Computers and Electronics in Agriculture 111 (2015) 48–54

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

journal homepage: www.elsevier.com/locate/compag

Development of mpi_EPIC model for global agroecosystem modeling

S. Kang [a,*], D. Wang [a], J.A. Nichols [b], J. Schuchart [c], K.L. Kline [a], Y. Wei [a], D.M. Ricciuto [a], S.D. Wullschleger [a], W.M. Post [a], R.C. Izaurralde [d]

[a] Climate Change Science Institute, Environmental Sciences Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA
[b] Computational Sciences and Engineering Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA
[c] Center for Information Services and High-Performance Computing (ZIH), The Dresden University of Technology, 01062 Dresden, Germany
[d] Joint Global Change Research Institute, Pacific Northwest National Laboratory, College Park, MD 20740, USA

ARTICLE INFO

Article history:
Received 27 June 2014
Received in revised form 14 November 2014
Accepted 8 December 2014

Keywords:
Bioenergy
Food
High performance computing (HPC)
Message passing interface (MPI)
Parallel design
Sustainability

ABSTRACT

Agroecosystem models that can incorporate management practices and quantify environmental effects are necessary to assess sustainability-associated food and bioenergy production across spatial scales. However, most agroecosystem models are designed for a plot scale. Tremendous computational capacity on simulations and datasets is needed when large scales of high-resolution spatial simulations are conducted. We used the message passing interface (MPI) parallel technique and developed a master–slave scheme for an agroecosystem model, EPIC on global food and bioenergy studies. Simulation performance was further enhanced by applying the Vampir framework. On a Linux-based supercomputer, Cray XT7 Titan, we used 2048 cores and successfully shortened the running time from days to 30 min for a global 30 years of modeling of a bioenergy crop at the resolution of half-degree (62,482 grids) with the message passing interface based EPIC (mpi_EPIC). The results illustrate that mpi_EPIC using parallel design can balance simulation workloads and facilitate large-scale, high-resolution analyses of agricultural production systems, management alternatives and environmental effects.

1. Introduction

The intertwined global issues of food security, bioenergy and climate change call for large scale high-resolution analysis of agricultural landscapes (Lynd et al., 2011; Lobell et al., 2011; Savage, 2013). Agroecosystem modeling provides opportunities for decision makers to evaluate trade-offs among various scenarios and management options. Spatial variability in climate (e.g., precipitation, temperature), soils, and crop management practices (e.g., fertilizer, irrigation, tillage, chemical control and harvest) affect crop productivity and environmental effects. Ideally, decision makers would prefer to guide management practices and production systems toward more sustainable options. But enormous data sets must be integrated and processed if modeling is to explore the implications of alternative land management across spatially heterogeneous scales. This is especially true when land managers or decision-makers want to evaluate trade-offs between management practices, environmental impacts and aspects of the sustainability of production systems in different locations.

The incorporation of spatially explicit data for climate, biophysical and crop management variables requires enormous computational capacity, particularly when the scope of analysis expands from plot-scale to larger scales, such as county, state or global. Traditional agroecosystem models, particularly biophysical process-based models, are designed for single-site agricultural research tests, production potential assessment and management practice optimization. They are valuable tools for studying novel cropping systems, new management practices at specific sites or local analyses. However, applying these process-based models at national and global scales is challenging due to the requirements for input data, calibration, validation and simulation setups appropriate for each of thousands to millions of spatial points (Nichols et al., 2011; Wang et al., 2012).

Super-computing advances have opened the door to detailed, global-scale modeling simulations that were previously considered impractical if not impossible. Several preliminary studies have demonstrated the potential of high-performance computing (HPC) application to large-scale data management, modeling and analysis of agroecosystem functions. For example, Hawick et al. (1997) proposed to use an HPC-based geographic information system for storage, access and processing of weather, soil, and land management data. Kuo et al. (2004) designed an HPC-visualization

* Corresponding author. Tel.: +1 865 574 5948.
E-mail address: kangs@ornl.gov (S. Kang).

http://dx.doi.org/10.1016/j.compag.2014.12.004
0168-1699/

tool for use in displaying a three-dimensional environment for farming system analyses. Wang et al. (2005) describe component-based software architecture for spatially explicit ecosystem models that could be adapted to address the large scale high-resolution ecosystem simulation challenge. High-resolution modeling with a parallel job launcher was applied to study climate change impacts on pasturelands on a 200-processors cluster machine (Vital et al., 2013). HPC can facilitate detailed management optimization at local scales, and can help aggregate regional, national and global distributions of production and environmental impacts as a tool for decision makings at larger spatial scales.

