Three-Dimensional Paleohydrologic Simulation using Parallel *geofe* on the Teragrid Resources

[1]Peng Wang, [2]Denis Cohen, [2]Mark A. Person, [3]Arvind Gopu, and [4]Carl W. Gable

Abstract—Parallel geofe, a parallel, three-dimensional finite element paleohydrologic modeling program, was developed based on the serial code, with the aim to simulate ground water flow on the Atlantic continental shelf in New England within the past two million years, and help to explain the large amount of unconventional freshwater plumes within offshore aquifer systems in the above area. We used *ParMetis* for mesh partitioning and *Aztec* to solve the sparse linear system for the modeled groundwater flow, heat transfer and solute transport processes. The resulting program demonstrates good scalability on the Teragrid systems.

Index Terms— Geology, Hydrology, Teragrid, groundwater flow, finite element method, simulation

◆

1 INTRODUCTION

One of the problems facing hydrogeologists is to better understand temporal and spatial patterns of recharge into permeable aquifer systems. Over geologic time scales (10^6 years), recharge rates are believed to have fluctuated significantly due to temperature fluctuations and the waxing and waning of ice sheets. Recharge fluctuations are particularly important in coastal areas where sea level has fluctuated by as much as 120 meters in the past 21 thousand years exposing large areas of the continental shelf to meteoric recharge [1]. Since processes that occur on these time scales are very complex, specifically, the sensitivity to various environmental parameters and boundary conditions must be evaluated to understand the overall processes, numerical modeling and simulation is a critical method for analysis in addition to field drilling and testing.

We are investigating how up to 2.8 x 10^3 km^3 of freshwater were emplaced within permeable sediments of the Atlantic continental shelf in New England's about 21 thousand years ago around the time of the last glacial maximum [6, 7, 10] (Figure 1). These unconventional freshwater plumes within offshore aquifer systems may represent an important untapped resource for large urban centers (e.g. New York City) in coastal regions [1]. They are likely related to numerous observations of on shore freshwater plumes that have been reported to have displaced oil field brines in sedimentary basins across North America [2]. Interestingly, these freshwater plumes have stimulated the formation of biogenic methane that is now being exploited as a viable petroleum exploration "play" concept by the oil industry [10]. A number of mechanisms have been proposed to explain the emplacement of these freshwater plumes on the continental shelf including meteoric recharge (i.e. precipitation infiltration) during periods of sea-level lowstands and ice sheet recharge with infiltration occurring beneath the terminus of the glacier or beneath pro-glacial lakes. Evaluation of these various mechanisms can be accomplished through numerical simulation of the groundwater flow beneath and beyond the ice sheet on the continental shelf within the past 2 million years, when the Laurentide ice sheet dominated the landscape of North America.

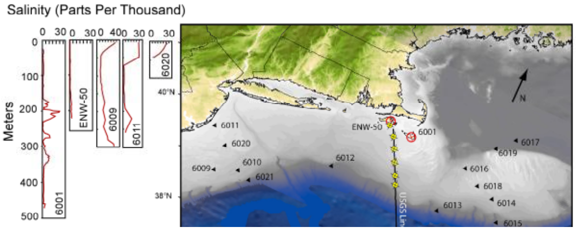

Fig. 1. Locations of offshore boreholes on the Atlantic continental shelf (triangles) and associated salinity profiles from selected wells. Wells not plotted have salinities near sea water [3]. Proposed International Drilling Program boreholes are shown as circles with cross.

Our numerical model is one of the first to consider the effects of ice sheet loading, and associated lithosphere flexure, on groundwater flow in basins [15]. Flexural adjustments to the lithosphere typically result in a deflection of the land surface that is equal to

---

1. *Peng Wang, ICTC Building, 535 West Michigan Street, Indianapolis, IN 46202 pewang@indiana.edu*
2. *Denis Cohen, Mark A. Person, Dept. of Geological Sciences, Indiana Univ., 1001 East 10th St., Bloomington, IN 47405-1405 denis.cohen@gmail.com, maperson@indiana.edu*
3. *Arvind Gopu, Indiana University, 2711 E. 10th St., Bloomington, IN 47408 agopu@Indiana.edu*
4. *Carl W. Gable, Earth and Environmental Sciences Division, Los Alamos National Laboratory, Los Alamos, NM 87545 gable@lanl.gov*

about one tenth the ice sheet thickness, although transient effects due to mantle flow can modify the magnitude of both a flexural moat and a forebulge.

