Raytheon

Raytheon Company
1151 E. Hermans Road, Bldg 805, M/S B3
Tucson, AZ 85756-9367 USA

Integrating Approximate Models into National Security Simulation Response Surface Analysis

Frederick A. Ahrens
80th MORS Symposium
12-14 June 2012

Abstract

Integration of an approximate model into the response surface analysis (RSA) of national security simulations can result in better-fitting surrogate models with fewer coefficients.

RSA is used to characterize the responses of simulations to multiple variables. It is particularly useful with lean designs of experiments (DOE) that do not evaluate all possible combinations of the variables. RSA can be used to developed fast-running surrogate models of simulations enabling dynamic "dashboard-like" presentations of results with the capability to explore multivariable trade spaces and multiple figures of merit. Surrogate models can also serve as objective functions in multi-objective optimization problems.

Simulations that support national security operations research often have highly nonlinear responses, causing undesirable behavior in the response surface estimates. Classical RSA with first order and polynomial terms was intended for local representations to identify gradients, ridges, saddle points and local extrema. The demand for global response surface representations led to advanced methods such as nonlinear models, neural network approximation functions and spatial correlation models. However, even with advanced methods, model misspecification can lead to response surfaces estimators with poor generalization.

An approximate analytical model, tailored to the analysis problem, and derived from the first principles of the problem can improve both the fit and generalization of response surface estimates for national security simulations. A trend model derived from first principles is able to account for known or hypothesized nonlinearities and interactions between variables, while a polynomial trend model may require many terms to represent the same features. There are several ways to employ an approximate model in RSA, one of the simplest being to use its output as a term in any of the popular RSA methods.

In an example of a site defense against cruise missiles using the Extended Air Defense Simulation (EADSIM), the presentation will show how incorporating an approximate model improved the fit and generalization of four RSA methods. A site defense scenario was simulated in EADSIM using a DOE of 1,200 trials varying 12 parameters, including sensor ranges, weapon range, reaction times, probability of kill and firing doctrine, and measuring the number of threat leakers. The data analysis compared least squares, stochastic kriging and neural network response estimates with and without an approximate spreadsheet model of the site defense. Fit was measured as the root sum squared error over all data used in the response surface estimate, and generalization was measured using the leave-one-out cross-validation method. In this example, integrating the approximate model also tended to reduce the number of coefficients, simplifying the screening task.

The presentation will briefly cite experiences with this method in air defense, homeland security ports of entry configuration analysis, and orbital analysis using data from STK®. Using approximate models in simulation analysis has other advantages over the "black-box" approach by providing a basis for theory and hypothesis testing, providing verification cross-checks during simulation development, providing explanations of the causal threads in the simulation response and providing a starting point for discovery when simulation results differ from analytical predictions.

Contents

Polynomial surrogate models will not fit many national security problems

The polynomial can show trend and curvature, but…

Response surfaces in simulation can have large ranges, changes in gradient, asymptotes and knees that are not well-represented in polynomials

Lack of fit error is often larger than the sampling error

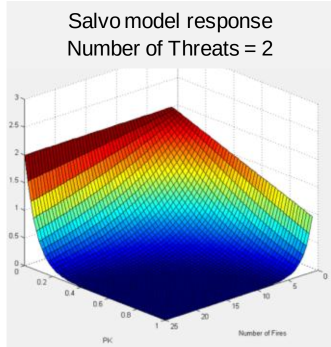
The example is a salvo model of expected kills with randomly coordinated interceptors (Waddell, 1961)

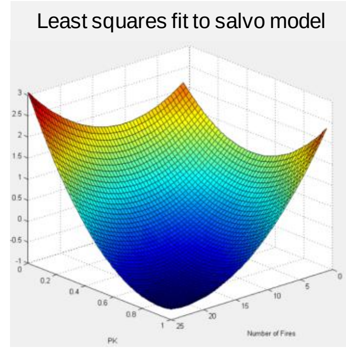
This example has errors of ~20% with some very large departures

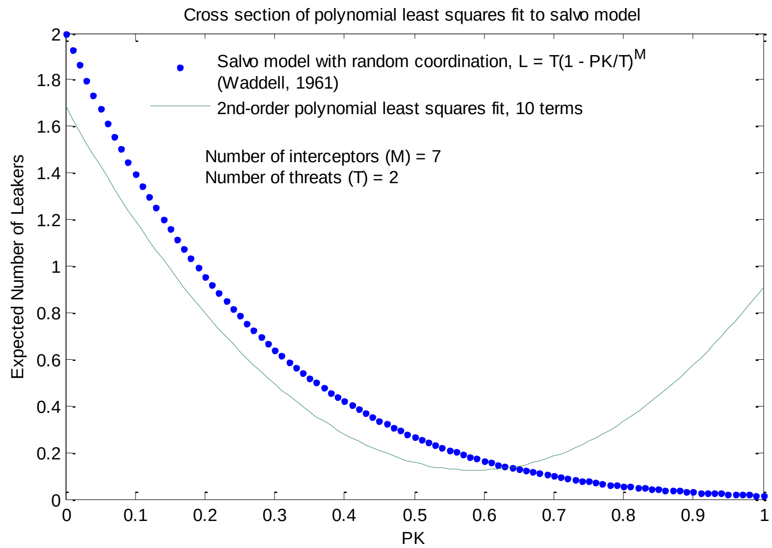
The example surrogate model exhibits false extrema and trend reversals

