text{Core}(x_k) \subset \text{Eval}_k$, then refine G_k.
   (d) Update $\hat{f}_k(x)$ and then $S_k(x)$.
   (e) Set $k = k + 1$.

An engineering control problem:

A model for low pressure chemical vapor deposition in a hot-wall tubular reactor*

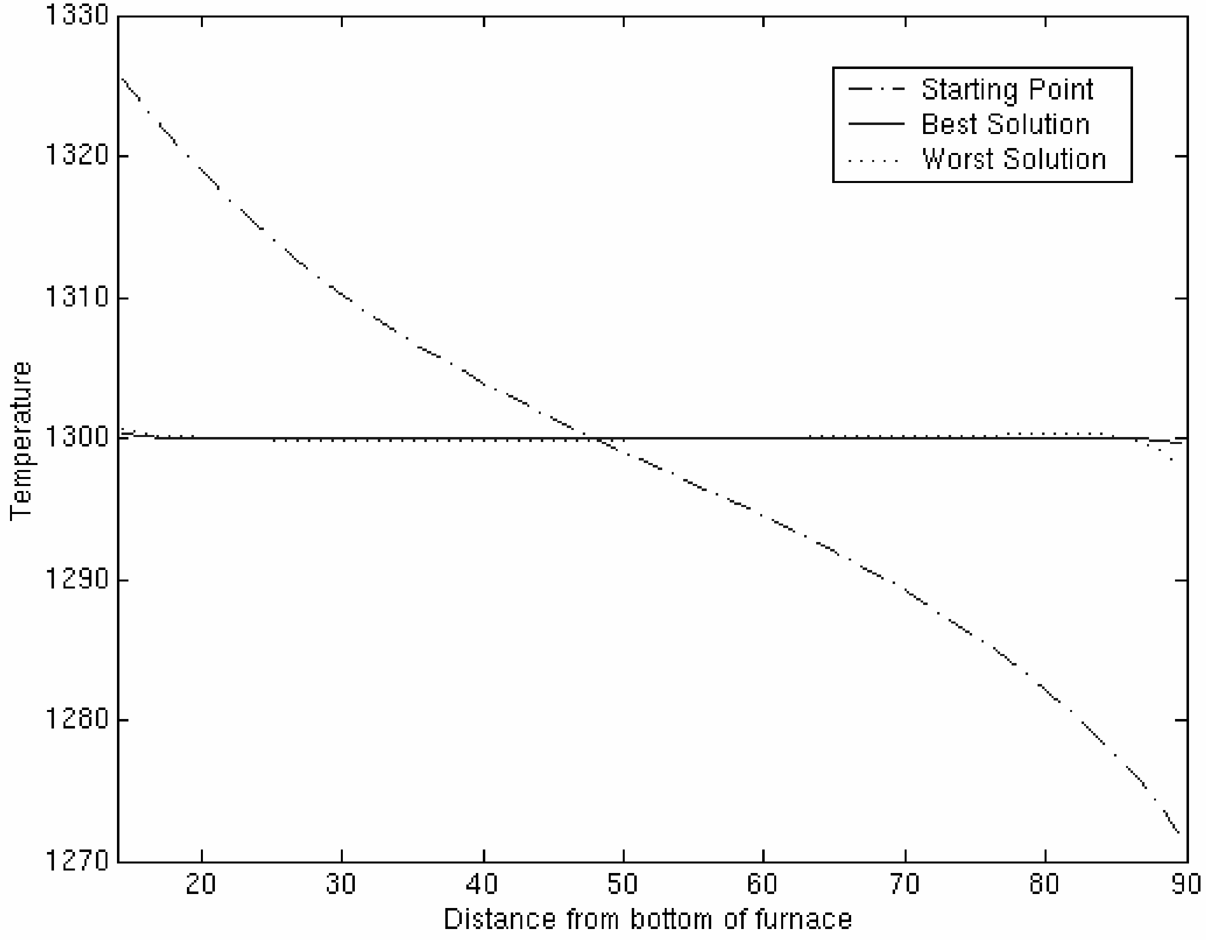
Materials Science and Engineering

W. G. Houf and J. F. Grcar
Sandia National Laboratories, Livermore

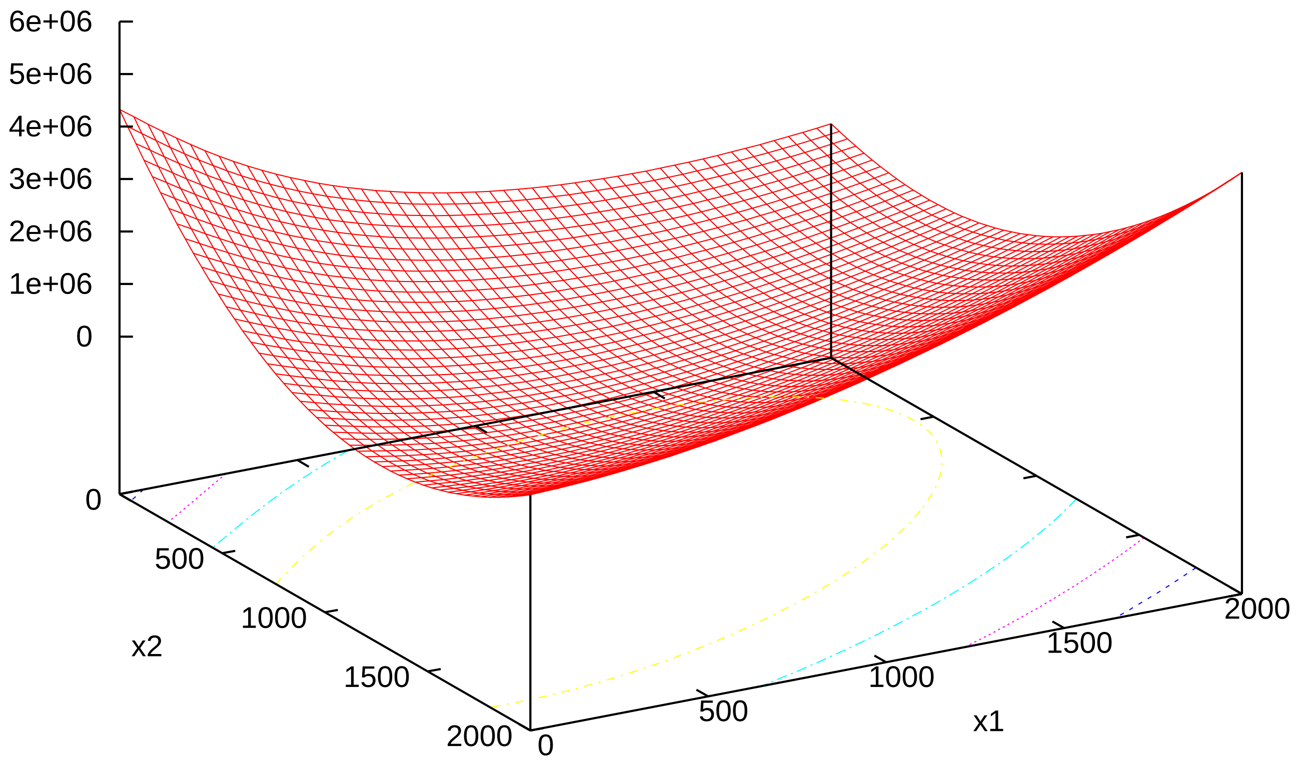
W. G. Breiland
Sandia National Laboratories, Albuquerque

*This work was supported by the U.S. Department of Energy and SEMATECH.

