Divergence of simulations due to machine dependence: Assessment and Implications on Climate Informatics

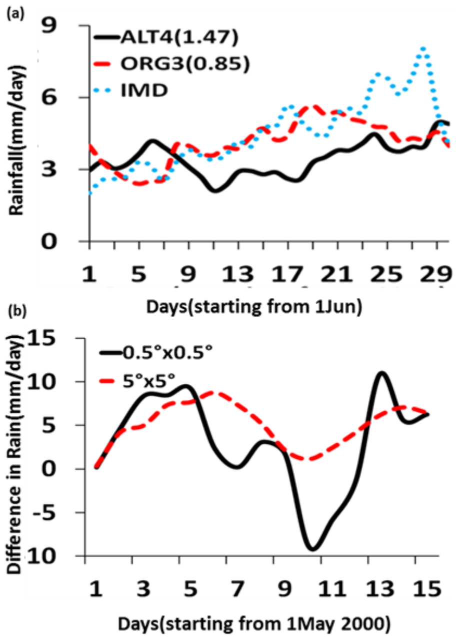
Kantha Rao B, Prashant Goswami

Climate and Environmental Modelling Programme, CSIR Centre for Mathematical Modelling and Computer Simulation, Wind Tunnel Road, Bangalore-560037, India.

Abstract

As the structure and scope of mathematical models become increasingly complex, computer simulations can no longer be validated easily; this, in turn, raises the criticality of the reliability of the computed results, especially in climate projections. The tremendous growth in high-performance computing has seen a corresponding increase in the scope and the complexity of weather and the climate models. An implicit assumption in designing a forecast or a projection system has been that the accuracy or skill of the simulation or the forecast is independent of the computing platform [comprising of compute server, compiler etc], at least in a statistical sense. In most cases, computing systems are chosen [benchmarked] as generic and passive tools, based on hardware specifications, performance benchmarking involving time, latency etc. Although careful consideration is given to data consistency (such as with reference to earlier simulations), accuracy (that is comparison with target data) is not generally considered; however, for end-use, and especially with respect to applications like weather forecasting, comparison with independent observations (accuracy) is critical. While the importance of the computing platform in the cycle of dynamical forecasting is well recognized, the emphasis so far has been primarily on aspects like scalability and interconnects or optimal hardware configuration. Although consistency of simulations by a new computing system is often considered, it is customary to compare forecasts from different computing platforms in experiments like model inter-comparison to determine the relative skill of forecast or simulation. In more software (model) centric benchmarking, every important model version is generally rerun on the new machine; these simulations are then compared with simulations on the old machine. It is not unknown that a new architecture (and/or compiler) can expose bugs in the code (often in the parallelization), but generally the basic characteristics of the climate of the models can be ascertained, and compared. The objective of the

present work is to examine and quantify the role of the computing platform in accuracy and the reliability of simulations, and especially weather forecasts. We show here, using simulations of increasing complexity, from basic library functions to the 3-variable Lorenz system to (internationally tested) models of weather forecasting, that simulations from two different computing platforms of the same brand, with identical initial conditions, can give rise to significantly different forecast skills, although both systems generate acceptable simulations. It is ensured that these differences cannot be attributed to numerical chaos, precision, compilation or dependence on particular initial condition. Implications of machine-dependence in areas like climate projections are discussed; an application-specific accuracy benchmarking is proposed. The primary conclusion from our results is that there is significant contribution from a computing platform to the quality and the accuracy of simulations, especially for complex systems. In addition to the differences in the simulations due to model formulations and sensitivity to initial conditions, a (residual) difference can remain even from the same model and initial condition simulated in different machine architecture. The machine dependence of accuracy is found to persist even beyond 30 days; it is reasonable to assume that such machine dependence will also remain in climate simulations and projections (since machine-dependence introduces error at every integration step). Considering the generally weak signal (such as trend) in climate change, projections of climate change are likely to have added and considerable uncertainty due to machine dependence. Many climate change effects have small signatures but large impact on policy; thus it is necessary to ensure that climate projections are free from uncertainties arising from machine dependence to improve the quality and the applicability of climate informatics.



Impact of computing platform on simulations different time scales and horizontal resolutions (degree of freedom).
(a) Area-averaged (Lon=75°-85°, Lat=8°-28°) daily rainfall averaged over 5 years (1982,1985, 1987, 1988, and 1990) from observation (IMD) and forecasts from the two computing systems; the number in the parentheses represents average (over 30 days) daily error for the respective forecasts.
(b) Impact of model configuration (horizontal resolution) on the difference in simulations with the two computing systems-ALTIX and ORIGIN. The solid line shows the difference in the simulations at horizontal resolution of 0.5°x0.5° of the model in area averaged daily rainfall for the month of May 2000; the dash line shows the corresponding results for a lower resolution (5°x5°).