While cross-sectional models have helped establish the relative importance of different recharge mechanisms [11], there is compelling geomorphic evidence that the Pleistocene flow system on the continental shelf was highly three-dimensional (Figure 2). That is, the spatial distribution of aquifers and confining units suggests that significant three-dimensional flow and freshwater patterns should occur. The advection dominated nature of solute transport necessitates using a highly resolved numerical grid with up to 5.8 million elements. To represent changes in sedimentary layer thickness and pinch outs requires a specialized mesh generator, LaGriT [9]. Both the time scale involved and magnitude of the grid require use of parallel processing and a high performance computing system.

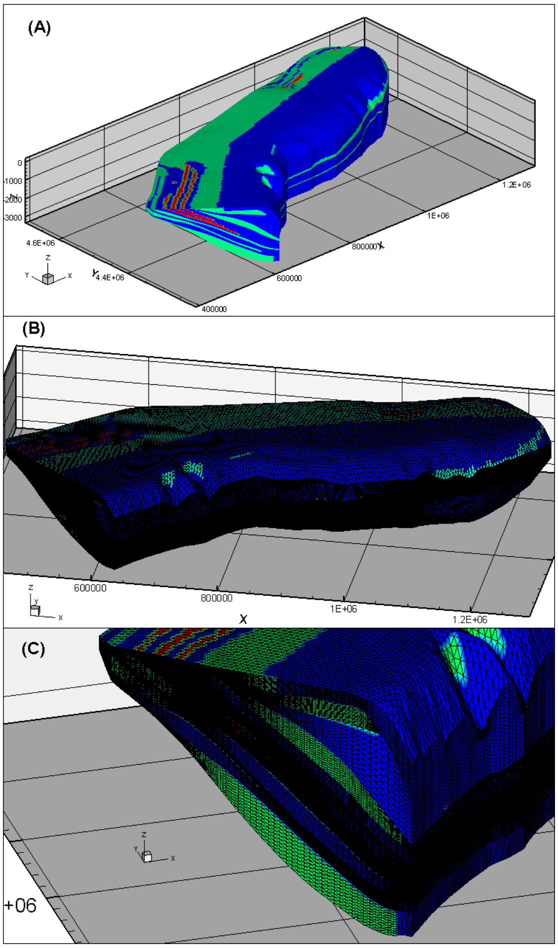

Fig. 2. (A) Lithologic units: green is marine clastic aquifer, blue is marine confining units, and red indicates other aquifers. (B) Three-dimensional numerical mesh of the Atlantic continental shelf in New England. The mesh extends laterally about 500 km from New Jersey to Maine and offshore about 250 km where the sediment thickness is 3 km. (C) Details of mesh showing levels of refinement in units. The mesh is vertically exaggerated for viewing convenience.

The Teragrid made the described simulation possible, by not only providing the necessary high performance computing resources, but also accelerating the interdisciplinary collaboration between geologists and IU Teragrid staff and scientific computing experts. The original serial application (*geofe*) was parallelized and run on the Teragrid resources (NCSA, SDSC, TACC, and IU) under a MRAC allocation. In the following sections, we discuss in more details of the modeling and simulation of the physical processes, the serial program *geofe*, the parallelization process, the scaling analysis of the resulting parallel program, preliminary results and ongoing work.

2 MODELING AND SIMULATION OF THE PHYSICAL PROCESSES

The physical processes we considered are variable density groundwater flow, conductive and advective heat transfer, and mass transfer and dispersion in porous media.

