Study of Machine Round-off response on Weather Forecasting Simulations using High Performance Computing Systems

S Janakiraman , J V Ratnam and Akshara Kaginalkar
Centre for Development of Advanced Computing
Pune University Campus
Pune - 411 007, India
jraman@cdac.ernet.in

Abstract

The weather forecasting model T80L18 is found to be sensitive to variations in the computing platform. The global spectral model simulations variation due to machine round-off is examined using rounding mode analysis and the perturbation methods. The perturbation simulations of machine round-off correctly estimates the error bounds of the T80L18 weather forecasting model.

keywords : machine rounding errors , floating-point arithmetic, global reduction calls, rounding modes, perturbations

1. Introduction

Numerical Weather Prediction is one of the computationally intensive scientific application. T80L18 weather model is a global spectral model used by National Centre for Medium Range Weather Forecasting (NCMRWF), India. It is based on the fluid dynamical equations of mass, momentum, vorticity, divergence, and thermodynamic state relation of the atmosphere. The sub grid scale phenomena such as clouds, radiation and soil vegetation are based on the empirical parameterizations. The model uses spectral transformation in the horizontal to solve the equations. It uses 80 spectral waves to resolve the quantities in the horizontal and it has 18 levels to represent the vertical. The model has been designed and optimized to run on CRAY-XMP system at NCMRWF. This model has been ported on PARAM 10000 parallel super computing system of Centre for Development of Advanced Computing (CDAC), India [4]. PARAM 10000 is a 36-node parallel computing system based on SUN Ultra 450 Enterprise Servers. Each Ultra 450 Enterprise is a SMP architecture with 4 CPUs [1].

NCMRWF T80L18 is found to be very sensitive to the computing environment. This paper attempts to study the effect of the computing platform on the medium-range weather forecasting simulation of the model. Medium Range forecast is obtained by integrating the forecast model for 5-days from the starting initial conditions. There is a pre-processor available to adjust the boundary datasets of the model. After, the forecast run, the model output is post-processed using the post-processors. In this study, few representative cases of variations in the computing platform which exhibit solution deviation to the weather forecasting model are taken in to consideration. This paper analyses the effect of machine round-off on the medium range forecasts of the T80L18 model. The sensitivity of the model is studied through dynamic rounding mode analysis and perturbation simulations of machine round-off . An attempt has been made to estimate the error growth in the solutions. Estimating the error growth has a practical utility for checking the implementation with respect to new compiling systems, new libraries and also on different computing platform.

2. Case Studies : Deviation of Weather forecasting simulations

The validated sequential version of the forecast model is compiled using SUN Fortran 77 version 4.2 with the Solaris 5.6 operating system. The simulations for the 5-day medium range forecast is conducted on a single node of SUN Ultra 450 Enterprise server of PARAM 10000. This run is referred hereafter as the 'std' run (i.e.standard configuration). It has been observed that whenever a smallest change introduced into the computing environment such as change of compiler, change in the computing platform, parallel codes etc. produce deviations in the medium range forecast as compared with the 'std' run. Here, few representative cases are presented. In all cases, daily rainfall averages has been taken for comparison. Medium range forecast is conducted for 120 hrs and the averaged rainfall values are computed at

03Z Hrs. The plots of average rainfall differences are presented for the south Asian region (0-40 N, 50 E - 110 E) for all the cases. All the simulations were conducted with the same input datasets.

2.1. Older Compiler Version

The medium range forecasts was repeated with the model compiled using the SUN Fortran 77 version 4.0 compiler. This run will be referred as 'sc4' run henceforth. The rest of the computing environment is unchanged for this forecast run. The only differences are the compiler and the associated intrinsic libraries. Examining the simulations, rainfall differences are found in the third and fourth day average rainfall. There is no difference in the rainfall averages for the first and second day forecasts. The differences in the fourth day forecast is shown in fig 1.

2.2. SGI Origin 200 platform

The sequential code was compiled using the MIPSpro Fortran 77 7.0 compiler on the SGI origin 200 system. The system is based on SMP architecture with 2 R10000 CPUs. the compilation was done with standard options (no other flags). The 5-day medium range forecast run will be referred as 'sgi' run. Like the 'sc4' run, this run showed deviation only on the third and fourth day rainfall averages. The difference in the average rainfall between 'std' and 'sgi' for the fourth day is shown in fig 2.

2.3. Parallel version of the weather model :

The parallel version of the model is designed from the base sequential model using the domain decomposition method. The parallel code employs MPI libraries (Version 1.1) for the communication and global summation operations. The parallel model was compiled using SUN Fortran 77 (version 4.2) compiler. The simulation was carried out on PARAM 10000 parallel computing system using eight processors. This run is referred henceforth as 'par8' run. 'par8' runs shows a larger deviation in the rainfall averages compared to the 'sc4' and 'sgi' runs (fig 3).

It was found that these large deviations were caused due to the use of MPI-ALREDUCE for the summation of the Legendre transform . The MPI-ALREDUCE does not preserve the order of summation operations. The parallel code was modified by replacing the MPI-ALREDUCE call with MPI-SEND and MPI-RECV calls and summing the Legendre transform in the proper order. The medium range forecast conducted with the proper order of summation is referred as 'par8sere' run. The deviation of the rainfall averages of the 'par8sere' is comparable with 'sc4' and 'sgi' runs (fig 4). the experiment conducted with the 4-processors showed similar trend but the deviations were smaller.

3. Study of Machine Round-off response for T80L18 model

All the cases discussed in the previous section can be generalized as the effect of round-off introduced by the computing environment on the simulations. In the case of 'sc4', 'sgi', the difference in the compiler and the library function could have introduced the round-off. In the case of 'par8', the round-off was introduced by the non-preservation of order of the global sums . To estimate the round off errors, the sequential version of the weather model is taken for the numerical experiments.

