UPDATE ON A NEW PARALLEL SPARSE CHEMISTRY SOLVER FOR CMAQ

George Delic*
HiPERiSM Consulting, LLC, P.O. Box 569, Chapel Hill, NC 27514, USA

1. INTRODUCTION

This presentation updates the previous year's report on the new parallel sparse matrix solver FSPARSE used for chemical species concentrations in CMAQ [Delic, 2012]. This version of the Rosenbrock (ROS3) chemistry solver in CMAQ 4.7.1 (hereafter ROS3-HC) distributes different blocks of cells to separate threads in the parallel thread team. This version supersedes an earlier hybrid parallel model with three levels of parallelism as described in previous reports at this meeting [Delic,2003-2010]. In the updated version several parallel bugs have been removed and performance results evaluated. Species concentration values are compared for original and ROS3-HC methods in a 24 hour simulation. A detailed performance and numerical analysis of the first simulation hour is presented.

2. TEST BED ENVIRONMENT

2.1 Hardware

The hardware systems chosen were the platforms at HiPERiSM Consulting, LLC, shown in Table 2.1. Each of the two platforms, Intel and AMD, has a total of 8 and 48 cores, respectively. This cluster is used for either MPI only, or hybrid thread-parallel OpenMP plus MPI execution. However, to focus analysis on the new modifications only results for a single MPI processes are discussed here since they apply to each separate MPI process.

2.2 Compilers

This report implemented the Portland Group® [PGI] compiler (release 13.4) for CMAQ 4.7.1 on 64-bit Linux operating systems and hardware from the Intel Corporation (Intel) and Advanced Micro Devices (AMD) shown in Table 2.1.

3. EPISODE STUDIED

For all CMAQ 4.7.1 results reported here the model episode selected was for August 09, 2006, using data provided by the U.S. EPA. This episode has the CB05 mechanism with Chlorine extensions and the Aero 4 version for PM modeling. The episode was run for a full 24 hour scenario on a 279 X 240 Eastern US domain at 12 Km grid spacing and 34 vertical layers for a total of 2.3 million grid cells.

Table 2.1. Platforms at HiPERiSM Consulting, LLC

Platform	AMD	Intel
Processor	AMD™ Opteron 6176SE	Intel™ IA32 W5590
Peak Gflops (SP/DP)	110.4 / 55.2	106.6 / 53.3
Power consumption	105 Watts	130 Watts
Cores per processor	12	4
Power consumption per core	8.75 Watts	32.5 Watts
Processor count	4	2
Total core count	48	8
Clock	2.3GHz	3.33GHz
Band-width	42.7 GB/sec	64.0 GB/sec
Bus speed	1333 MHz	1333 MHz
L1 cache	64KB	64KB
L2 cache	512 KB(1)	256MB
L3 cache(2)	12MB	8MB
Total memory	128GB	96GB
(1) Per core, (2) Per socket		

4. PERFORMANCE TIMES

4.1 Speedup and scaling for 1 Day runs

In this section, and the next, two performance metrics are introduced to assess thread parallel performance in the ROS3-HC modified code:

(a) *Speedup* is the gain in runtime over the standard U.S. EPA runtime,
(b) *Scaling* is the gain in runtime with thread counts larger than 1, relative to the result for a single thread.

For the CMAQ chemistry solver the grid of 2.3 million grid cells is partitioned into blocks of size BLKSIZE and these blocks are distributed to threads in a thread team in ROS3-HC. In the standard U.S. EPA distribution BLKSIZE=50 is fixed, but in ROS3-HC values of 16, 32, 48, and 64 are investigated for cache effects. Table 4.1 shows wall clock times for completion of the 24 hour simulation with the U.S. EPA (ROS3-EPA) and ROS3-HC versions of the Rosenbrock solver.

* *Corresponding author:* George Delic, george@hiperism.com.

