

Reply to Referee 3:

Thank you for the suggestions you provided. They are detailed and crucial for us to improve the manuscript.

General comments

Referee 3: The research paper focuses on the utilization of MIPS CPUs, particularly the Loongson 3A4000 CPU platform, in air quality prediction models. It evaluates the performance of the WRF-CAMx air quality modelling system in the Beijing-Tianjin-Hebei region using this platform. The study compares the MIPS CPU platform's performance with a benchmark X86 platform, analyzing various aspects like relative errors for major species, computational efficiency, and energy consumption. The results indicate the feasibility and efficiency of using MIPS architecture for such applications. This work has the potential to offer valuable guidance for using the MIPS platform for geoscientific modeling. I would suggest that the authors provide a more in-depth discussion on how to exploit the advantages of the MIPS platform. The structure of the paper could be improved, and additional tests are necessary to demonstrate the MIPS platform's performance.

Reply: Thank you very much for your appreciation of the potential practical value of this research. The MIPS platform with Loongson 3A4000 CPU in the manuscript has limitations in hardware performance and software ecosystem. Firstly, the platform only supports the GNU compiler and specific Linux operating system currently. Therefore, the performance evaluations can only be conducted in a certain environment. Secondly, due to the limited number of CPUs and cores on the platform used in this work, the maximum parallelism supported by the model system is 4. When the number of parallel processes exceeds 4, competition for computational resources restricts further performance evaluation of the platform. Additionally, the simulation case set up in this study fully utilize the computational resources of the platform, with the CPU utilization approaching 100% for each core during the simulation. Therefore, performance evaluation with higher parallelism and larger computational scales is challenging on this platform, but needs more computational resources, which exceeded our capability. However, the process-based MPI and the thread-based OpenMP were evaluated in the study. This provides a certain level of representativeness, as illustrated in “Figure 9” of the manuscript, as shown as Figure 1 in the followed. What’s more, the lack of the effective software tools to provide more detailed assessments of the scientific computing capability on the platform is also an objectively existing issue. In the future, with the improvement of the ecosystem and performance of the MIPS and LoongArch platforms, additional tests will be considered to further demonstrate the platform's performance and discuss how to exploit the advantages. Recently, we acquired a platform equipped with the Loongson 3A6000 CPU, which is the latest product released by Loongson Technology. In order to enhance the universality of our research, we conducted the same tests based on this platform, the CAMx modeling system can run

stably on the LoongArch platform, and the performance evaluation is shown in Figure 2 in the followed. Relevant descriptions and results will be added to the revised manuscript. Additionally, the structure of the paper will be improved.