Polynomial models have many terms. They have to be screened to prevent over fitting

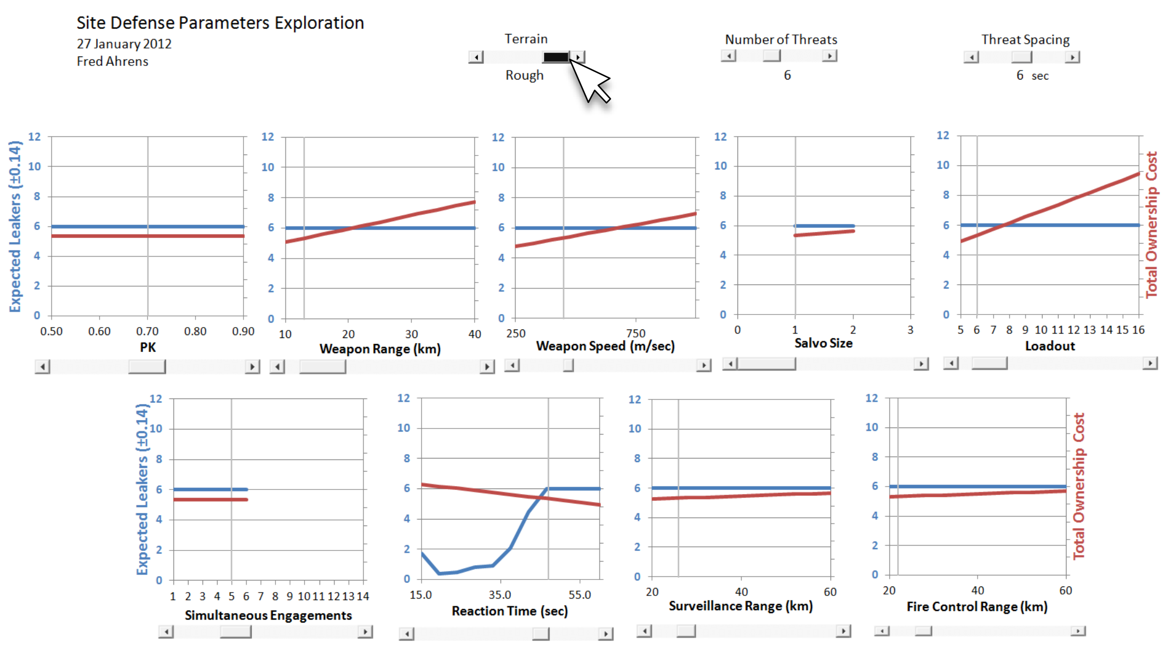
While a polynomial surrogate model can give local properties (gradient and curvature), it is usually not suitable for a global representation