Table 4.1. Wall clock times (in hours) for the U.S. EPA (ROS3-EPA) and ROS3-HC versions of CMAQ on Intel and AMD platforms for the Portland compiler.

BLKSIZE	Platform	ROS3-EPA	ROS3-HC Time in hours by thread count				
			1	2	4	6	8
16	Intel	28.4	39.7	28.4	22.9	21.2	20.6
32		27.2	35.5	27.5	22.6	20.9	19.7
48		27.0	35.8	27.2	22.4	21.0	20.5
64		27.7	37.8	27.9	22.8	20.8	20.4
16	AMD	54.7	73.6	53.0	43.8	40.4	38.9
32		52.8	70.5	51.7	42.9	39.9	38.9
48		54.5	71.4	52.8	43.4	39.8	40.3
64		55.1	71.8	53.0	44.2	41.2	39.1

Figs. 4.1 and 4.2, respectively show the speedup and scaling results on the Intel platform, while Figs. 4.3 and 4.4 show the corresponding results on the AMD platform. On both platforms the effects of varying the BLKSIZE parameter are mild but not negligible. The choice BLKSIZE=32 is optimal for both platforms. A 1.3 speedup over the ROS3-EPA version is reached (or exceeded) with 6 threads on either platform, with small gains at a thread count of 8. However, since these performance results are for the entire CMAQ code, it is appropriate to delve further into performance in the parallel region.

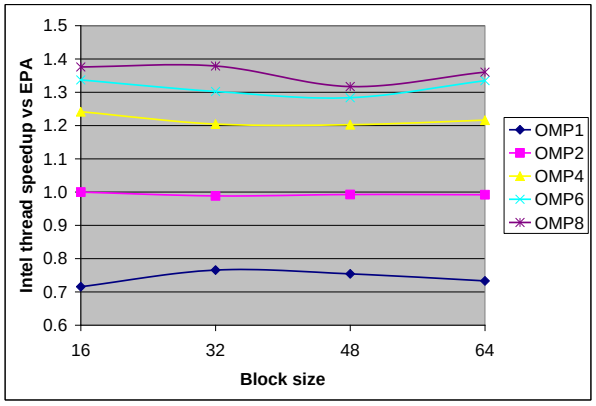

Fig 4.1: Speed up of ROS3-HC for CMAQ4.7.1 over the serial ROS3-EPA version with four different choices of the BLKSIZE parameter. The ROS3-HC results correspond to the thread counts shown in the legend on the Intel platform.

4.2 Chemistry time step iteration counts

The CMAQ ROS3 solver performs chemistry time step iterations for each block in the grid domain. Each iteration requires three solves for a linear system Ax = y, where A is sparse. The standard U.S. EPA version (ROS3-EPA) uses the Jacobson and Turco method [Jacobson, 1994] whereas ROS3-HC uses a new sparse solver method [Delic, 2012] based on Davis [Davis, 2006].

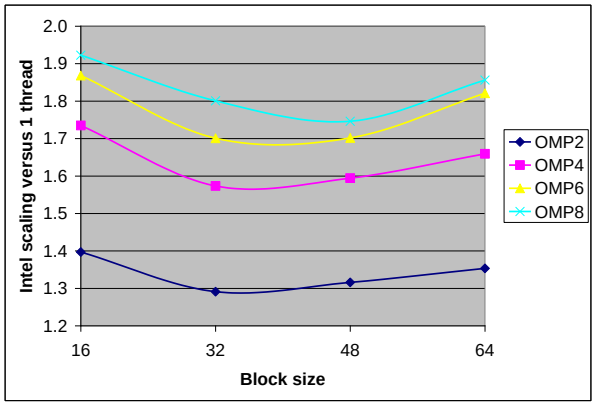

Fig 4.2: Scaling of ROS3-HC for CMAQ4.7.1 over the single thread result with four different choices of the BLKSIZE parameter. The ROS3-HC results correspond to the thread counts shown in the legend on the Intel platform.

