SAND2008-2486P

CIS External Review 2008
Overview of Algorithms and Enabling Technology

David Womble, Senior Manager
Computer Science and Mathematics

1. Introduction. Algorithms and enabling technology (A&ET) is a key part of the CIS efforts to develop and support a high performance modeling and simulation (M&S) capability at Sandia National Laboratories. This area complements the other areas of CIS investment, Applications and Systems, and these three areas together form a coordinated, strategic approach towards developing an effective M&S capability. Investment areas are determined based on the CIS strategic plan, and established mission needs of the laboratories.

The CIS strategic plan includes three focus areas (scalability, transformation and informatics) and was briefed in last year's external review. The emphasis is on building a strong A&ET capability relevant to Sandia's mission. Projects are developed by staff and managers and are funded primarily through the computer science research foundation (CSRF), laboratory directed research and development (LDRD) and DOE's Office of Advanced Scientific Computing Research (ASCR).

Sandia's mission areas include nuclear weapons (NW), defense systems and assessment (DSA), energy resources and nonproliferation (ERN), and homeland security and defense (HSD). The mission areas typically fund the development of specific libraries and the application of advanced algorithms to specific applications. Examples include the funding for the Trilinos and Dakota projects through the advanced simulation and computing (ASC) program (NW) and the funding of specific informatics projects for the intelligence community (DSA).

2. Organization. There are two organizations with significant effort in algorithms and enabling technologies.

2.1. The Computer Science and Mathematics Group (1410). This group was reorganized in 2007 to align with the strategic plan. The departments and their foci are as follows.

1. Optimization and Uncertainty Quantification (1411). This department is managed by Jim Stewart and focuses on transformation of modeling and simulation through the development of new tools for optimization and uncertainty quantification and their application to engineering design and analysis. This group is responsible for the Dakota framework, which is the "delivery vehicle" for most of these tools.

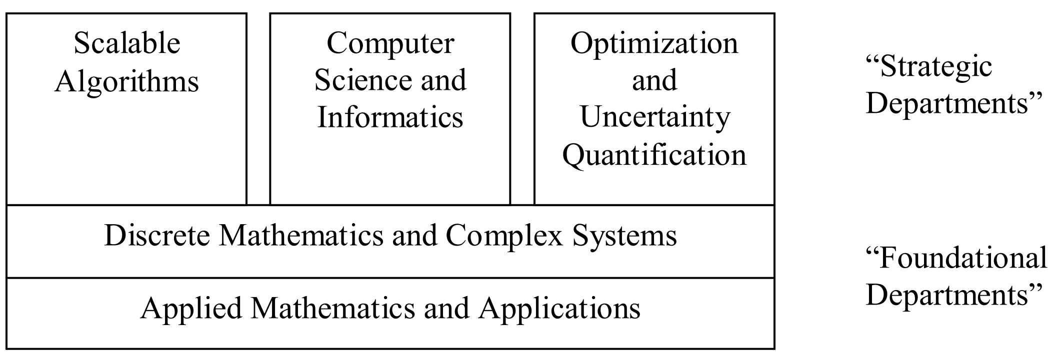
2. Computer Science and Informatics (1415). This department is managed by Suzanne Rountree and focuses on research and development in informatics and the application of these algorithms to national security and problems in the intelligence community. Areas of research include discrete-math-based

informatics, graph algorithms, multi-linear algebra methods with link analysis, and scalable information analysis.

3. Scalable Algorithms (1416). This department is managed by Scott Collis and focuses on the development and implementation of algorithms that scale to large problem size, large number of processors, or both. The department has technical leadership of the Trilinos framework that is the delivery vehicle for Sandia's scalable algorithms and solvers. The department is also responsible for work in performance modeling and prediction for applications and algorithms on new architectures.

4. Applied Mathematics and Applications (1414). This department is managed by Ken Alvin and focuses on foundational work in numerical methods, analysis and classical applied mathematics. Areas of interest include discretizations, multiscale methods, multiphysics couplings, and stability analysis.

