The Role of Massively Parallel Computing in Developing a Practical, Maximally Impactful Validated Predictive Modeling Capability

C. Holland* (UCSD), W. Guttenfelder (PPPL)
*corresponding author: chholland@ucsd.edu

Request Oral Presentation: No
Primary Panel: Whole-Device Modeling (C) **Secondary Panels:** D-G

Motivation

While rapid advances in high-performance computing in the last two decades have facilitated significant new insights and understanding of the rich nonlinear, multiscale dynamics of MFE-relevant plasmas, the uncomfortable truth is that virtually none of these insights have been incorporated into any practically viable or useful "whole device modeling" (WDM) predictive tools. By "practically viable or useful," we mean a tool which can be routinely utilized by multiple analysts to provide clear and timely results relevant to a wide range of stakeholders (including not just scientific managers but also funding and regulatory agencies). If development of such a practical predictive WDM capability through massively parallel computing is to be a goal of the US MFE program (as has been articulated in the guidance to this activity), then there must be a clear plan as to how this goal will be achieved, particularly under the significant resource constraints of any plausible funding scenario. We believe achieving this goal will require not only new research initiatives but also cultural and structural changes in both OFES and the MFE community. Details of the key required initiatives and changes are discussed below. *However, the bottom-line message is that both OFES and the MFE community must be vigilant in ensuring that the development process is driven by the needs of the fusion program, and that massively parallel computing is used as a means to an end, and not an end in and of itself.*

Approach

In this discussion, we use the HPC (high performance computing) acronym to denote any simulation requiring computational resources beyond those available from a typical local university or national fusion laboratory cluster, and the term "extreme-scale" to denote capability-level computing on petascale or larger systems. The two key initiatives are:

1) Increased support for use of massively parallel first-principles simulations to build "intermediate-complexity" WDM components. A maximally impactful predictive WDM capability will be one which can be routinely and rapidly exercised over a wide range of parameters, and provide actionable results (e.g. for experimental planning) in a timely fashion. Therefore, it cannot require or rely upon purely first-principles nonlinear simulations or "extreme-scale" HPC resources for routine operation. Rather, the appropriate primary role for HPC simulations here is to provide the basis for development of "intermediate complexity" reduced models (ICRM). ICRMs combine physical understanding (often gained from first-principles simulation studies) with case-dependent quantitative scalings derived from first principles simulations to yield acceptably accurate approximations. Notable examples of this approach include TGLF and EPED. Looking to the future, focused refinements to existing ICRMs will be needed to extend and improve their fidelity in reactor-relevant regimes, and new ICRMs developed to address key gaps in areas such such as edge/SOL transport and plasma-material interactions. An even more substantial challenge will be to develop the physics understanding to build practical ICRMs that enable the integration needed to address outstanding issues such as the interplay of microturbulence driven transport, tearing modes, and equilibrium profile evolution. Development of these ICRMs will require increased support for collaboration and new research initiatives which combine the expertise of the fusion "base theory" program with computational plasma physicists and applied mathematicians. Increased collaboration will be particularly needed for developing the physics understanding required for addressing integration gaps, which in many cases will require extensions of existing theoretical frameworks. As a specific example of the kinds of changes required, we propose that one measure of success for the next round of fusion SciDAC centers should be to what extent they yield new and/or significantly improved such ICRMs.

2) Increase support and utilization of capacity computing for uncertainty and margin quantification (UMQ). Virtually all MFE codes and tools (particularly those used in current WDM activities) are deterministic- a single, repeatable set of outputs for a given set of inputs. However, quantification of experimental and model uncertainties is an essential aspect of robust validation, and so a validated predictive capability will require routine UQ. Moreover, for many potential WDM applications quantification of key safety margins will be essential. In practice, both uncertainty and margin quantification effectively require ensembles of deterministic simulations, which presents a variety of new challenges. Most importantly, the computational infrastructure (including both data management and analysis tools) to support ensemble-based approaches must be developed and deployed. A particular challenge for MFE-relevant UQ analysis will be the application of more advanced techniques such as polynomial chaos expansions which go beyond "brute-force" Monte Carlo approaches to the multi-dimensional parameter space most MFE models operate in. A closely related challenge will be determining how to be adapt or tune these new approaches to most efficiently operate for both reduced models and first-principles simulations. Additionally, the data management challenges presented by ensembles of first-principles simulations must be addressed. These simulations are often posed as nonlinear initial-value problems, and so significant data reduction will likely be needed for archiving many ensembles of such simulations. Finally, we note that the proposed increased emphasis on ensemble calculations pushes modelers towards capacity rather than capability computations, which often run counter to the desired goals of OASCR and HPC facilities. While in a number of cases, intelligent design of job submission can effectively transform an ensemble of calculations into a capability-level simulation, it is important that OFES identify and vigorously support resources for significant capacity-level computing, and avoid pushing the modeling community towards artificial milestones and compute job prioritizations driven by considerations other than the underlying plasma physics.

Impact

These initiatives will enable the US MFE program to achieve its goal of successfully building a validated predictive modeling capability that is both practical for widespread use and grounded in physics understanding.