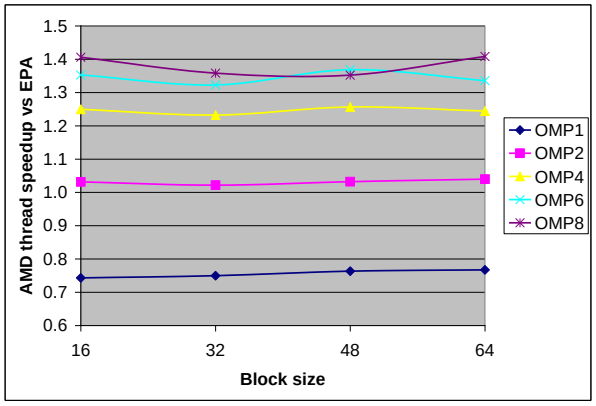

Fig 4.3: Speed up of ROS3-HC for CMAQ4.7.1 over the serial ROS3-EPA version with four different choices of the BLKSIZE parameter. The ROS3-HC results correspond to the thread counts shown in the legend on the AMD platform.

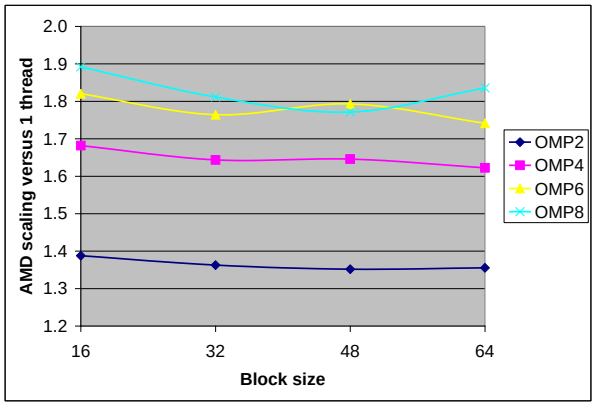

Fig 4.4: Scaling of ROS3-HC for CMAQ4.7.1 over the single thread result with four different choices of the BLKSIZE parameter. The ROS3-HC results correspond to the thread counts shown in the legend on the AMD platform.

The two sparse solver methods solve the same sparse system, but differ in the number of chemistry time steps required to solution. Specifically, the ROS3-HC FSPARSE method performs more time step iterations to reach convergence. Convergence is controlled in both methods by accuracy parameters ATOL and RTOL. The total iteration counts are compared in Fig 4.5 for the 12 calls to the sparse solver in the first hour of the simulation. Inspection of the convergence pattern shows that ROS3-EPA repeatedly exceeds the upper time step bound and terminates earlier. This is the result of less accuracy in ROS3-EPA because of numerical precision issues (see Section 5 below).

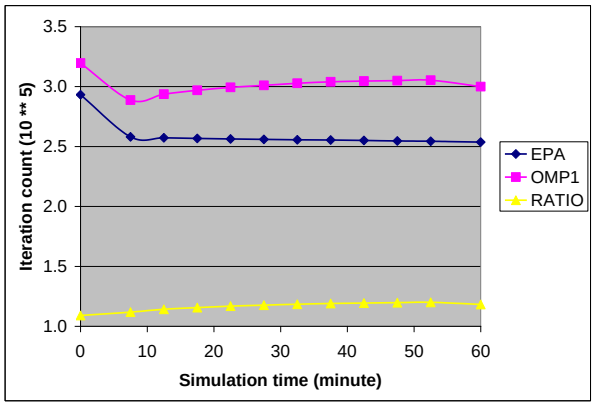

Fig 4.5: This shows the chemistry times step iteration count at each of 12 calls to the CMAQ chemistry solver. The legend indicates counts for ROS3-EPA and ROS3-HC (1 thread case) and the ratio for BLKSIZE=48.