Demonstrations of regional and national high-resolution agroecosystem modeling under an HPC environment are limited (Nichols et al., 2011; Zhao et al., 2012). Nichols et al. (2011) constructed a modeling system to conduct regional high-resolution (30 m) assessment of production and environmental effects for cropping systems in the Midwest US with the high-performance computing Environmental Policy and Integrated Climate (HPC-EPIC) model. The study assigned jobs or simulation packages concurrently to multiple processors by "embarrassingly parallel" design – also referred to as "pleasingly parallel" for work that can be distributed because it requires little or no communication of results among tasks. This approach used Linux-based computing cluster and achieved a 40-fold reduction in the time required to run 140,000 EPIC simulations. A hybrid HPC approach was adopted to simulate national agricultural systems of Australia on a heterogonous distributed computing grid (Zhao et al., 2012). Similar to the embarrassingly parallel design in HPC-EPIC, jobs or simulation packages were sent to the individual computers on the grid. The hybrid HPC approach accelerates most processing by 1000-fold and completes jobs within a few days rather than months. However, both approaches are challenged with simulation load balancing which can delay overall execution time due to just a few extremely slow or failed executions. Slow or failed executions on a few nodes in clusters or an individual computer on a hybrid computing grid often occur (Nichols et al., 2011; Zhao et al., 2012). A few failed executions can cause significant processing delays when millions of simulations are conducted for large scale of high-resolution ecosystem studies involving massive data management (I/O processing) and bandwidth/memory.

To meet the heavy computational load of large-scale high-resolution agroecosystem simulations, we developed a design to address the problem of simulation load balancing using the message passing interface (MPI) on clusters or supercomputers. We directly coded MPI routines into a widely applied agroecosystem model, EPIC, using a master/slave scheme to dynamically manage model input, outputs and simulations. An application of global half-degree biomass productivity analysis was used to test and evaluate this new mpi_EPIC approach to agroecosystem modeling.

2. mpi_EPIC development and computational improvement

2.1. Description of the EPIC model

The EPIC model is a process-based agroecosystem model representing soil–crop–atmosphere interrelationships with various management scenarios (Williams et al., 1984). Major processes in the model include plant growth, development and production, nutrient cycling and nonpoint sources pollution, emissions of greenhouse gases, and plant management practices. With continuing testing and improvement, the EPIC model has become a widely applied tool designed to meet needs for production estimation, environmental effect assessment, and climate change studies (Easterling et al., 2001; Mausbach and Dedrick, 2004; Gassman et al., 2005; Liu et al., 2008; Balkovič et al., 2013; Rosensweig et al., 2014). Currently, EPIC is able to simulate over 100 crops and has been calibrated and applied in over 30 countries and regions (Gassman et al., 2005).

2.2. mpi_EPIC design

Each EPIC simulation represents a single location or point, described by different input variables such as site characteristics (e.g. location, slope, and soil), crop rotation, management practices, and weather. Usually, the individual simulations are of short duration, typically 20–30 s for a 30 year daily time step simulation. However, the computational time can vary widely depending on the complexity of involved biogeochemical processes. For example, the numerical processes simulating soil water and nutrient transport, and gas diffusion in soils can be time-consuming due to multiple layers of soils and management practices. Since some points are more computationally demanding than others, the total time required to complete a full set of simulations can be reduced by distributing the work load to processors that have completed one job, and are available for another. This "load balancing" can be done under parallelization schemes, for example, by managing jobs among available cores on a supercomputer.

To accelerate high-resolution modeling at large scales, e.g., national and global ecosystem studies, we redesigned the EPIC module structure and calling sequence to be capable of dynamically balancing simulations across multiple cores of clusters or supercomputers. In this study, we adopt MPI to distribute site-specific simulation jobs among cores. A master/slave paradigm is applied to balancing job assignments and input control (Fig. 1). The master process generates and schedules thousands of jobs and sends them to each slave process for execution. The slave processes dynamically communicate with the master process for job completion and new job assignments. If a simulation fails, the simulation ID is sent to master process for check by users, and the next simulation will be continued. The failed simulations can be further examined through output log information.

2.3. Titan and Lustre file system advantages associated with high-resolution simulations

The computational platform used in this study is the Cray XT7 Titan supercomputer at the National Center for Computational Sciences (NCCS) at Oak Ridge National Laboratory (ORNL). Each Titan node consists of a 16-core AMD 6274 Opteron processor in conjunction with Nvidia Tesla K20X GPUs to boost its computational capabilities and improve energy efficiency. The 18,688 of these nodes perform at a theoretical peak of 27 PetaFLOPS (27 quadrillion floating point operations per second). A center-wide Lustre file-system provides 5 PB of disk space for all NCCS computing resources. The NCCS has implemented customized queue policies to encourage large jobs to run in a timely fashion on the systems at ORNL.

In order to obtain detailed information about how data are processed in our global simulation using mpi_EPIC in a parallelization design, the Vampir framework (Knüpfer et al., 2006) is used to gather information on simulation characteristics for diagnosis. The Vampir framework consists of two components, VampirTrace and Vampir. VampirTrace is used to instrument the source code of the program and to manage the recording of events while the instrumented program is running. Vampir is used to visualize the data/events gathered during the program execution. Within the Vampir framework, the instrumentation could be done via compiler instrumentation or using the Tuning Analysis Utility (TAU) instrumentor (Shende and Malony, 2006). A detailed explanation on how to configure the Vampir framework and TAU for environmental simulation can be found in Domke and Wang (2012).