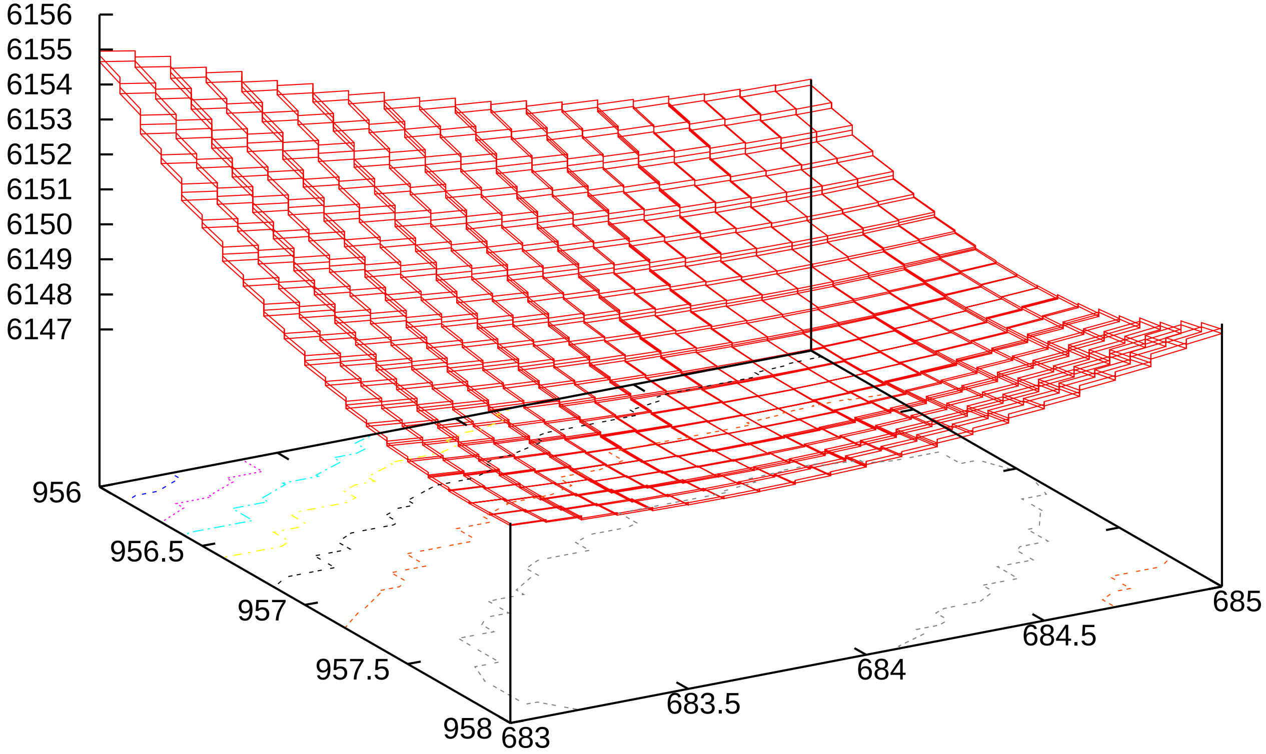
The goal: choose power settings that minimize the mismatch between the measured and the target temperature in the reactor



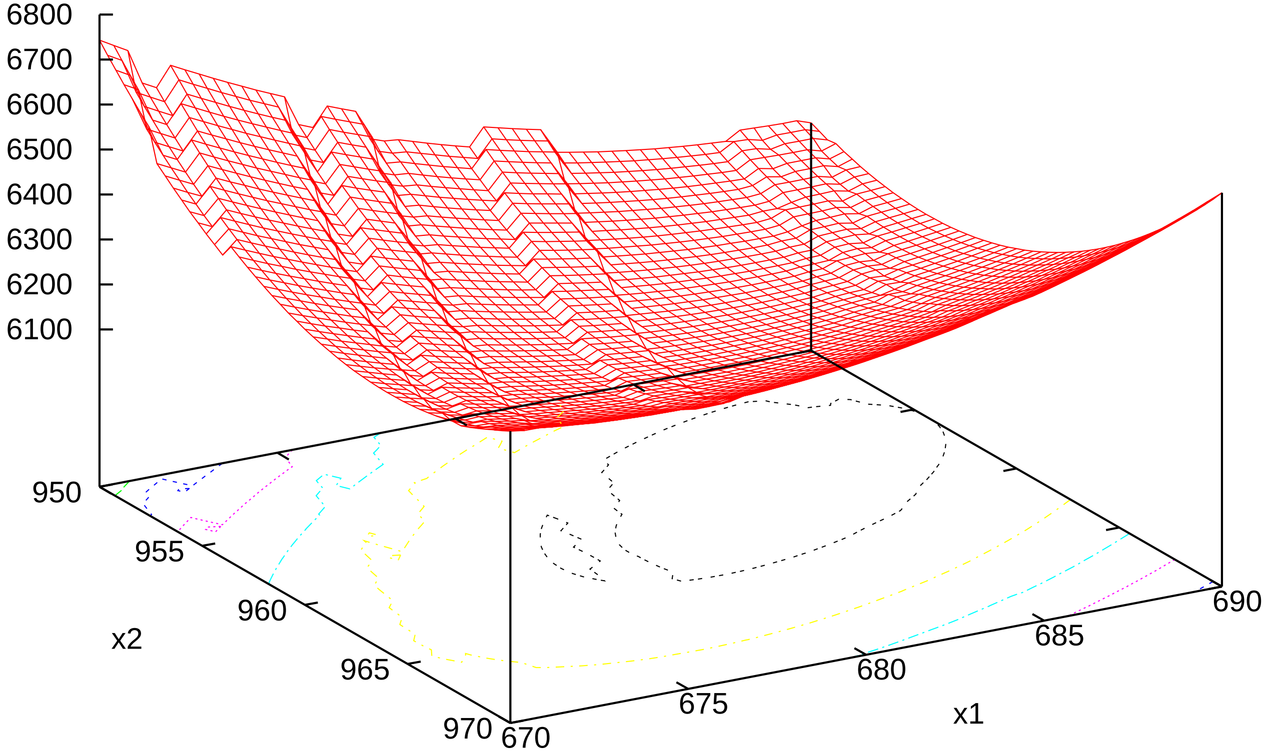
The objective f graphed over the feasible region



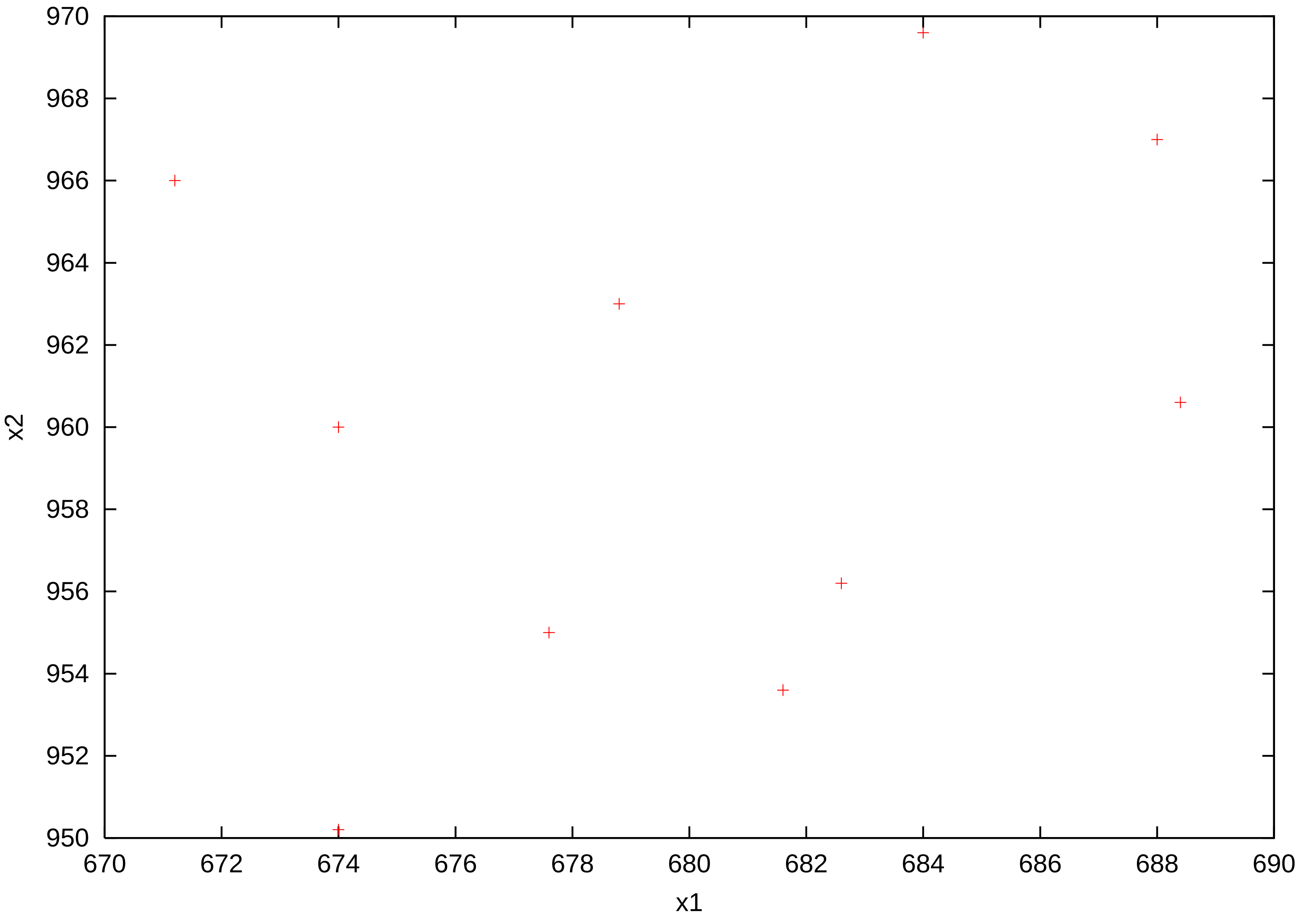
The objective f (with `tol = 10e-12`) graphed near the solution