4.3 Speedup and scaling for 1 hour runs

Fig. 4.6 shows the parallel speedup of ROS3-HC over the ROS3-EPA chemistry solver. These results are for the timings bracketing the serial and parallel regions of the chemistry solver in the respective methods. Speedup is, respectively, ~2 (4 threads), ~ 2.7 (6 threads), and 3.2 (8 threads). This is despite of the 15%-18% higher iteration count for ROS3-HC algorithm observed in Fig. 4.5. Thread parallel scaling results are shown in Fig. 4.7 and reach an outstanding value close to 5.5 with 6 threads, and slightly higher with 8 threads. Parallel efficiency is typically above 88% with 6 (or fewer) threads, and > 70% with 8 threads.

5. NUMERICAL ANALYSIS

5.1 Norms in the chemistry solver

The differences in chemistry time step iterations counts noted in the previous section suggest differing numerical results for ROS3-EPA and ROS3-HC. To understand these differences this section examines numerical metrics. These metrics show precision after the decomposition and solve steps of the sparse linear system Ax = y. Such metrics are easily monitored in ROS3-HC with an option to calculate several types of norms including |A|, |x|, and |Ax-y|. They are also extracted for the Jacobson and Turco method after a suitable modification of ROS3-EPA.

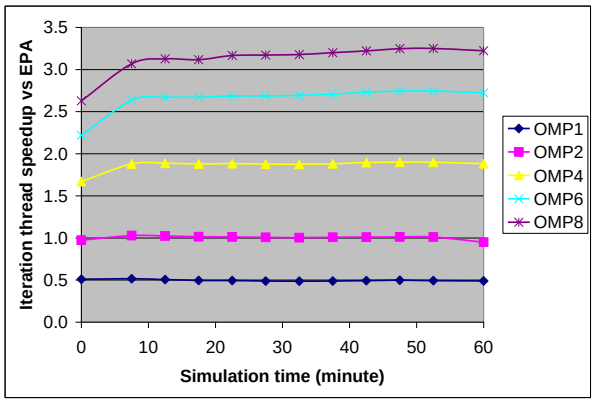

Fig 4.6: Speedup of ROS3-HC for CMAQ4.7.1 over the ROS3-EPA result for BLKSIZE=48. The ROS3-HC results correspond to the thread counts shown in the legend on the Intel platform

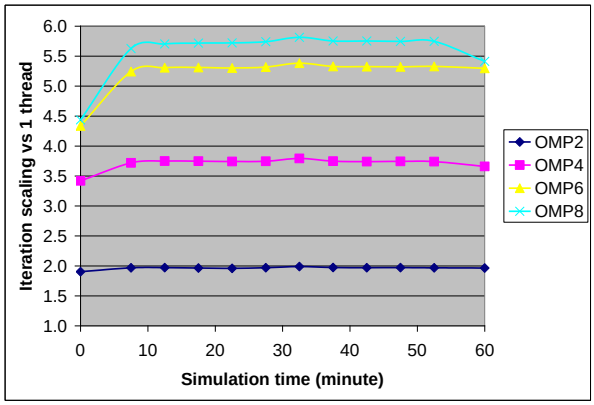

Fig 4.7: Scaling of ROS3-HC for CMAQ4.7.1 over the single thread result for BLKSIZE=48. The ROS3-HC results correspond to the thread counts shown in the legend on the Intel platform.

Table 5.1 summarizes metrics discussed here. Statistics are computed for each of 47,430 blocks (for BLKSIZE=48) after the chemistry time step iteration is completed on each call, and species concentrations saved to global storage. The length of the vector (Ax-y, or x) is the number of species. The “inf” norm selects the maximum value of each vector, Ax-y (residual), or x (solution), respectively. The statistic of Table 5.1 is then computed as

either the mean over all 48 cells in a block, or sampled for a specific cell number in the block.