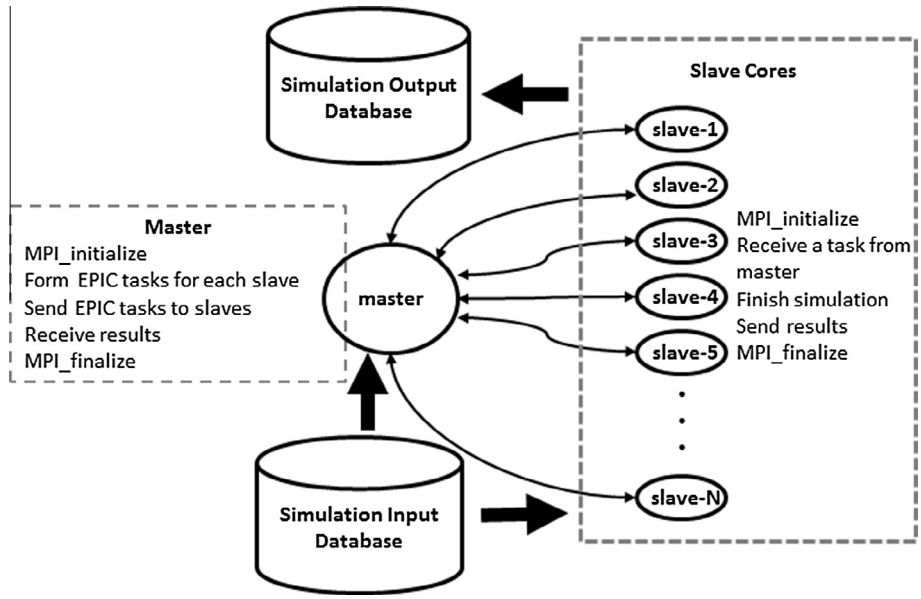

Fig. 1. Message passing interface (MPI) parallel design of the mpi_EPIC model for large scale high-resolution agroecosystem modeling. Through a mast/slave scheme, jobs are dynamically assigned to slave cores by the master core on computing clusters or supercomputers.

2.4. mpi_EPIC parallel coding, testing and computational improvement

Our tests indicate that introducing a MPI parallel scheme to the existing sequential Fortran-based code of EPIC represents an important improvement in computational efficiency. This was achieved by splitting the standard EPIC model into two sub-modules. A new main routine (master) handles inputs and outputs (e.g., site, weather, parameters) and coordinates task assignments. The master also includes the Fortran-based MPI parallel scheme, which employs a master/slave algorithm. The slave module receives information from the master and reports results back. The biophysical processes simulations run by each slave are unchanged from the original EPIC.

After the coding was completed, a series of tests were conducted to examine the effectiveness of the parallel design, outputs, and load balancing. We used the parallel debug tool, Distributed Debugging Tool (DDT) to diagnose the performance and debug the errors on Titan. The outputs including format and values from mpi_EPIC were carefully compared with the outputs from the EPIC model (without the MPI modification) when processing the same simulation input data. We found that mpi_EPIC was able to duplicate the format and values of EPIC without error.

We examined load balancing (Fig. 2a) through a series of standard test jobs on mpi_EPIC. The tests illustrated that it took significantly different amounts of time to execute the same number but different simulations. For example, a job on processor 175 used over 120 s to complete, but it took under 40 s for another job with the same number of simulations on processor 170 (Fig. 2a). After we examined the execution time of those model components that were most time-consuming, we found that some simulations took longer because of numerical procedures used to solve subsurface biogeochemical processes. The simulations illustrated in Fig. 2a applied a standard job size of 20. Through a series of tests considering two to one hundred simulations per job, we found that a job size with five was the optimum for the mpi_EPIC simulations on Titan. We therefore adjusted the job size (number of simulations sent to a slave) to five when the master/slave scheme was applied and found similar execution times across cores (Fig. 2b). The mpi_EPIC using parallel design and optimized job size required 28 min to complete the same global simulation that required over two weeks using a sequential execution on a single core (Table 1). We also documented the time required to complete the same simulation test package with HPC-EPIC (embarrassingly parallel design) (Nichols et al., 2011). Table 1 illustrates that mpi_EPIC also reflects an improvement relative to HPC-EPIC, despite the different types of cores and I/O systems used in Titan and Oak Ridge Institute Cluster (OIC). Of the approaches tested, mpi_EPIC offers the greatest capacity to efficiently complete a large number of agroecosystem simulations when using computing clusters or supercomputers.

