Applying Computationally Efficient Schemes for BioGeochemical Cycles (ACES4BGC)

Forrest M. Hoffman[5,α], Pavel B. Bochev[7], Philip J. Cameron-Smith[2], Richard C. Easter, Jr.[6], Scott M. Elliot[3], Steven J. Ghan[6], Iulian R. Grindeanu[1], Robert B. Lowrie[3], Donald D. Lucas[2], Vijay S. Mahadevan[1,β], Kara J. Peterson[7], Bill Sacks[4], Manishkumar B. Shrivastava[6], Mark A. Taylor[7], Mariana Vertenstein[4], and Patrick H. Worley[5,β];

External Collaborators:

Steven Goldhaber[4], Alex Guenther[6], Peter H. Lauritzen[4], and Xiaohong Liu[8]

[1]Argonne National Laboratory, [2]Lawrence Livermore National Laboratory, [3]Los Alamos National Laboratory, [4]National Center for Atmospheric Research, [5]Oak Ridge National Laboratory, [6]Pacific Northwest National Laboratory, [7]Sandia National Laboratories, and [8]University of Wyoming;

[α]Principal Investigator, [β]SciDAC Institute Liaison

July 31, 2014

SciDAC-3 Principal Investigator Meeting

Omni Shoreham Hotel, Washington, DC, USA

U.S. DEPARTMENT OF ENERGY | Office of Science

Project Goal and Objective

- **Goal:** Advance predictive capabilities of Earth System Models (ESMs) by reducing two of the largest sources of uncertainty, *aerosols and biospheric feedbacks*, utilizing a *highly efficient computational approach*.
- **Objective:** Deliver a second-generation ESM with improved representation of biogeochemical interactions at the canopy-to-atmosphere, river-to-coastal ocean, and open ocean-to-atmosphere interfaces.
- **ACES4BGC** is
  - implementing and optimizing new computationally efficient tracer advection algorithms for large numbers of tracer species;
  - adding important biogeochemical interactions between the atmosphere, land, and ocean models; and
  - applying uncertainty quantification (UQ) techniques to constrain process parameters and evaluate feedback uncertainties.

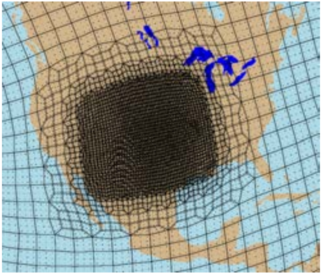
Research Team

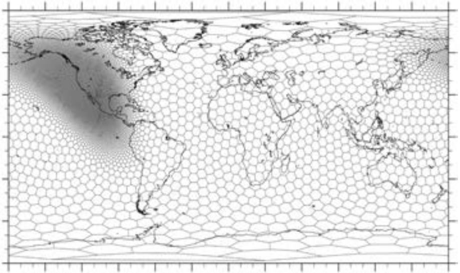
Name	Lab	Science Team	Topic
Pavel B. Bochev	SNL	Atmosphere	Advection
Philip J. Cameron-Smith†	LLNL	Atmosphere	Atm. BGC
Richard C. Easter, Jr.	PNNL	Atmosphere	Aerosols
Scott M. Elliott†	LANL	Ocean	Ocean BGC
Steven J. Ghan	PNNL	Atmosphere	Aerosols
Iulian R. Grindeanu	ANL	Comp. Tools & Perf.	Mesh Tools
Forrest M. Hoffman†	ORNL	Land	Land BGC
Robert B. Lowrie	LANL	Ocean	Advection
Donald D. Lucas	LLNL	Atmosphere	UQ
Vijay S. Mahadevan‡	ANL	Comp. Tools & Perf.	Mesh Tools
Kara J. Peterson	SNL	Atmosphere	Advection
Bill Sacks	NCAR	Comp. Tools & Perf.	SE
Manishkumar B. Shrivastava	PNNL	Atmosphere	Aerosols
Mark A. Taylor	SNL	Atmosphere	Advection
Mariana Vertenstein	NCAR	Comp. Tools & Perf.	SE
Patrick H. Worley†‡	ORNL	Comp. Tools & Perf.	Performance

†Science Team Lead; ‡SciDAC Institute Liaison

ACES4BGC Tracer Advection Research

Climate Science Needs for Reducing Biogeochemical and Aerosol Feedback Uncertainties





- Faster tracer transport methods for CESM and ACME atmosphere and ocean components
- Accurate on fully unstructured grids needed for next generation models
- Transport hundreds to a thousand of reactive and non-reactive biogeochemical species (trace gases, aerosols, dust, etc.)

Tracer Transport

A tracer, represented by its mixing ratio q and mass ρq, is transported in the flow with velocity $\mathbf{u}$ as

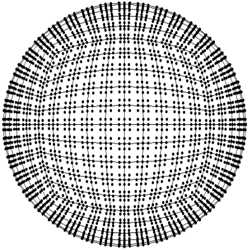
$$\left.\begin{aligned} \frac{\partial \rho}{\partial t} + \nabla \cdot \rho \mathbf{u} = 0 \\ \frac{\partial \rho q}{\partial t} + \nabla \cdot \rho q \mathbf{u} = 0 \end{aligned}\right\} \rightarrow \frac{Dq}{Dt} = 0.$$

Solution methods should satisfy

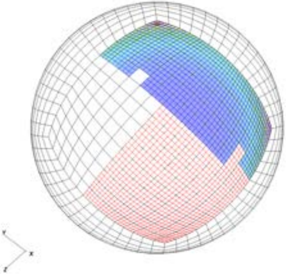
- local conservation of ρq,
- monotonicity or bounds preservation of q, and
- consistency between q and ρ (free stream preserving).