5. Discrete Mathematics and Complex Systems (1412). This department is managed by Danny Rintoul and focuses on foundational work in discrete mathematics and computer science. Areas of interest include graph algorithms, discrete optimization, computational biology and complex system modeling.



2.2. The Computer Sciences Group (8960). This group is physically located at Sandia/California and focuses on cross-cutting research in algorithms and enabling technology, systems, and applications. The group is composed of five departments:

1. Scalable Computing Research and Development (8961), managed by Damian Rouson (acting)
2. Informatics and Decision Sciences (8962), managed by Heidi Ammerlahn.
3. Visualization and Scientific Computing (8963), managed by Howard Hirano (acting).
4. Advanced Software Research and Development (8964), managed by Mike Hardwick.
5. Computer and Network Security (8965), managed by Edward Talbot.

Algorithm and enabling technology foci within these departments include research and development of component architectures for scientific computing; algorithms and methods for optimization, uncertainty quantification, and verification and validation; flexible frameworks for solving PDEs, multi-physics modeling, and computational chemistry; mathematical and computer science approaches to advanced multi-dimensional data analysis including inference, forecasting, classification, and visualization; and foundational approaches to modeling complex systems and analyzing limits of predictivity.

3. Technical Investment areas. The technical areas for A&ET are listed below.

3.1 Integrated Capabilities. The group continues to place a strong emphasis on providing an integrated capability for solvers, optimization and other tools required for efficient modeling and simulation. Much of the integration is accomplished though deployment in frameworks with common interfaces, data structures and software quality standards. (Of course, even with these tools, CIS research must explicitly keep a focus on integration.) Consistent with this vision, our dynamic load balancing tools are being brought into the Trilinos framework, significant algorithmic ties are being established between the discrete and continuous optimizations efforts, and interoperability between Trilinos and Dakota is being designed.

One of this year's highlights was the demonstration of improved integration between components within the Trilinos framework in an ASC level 2 milestone.

3.2. Solvers and solver frameworks. This has traditionally been a strong investment area for Sandia. Solvers are the key to most scientific modeling and simulation efforts and often more than one half of the computational time is spent in the solvers. Most of the solver research is driven by finite element simulation codes which are being developed within the ASC program, although applications such as optimization and circuit simulation are also being addressed. We are also continuing our efforts in time integrators and nonlinear solvers.

Recent accomplishments include a new multigrid solver that scales effectively to over 25,000 processors on Red Storm and a new preconditioner for circuit simulations that, combined with hypergraph partitioning, reduces solve times by up to 2000 times.

3.3. Finite element methods and numerical analysis. Finite element methods are widely used in engineering simulations, particularly at Sandia, and we have continued our work on the development of the Intrepid library for compatible discretizations, which will be deployed through the Trilinos framework. Discretization establishes the essential link between continuum models and their algebraic representations. The algebraic models must provide accurate and physically correct solutions so as to enable validated computer simulations of the actual process. Design of such models requires careful analysis of the attending mathematical structure of the PDE model. This structure encodes intrinsic physical properties of the process that is being modeled, such as conservation laws, solution symmetries, positivity, and maximum principles, to name just few.

We have also continued our work in automatic differentiation, stability analysis and sensitivity analysis with particular emphasis on providing an infrastructure for "predictive simulation." We recently completed an ASC level 2 milestone that demonstrated many of these capabilities in Charon, the device simulation code used in the QASPR project.

3.4. Optimization and Uncertainty Quantification (UQ). Optimization is a major thrust area for the algorithms and enabling technology effort. The Design and Optimization project focuses on three interrelated areas: (1) the development of fundamental optimization algorithms, (2) the development of frameworks for optimization, and (3) the development of tools and techniques for the quantification of uncertainty.