Next, we conducted a scalability analysis of mpi_EPIC to determine to what extent we could reduce the execution time by adding computing power (additional cores) to complete simulations. The total computational time for identical global simulation setups with computing cores ranging from 16 to 2048 was measured (Fig. 3). We observed nearly linear scalability in terms of proportional time reduction with each additional core when 128 or fewer cores were applied in the tests. This scalability fell slightly when more computing cores beyond 128 cores were used (Fig. 3). As more cores or computing resources are available, the speed of global simulation can be considerably increased. Although these tests were conducted on Titan, a super computer at ORNL, the model can be used similarly in the other computing clusters for applications.

3. mpi_EPIC application to large-scale agroecosystem modeling

3.1. Modeling workflow

In order to test mpi_EPIC, we conducted a study of global biomass productivity on Titan. The workflow was applied for a global half-degree assessment of production and environmental effects (Fig. 4). This design integrates global natural resources data sets, local management scenarios, model calibration and validation, simulation, and simulation output processing, and result analysis. The database of natural resources and management provided model inputs of weather, soil property, slope, land use and various management practices commonly used by biophysical crop models. This workflow can also be applied to other large-scale, high-resolution crop modeling assessment to consider effects on agroecosystems.

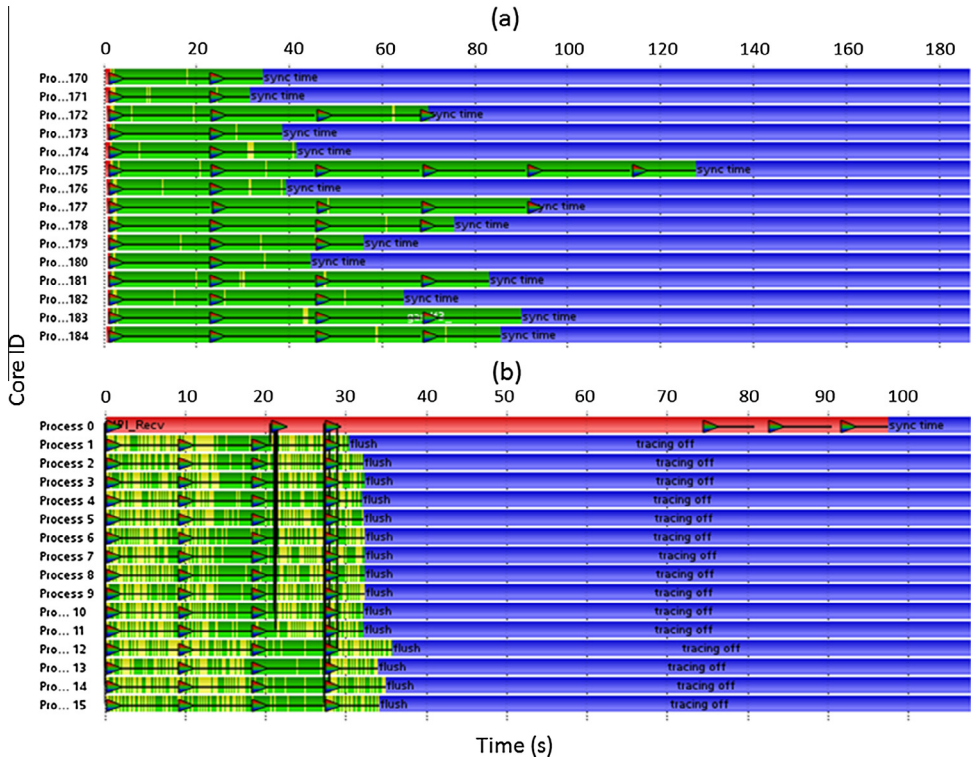

Fig. 2. Diagnosis and improvement of parallel design of the mpi_EPIC model using Vampir tools on the Titan, a supercomputer at Oak Ridge National Laboratory: (a) execution time (in seconds) for each job in different cores (processers) before the improved master–slave design, and (b) execution time (in seconds) in different cores (processers) after the improved master–slave design. Note that a different time scale is used in (b). The green represents physical-process simulation, the yellow for I/O, and the red for master core communication processing. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 1
Execution time (min) required to model global switchgrass productivity and environmental effects using mpi_EPIC on the Titan supercomputer and HPC-EPIC on the OIC computing cluster at Oak Ridge National Laboratory (Nichols et al., 2011), USA. The standard test involved over 1500 individual EPIC simulations. The times reported here for mpi_EPIC reflect the optimized job-size of five simulation routines per job.