Following the principle of conservation of mass and Darcy's Law, which states that the specific discharge through a porous medium is proportional to the intrinsic permeability of the porous medium and the hydraulic gradient and inversely proportional to the absolute viscosity of the fluid, the governing equation of the groundwater flow of a variable density fluid in an anisotropic porous medium can be derived as:

$$\nabla \cdot \left[\overline{K}_0 \rho_f \mu_r \left(\nabla h_{fw} + \rho_r \nabla z \right) \right] = S_s \rho_0 \frac{\partial h_{fw}}{\partial t} \qquad (1)$$

where:

$\overline{K}_0$	is the hydraulic conductivity tensor for water at standard pressure, salinity, and temperature
ρ_f	is the fluid density at its current pressure and temperature
μ_r	is the relative fluid viscosity
h_{fw}	is the equivalent fresh water head
ρ_r	is the relative fluid density
S_s	is the specific storage
ρ_0	is the fluid density at standard pressure and temperature
t	is time

Similar to groundwater flow being driven by pressure gradient, diffusion of solute is driven by concentration gradient. In addition, heterogeneities in the porous medium at different scales can cause dispersion. Considering above effects, the governing equation for solute transport is:

$$\nabla \cdot \left(\overline{D}_M \nabla C\right) = \frac{\partial \phi C}{\partial t} + \nabla \cdot \left(qC\right) \qquad (2)$$

where:

ϕ	is porosity
$\overline{D}_M$	is the hydrodynamic dispersion tensor
C	is the dissolved concentration of the chemical species of interest
q	is the specific discharge vector

The modeled heat transfer process consists of heat conduction and mechanical dispersion. The conductive heat transfer is analogous to the groundwater flow, while the mechanical dispersion is mathematically equivalent to a molecular diffusion process. The governing equation for the heat transfer is:

$$\nabla \cdot \left(\overline{D} \nabla T\right) = C_e \frac{\partial T}{\partial t} + \nabla \cdot \left(\rho_f c_f q T\right) \qquad (3)$$

where:

$\overline{D}$	is the total thermal diffusivity tensor
T	is temperature
C_e	is the effective heat capacity
c_f	is the specific heat capacity of the fluid at constant pressure

The first term on the right hand side accounts for the ability of porous medium to store heat, while the second term represents the amount of energy moving through the porous medium within the groundwater discharge.

Tetrahedra elements are used to approximate equivalent fresh water head, temperature and solute concentration throughout the model domain. We applied the standard Galerkin approach and temporal discretization to formulate implicit solutions for the three attributes.

Moreover, for solute transport, in addition to equation 2, we need to consider advection effect where the compounds of interest are carried along with the groundwater. The modified method of characteristics (MMOC) is applied to address this issue. The MMOC algorithm consists of placing particles of a given initial concentration at the nodes of the finite element mesh, tracing them backward in the velocity field, and reinterpolating a concentration distribution at their new locations to simulate the effect of advection. After the concentrations at the nodes have been updated to simulate advective effects, the finite element approach is applied to account for the effects of diffusion and dispersion.

In summary, the overall simulation process follows algorithm in Figure 3.

3 PARALLELIZATION AND SCALING ANALYSIS

As seen in Figure 3, our *geofe* paleohydrologic model runs in a transient mode within which groundwater flow, heat, and solute transport are simulated sequentially. Each time step depends on the results of the previous time step. Thus, the only way to effective parallelization is to perform domain decomposition on the grid and parallelize the simulation process according to the partition. To accurately represent time dependent fluctuations in sea level and ice sheet loading, a large number of time steps are required. Therefore, it is necessary to have a load-balanced mesh partition and an efficient sparse linear system solver, to allow the simulation to be completed in a reasonable amount of time.

```
read_initial_condition
calculate_density
calculate_viscosity
do t=start_time, end_time, time_step_size
  flow_eqn   !solve for the head, eqn 1
  calculate_q_using_head
  heat_eqn   !solve for the temperature, eqn 2
  mmoc_particles_generate
  compute_num_of_mmoc_steps
  do i=1,num_of_mmoc_steps
    if (i .eq. 1) move_particles
    calculate_nodal_concentration
    solute_eqn !solve for the concentration, eqn 3
  enddo
  calculate_density
  calculate_viscosity
enddo
```