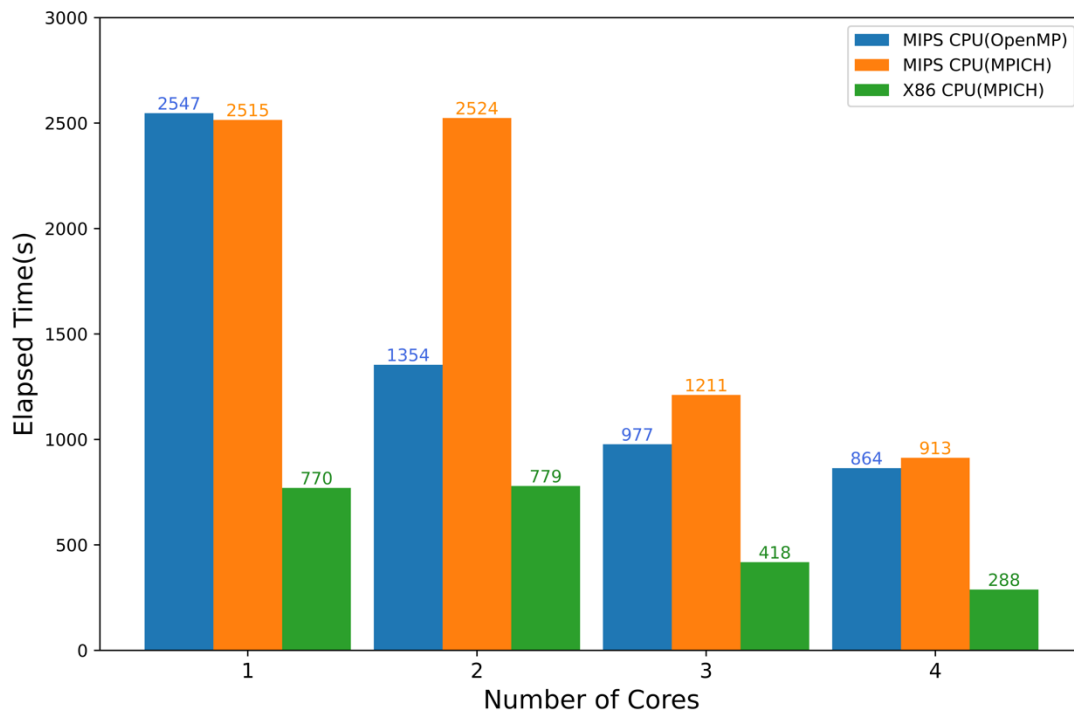


Figure 1. Elapsed time of CAMx model running simulation case for 2 hours on the MIPS and benchmark platforms.

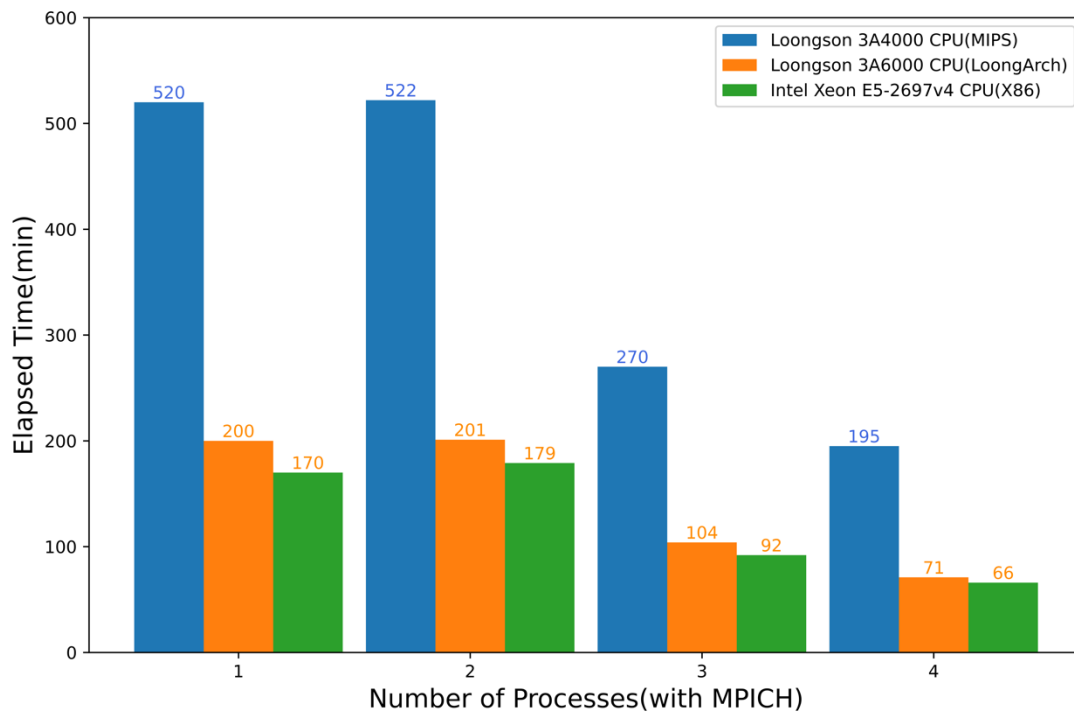


Figure 2. Elapsed time of CAMx model running the same simulation case with MPICH

for 24 hours on the MIPS, LoongArch and benchmark platforms.

Major comments and questions

Referee 3: Every abbreviation that appears in the paper, including in the abstract and the main body, should be spelled out in full the first time it is used. For example, the abbreviations 'MIPS' and 'WRF-CAMx' are not spelled out in the abstract.

Reply: Thank you for the issues you pointed out. All abbreviations which appear for the first time will be carefully reviewed and corrected in the revised manuscript.

Referee 3: The model setups and analysis methods used in this paper should be presented prior to the results in Section 4. The content in Lines 305-309, 323-325, and 404-411 should be consolidated in Sections 2 or 3 as part of the methodology.