Much of our recent work focuses on the third item and supports the verification and validation efforts at Sandia. Algorithmic advances are deployed through the Dakota framework and include work in polynomial chaos methods and stochastic sampling methods, as well as surrogate and multiscale methods. One of the primary areas of interest is UQ for problems for which a large number of function evaluations is not practical.

Sandia also has unique capabilities in combinatorial and global optimization within DOE. Problems with explicit combinatorial structure often require discrete analysis to develop efficient solutions. We are developing techniques for specific applications like sensor placement in water networks and node allocation in parallel clusters. Additionally, we are developing general techniques like graph algorithms, online algorithms and polyhedral combinatorics. We are also developing global search methods for applications in which (near-)optimal solutions cannot be efficiently identified. We are developing parallel global optimizers that rigorously guarantee that a globally optimal solution is generated (e.g. branch and bound), as well as practical heuristics that balance global searching with efficient generation of near-optimal solutions. Current applications include network analysis and design and logistics simulation and optimization.

3.5. Informatics: The field of informatics has recently become a key part of the work done in the CIS. There are many reasons for this rapid rise: there is a large amount of data combined with an increasing ability to store it. Continually increasing computer power offers the possibility of analyzing the data on a large scale, discovering trends and data relationships, and developing predictive analysis capabilities. There is a compelling national security driver for informatics and the analysis of large data sets, especially in the intelligence community and national security-related areas. Our research in new analysis methodologies and algorithms is being complemented with architectural research motivated by these challenging large-data problems. Unstructured data analysis and graph-based problems have different data access and computation needs from traditional scientific computing.

Our current work leverages research across multiple areas, including architectures, algorithms, high-performance data techniques, and visualization. The implementation of the Multi-Threaded Graph Library (MTGL) on the Cray Eldorado machine has been particularly well received and is continuing under several sources of research and development funding. New informatics research is being done under a “grand challenge” LDRD on the discovery, characterization, and prediction of adversarial networks.

Work has focused on multi-linear algebra based methods (e.g., multi-way tensor methods) and graph-based algorithms, coupled with information visualization and an underlying analysis framework. The group also has research strengths in scalable supervised machine learning methods and novel Bayesian analysis algorithms. Longer-term research is expected to grow from statistics-based techniques underlying complex network analysis and prediction, along with the quantification of uncertainty in informatics-based problems. Our research is conducted in collaboration with the architectures group, visualization group, scalable algorithms groups, and several partners throughout Sandia.

4. Programs. Many programs provide funding to our algorithms and enabling technologies efforts. Each has its own objectives, constraints and funding mechanisms. These programs are summarized briefly below.

4.1. ASCR.
Sandia receives significant research funding from the Office of Advanced Scientific Computing Research (ASCR) program in DOE’s Office of Science in several research areas. The funding levels given below are for all of Sandia, not just for the CIS (although the majority of ASCR funding supports CIS staff). Sandia has the largest Applied Math Research (AMR) program funding of all DOE Laboratories in FY08. A current focus of ASCR is the deployment of advanced algorithms to enable petascale computing for science applications through the SciDAC program. Sandia has strong participation within SciDAC algorithms activities including the ITAPS and TOPS Centers and the CSCAPES Institute. Currently, Sandia researchers are responding to an ASCR call on multiscale mathematics and optimization of complex systems. Future ASCR programs are planned in V&V/UQ in the FY09/FY10 timeframe. Through both the complex systems call and future V&V/UQ calls, continued growth in ASCR algorithms funding is anticipated.

4.2. ASC. The Advanced Simulation and Computing Program is DOE’s program to develop and apply advanced modeling and simulation capabilities in support of the nation’s nuclear weapons stockpile. ASC includes a $7M Algorithms effort to develop and deploy advanced algorithms and enabling technologies to codes being developed within ASC. Projects are selected by program managers within the ASC applications program based on prioritized strategic needs and reviewed annually. This is considered a development program, not a research program, and supports both Trilinos and Dakota development. The ASC budget is ultimately tied to nuclear weapons (NW) funding levels for the laboratories.