Fig. 3. algorithm for *geofe* simulation

We chose ParMETIS to perform mesh partitioning and the Aztec library to solve the sparse linear system, both in parallel. Both of these toolkits have been used in large numerical simulations and known to scale up to hundreds of processors [5, 6, 12, 13, 14]. ParMETIS is able to generate mesh partitions that have good load balance and theAztec library has been demonstrated to solve linear systems up to millions of unknowns on up to 1024 processors.

To verify the scalability of our parallel program *pgeofe*, we performed a scalability analysis using our DAC allocation on the NCSA IA64 Teragrid cluster. Parallel speedup with respect to the optimal serial version is shown in Figure 4 A and B for grids with 104,758 and 1,277,068 finite element nodes (577,993 and 7,489,503 tetrahedra elements respectively). The serial version was implemented with splib iterative solver and the Goto BLAS library. We enforced consistent parameters in both the serial and parallel versions, e.g., ILUT preconditioner and CGSTAB sparse linear system solver with the same convergence criteria in both versions. The parallel code was run with up to 256 processors on both meshes for comparison purposes illustrated below. Although the finer mesh was generated for a different problem, it does reflect the scaling behavior of the *pgeofe* code.

We can observe the different scaling behavior of *pgeofe* for the relatively coarse and more refined mesh. For the relatively coarse mesh, the speedup increases linearly up to 16 processors but plateaus around 128 processors; for the more refined mesh, we found the linear speedup prevailed even using 256 processors. The reason of this behavior, we believe, is due to the effect of communication overhead with respect to the computation. For the same mesh, when a larger number of processing elements (PEs) are used, less number of finite element nodes will be assigned to each PE, which leads to more communication overall and less computation on each PE. Communication is going to be the dominant factor and parallel speedup will suffer when nodes per PE fall below certain threshold, which is the case in using 256 PEs on the coarse mesh. We project that our final mesh of 1 million nodes will scale well beyond 256 processors.

4 PRELIMINARY RESULTS AND ON-GOING WORK

pgeofe was tested on a three-dimensional representation of New England's continental shelf between New Jersey and Nantucket Island, Massachusetts. Figure 5 shows the discharge computed from Darcy's law after two million years for a constant sea level followed by twice 780,000 years of sea-level fluctuations given by the SPECMAP record [4]. In this simulation, no ice sheet is present. Discharge areas (upward velocities) are found on the continental shelf beneath sea level, indicating that fresh water can resurface away from shore. Ice sheets and proglacial lakes will enhance this effect because higher fluid heads underneath ice or lakes will drive water further offshore. In addition, permafrost around the ice sheet margin will prevent fluids from resurfacing near the margin, enhancing discharge further away on the continental shelf. Figure 6 presents spatial variations in dissolved solids concentration on the continental shelf at the end of the run. Significant spatial variations in salinity are observed. Sea-level fluctuations expose recharge areas on the continental shelf that will enhance migration of fluid off shore. The new elements (ice sheet, proglacial lake, and permafrost) are being tested in a new version of *pgeofe*.

We are working on profiling *pgeofe* on both BigRed at IU and lonestar system at TACC. Further scaling analysis with 512 processors and more is also being planned. We encountered some problem using *ParMetis* with 256 PEs, and switched to serial *METIS* for now, but that did not have any significant impact performance-wise since mesh partitioning is only performed once at the beginning of the program.