Reply: Thank you for the suggestions about the organization and structure of the article. These will be carefully considered and reflected in the improvement of the revised manuscript.

Referee 3: I am curious about the number of sockets available on the motherboard for the Loongson 3A4000 platform. Could the author conduct a larger-scale comparison using more Loongson 3A4000 CPUs compared to the X86 platform as shown in Figure 9?

Reply: The motherboard of the Loongson 3A4000 platform used in this study supports only one CPU with four physical cores. We will consider seeking platforms with more Loongson 3A4000 CPUs for larger-scale tests in the future.

Referee 3: Could the author investigate the impact of using different compilers or different compiler parameters on computational performance, in addition to the GNU?

Reply: Currently, the MIPS platform with Loongson 3A4000 CPU only supports the GNU compiler. Other compilers have not been adapted for the MIPS architecture. As for compiler parameters, the default options for MIPS GNU are fixed. They are used to specify the architecture of the target platform and optimize the target program based on specific instruction sets. There are no compiler parameters that significantly impact computational performance.

Minor comments

Referee 3: Line 92: Remove “The remainder is organized as follows.”

Reply: This statement will be removed in the revised manuscript.

Referee 3: Line 113-115: Rephrase this sentence. The WRF is developed by NCAR, and CAMx is developed by Rambell. WRF-CAMx is just applied in Xi'an, China and Milan, Italy (not Europe).

Reply: This part of the statement is not clear and accurate enough. The intended meaning here is that the WRF-CAMx modeling system suitable for specific regions (such as Xi'an and Milan) has been applied in the research cited in the references. Relevant statements will be refined in the revised manuscript.

Referee 3: Line 123-126: The introduction for WRF is not professional. WRF is a meso-scale meteorology model, and you can use it for weather research and prediction. It can be used with a data assimilation technique, and testing its parameterization schemes is a way to improve WRF.

Reply: The description about the WRF model is not detailed and accurate enough, and it will be improved based on your suggestions in the revised manuscript.

Referee 3: Line 152-154: Why you used 14 layers not original 34 layers?

Reply: In practical applications of air quality simulation, we primarily focus on the concentration of air pollution species near the land surface. Therefore, the analysis of simulation results often involves extracting data from the near-surface layer for the CAMx model with wrfcamx module. This approach can significantly save computational resources in the air quality simulation.

Referee 3: Line 195-196: Was FFT used or related to this paper? If not, please remove it.

Reply: FFT was not used in this paper. The citation is intended to illustrate that there has been previous research on the application of scientific computing on MIPS platforms with Loongson CPUs. However, there is scarce research specifically focusing on large-scale applications like numerical models, with most studies centered around specific algorithms or programs (such as FFT). Therefore, we can only reference these studies. This part will be considered to remove in the revised manuscript.

Referee 3: Line 436: Give full names to RMSE, std. Why std are in lowercase but RMSE is not? Also, what's the statistic meaning or purpose of using the ratio of RMSE/STD?

Reply: The full names of RMSE (Root Mean Square Error) and std (standard deviation) will be added to the revised manuscript. The lowercase of std does not have a specific meaning. It is only used to distinguish it from statistical indicators such as MAE and RMSE, which are applied to analyze simulation differences between platforms. 'std' is only used to describe the dispersion of a certain species' concentration in the simulation

results. The ratio of RMSE/ STD does not have a specific statistical meaning. Wang et al. (2021) introduced it as a metric to assess the scientific usability of simulation results after model improvement. This has been recognized in academic research. In our study, the extremely small ratio of RMSE/STD means the simulation differences between the two platforms are negligible compared to the spatial variations of simulation results.

Citation: Wang, P., Jiang, J., Lin, P., Ding, M., Wei, J., Zhang, F., Zhao, L., Li, Y., Yu, Z., Zheng, W., Yu, Y., Chi, X., and Liu, H.: *The GPU version of LASG/IAP Climate System Ocean Model version 3 (LICOM3) under the heterogeneous-compute interface for portability (HIP) framework and its large-scale application*, *Geosci. Model Dev.*, 14, 2781–2799, <https://doi.org/10.5194/gmd-14-2781-2021>, 2021.