# Cores	Titan (mpi_EPIC)					OIC (HPC-EPIC)	
	1	16	64	128	2048	1	50
Execution time (min)	18,184	1221	308	156	28	19,527	935

3.2. mpi_EPIC model input processing and calibration

Model inputs used for mpi_EPIC include three major categories of data (weather datasets, soil properties, and landscape attributes). We generated 62,482 simulation units using the half-degree global land mask from the Climate Research Unit – National Centers for Environmental Protection (CRU–NCEP) (Viovy, 2008) and 30-year daily weather input data (1980–2010) for each simulation unit. Daily CRU-NECP data include four sets of 6-h weather data, total shortwave solar radiation (W m^{-2}), air pressure (Pa), temperature (K), U-wind (m s^{-1}), V-wind (m s^{-1}), and precipitation (mm). The daily radiation (MJ m^{-2} d^{-1}), maximum temperature and minimum temperature (°C), precipitation (mm), average relative humidity (%) and average wind speed (m s^{-1}) were calculated from CRU-NECP data. Next, we processed soil property data for each simulation unit using the Harmonized World Soil Database (HWSD) (FAO/IIASA/ISRIC/ISSCAS/JRC, 2008). Contents of sand, silt, clay, gravel content, bulk density, and soil organic carbon as well as cation exchange capacity, texture, and pH of two soil layers (0–50 cm, 50–100 cm) from the dominant soils were assigned to each simulation unit from the polygon-based HWSD. Slope data are also needed in the EPIC model and were calculated from 30 arc-second global elevation data, GTOPO30 (USGS, 2008) using ArcGIS10 (ESRI, 2011). The first simulation year, 1980 was treated as a spin-up or initialization to balance the soil variables that are the most sensitive to physiological processes in EPIC such as soil water and nitrogen. Therefore, this year is excluded from simulation analysis.

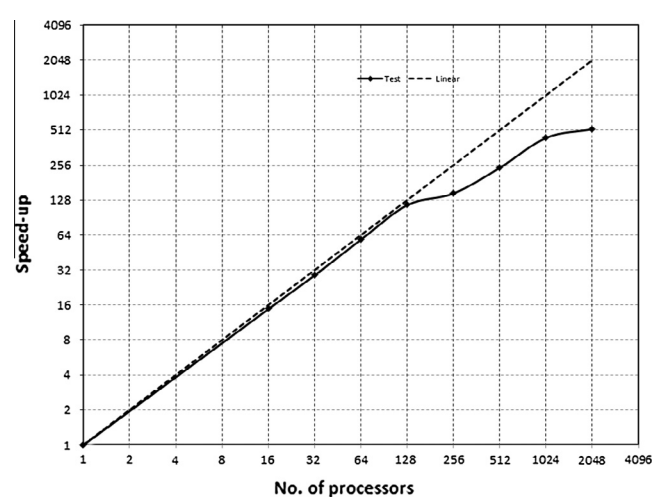

Fig. 3. Parallel computing scalability analysis of the mpi_EPIC model on Titan. Speedup refers to how much a parallel design is faster than a corresponding sequential algorithm when processing the same workload.

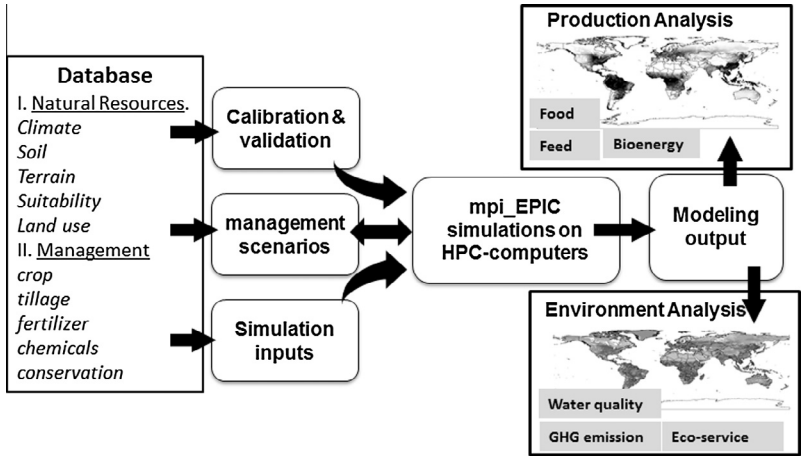

Fig. 4. Workflow of data management, calibration–validation, and simulation and result analysis of the mpi_EPIC model.

In this study, we used mpi_EPIC to assess global biomass productivity and environmental effect of a bioenergy crop, switchgrass (*Panicum virgatum* L.). This bioenergy crop has been identified as one of the potential crops for biomass production in the US (Parrish and Fike, 2005). Agronomic characteristics and management for switchgrass cultivars in different regions were obtained from the field trial database developed at https://bioenergykdf.net/sites/default/files/Introduction_global_switchgrass_trial_dataset.pdf. The database consists of over 2000 global observations from different continents, including the detailed information of site, cultivar, management practices, and biomass production. We designed 80 management files representing the various switchgrass cultivars and management practices (e.g., planting dates, fertilizer rates and harvesting dates) based on the studies from Kiniry et al. (2006), Wullschleger et al. (2010) and Nichols et al. (2011).

mpi_EPIC was carefully calibrated with the global switchgrass agronomic databases. Over 50 crop parameters would be considered for modification to different switchgrass cultivars in different ecological zones. Among them, the crop physiological parameters (radiation use efficiency, leaf area indices, harvest index, base and optimal temperature, nutrient parameters, root ratio) are critical to calibration. In this study, we used two common indicators of model performance evaluation using the REG procedure in SAS (SAS Institute, 2012). Our calibration results indicated that mpi_EPIC was able to simulate switchgrass productivity well with $r^2 = 0.67$ and root mean square errors (RMSE) less than 3.69 Mg ha^{-1} (Fig. 5). Most outliers of calibration for switchgrass cultivars were within one standard deviation of the observed values. Overall, the calibrations for the major ecological zones were acceptable although further improvements in the continental calibration are necessary for some sub-continental regions and can be made when more experimental data are available.