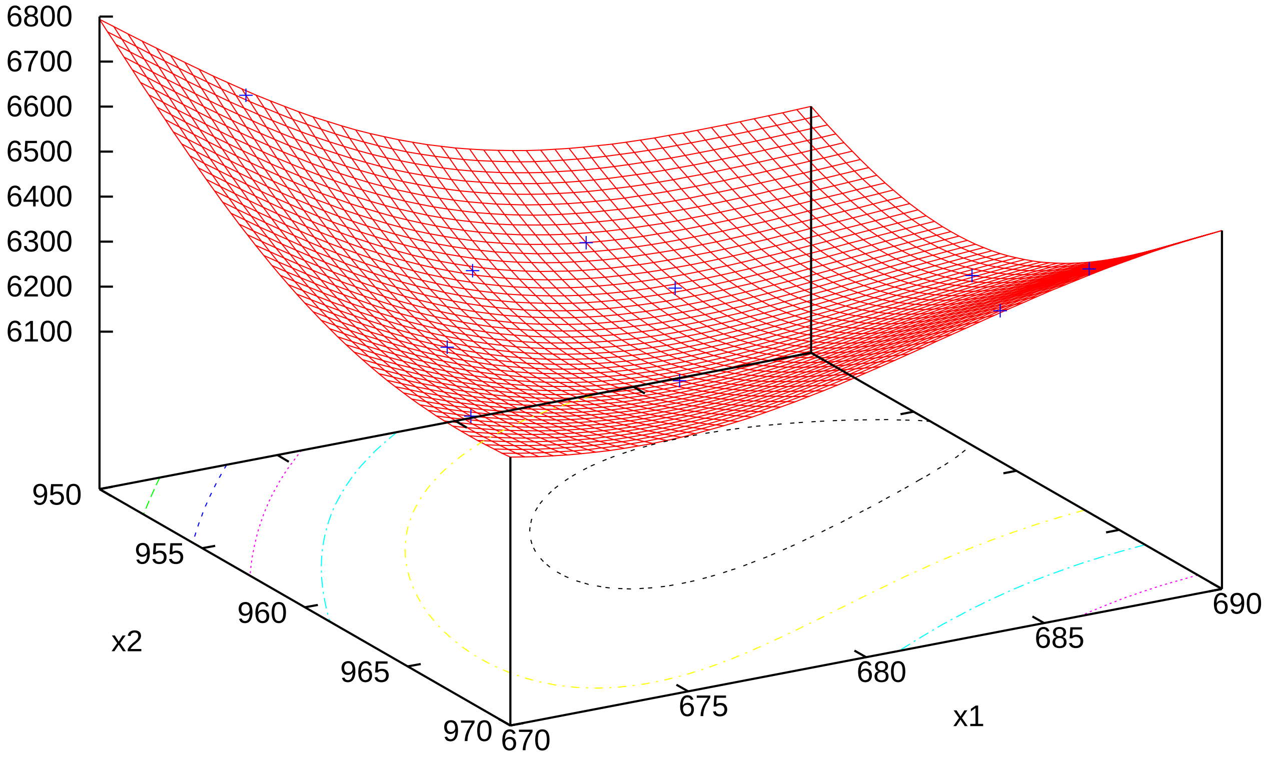
The objective f (with `tol= 10e-3`) graphed near the solution



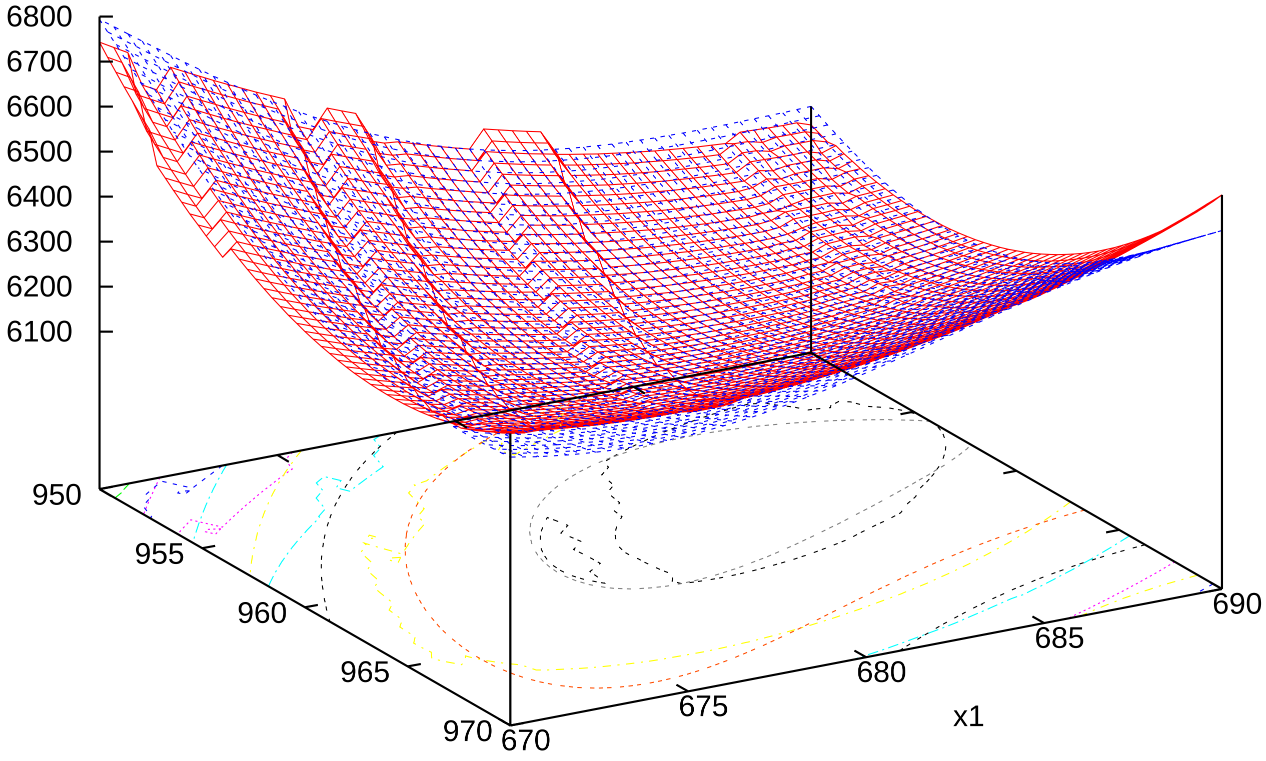
The initial design sites $x^1, \ldots, x^{10}$