Table 5.1. Metric for chemistry solver ROS3-HC of CMAQ 4.7.1 for each block of the entire domain.

value	metric	
	norm	Statistic (calculated over all cells in a block)
Residual	norm(Ax – y, inf)	mean, standard deviation, coefficient of variation.
Solution	norm(x, inf)	mean

5.2 Statistics for the residual

Fig. 5.1 shows the residual sampled for cell 48 in each of 47,430 blocks. This is for the last of three Rosenbrock solve stages at the last call to the chemistry solver in the first simulation hour. Both methods used control parameters RTOL=1.E-03 and ATOL=1.E-07. In the FSPARSE method (HC) this statistic for the residual does not exceed 1.E-18 for any of the 47,430 blocks in the entire cell domain. These results suggest that the residual remains very small indeed in the FSPARSE algorithm for the chemistry solver.

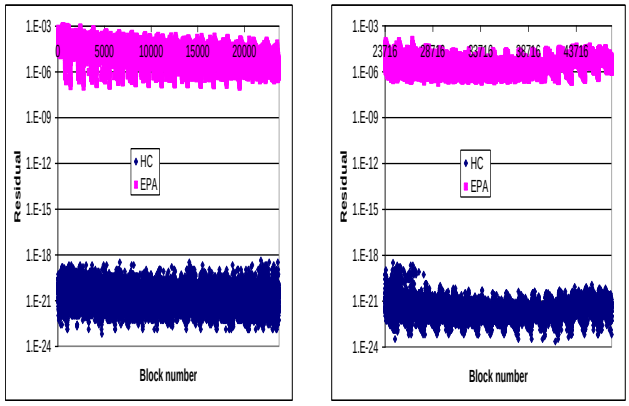

Fig 5.1: At the end of the first simulation hour this shows the norm of the residual Ax-y at the last call to the CMAQ chemistry solver sampled in cell 48 for each of 47,430 blocks (with 47,430 divided into two halves) for FSPARSE (HC) and JSPARSE (EPA) methods.

The situation is quite different in the JSPARSE (EPA) algorithm where the residual value is between 1.E-07 and 1.E-03 in Fig. 5.1. The discrepancy between the two methods is explained by the way precision is treated in each. The Chemistry solver uses double precision arithmetic but accepts some input data from single precision variables (temperature, pressure, photolysis rates, reaction rates, etc.). The U.S. EPA code implements mixed mode arithmetic and is not consistent in promoting constants or variables from single to double precision arithmetic. This is particularly egregious in the CALCKS procedure where thermal and photolytic reaction rates are computed using single precision arithmetic. As a consequence, in the U.S. EPA version of CMAQ, matrix values in the sparse matrix, A, loose precision when inheriting single precision values. This loss in precision is amplified as the solution progresses in the three Rosenbrock solve stages. Inspection showed that, at worst, 5 significant figures are achieved for the largest elements of A, and considerably less for the smallest matrix elements. Therefore using an ATOL=1.E-07 is moot for the ROS3-EPA JSPARSE method.

In the ROS3-HC FSPARSE algorithm arithmetic consistency has been implemented throughout the chemistry solver and because of the higher accuracy, a reduced ATOL value could be tolerated, to reduce the number of chemistry time step iterations in ROS3-HC. This is confirmed by comparing results in Fig. 5.2 (ATOL=1.E-05) with Fig. 4.6 (ATOL=1.E-07 the default used in ROS3-EPA). Note the shift in scale of the Y-axis.