3.3. Simulation analysis of switchgrass biomass production

We illustrate the simulated biomass production potential of switchgrass in Fig. 6. The average switchgrass productivity ranges from near zero in boreal and desert areas to a maximum exceeding 30 Mg ha^{-1} in moist tropics. However, these extreme low and high values are uncertain because they are simulated for areas where reliable field-trial data were not available to support crop parameter calibration.

The mpi_EPIC model simulates production potential based on biophysical factors across the globe, regardless of current land cover, markets and zoning. High simulated productivity does not necessarily identify the best opportunities for cultivating switchgrass. This case study simulates high potential productivity in areas with favorable climate and soils even though other local conditions and current land cover may limit or prohibit cultivation. Areas with high productive potential also include lands that have been under cultivation for food crop production. This is not surprising because agricultural land is expected to represent relatively fertile areas in the ecological zones. Furthermore, environmental variables such as soil organic carbon and nonpoint source pollutions are not included in this discussion because they cannot be calibrated and validated due to limited data. Therefore, enhanced global field experiment data for calibration and validation will be critical to further improve the global applications to food and bioenergy ecosystem studies.

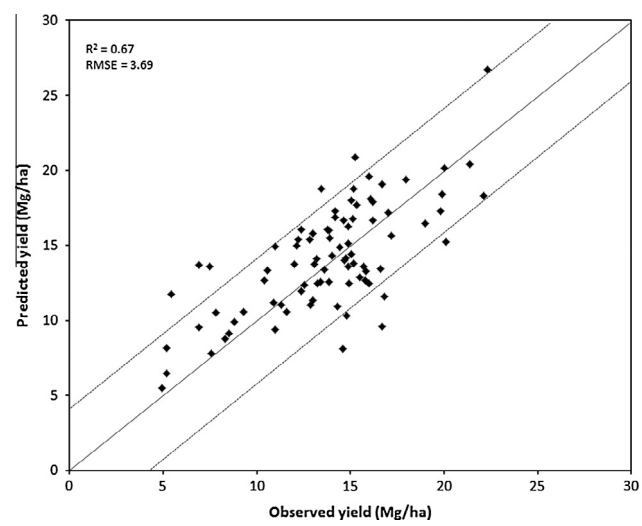

Fig. 5. Calibration and validation of the mpi_EPIC model for large-scale high-resolution agroecosystem analysis. The solid line is 1:1 line, and the dotted lines show one standard deviation from the 1:1 line.

4. Conclusion

To advance capabilities for analyses of global productivity and sustainability of food and bioenergy under climate change, we used the MPI parallelization technique and applied it to the mpi_EPIC model. The master/slave scheme in the model was capable of managing job allocation and execution on an ORNL supercomputer. A case study to test capacity to complete high-resolution