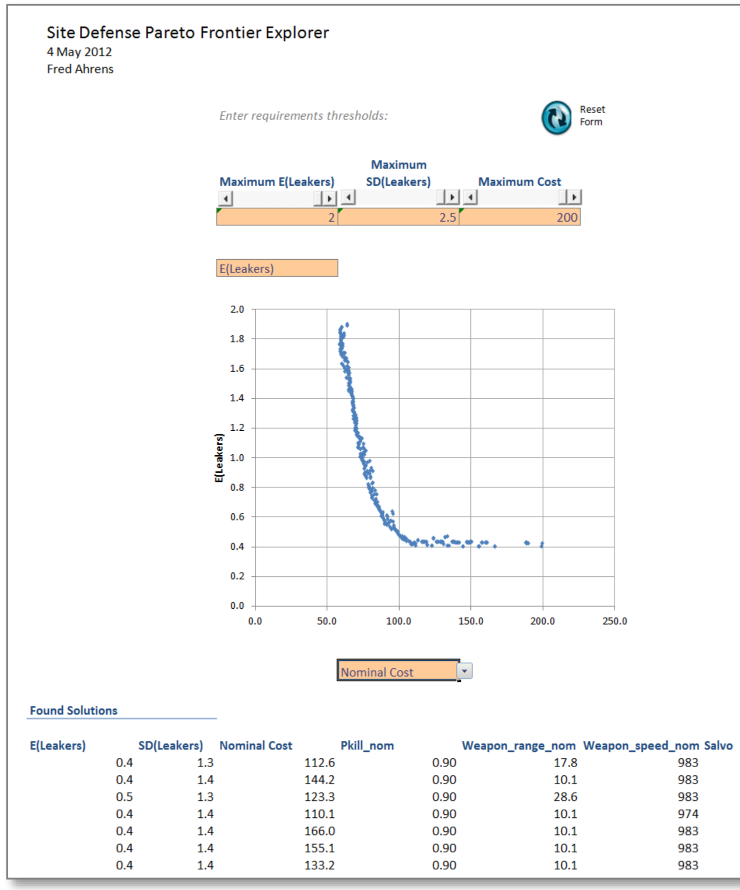


Desirability of a global representation of the simulation response



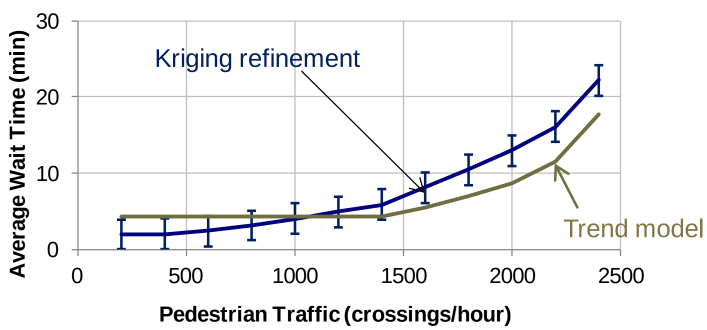
Interactive Solutions Exploration Example: A surrogate model provides a fast-running representation suitable for collaborative exploration of the design space

Optimization Example: Surrogate models can be combined to generate hundreds of non-dominated solutions for further screening and exploration



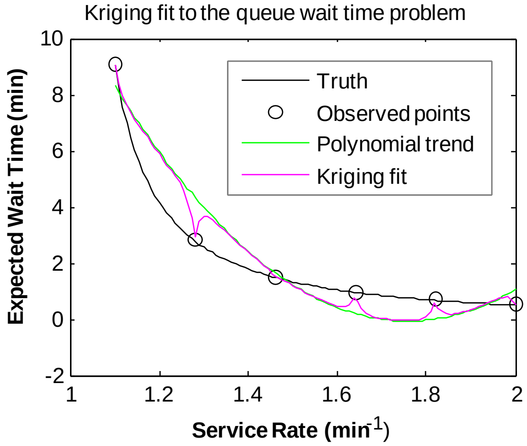
E(Leakers)	SD(Leakers)	Nominal Cost	Pkill_nom	Weapon_range_nom	Weapon_speed_nom	Salvo
0.4	1.3	112.6	0.90	17.8	983	
0.4	1.4	144.2	0.90	10.1	983	
0.5	1.3	123.3	0.90	28.6	983	
0.4	1.4	110.1	0.90	10.1	974	
0.4	1.4	166.0	0.90	10.1	983	
0.4	1.4	155.1	0.90	10.1	983	
0.4	1.4	133.2	0.90	10.1	983	

These types of applications require global surrogate models with acceptable errors



Advanced methods allow us to keep a trend model to describe large movements and to use a refiner to reproduce observed responses

Kriging fit to the queue wait time problem



Even with an advanced method, the misspecification of the trend model can result in poor interpolation (Staum, 2009)

Advances in response surface modeling aimed at a global representation

Linear regression using non-polynomial functional forms, such as exp, sin, 1/x

Generalized linear models transform the output of a linear regression to improve the behavior of the response surface (such as a bounded domain) (Staum, 2009)

Non-linear functional forms have a general relationship between the coefficients and the predictors (Santos & Porta Nova, 1999), such as a(1 – exp(-bx))

Neural networks are a generalized nonlinear form with layered parallel architectures (Swingler, 1996)

Spatial correlation methods, such as kriging, estimate new points using weighted averages of observed points (Sacks et al, 1989)

Stochastic kriging is an extension that accepts simulation data with sampling errors (Ankenman et al, 2008)