4.3. CSRF. The Computer Science Research Foundation program funds research into computer science, computational science, and enabling technologies in support of Sandia's high performance computing effort. These four projects are focused around the following areas with the approximate (yet to be determined) levels of funding

- Simulation Computing ($3.5M)
- Next-Generation Systems ($2.5M)
- Information Sciences ($1M)
- Disruptive Technologies ($1M)

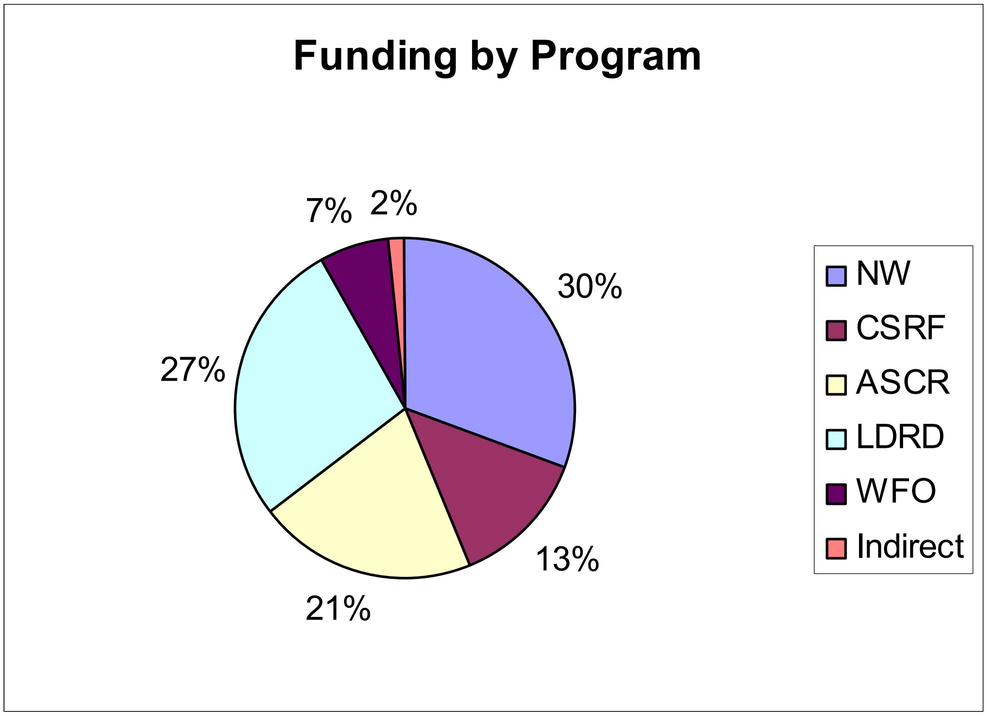
CSRF is formally part of the ASC program and supports research that will impact the nuclear weapons program. This budget, too, is ultimately tied to the NW funding levels for the laboratory.