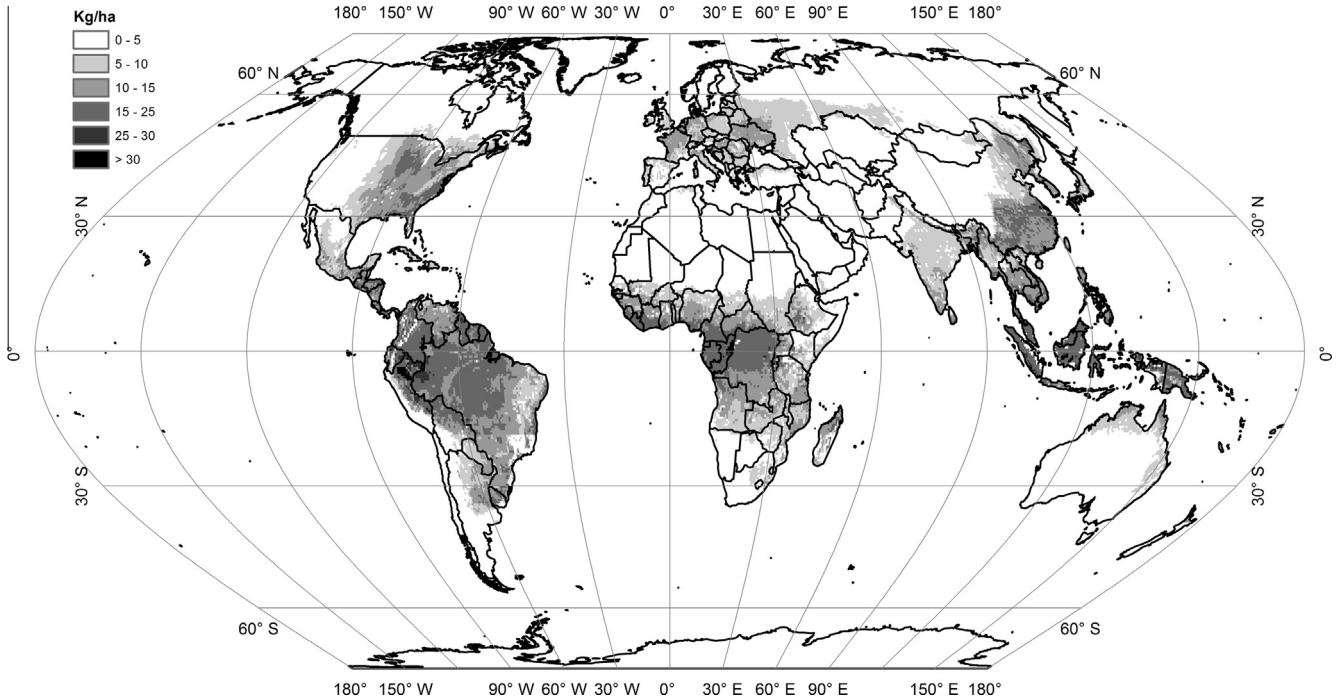

Fig. 6. Potential global biomass productivity of switchgrass simulated by the mpi_EPIC model.

agroecosystem simulations using the mpi_EPIC model was conducted for a bioenergy crop, switchgrass. A global 0.5-degree resolution crop production assessment involved 62,482 simulations and was completed in less than 30 min. The simulation processed approximately 11 gigabytes of information in the model database. The results show that the mpi_EPIC model is capable of serving a valuable role in exploring production and management scenarios and corresponding environmental effects for food and bioenergy production under changing climate conditions. Indeed, the mpi_EPIC modeling platform demonstrated that it can effectively manage the 2.2 billion 1-km scale agroecosystem simulations projected by Wang and Kang (2013) as required to assess global effects of different management options and production scenarios. This study provides an insight of how traditional sequential ecosystem models can effectively be integrated into available HPC systems for various ecosystem assessments.

This study targets large scale of high-resolution agroecosystem modeling for regional and global agricultural production and sustainability analysis, but the demand on intensive computational resources, large dataset management and knowledge of HPC application on clusters or supercomputers make its wide applications challenging. Some HPC-based code in mpi_EPIC associated with different module versions and libraries on machines or platforms may have to be adjusted and validated. Limited training on parallel programming and experience on clusters will facilitate solving the problem of applications. Additionally, detailed model input data and experiment data for calibration and validation are other challenges that limit reliable regional and global applications.

Acknowledgements

This research was thanks to computational resources made available through the Oak Ridge Leadership Computing Facility, located in the National Center for Computational Sciences at Oak Ridge National Laboratory. Contributions were also supported by the U.S. Department of Energy (DOE) Bioenergy Technologies Office. Oak Ridge National Laboratory is managed by UT-Battelle LLC for the Department of Energy under Contract DE-AC05-00OR22725.

References

Balkovič, J., van der Velde, M., Schmid, E., Skalsky, R., Khabarov, N., Obersteiner, M., Sturmer, B., Xiong, W., 2013. Pan-European crop modelling with EPIC: implementation, up-scaling and regional crop yield validation. Agric. Syst. 120, 61–75.

Domke, J., Wang, D., 2012. Runtime tracing of the community earth system model: feasibility study and benefits. In: 12th Workshop on Tools for Program Development and Analysis in Computational Science, Omaha, Nebraska, Procedia CS 9, pp. 1950–1958.

Easterling, W.E., Mearns, L.O., Hays, C.J., Marx, D., 2001. Comparison of agricultural impacts of climate change calculated from high and low resolution climate change scenarios. Clim. Change 51, 173–197.

ESRI, 2011. ArcGIS Desktop: Release 10. Environmental Systems Research Institute, Redlands, CA.

FAO/IIASA/ISRIC/ISSCAS/JRC, 2008. Harmonized World Soil Database (Version 1.2). FAO, Rome, IT and IIASA, Laxenburg, AT.

Gassman, P.W., Williams, J.R. Benson, V.W., Izaurralde, R.C., 2005. Historical Development and Applications of the EPIC and APEX Models. CARD Working Paper 05-WP. Center for Agricultural and Rural Development Iowa, US.

Hawick, K.A., James, H.A., Maciunas, K.J., Vaughan, F.A., Wendelborn, A.L., Buchhorn, M., Rezny, M., Taylor, S.R., Wilson, M.D., 1997. Geographic information systems applications on an ATM-based distributed high high-performance computing system. High-perform. Comput. Netw. 1225, 1035–1037.