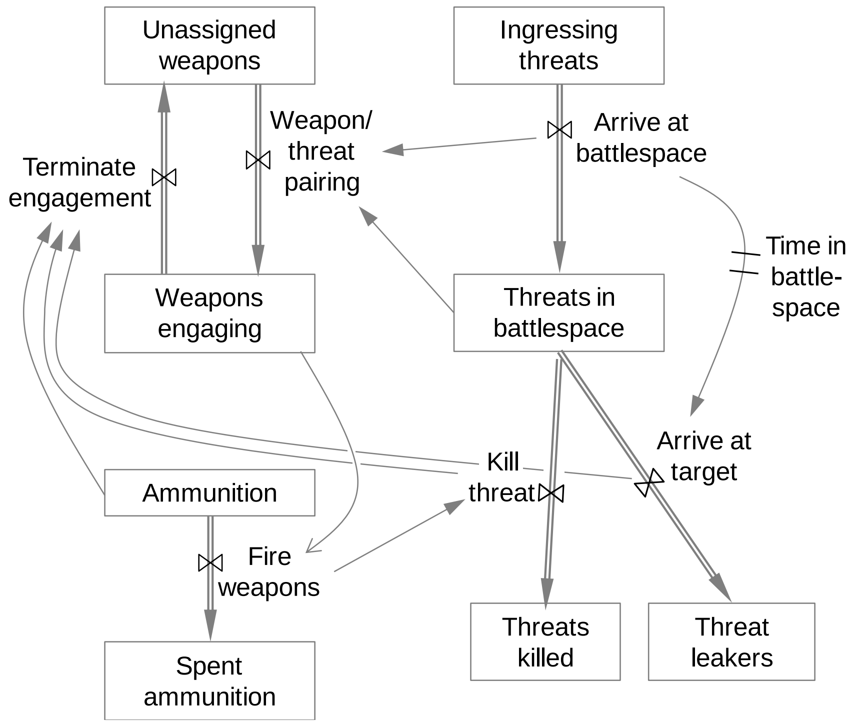
Approximate models

Approximate models (AM) are simplifications of detailed models that trade off accuracy to reduce computation time (Qian et al, 2006)

AM can be used in early project exploration to frame the problem domain, gain initial insight and identify knowledge gaps

Experienced simulation developers use AM to cross-verify results (author's personal experience)

For our purposes, an approximate model is any representation simple enough to update in real time ("slider-bar" fast!)



A system dynamics model of site defense using stocks and flows

Aggregation at force level is an approximation that ignores phenomena occurring in an entity-based Monte Carlo simulation

Incorporating approximate models into response surface models

Multiple linear regression

$$y = \left[\mathbf{p}(\mathbf{x})^T \quad \mathbf{f}(\mathbf{x})^T\right]\boldsymbol{\beta} + \varepsilon$$

Generalized linear model

$$y = g\left(\left[\mathbf{p}(\mathbf{x})^T \quad \mathbf{f}(\mathbf{x})^T\right]\boldsymbol{\beta}\right) + \varepsilon$$

Kriging with FPM trend function (Qian, et al., 2006)

$$y = \left[1 \quad \mathbf{f}(\mathbf{x})^T\right]\boldsymbol{\beta} + m(\mathbf{x};\tau^2,\boldsymbol{\theta}) + \varepsilon$$

Neural net with FPM predictors

$$y = \gamma_0 + \sum_i \gamma_i S\left(\left[\mathbf{x} \quad \mathbf{f}(\mathbf{x})\right]\boldsymbol{\beta}_i\right) + \varepsilon$$

Note that AM predictors will probably be covariates with some members of x

Therefore, consider dropping polynomial or even first-order x terms if an AM output is in the model

Definitions:

$\mathbf{x}$	Design point (vector of the independent variables)
$\mathbf{f}(\mathbf{x})$	Vector of predictors from the AM
$\mathbf{p}(\mathbf{x})$	Polynomial terms
ε	Sampling error
$g(\cdot)$	GLM link function
$m(\mathbf{x};\tau^2,\boldsymbol{\theta})$	Gaussian process with parameters τ^2 and $\boldsymbol{\theta}$
$\gamma_0, \gamma_i, \boldsymbol{\beta}_i$	Neural net coefficients
$S(\cdot)$	Neural net activation function

Case study: a site cruise missile defense problem was modeled both with AM and entity-based Monte Carlo simulation

Where should the materiel developer invest to achieve the best end-user value?

How do system-level performance parameters relate to site defense effectiveness?

12 independent variables

Two MOEs

Latin-hypercube DOE with 120 unique cases

Scenario modeled in EADSIM (U.S. Army Space and Missile Defense Command, 2010)

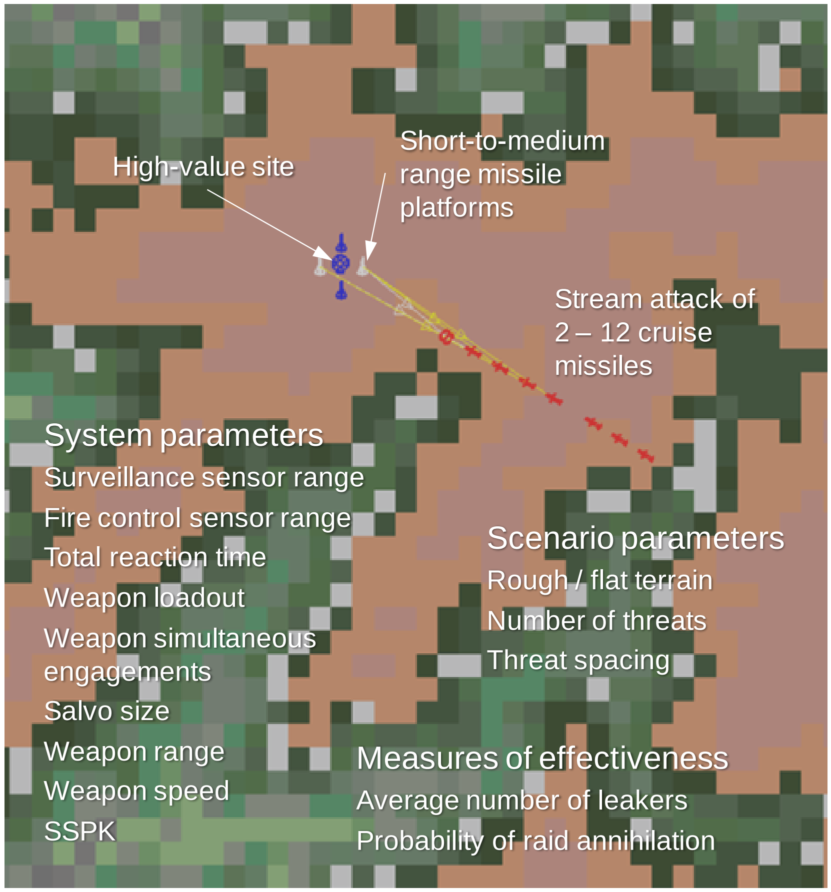
1,200 total (10 per case) Monte Carlo simulations