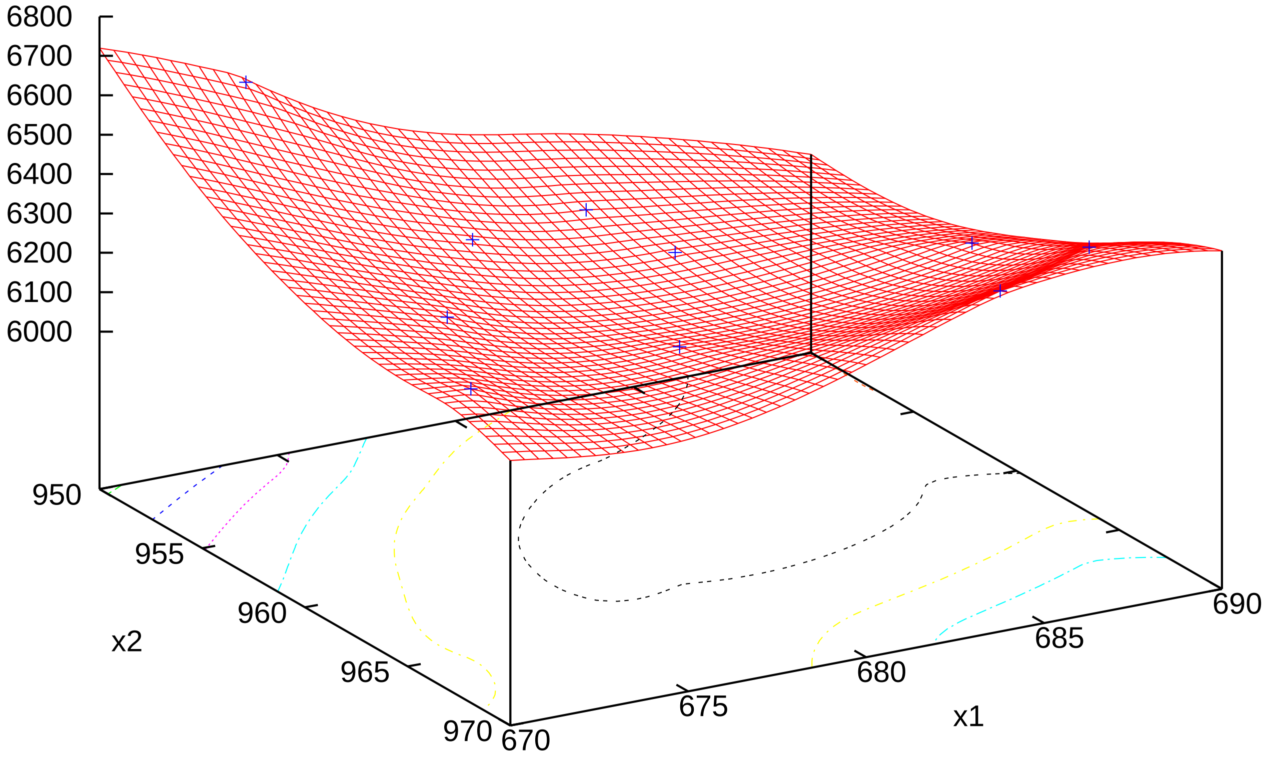
The initial approximation $\hat{f}_0$



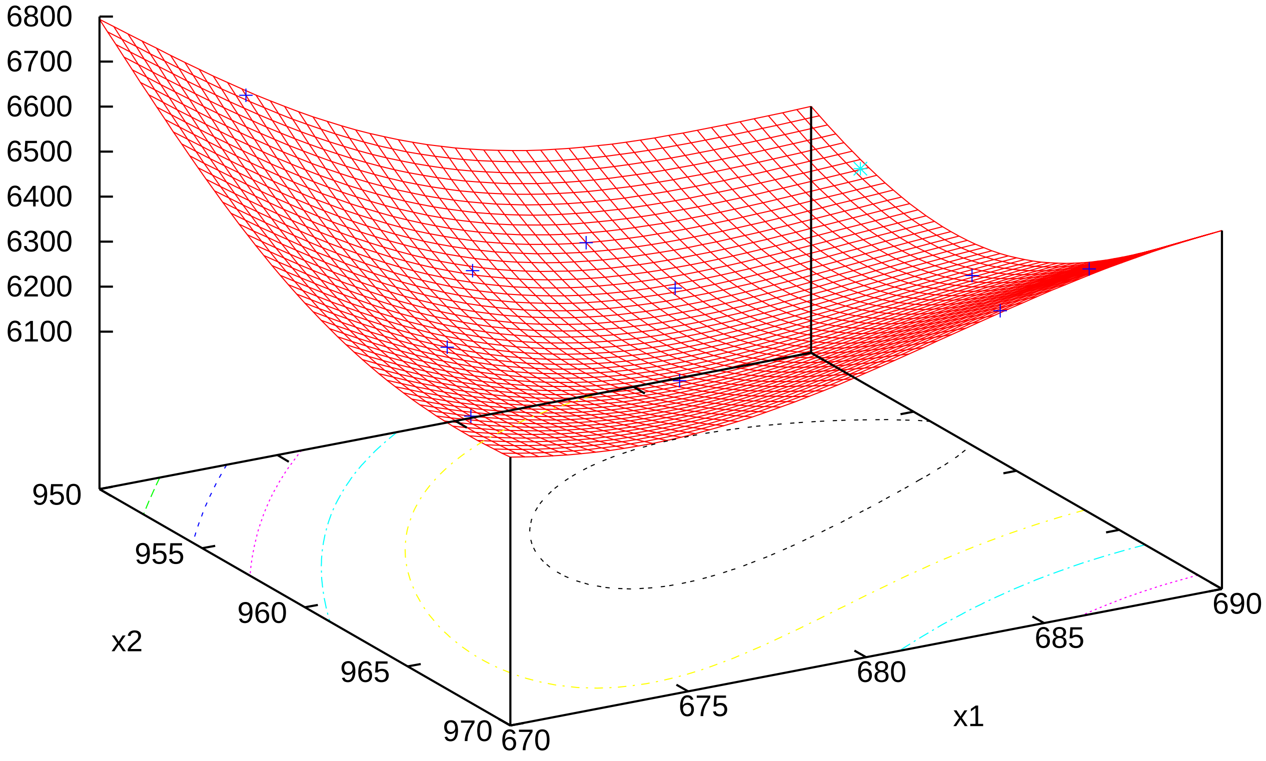
The objective f and the initial approximation $\hat{f}_0$