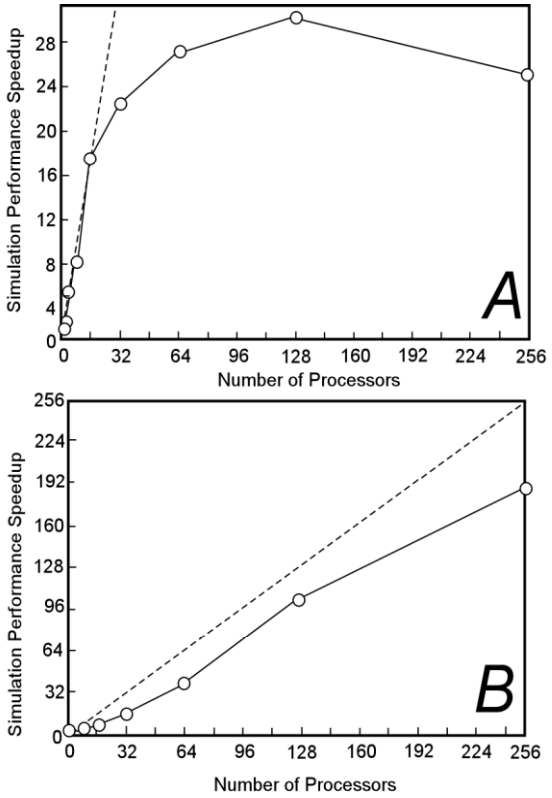

Fig. 4. (A) Speedup of pgeofe code against serial geofe with a coarse mesh of New England Continental Shelf containing 104,758 nodes and 577,993 elements. The dashed line indicates ideal scaling. (B) Similar scalability tests comparison but using a more refined mesh containing 1,277,068 nodes and 7,489,503 elements of a hydrothermal flow system. Simulation was performed on the NCSA IA64 cluster.

5 CONCLUSION

The analysis of the three-dimensional distribution of freshwater on the Atlantic continental shelf can only be accomplished using parallel models such as *pgeofe* due to the advection dominated nature of solute transport through porous media. Our preliminary results suggest that significant spatial variations in the degree of flushing of the continental shelf aquifers due to the effects of topography, sea level, and ice sheet loading. We expect to elucidate these processes in the future by including the effects of isotope transport.

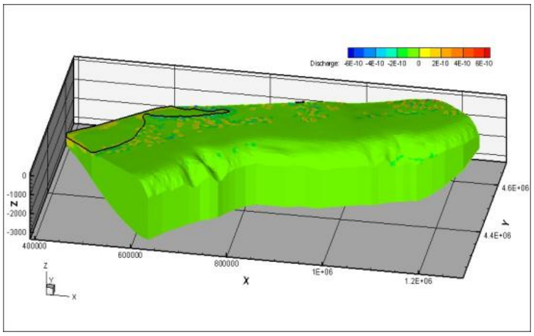

Fig. 5. Three-dimensional plot of discharge (upward vertical velocity; m/s) on New England's continental shelf. Colors yellow to red indicate discharge, green and blue indicate recharge into the aquifer.

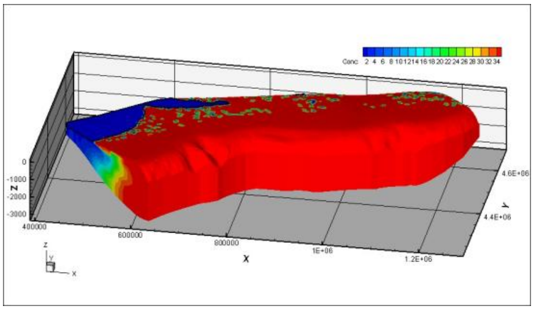

Fig. 6. Simulated total dissolved solids concentration (red: saltwater; blue: freshwater) on New England's continental shelf at the end of the run.

ACKNOWLEDGMENT

We would like to thank Randy Bramley for several fruitful discussions regarding parallelization approaches and fortran90 features. We would also like to thank Jim Dietz of UITS for the effort to acquire Vampir license for our profiling work. This research was supported, in part, by a grant from the National Science Foundation (EAR-0337634) to Mark Person. IU's involvement in the Teragrid is supported by the NSF grants No. 0833618, SCI451237, SCI535258 and SCI504075.