The standard errors of the mean varied by case, with overall root-mean-squares:

0.21 for Average leakers

0.09 for P(raid annihilation)

A stock and flow model (see below) provides an AM predictor



The stock & flow AM correlates with outputs from EADSIM

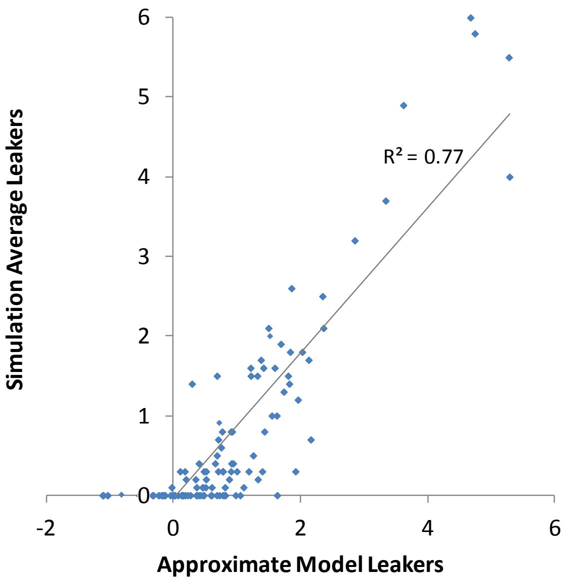
The AM is a spreadsheet model using the stock and flow principles of slide 6

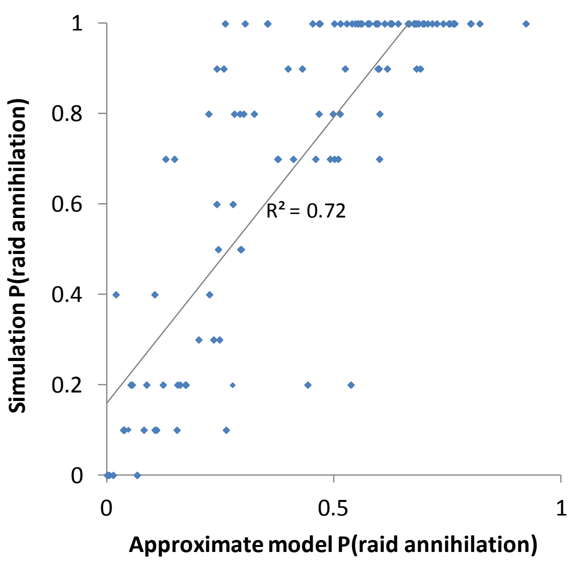
The AM for leakers sometimes predicts negative leakers

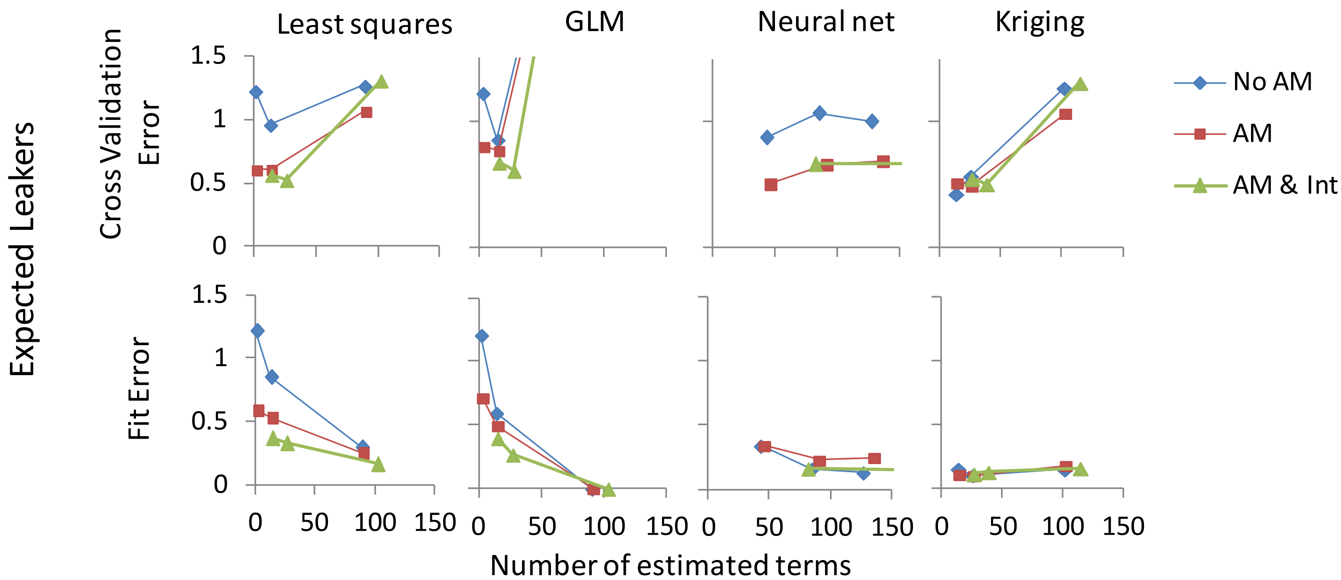
The AM for P(raid annihilation) applies an independence assumption to compute a binomial approximation from individual threat survival probability

