Beyond Benchmarking - How Experiments and Simulations Can Work Together

Martin Greenwald – *MIT Plasma Science & Fusion Center*

There has been dramatic progress in the scope and power of plasma simulations in recent years driven by better science, better numerics and the exponential increase in computing speed. Physical parameters can easily be varied allowing researchers to model experiments which have not yet been built (or which, due to engineering constraints, could not be built.) Because codes are generally cheaper to write, to run and to diagnose than experiments, they have a well-recognized potential to extend our understanding of complex phenomena like plasma turbulence. However, simulations are only imperfect models for physical reality and can be trusted only so far as they demonstrate agreement, without bias, with experimental results. If our goal is to build models capable of prediction we must be in a position to assess the reliability or accuracy of these models. It is worth noting that in many instances, for example in fusion energy, the accuracy of predictions have significant economic consequences. Confidence in the ability to predict is ultimately based on code performance against experimental data. This process, often referred to as "benchmarking" or "validation" tests the correctness and completeness of the physical model used by the code and the validity of assumptions and simplifications required for solution on existing computer hardware. Experiments can be thought of as analog computations that always get the "right" answer. However experimental measurements are almost always incomplete and subject to significant errors. Further, we may not be able to ask the right questions – that is do the right experiments due to physical constraints or those of budget and time. The difficult questions are how to characterize and measure "agreement" and to assess confidence in a code to extrapolate into new untested regimes. How do we make meaningful comparisons when calculations are incomplete and where critical measurements may not be available? Falsification is not necessarily a easy concept to apply – some degree of non-agreement is to be expected in these cases. To reach a successful outcome, simulations and experiments must be seen as complementary not competitive approaches. We must move beyond simple benchmarking or comparison by vugraph where a few points from each are placed on the same graph with no assessment of errors, validity of assumptions or uniqueness of solutions. Comparisons must be thorough and quantitative and a premium should be placed on close and ongoing collaborations which are open and candid about the sources of error and the strengths and weaknesses of each approach. We need experiments dedicated to answering critical questions raised by the simulations and which examine the validity of models and which explicitly test their assumptions. Common sets of tools including synthetic diagnostics and the ability to easily share and compare data are essential components of this enterprise. Ultimately both experiments and simulation have much to gain by adopting a an approach of co-development, where simulations are continuously and carefully compared to experimental data and where experiments are guided by the results of simulations.