REFERENCES

[1] Edmunds, W M (2001), Palaeowaters in European coastal aquifers; the goals and main conclusions of the PALAEAUX Project, Geological Society Special Publications, v.189, p.1-16.

[2] Grasby, S., K. Osadetz, R. Betcher, and F. Render (2000), "Reversal of the regional-scale flow system of the Williston Basin in response to Pleistocene Glaciation", Geology, v. 28(7), p. 635-638.

[3] Hathaway, J. C., C. W. Poag, P. C. Valentine, R. E. Miller, D. M. Schultz, F. T. Manhiem, F. A. Kohout, M. H. Bothner, and D. A. Sangrey (1979), U.S. Geological Survey core drilling on the Atlantic Shelf, 1979, Science, v. 206(4418), p. 515-527.

[4] Imbrie, J., J. D. Hays, D. G. Martinson, A. McIntyre, A. C. Mix, J. J. Morley, N. G. Pisias, W. L. Prell, and N. J. Shackleton (1984), The orbital theory of Pleistocene climate: Support from a revised chronology of the marine d18O record. In: (Berger, A., et al. eds.) Milankovitch and Climate,D. Reidel Publishing Company, Part I, p. 269-305.

[5] Karypis, G. and V. Kumar (1997), "A Coarse-Grain Parallel Formulation of Multilevel k-way Graph Partitioning Algorithm", 8th SIAM Conference on Parallel Processing for Scientific Computation.

[6] Karypis, G. and V. Kumar (1996), "A Parallel Algorithm for Multilevel Mesh Partitioning and Sparse Matrix Ordering", 10th International Parallel Processing Symposium, p. 314-319.

[7] Kohout, F. A., J. C. Hathaway, D. W. Folger, M. H. Bothner, E. H. Walker, D. F. Delaney, M. H. Frimpter, E. G. A. Weed, and E. V. C. Rhodehamel (1977), "Fresh groundwater stored in aquifers under the continental shelf, Implications from a deep test, Nantucket Island, Massachussets", Water Resources Bulletin, 13(2), p. 373-386.

[8] Kohout, F. A., H. Meisler, F. W. Meyer, R. H. Johnston, G. W. Leve, and R. L. Wait (1988) "Hydrogeology of the Atlantic continental margin, in The Atlantic continental margin", U.S, edited by R. E. Sheridan and J. A. Grow, 463-480, The Geological Society of America, Boulder, CO.

[9] Los Alamos Grid Toolbox, LaGriT, Los Alamos National Laboratory, <http://lagrit.lanl.gov>, (2007)

[10] McIntosh, J. C., and L. M. Walter (2006), "Paleowaters in Silurian-Devonian carbonate aquifers: geochemical evolution of groundwater in the Great Lakes region since the Late Pleistocene", Geochimica et Cosmochimica Acta, v. 70, p. 2454-2479

[11] Person, M., B. Dugan, J. B. Swenson, L. Urbano, C. Stott, J. Taylor, and M. Willett, "Pleistocene hydrogeology of the Atlantic continental shelf", New England (2003), Geological Society of America Bulletin, 115, 1324-1343

[12] Shadid, J. N., S. A. Hutchinson and H. K. Moffat (1994), Proceedings of Colorado Conference on Iterative Methods, Breckenridge, Colorado, April 5-9.

[13] Shadid J. N. and R. S. Tuminaro, "Sparse Iterative Algorithm Software for Large-Scale MIMD Machines: An Initial Discussion and Implementation", Concurrency: Practice and Experience, Vol. 4, No. 6, pp 481-497, Sep., 1992

[14] Tuminaro R. S., M. Heroux, S. A. Hutchinson, and J. N. Shadid, "Official Aztec User's Guide: Version 2.1", December, 1999.

[15] Van Der Veen, C. J. (1999), Fundamentals of Glacier Dynamics, A. A. Balkema, Rotterdam, p. 462.