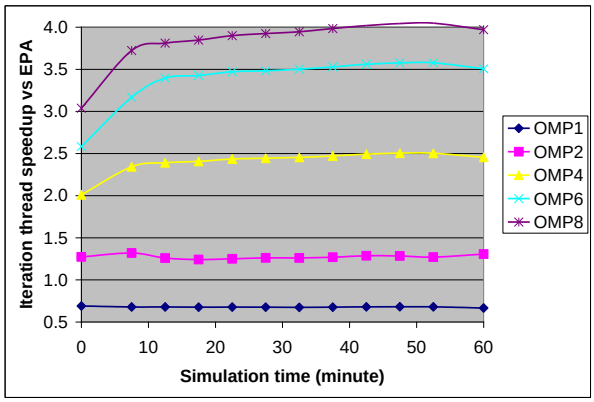

Fig 5.2: Speedup of ROS3-HC (ATOL=1.0E-05) for CMAQ4.7.1 over the ROS3-EPA (ATOL=1.0E-07) result for BLKSIZE=48. The ROS3-HC results correspond to the thread counts shown in the legend on the Intel platform.

Table 5.2. Wall clock times (in hours) for the U.S. EPA (ROS3-EPA) and ROS3-HC versions of CMAQ on Intel and AMD platforms for the Portland compiler. Numbers in italics show the speed up versus ROS3-EPA.

BLKSIZE	Platform	ROS3-EPA	ROS3-HC (ATOL=1.E-05) Time in hours by thread count				
			1	2	4	6	8
48	Intel	27.0 *1.0*	33.5 *0.81*	26.0 *1.04*	21.3 *1.27*	20.1 *1.34*	20.0 *1.35*
48	AMD	54.5 *1.0*	62.6 *0.87*	47.8 *1.14*	41.2 *1.32*	38.3 *1.42*	36.9 *1.48*

Sample times for the 24 hour simulation with ATOL=1.E-05 are shown in Table 5.2 (compare with Table 4.1 for ATOL=1.E-07). Speed up versus ROS3-EPA reaches 1.35 (Intel) and 1.48 (AMD).

5.3 Species concentrations at hour 23

A direct comparison of accuracy is for species concentration values predicted by the U.S. EPA and HiPERiSM versions of CMAQ. For this purpose the respective output files were used to extract all 66,960 concentration values of each selected species in layer 1 at the 23rd hour of a 24 hour simulation. The concentration values and their differences (ROS3-EPA – ROS3-HC) were sorted into histograms with 7 decade bins at boundaries between [0,1.E-07] and [1.E-02 ,1.E-01]. Such histograms show the frequency of occurrences as a percentage of the 66,960 values in a species sample. Species selected for this investigation included O3, NO2, NO, ISOP, H2O2, FORM, CO, ASO4J, ASO4I, ANO3J, and ANO3I.

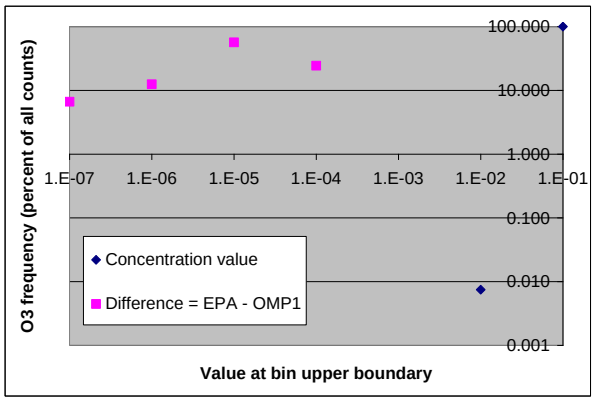

Fig 5.3: For hour 23 of a 24 hour simulation this shows the O3 concentration value and difference in values (EPA – OMP1) between ROS3-EPA and ROS3-HC.