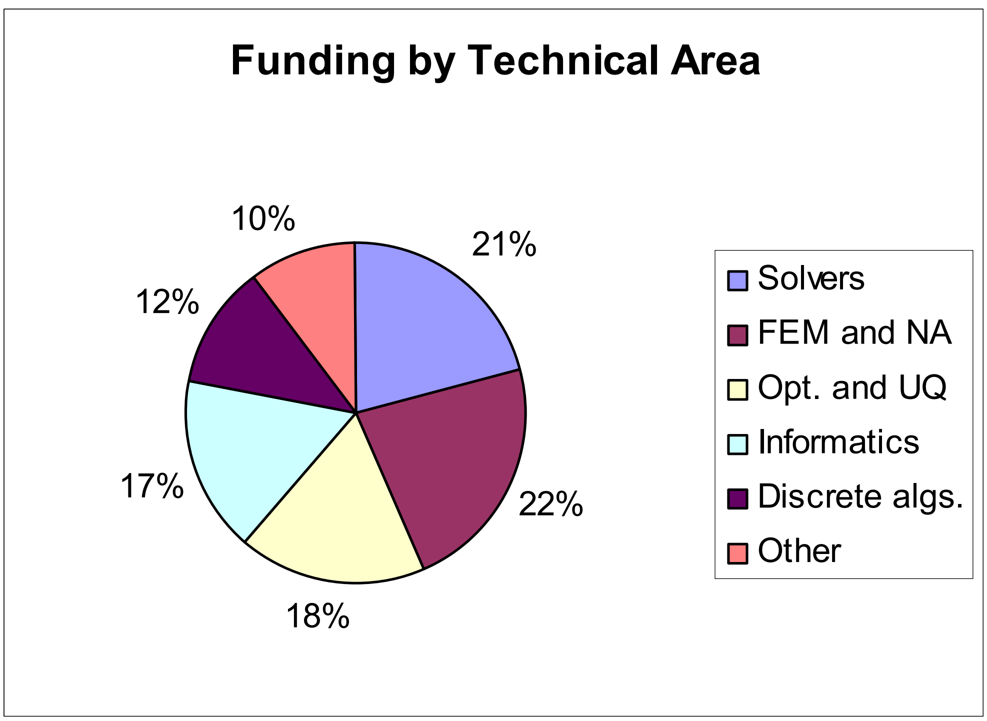
4.4. LDRD. Sandia's Laboratory-Directed Research and Development program is funded through a 6% tax on laboratory direct spending. Projects are selected based on an internal review and are reviewed annually. CIS researchers participate in the regular LDRD call, primarily through the Enabling Predictive Science Investment Area (projects less than $1M/year) and in the Grand Challenge Call (larger projects). CIS researchers also receive LDRD funding from other investment areas, including energy and national security. The level of funding in the Computer Science investment area is set by Sandia's vice presidents. Funding seems to have leveled off.

4.5. CSRI. The Computer Science Research Institute is the program through which we interact with the external research community in mathematics and computer science. The CSRI provides space and infrastructure for our visitor and student programs, hosts workshops and support technical collaborations with the external community. The CSRI is no longer funded by the ASC program, although we are working to restore a small amount of funding for infrastructure and administration.

4.6. Work For Others (WFO). This is the formal designation for work outside of the NW program and not funded through internal taxes (e.g., LDRD). The algorithms and enabling technology area has projects with various organizations or agencies, including the DoD (Department of Defense), NIH (National Institutes of Health), EPA (Environmental Protection Agency), DARPA (Defense Advanced Research Project Agency), U.S. Industry (especially Goodyear, Boeing and ExxonMobil) and various agencies associated with the intelligence community.

5. Funding. The following table shows the approximate amount of funding in each of the technical investment areas and funding areas for FY08. All funding is for staff labor. The approximate conversion is $300K/staff member. The total funding for AET programs is just under $20M.





6. Response to previous reviews. In this section, we respond to specific comments and recommendations directed towards algorithms and enabling technologies from reviews in 2005, 2006 and 2007. We appreciate the many positive comments received from the review committee; however in this section, only those comments and recommendations specifically suggesting actions or otherwise requiring a response are included.

6.1. 2005.