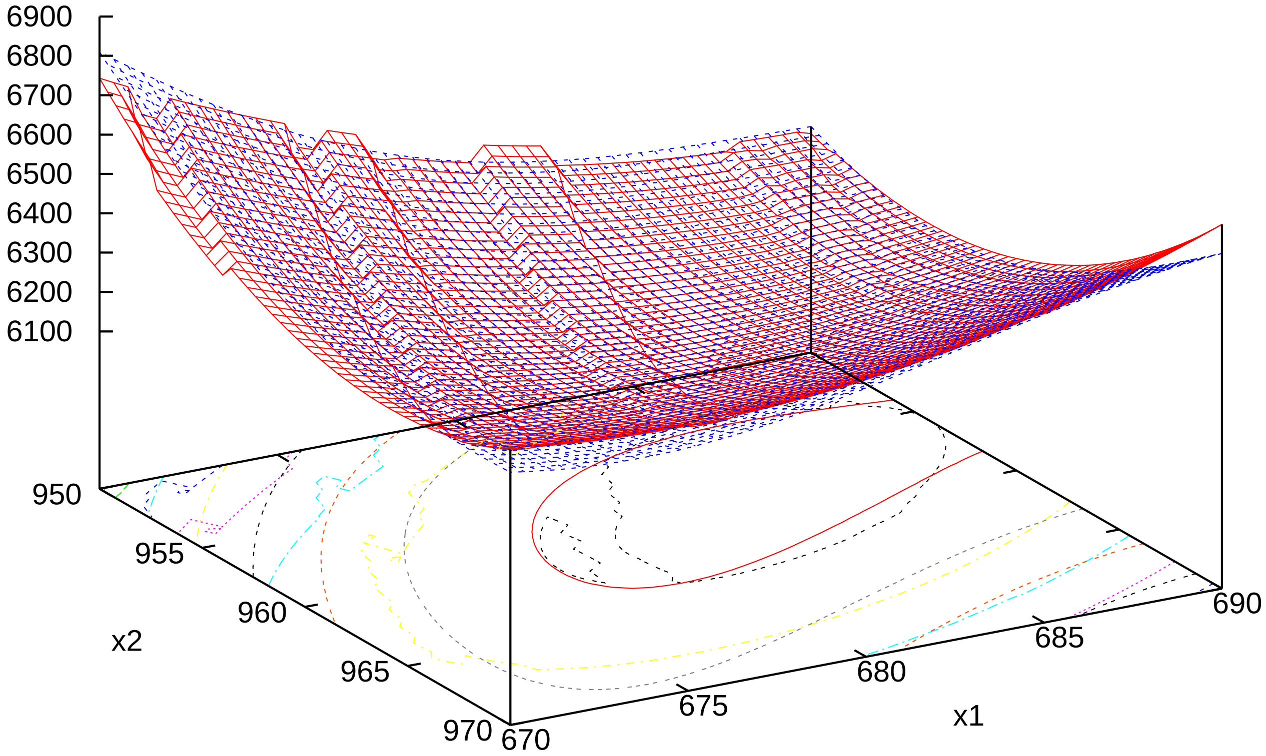
The initial search criterion S_0, with the original design sites shown



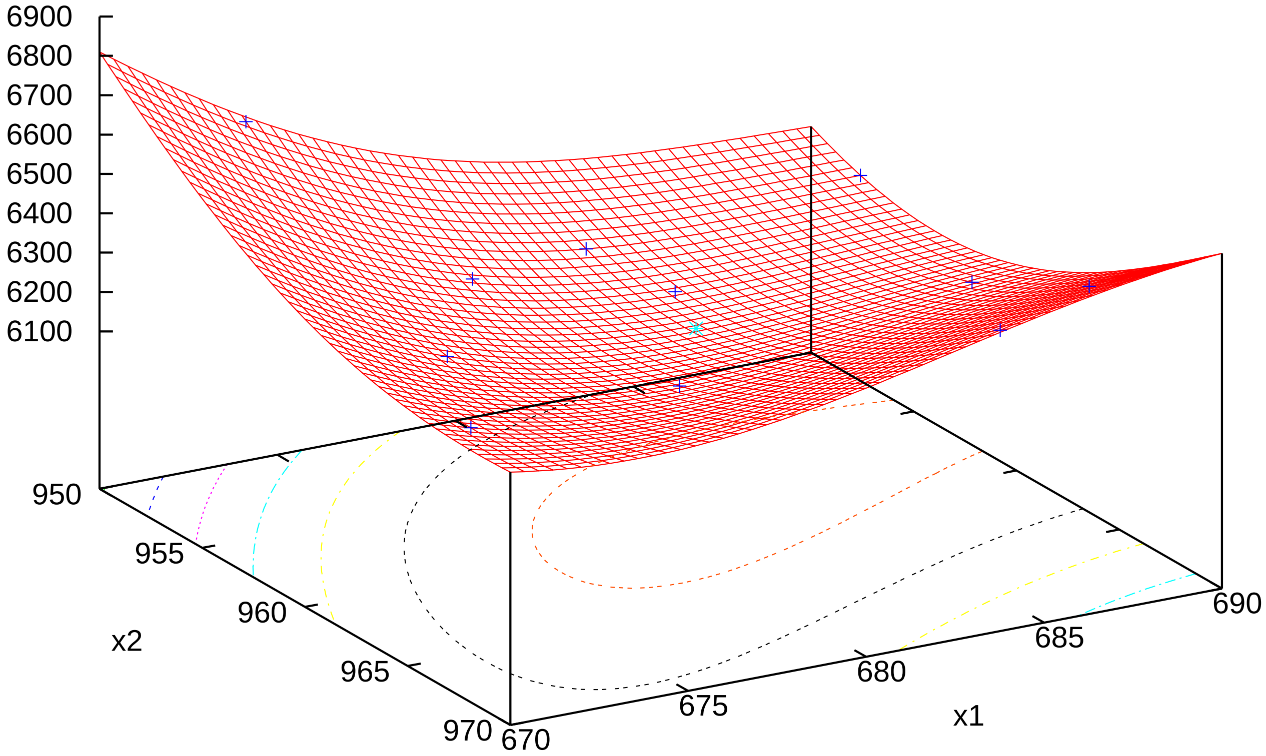
Approximation $\hat{f}_0$, with x_1 and original design sites shown



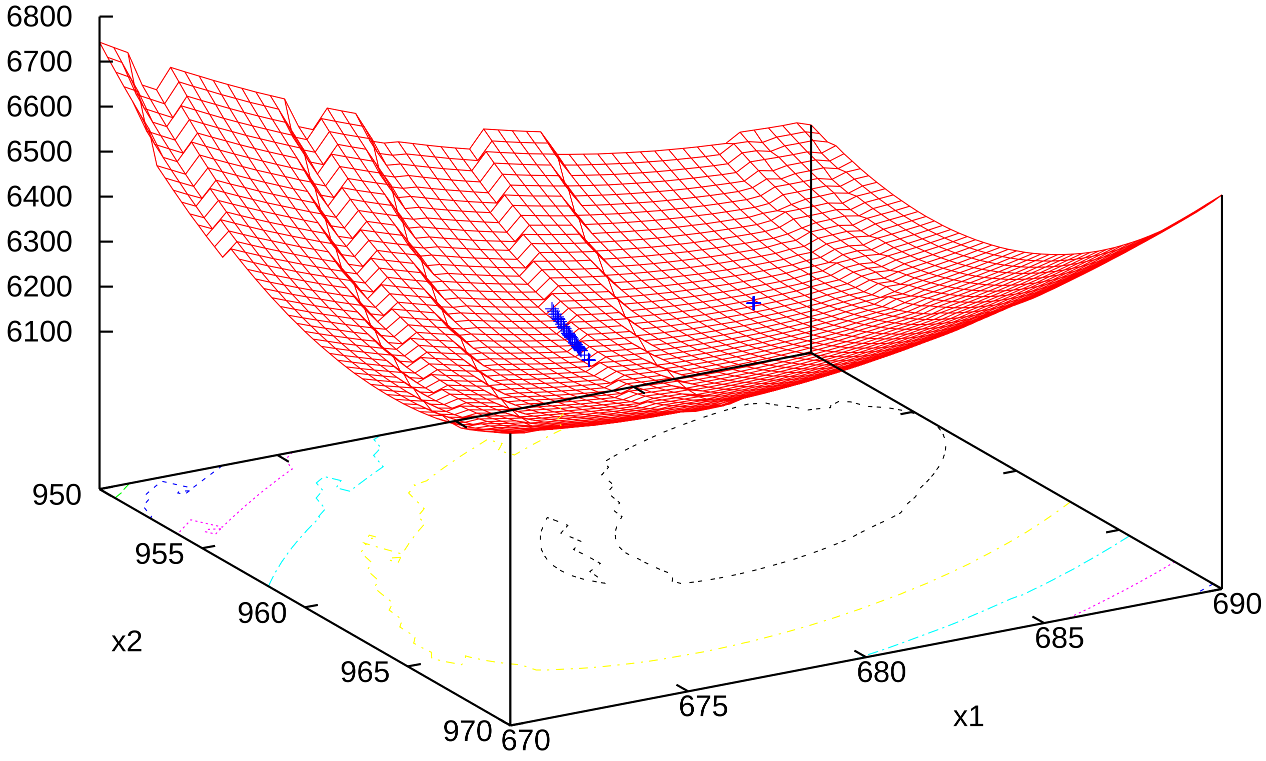
The objective f and the first update of the approximation $\hat{f}_1$