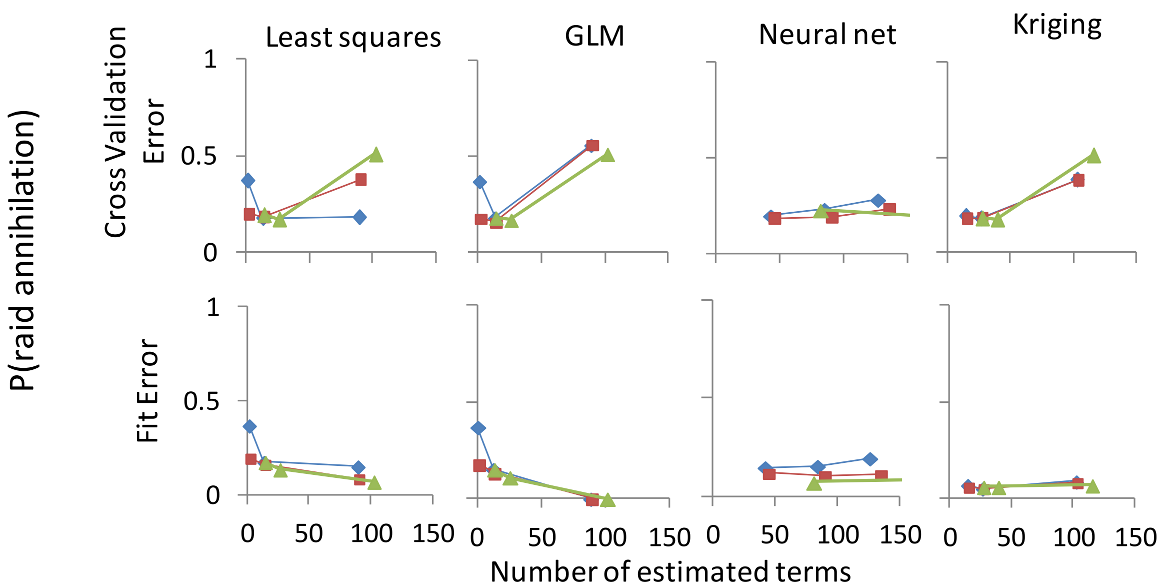
While the AMs are predictors of simulation data, they fail lack of fit tests (p < 10-4)

Averaged MOEs from the site air defense data set (120 design points)









Fit and generalization assessed for four methods with and without AM

These graphs compare the RMS errors for four types of response surfaces, first done traditionally without the AM, and then with the AM in the manner shown on page 4

AM & Int also includes interactions of the AM with first-order design inputs ($\mathbf{x}\cdot f(\mathbf{x})$), per Qian *et al* (2006)