- **[Mixed-integer programming] "Future efforts in mixed integer-nonlinear programming are encouraged."** [Response] There has been strong work in maturing the serial version of PICO in order to gain a broader acceptance within the user community. In addition, these improvements generally translate to better parallel performance. From a mathematical perspective we have pursued work on finding feasible solutions to IP's given a matrix representation of the problem. We are currently developing techniques to integrate constraint programming methods with the classical branch-and-bound methods that we have traditionally used. From the applications perspective, we have continued to work on more robust solutions to the sensor placement problem, and have begun to attack the problem of data gathering in wireless sensor networks where the sensors have a finite battery life. Finally, we have made significant new thrusts in the areas of logistics and quantum computer architecture design.
- **[Reconstruction of contamination events] "The panel is unable to evaluate future direction of this general effort given that only a limited follow-on plan was presented. An obvious follow-on might be consideration of the ramifications of imperfect sensors and measurements as part of their initiative in uncertainty quantification."** [Response] This work was a proof of concept that inversion of air sensor data is possible with embedded optimization algorithms. We are disappointed that DHS chose not to continue this work. However, the committee's observation that it will be necessary to deal with uncertainties in the sensor data has become an important part of several other efforts, including the sensor placement work for the EPA.
- **[Sundance] "It would be useful to compare timings and quality of results of application done with this approach versus an existing application. It is not clear how numerically robust is this symbolic differentiator. The Panel could not understand what was under the hood."** [Response] The Sundance package has proven to be an excellent platform in which to prototype simulations of small problems; however, it has not been successful as a production tool for many reasons, including existing software infrastructure, the complexity of the applications, the ability to incorporate models and the current focus on verification and validation. We are using some of the ideas in current research projects, and the software has been released as open-source. Sandi
- **[CCA] "How successful is the CCA community effort in creating broadly used standards and software?"** [Response] The CCA has been successful on four fronts primarily. It has been accepted as a valid standard/architecture for writing scientific computing frameworks; there are currently three competing frameworks in the USA and a few prototypical ones in Europe. Further, CCA-related ideas have been adopted by NASA's Earth System Modeling Framework. CCA has also be adopted by computational scientists as an enabling technology for simulations; at last count there were at least 15 separate scientific projects

using it (two in Europe, as well as a few funded by NSF), and they spanned the gamut from accelerator physics to space weather. The main use of CCA in such scientific projects has been to tame complexity. In some cases it has enabled a small team to integrate multiple codes of varied legacies and programming approaches; in others, it has helped with the churn of team-members (i.e. short term contributions by students and post-doctoral researchers). Finally, CCA's approach to inter-language interoperability (Babel) has been widely adopted, even in projects that have decided not to use CCA. It has also been exploited, somewhat unconventionally, to perform ensemble runs on processor farms for multiscale simulations. Babel has been downloaded multiple (O(1000)) times.

6.2. 2006

- **[General] "The distribution of FTE's across the efforts in 1410 appears to be reasonable although we encourage additional FTE's for the V&V work."** [Response] The emphasis on V&V, including the number of FTE's, in CIS continues to increase. Much of the work in our Optimization and Uncertainty Quantification department is now focused on research and development explicitly supporting the V&V effort. Also, most of the code verification and validation work at Sandia uses Dakota. The overall research effort at Sandia has not increase as much; most of the increase outside of the CIS organizations is focused on the application of V&V to the modeling and simulation codes.
- **[Pattern recognition in massive, messy data] "Good example of applying modern ML techniques to real problem - Advance in training data generation – how general is this?"** [Response] Very general. Below are a subset of the disparate Sandia and other national lab applications to which these techniques have been applied:
  - LLNL: The Avatar Tools code is part of the production operation of the NIF optics inspection system, which means it is being used multiple times a day to look for emerging and dangerous flaws in NIF optics.
  - LLNL: Categorization of failure modes in implosion simulations
  - LBL: Detection of supernova candidates in nightly scans
  - LBL: Machine learning for selection of the best eigensolver for given problem
  - SNL: HCARS, High Consequence Automatic Response Systems, which is essentially "identify friend or foe" from body movements
  - SNL: Detection and identification of ideology in text. Recent paper presented on results at a computational linguistics conference in Israel.
  - SNL: Analysis of Border Gateway Protocol (BGP) data for situational awareness of the internet.
  - SNL: Text classification, predicting which journal published which paper. Recently resulted in co-authorship of an accepted book chapter.
  - SNL: Malware classification for cyber security applications.
  - SNL: Detection of steganographic content in audio signals.
- **[Pattern recognition in massive, messy data] "Not clear that a black box approach can work for target application."** [Response] The "black box"

approach has been shown to work for many applications (see above). Within ASC, in the last year this approach has been used, for example, to find material tears in NW component simulations, the likely future location of tears, broken and near broken bolts, and what turned out to be malformed mesh connections, all without any specific tuning to those applications.