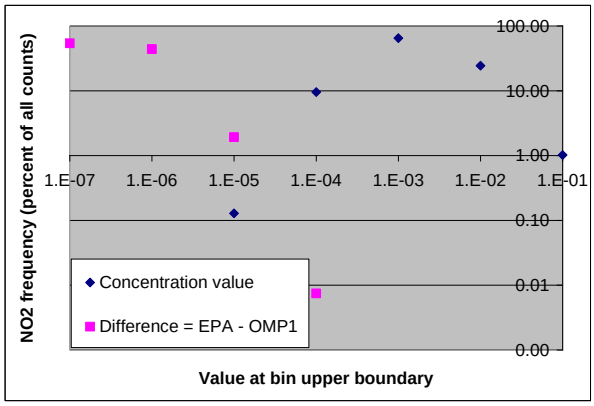

Fig 5.4: For hour 23 of a 24 hour simulation this shows the NO2 concentration value and difference in values (EPA – OMP1) between ROS3-EPA and ROS3-HC.

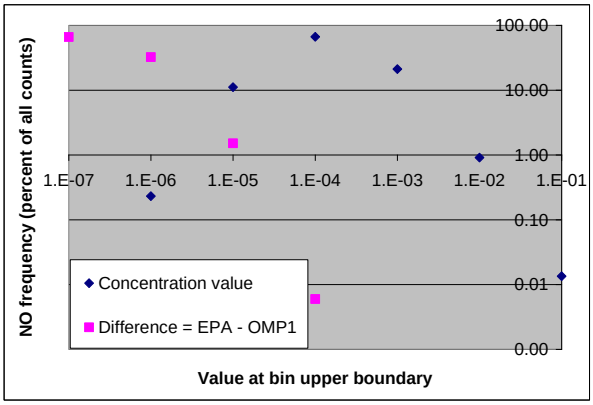

Fig 5.5: For hour 23 of a 24 hour simulation this shows the NO concentration value and difference in values (EPA – OMP1) between ROS3-EPA and ROS3-HC.

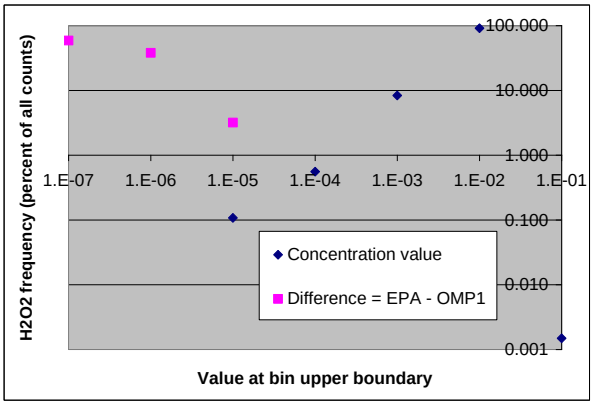

Fig 5.6: For hour 23 of a 24 hour simulation this shows the H2O2 concentration value and difference in values (EPA – OMP1) between ROS3-EPA and ROS3-HC.

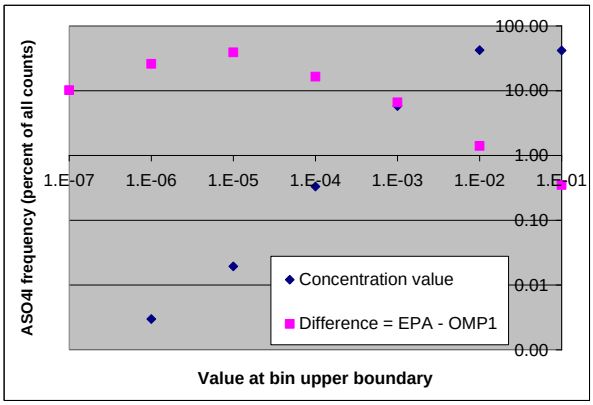

Fig 5.7: For hour 23 of a 24 hour simulation this shows the ASO4I concentration value and difference in values (EPA – OMP1) between ROS3-EPA and ROS3-HC.

Selected results are presented in Figs. 5.3-5.8 with the vertical scale as the frequency (fraction of sample) falling into the decade bin below the upper limit shown on the horizontal scale.