Within each method are three variations that vary the number of estimated terms

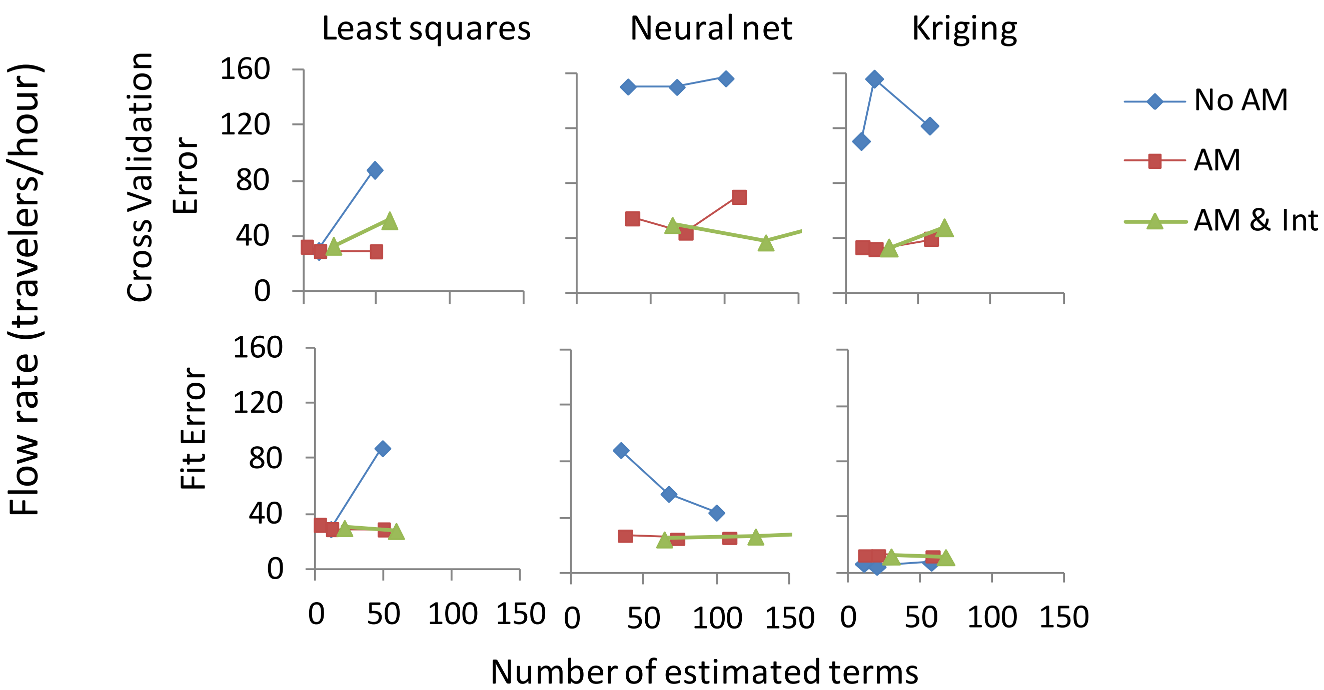
> Least squares, GLM and kriging: polynomial terms up to order 2.
>
> Neural net: three, six and nine nodes

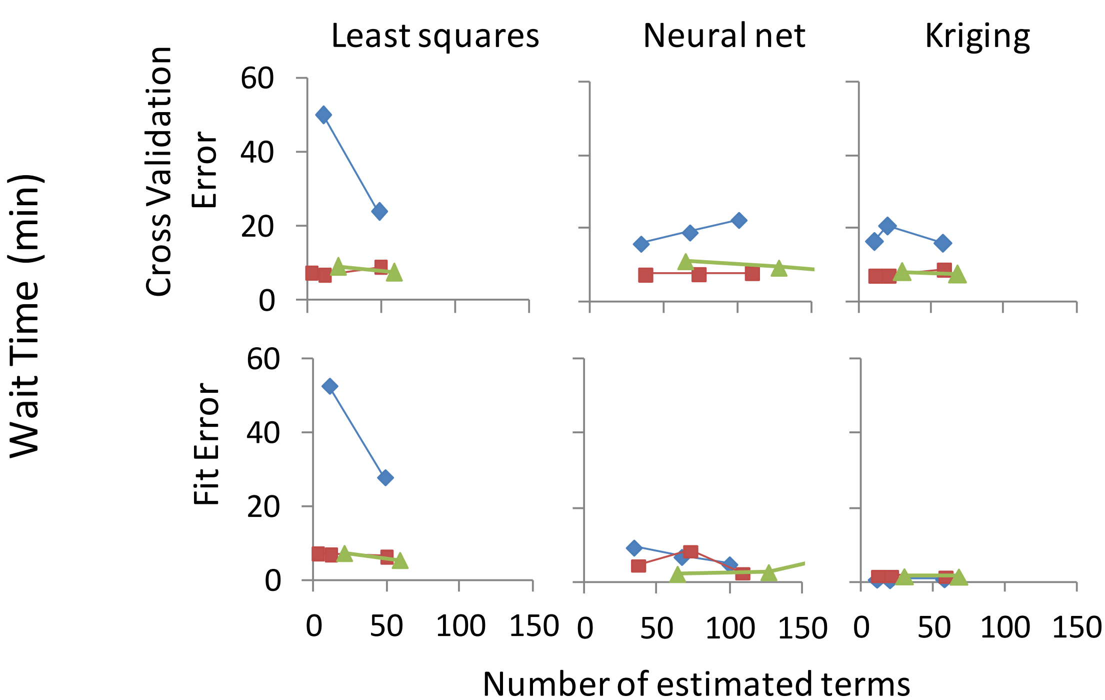
Fit error is measured as the RMS error of the predicted against the simulation means for the whole data set (120 means).