- **[Pattern recognition in massive, messy data] "Is there work on unsupervised learning as well?"** [Response] There is indeed. In fact, probably the majority of Sandia's data mining research is in unsupervised learning. These include a "Grand Challenge" LDRD on network discovery and prediction, CIS research LDRDs on multi-way graph analysis using tensor methods, frameworks for analyzing text, and scalable algorithms for community-finding. Other areas include ASC and intelligence community-funded research in unsupervised machine learning on graphs.
- **[Automatic differentiation] "Should think of refactoring tool to facilitate 'templatization'."** [Response] We have considered this approach, including looking at Eclipse, but have concluded that there is not a feasible, robust C++ refactoring tool for our environment. We have found that the model of having a Sacado expert user consult with code teams scales well to Sandia development efforts, and that some teams have already successfully picked up the capability with very little help from Sacado developers. Finally, we have included a demonstration FEM code with the Trilinos/Sacado distribution that models the use of Sacado for codes to emulate.
- **[V&V] "To fully realize potential, work needs to remain unclassified."** [Response] We agree completely! So far classification issues have not been a problem for the research and development efforts in V&V, although some of the application of V&V to specific codes and models is restricted at the OUO (Official Use Only) level or higher. A positive development in keeping the V&V work unclassified is the announcement of the ASC Predictive Science Academic Alliance Program (PSAAP) which will explicitly engage universities in predictive science research, including V&V research.

6.3. 2007

- **[General] "Committee looks forward to hearing about the impact of the work in future reviews."** [Response] We consider "impact" to be extremely important in a mission laboratory like Sandia and are disappointed that we did not do a better job of conveying this in the 2007 review. We will ask each of the speakers this year to address the question of impact directly.
- **[Mathematics of computation (Intrepid)] "We hope to see examples of how it is being applied to complex, "real-world" physical systems in the future."** [Response] As the panel observed last year, the Intrepid development work was in relatively early stages. Nevertheless, Intrepid is based on research that we have been doing over the past couple of years and there have been some significant impacts demonstrated. For example, compatible discretizations have been deployed in Alegra (MHD) and Charon (device simulation). In both cases,

compatible discretizations have significantly improved the accuracy of the simulations (one specific example was given in the 2007 presentation) and have significantly improved the efficiencies of the solvers. For example in MHD simulations in Alegra, the ability of the improved discretization to capture the discrete null space accurately resulted in more than an order of magnitude speedup in the multilevel solver. The Intrepid library will be released this year (in the September, 2008 release of Trilinos) and will be interoperable with other packages, e.g. automatic differentiation, which will result in a significant increase in capabilities. Major production codes that have already committed to use the new package include Alegra, Charon and Kull (the LLNL MHD code). In addition, any code that links Trilinos will have easy access to compatible discretizations.

- **[Verification and Validation Processes] "Further work is needed to establish the general applicability and advantages of the treed Gaussian methods."** [Response] We are sorry that the general applicability and advantages of Treed Gaussian Process (TGP) were not more clearly communicated. TGP was developed at UC Santa Cruz by H. Lee and his graduate students and has a proven track record of application at NASA, DOE laboratories, and elsewhere. In 2006, for his work on TGP, R. Gramacy was chosen as the recipient for the Savage Award "for a dissertation that makes an outstanding contribution with novel Bayesian analysis of a substantive problem that has potential to impact statistical practice in a field of application." Sandia's own application of these techniques to reducing experimental data uncertainty due to sampling processes and improving the efficiency of and managing the uncertainty inherent in model calibration has also been very successful. Currently, the DAKOTA team is investigating the differences between TGP and other Gaussian Process methods to aid users in selecting the best approach for different classes of problems.