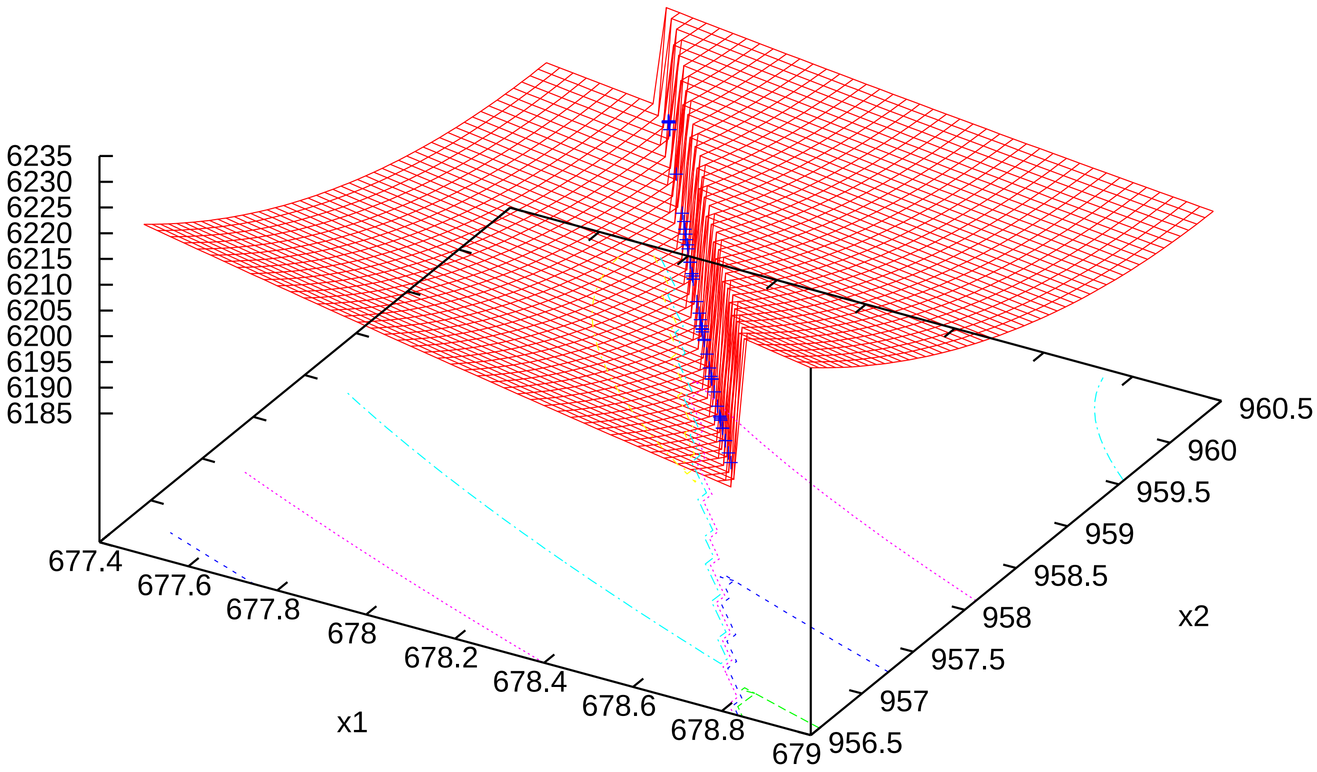
The first update of the search criterion S_1 and the choice of x_2



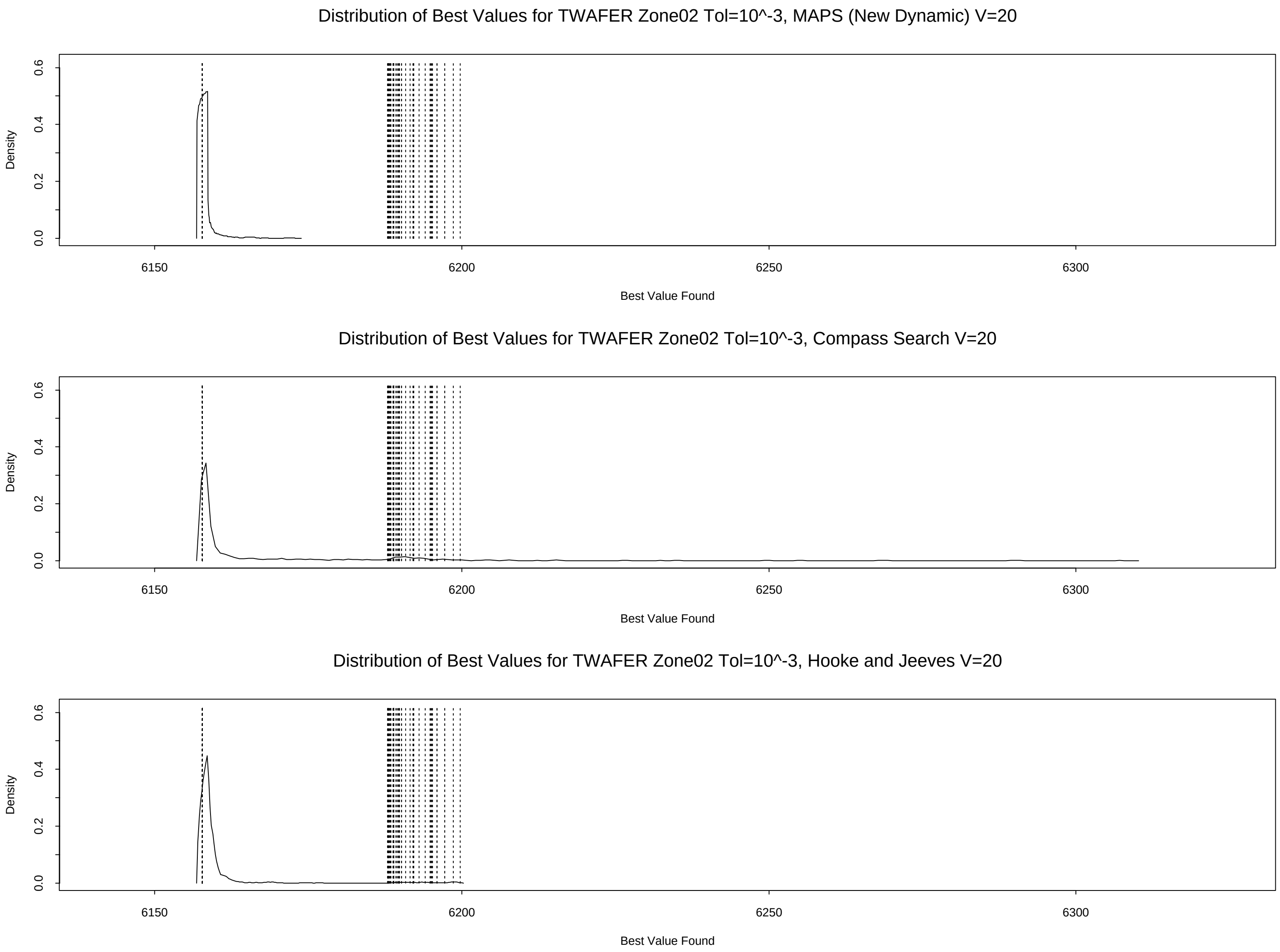
Compass search results for 1000 different sets of initial design sites