3.1. IEEE dynamic rounding mode analysis

The forecast model was compiled with 3 different IEEE dynamic rounding modes. The IEEE dynamic rounding modes introduce rounding errors at every computation of the weather model. The simulation made with rounding towards positive infinity is named as 'posr' run, the simulation made with the rounding towards negative infinity direction is named as 'negr' run and finally the model run with the chopped floating-point operation is named as 'zerr run. These simulations show a larger deviation in the rainfall starting from the second day simulation itself. The deviation of rainfall averages are shown in the fig 5, 6 and 7. They show a larger deviations compared with the cases shown in the previous section.

We see from the figures, that the solution produced by model running on a different platform(sgi), with old compiler version(sc4) and the parallel model(par8scre) produced much smaller deviations than the IEEE dynamic rounding mode runs. Thus rounding modes give the larger bound for the solution deviation. so this method is not useful for a numerically sensitive application such as weather models [3].

3.2. Perturbation simulations of machine-level rounding errors

Whenever a computing platform is changed or a compiler is changed, the floating-point environment introduces machine-level rounding errors at the least significant bits of the computations[5], [6]. This should be causing the deviation in the forecasting simulations when the compiler is changed or the machine platform is changed. In a validation study made by Rosinski et al [5] introduced a machine level perturbation into the input temperature field to study the error propagation in the climate model CCM2. In this study, machine rounding level perturbations are introduced into all the variables of non-static datasets. This type of perturbation is more realistic to represent the uncertainties of round off the computing platform. Such type of perturbation methods are found useful in scientific computing [2]. We attempt to simulate such rounding errors through the introduction of perturbations of the size of machine precision in to the non-static input datasets of the forecasting model. The original datasets of the model an and bd datasets are taken and the rounding errors of machine precision are introduced in to them. This is achieved by just using the directed IEEE rounding modes on the non-static an and bd datasets.

The resultant data sets will be modified by the order of machine rounding level. i.e. The precision of the data will be affected in the least significant decimal places of the datasets(see table 1). Thus the magnitude of the dataset will be preserved, but only the last few digits will be affected. Whenever there is change in the platform, the last digit of the calculation is always affected [6], This is the actual source of deviation the floating-point environment which implement IEEE 754 standard. The perturbations are created in such a way that it represents the uncertainty in the computations whenever a computing platform is changed. Also, the perturbation which results of these machine -level rounding introduces a non - linear type of variation in to the quantities. That is when the values are small it will introduce only small variation so that the magnitude is never altered. This is unlike employing the uniform perturbation function which constantly alter all the values. Also, we can notice the floating-point number distribution is not uniform. The density of the numbers for the values between 0 and 1 is very high and the number distribution decreases as the magnitude of the numbers increase. This is the property of any finite representation of real numbers . When the perturbations are introduced at the machine rounding level, this will introduce an rounding error of the order at most O(e). where e is the machine epsilon.

The dataset created by the rounding towards negative infinity direction is called as 'pert1' datasets. The dataset created using the rounding towards positive infinity direction is named as 'pert2' datasets. The datasets created by chopping is called as 'pert3' datasets. The medium range forecasts made with respective datasets are named with the name of respective dataset names. i.e. pert1, pert2, pert3 etc. the deviation of rainfall produced by the perturbation simulations are shown in fig 8, 9, 10 It clearly estimates the rainfall deviations produced by the representing cases of 'sc4', 'sgi'and 'par8sere' run.

4. Conclusions

T80L18 weather forecasting model is numerically very sensitive. This is evident from the representative cases discussed in the paper. A change of compiler or computing platform introduces deviation in the forecast simulations. Even the cause of solution deviation of the parallel model is identified. Table 2 presents the range of rainfall differences (in cm) for the entire globe for different forecast runs. The simulations conducted with the directed IEEE rounding modes introduce stronger deviations in the medium range forecasts than the deviation in the simulations for different compilers, or different computing platform. The error estimate given by the directed IEEE rounding modes is larger and hence considered to be worst case of round off propagation. The perturbation simulations clearly provides a reasonable estimate of the errors if the simulation produced by the different compilers and the associated intrinsic library or even the change in computing platform. This study proposes that the perturbation simulations at machine-level rounding errors can be a correct procedure to estimate error bounds for a sensitive applications such as weather forecasting models.

Acknowledgements : We gratefully acknowledge the scientists of NCMRWF for providing us the validated sequential code of T80L18. We thank Prof. Kanhere ,University of Pune for providing SGI Origin 200 system for conducting the simulations. An useful discussion with Prof. James Demmel, University of California, Berkeley is fondly remembered. We also thank Mr. Chetan Kumar for his help on LaTeX. Finally, we thank Dr. Purohit ,CDAC for his constant encouragement throughout this study.

References

[1] *PARAM 10000.* CDAC Technical Report, CDAC, INDIA, 1998.

[2] R. Barrio and J.-C. Berges. Perturbation simulation of rounding errors in the evaluation of chebyshev series. *Journal of Universal Computer Science*, 4(6):561–573, June 1998.

[3] D. Hough. *Personal Communication.* Consultant, Validgh, 1999.

[4] S. C. Purohit, T. V. Singh, P. S. Narayanan, and A. Kaginalkar. Global spectral medium range weather forecasting model on param. *Supercomputer*, 13(3), August 1996.

[5] J. M. Rosinski and D. L. Williamson. The accumulation of rounding errors and port validation for global atmospheric models. *SIAM Journal of Scientific Computing*, 18(2):552–564, March 1997.

[6] S. Stevenson. *Personal Communication.* Clemson University, 1999.