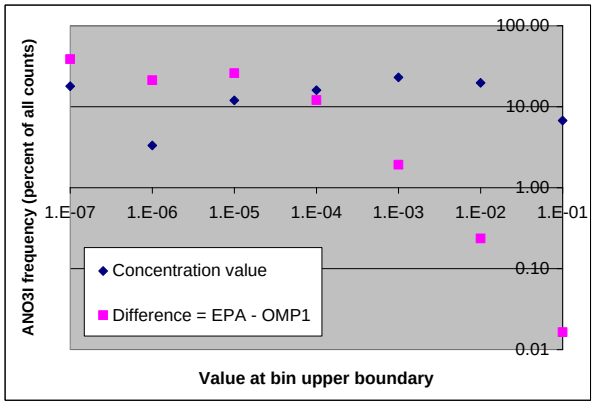

Fig 5.8: For hour 23 of a 24 hour simulation this shows the ANO3I concentration value and difference in values (EPA – OMP1) between ROS3-EPA and ROS3-HC.

Concentration values have higher frequency in the upper bin ranges and the difference in predicted values cluster in the lower bin ranges. The trend is that the largest difference in predicted concentrations occurs when the concentration value is smallest. The widest separation in the two trends is for O3, with somewhat closer overlap for NO2, NO, and H2O2. However, differences only occur for fractions of a percent of the species sample size. The worst cases are for ANO3I where the concentration value histogram is broadly scattered over many orders of magnitude and some 25% of the population sample could have errors in the 2nd significant figure.

6. LESSONS LEARNED

6.1 Benefits of the FSPARSE method

Comparing runtime performance for CMAQ 4.7.1 in the new OpenMP parallel version with the U.S. EPA release showed benefits such as:

- A speedup ~1.4 with 8 parallel threads.
- Parallel efficiency that was ~88% with 6 threads, and > 70% with 8 threads.

6.2 Numerical precision issues

A comparison of numerical precision for CMAQ 4.7.1 in the new OpenMP parallel version with the U.S. EPA release used metrics of the residual in the Rosenbrock solver to show:

- Limitations due to the EPA method's inconsistent use of mixed mode arithmetic.
- The FSPARSE method was more precise by many orders of magnitude.
- The FSPARSE method allows a relaxation of the chemistry time step convergence error criterion that also reduces runtime.

6.3 Comparing species concentrations

A comparison of species concentration values predicted by ROS3-EPA and ROS3-HC showed:

- Good agreement for species such as O3, NO2, NO, H2O2.
- Degraded agreement for species such as ASO4I.

Such differences in species concentration values could be due to cumulative error propagation in the U.S. EPA method.

7. CONCLUSIONS

This report has described an analysis of CMAQ 4.7.1 behavior in the standard U.S. EPA release and a new thread parallel version of CMAQ for the Rosenbrock solver. Opportunities exist for speedup with an increasing number of parallel threads that reaches the range 1.4-1.5 over the standard CMAQ release with 8 threads on Intel or AMD platforms. However, numerical precision issues were observed and are due in part to the way arithmetic precision is treated in the U.S. EPA method.

Further opportunities remain for thread parallelism in other parts of the CMAQ model outside of the solver and work in this direction continues at HiPERiSM Consulting, LLC. The new (second) version of ROS3-HC offers layers of parallelism not available in the standard U.S. EPA release and is portable across hardware and compilers that support thread parallelism. Results with the Intel compiler [INTEL] are forthcoming.

REFERENCES

Davis, T.A., Direct Methods for Sparse Linear Systems, SIAM, Philadelphia, 2006.

Delic, G., 2003-2010: see presentations at the Annual CMAS meetings (http://www.cmasecenter.org).

Delic, G., 2012: see presentation at the Annual CMAS meetings (http://www.cmasecenter.org).

INTEL: Intel Corporation, http://www.intel.com

Jacobson, M. and Turco, R.P., (1994), Atmos. Environ. 28, 273-284

PGI: The Portland Group http://www.pgroup.com