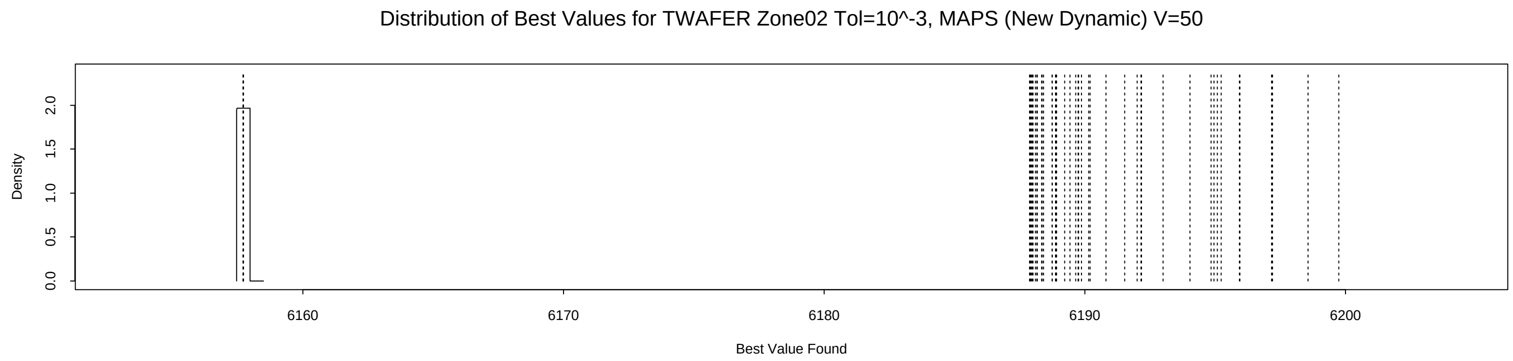
Solutions along the valley

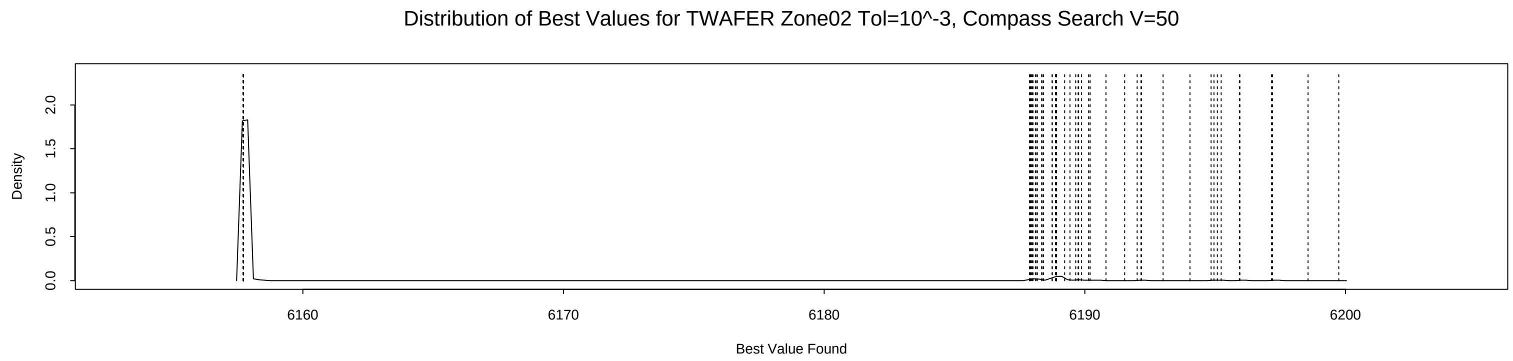


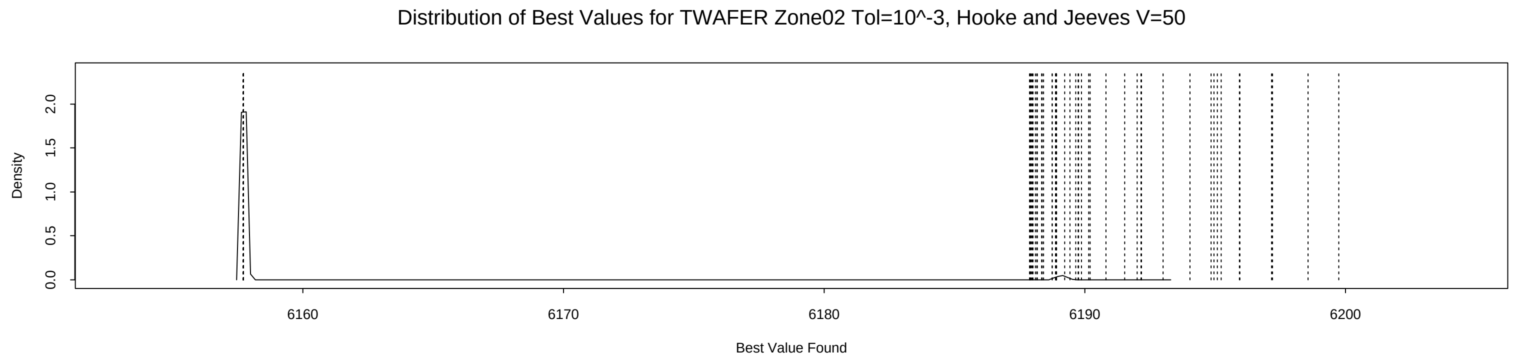
MAPS identifies the solution with only 20 evaluations of f



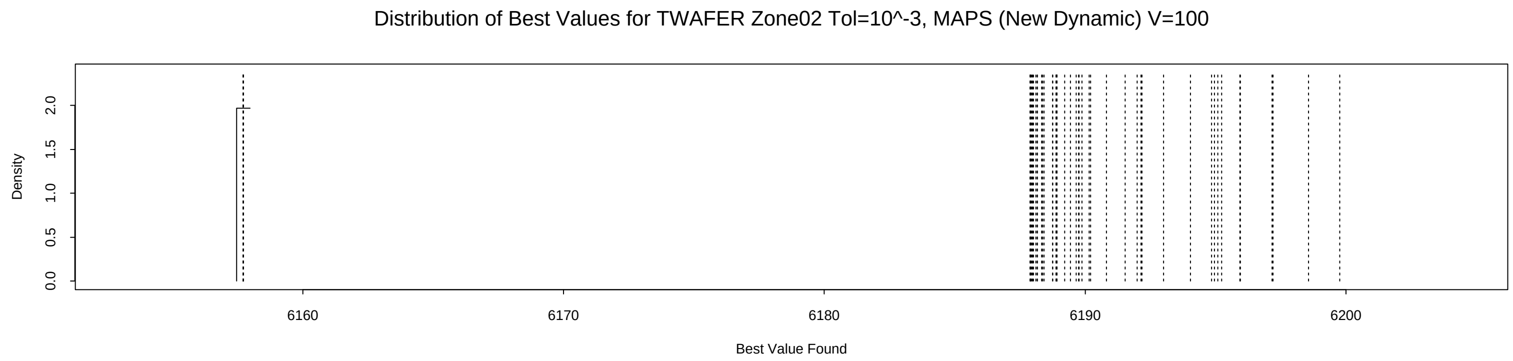
Pattern searches still trapped with 50 evaluations of f