Generalization is measured using “leave-one-out” cross validation (Refaeilzadeh, Tang, & Liu, 2009)

Models with AM generally have lower fit and cross validation errors than those without while using fewer terms

Increasing number of estimated terms improves fit at the expense of generalization





AMs improved fit and generalization in a homeland security application

These graphs show the effect of incorporating a spreadsheet AM into response surface models for a simulation of primary inspection operations in a land port of entry

Simulations measuring service flow rate and traveler wait time

Discrete event simulation in Extend®

(Krahl, 2003)

The standard errors varied by case, with overall root-mean-squares:

20 travelers/hour for flow rate

2.1 min for wait time

Generalization was measured using "k-fold" cross validation (Refaeilzadeh, Tang, & Liu, 2009) with k=2

Other applications of AM in security modeling and simulation

Simple orbital models in MATLAB can be recalibrated to STK outputs

- P(Access)
- Access time probability distribution function (PDF)
- Mean revisit time
- Revisit time PDF

In the top right example, high rate of intercept failures caused by premature masking

In the bottom right example, the simulation was dropping travelers before they exited the system, discounting them from the flow rate measurement

The AM can be a useful tool for explaining simulation outcomes. Some outcomes are understandable from first principles and some are more nuanced

"Pulling the thread" on valid differences leads to discovery: how the actual response differs from our preconceptions

AMs from an earlier phase of analysis can be useful for connecting the prior understanding to current understanding

A proven AM can be useful in design of experiments to avoid degenerate or uninteresting cases

Summary of Key Points

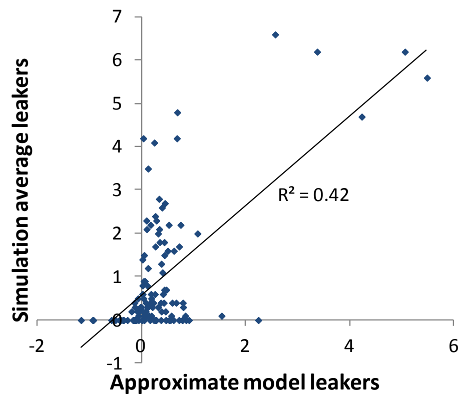
Consider alternatives to polynomial surrogate models: linear regression with non-polynomial forms, generalized linear models, non-linear models

Understand problem from a first-principles theory before fitting surrogate models

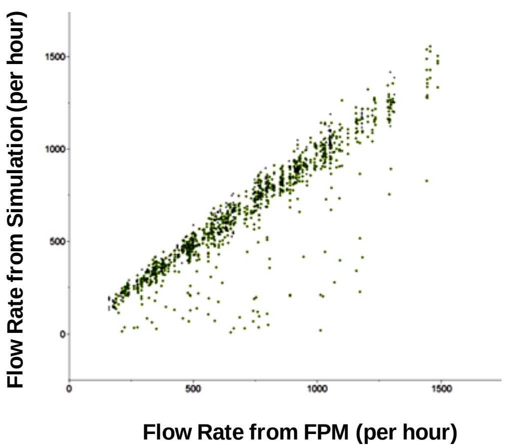
An AM can be made a predictor term in classic and advanced response surface modeling

A very good AM can be the sole trend predictor, reducing the error of fit, generalization error and number of terms in the surrogate model

Use of AMs in security simulations can have benefits in DOE, verification, debugging, understanding causal threads and in communicating outcomes



A discrepancy from development of site defense scenario



A bug from development of a PoE simulation

References

Ankenman, B., Nelson, B. L., & Staum, J. (2008). Stochastic Kriging for Simulation Metamodeling. *Proceedings of the 2008 Winter Simulation Conference* (pp. 362-370). IEEE.

Krahl, D. (2003). Extend: an Interactive Simulation Tool. *Proceedings of the 2003 Winter Simulation Conference* (pp. 188-196). New Orleans, LA: IEEE.

Refaeilzadeh, P., Tang, L., & Liu, H. (2009). Cross Validation. In M. Tamer Ã–zsu, & L. Liu, Encyclopedia of Database Systems (pp. 532-537). Springer.

Sacks, J., Welch, W. J., Mitchell, T. J., & Wynn, H. P. (1989). Design and Analysis of Computer Experiments. *Statistical Science* , *4* (4), 409-423.

Santos, M. I., & Porta Nova, A. M. (1999). The main issues in nonlinear simulation metamodel estimation. *Proceedings of the 1999 Winter Simulation Conference* (pp. 502-509). Phoenix, AZ: IEEE.

Staum, J. (2009). Better Simulation Metamodeling: the Why, What, and How of Stochastic Kriging. Proceedings of the 2009 Winter Simulation Conference (pp. 120-113). IEEE.

Swingler, K. (1996). *Applying Neural Networks: a Practical Guide.* San Francisco: Morgan-Kaufman.

Qian, Z., Seepersad, C. C., & Joseph, V. R. (2006). Building Surrogate Models Based on Detailed and Approximate Simulations. Journal of Mechanical Design , 128 (4),

Qian, P. Z., Wu, H., & Wu, C. F. (2007). *Gaussian Process Models for Computer Experiments With Qualitative and Quantitative Factors.* University of Wisconsin, Department of Statistics.

U.S. Army Space and Missile Defense Command. (2010, January 13). EADSIM Executive Summary. Retrieved December 8, 2011, from Extended Air Defense Simulation (EADSIM): http://www.eadsim.com/EADSIMExecSum.pdf

Waddell, M. C. (1961). *Surface-to-Air Guided Missile Systems Methods of Tactical Analysis.* Johns Hopkins University Applied Physics Laboratory.