Kiniry, J.R., Sanderson, M.A., Williams, J.R., 2006. Simulating Alamo switchgrass with the ALMANAC model. Agron. J. 88, 602–606.

Knüpfer, A., Brendel, R., Brunst, H., Mix, H., Nagel, W.E., 2006. Introducing the Open Trace Format (OTF). In: Alexandrov, V.N., van Albada, G.D., Sloot, P.M.A., Dongarra, J. (Eds.), International Conference on Computational Science, Lecture Notes in Computer Science, 3992. Springer, pp. 526–533 (2).

Kuo, C., Shiau, Y., Huang, C., Shen, C., Tsai, C., 2004. Application of virtual reality in ecological farmland navigating system. In: Proceedings of the Seventh International Conference on High Computing and Grid in Asia Pacific Region, 0-7695-2138-X/04.

Liu, J., Fritz, S., van Wesenbeeck, C.F.A., Fuchs, M., You, L., Obserseiner, M., Yang, H., 2008. A spatially explicit assessment of current and future hotspots of hunger in Sub-Saharan Africa in the context of global change. Global Planet. Change 64, 222–235.

Lobell, D.B., Schlenker, W., Costa-Roberts, J., 2011. Climate trends and global crop production since 1980. Science 333, 616–620.

Lynd, L.R., Aziz, R.A., de Brito Cruz, C.H., Chimphango, A.F.A., Cortez, L.A.B., Faaij, A., Greene, N., Keller, M., Osseweijer, P., Richard, T.L., Sheehan, J., Chugh, A., van der Wielen, L., Woods, J., van Zyl, W.H., 2011. A global conversation about energy

from biomass: the continental conventions of the global sustainable bioenergy project. Interface Focus 1, 271–279.
Mausbach, J.M., Dedrick, A.R., 2004. The length we go: measuring environmental benefits of conservation practices in the CEAP. J. Soil Water Conserv. 59, 96.
Nichols, J., Kang, S., Post, M., Wang, D., Bandaru, V., Manowitz, D., Zhang, X., Izaurralde, R., 2011. HPC-EPIC for high resolution simulations of environmental and sustainability assessment. Comput. Electron. Agric. 79, 112–115.
Parrish, D.J., Fike, J.H., 2005. The biology and agronomy of switchgrass for biofuels. Crit. Rev. Plant Sci. 24, 423–459.
Rosensweig, C., Elliott, J., Deryng, D., Ruane, A.C., Muller, C., Arneth, A., Boote, K.J., Folberth, C., Glotter, M., Khabarov, N., Neumann, K., Pionteck, F., Pugh, T., Schmid, E., Stehfest, E., Yang, H., Jones, J.W., 2014. Assessing agricultural risks of climate change in the 21st century in a global gridded crop model intercomparison. Proc. Natl. Acad. Sci. USA 111, 3268–3273.
SAS Institute, 2012. The SAS System for Windows. Release 9.2. SAS Inst., Cary, NC.
Savage, N., 2013. Modeling: predictive yield. Nature 501, S10–S11.
Shende, S.S., Malony, A.D., 2006. The TAU parallel performance system. Int. J. High Perform. Comput. Appl. 20, 287–331.
United States Geological Survey (USGS), 2008. GTOPO30, Earth Resources Observation and Science (EROS) Center, USGS.
Viovy, N., 2008. CRU–NCEP Reanalysis Data. <http://www.cru.uea.ac.uk/cru/data/ncep/>.
Vital, J., Gaurut, M., Lardy, R., Viovy, N., Soussana, J., Bellocchi, G., Martin, R., 2013. High-performance computing for climate change impact studies with the pasture simulation model. Comput. Electron. Agric. 98, 131–135.
Wang, D., Kang, S., 2013. Modern accelerator technologies for spatially-explicit integrated environmental modeling. In: Shi, X. et al. (Eds.), Modern Accelerator Technologies for Geographic information Science. Springer Science Business Media, New York, pp. 237–251.
Wang, D., Carr, E., Gross, L.J., Berry, M.W., 2005. Toward ecosystem modeling on computing grids. Comput. Sci. Eng. 7, 44–52.
Wang, D., Kang, S., Nichols, J., Post, W., Liu, S., Zhao, Z., 2012. A computational framework for spatially-explicit agroecosystem modeling: application to regional simulation. J. Comput. Sci. 4, 386–392.
Williams, J.R., Jones, C.A., Dyke, P.T., 1984. A modeling approach to determining the relationship between erosion and soil productivity. Trans. ASAE 27, 129–144.
Wullschleger, S.D., Davis, E.B., Borsuk, M.E., Gunderson, C.A., Lynd, L.R., 2010. Biomass production in switchgrass across the United States: database description and determinants of yield. Agron. J. 102, 1158–1168.
Zhao, G., Bryan, B.A., King, D., Song, X., Yu, Q., 2012. Parallelization and optimization of spatial analysis for large scale environmental model data assembly. Comput. Electron. Agric. 89, 94–99.