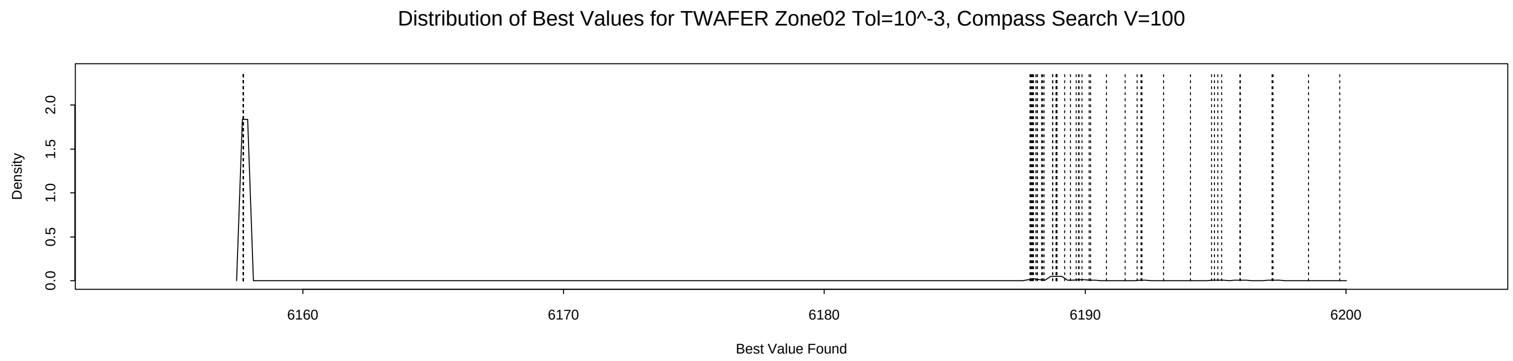


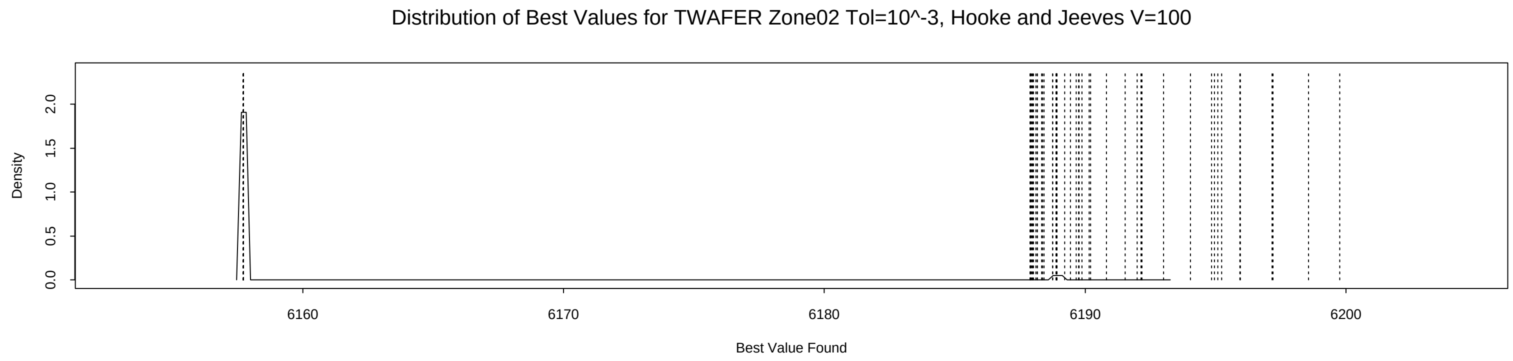




No improvement for pattern search even with 100 evaluations of f







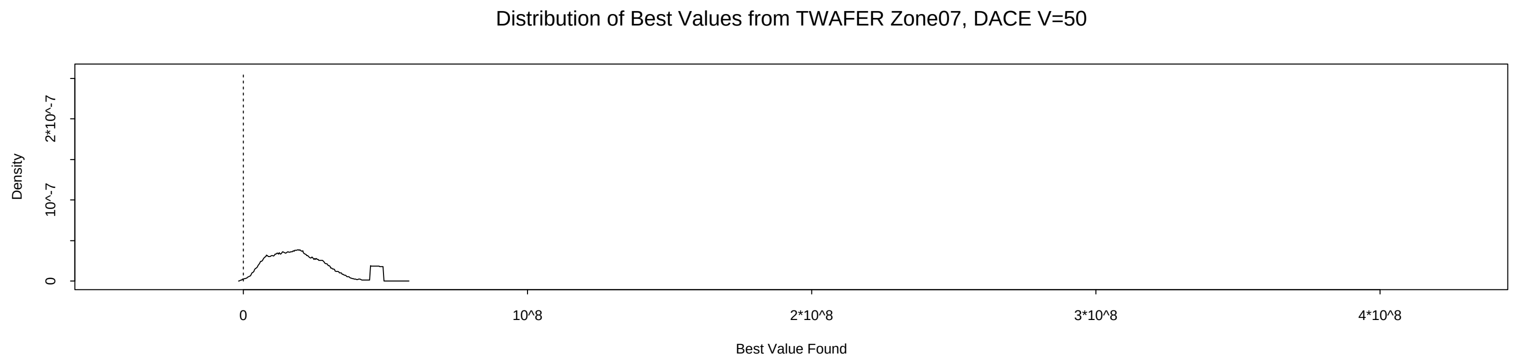
MAPS also works well with the more realistic 7-zone problem

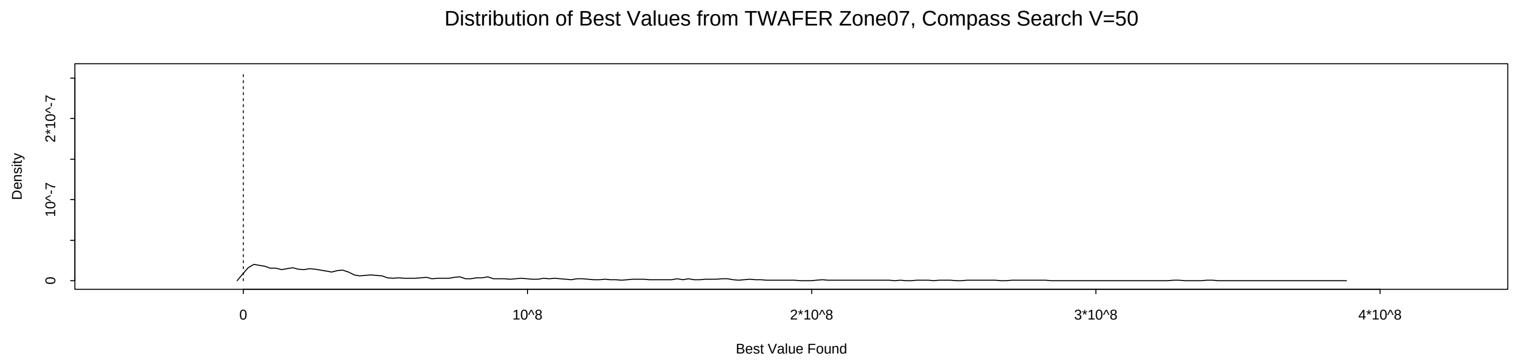
Distribution of Best Values for TWAFER Zone07, MAPS (Dynamic) V=50

Distribution of Best Values for TWAFER Zone07, Original MAGS V=50

Distribution of Best Values for TWAFER Zone07, MAPS-like-SDO V=50

Compass search and DACE are not as successful





Conclusions:

- We have now run literally tens of thousands of experiments using MAPS on a variety of low-dimensional problems and it consistently identifies stationary points of the objective with a modest number of evaluations.

- There is still much interesting work to be done.

References for managing approximation models:

1997 J. E. Dennis and Virginia Torczon
Managing approximation models in optimization
Multidisciplinary Design Optimization: State of the Art,
edited by Natalia M. Alexandrov and M. Y. Hussaini.

1997 Michael W. Trosset and Virginia Torczon
Numerical optimization using computer experiments
ICASE Technical Report 97–38.

1998 Virginia Torczon and Michael W. Trosset
Using approximations to accelerate engineering design optimization
Proceedings of the 7th AIAA/USAF/NASA/ISSMO Symposium on Multidisciplinary Analysis and Optimization.

1998 Michael W. Trosset
Optimization on a limited budget
Proceedings of the Section on Physical and Engineering Sciences, American Statistical Association.

1998 Andrew J. Booker, J. E. Dennis, Jr., Paul D. Frank, David B. Serafini, and Virginia Torczon
Optimization using surrogate objectives on a helicopter test example
Structural Optimization.

1999 Andrew J. Booker, J. E. Dennis, Jr., Paul D. Frank, David B. Serafini, Virginia Torczon, and Michael W. Trosset
A rigorous framework for optimization of expensive functions by surrogates
Structural Optimization.

2000 Christopher M. Siefert
Model-Assisted Pattern Search
Honors Thesis, The College of William & Mary.

Web Resources:

For PostScript files of related papers and the MAPS software, visit:

`http://www.cs.wm.edu/~va`

and click on either:

- `Research publications` or
- `software`.