Spectral Element Dynamical Core

The **High-Order Method Modeling Environment (HOMME)** within CAM-SE provides a continuous Galerkin finite element method using Gauss-Lobatto quadrature.
Advection using the standard spectral element method with high-degree polynomials is accurate, but expensive due to time step restrictions.



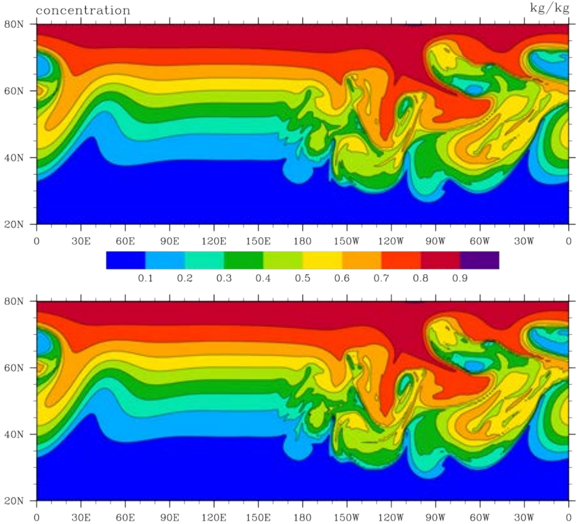
(a) Original CAM-SE mesh



(b) CSLAM mesh (refined, 2/16 proc)

Conservative Semi-LAgrangian Multi-tracer (CSLAM) transport scheme (Lauritzen et al., 2010) offers one approach for improving efficiency for large tracer counts. Stable method for large time steps with CFL ~ 5.

Early CSLAM Performance



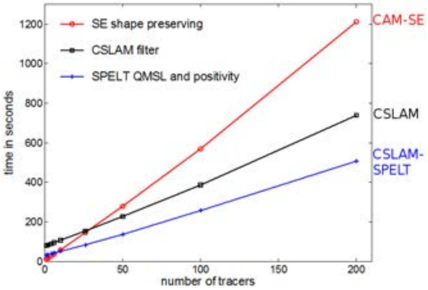
CAM-SE

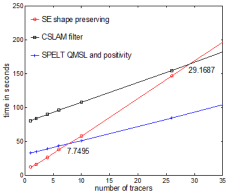
CSLAM

No obvious loss of accuracy.

Erath, Taylor, Lauritzen (SNL & NCAR)

Early CSLAM Performance

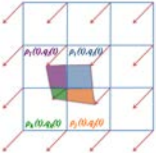




- CSLAM is expensive for one tracer, but breaks even with CAM-SE at $\sim$30 tracers.
- A fast, scalable method for generating grid intersections allowing for CFL$\geq$5 should improve performance significantly.

Erath, Taylor, Lauritzen (SNL & NCAR)

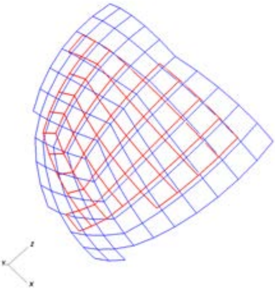
Eulerian/Lagrangian Grid Intersection

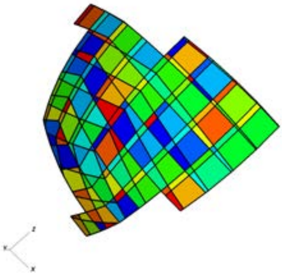


The CSLAM scheme uses an Eulerian arrival grid and integrates backwards to a Lagrangian departure grid.

An intersection of these two grids is required to compute the disposition of tracers through time.

The Mesh-Oriented data Base (MOAB) from the **FASTMath Institute** provides a scalable, parallel algorithm for intersecting arbitrary meshes.





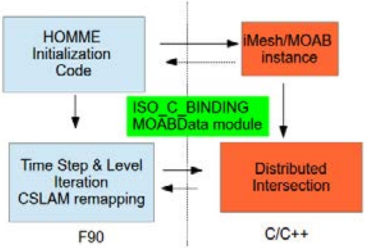
The arrival (red) and departure (blue) meshes are intersected to produce a final set of polygons for reconstruction of all tracer mole fractions.

Argonne National Laboratory · Lawrence Livermore National Laboratory · Los Alamos National Laboratory · NCAR · Oak Ridge National Laboratory · Pacific Northwest National Laboratory · Sandia National Laboratories

MOAB Now Coupled to HOMME

Initialization:

- MOAB infrastructure is instanced on every task,
- arrival mesh is initialized from HOMME (CAM-SE and refined CSLAM meshes), and
- parallel infrastructure is established.



Software linkage between HOMME (CAM-SE) and MOAB

At every time step:

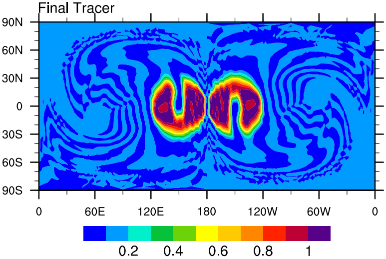
- departure point positions are passed to MOAB,
- MOAB computes the intersections (communicating as needed), and
- MOAB returns the intersections to HOMME for reconstruction.

Next Steps:
Complete coupling (return of intersections), verify, and compare results with current CSLAM (cubed-sphere, regular mesh) implementation.

Additional Approaches

- Characteristic Discontinuous-Galerkin (CDG) represents tracers by discontinuous modal expansion within each element. Modal expansion provides compact, numerically efficient high-order reconstruction "built it", but difficult to impose exact monotonicity (Lowrie and Ringler, 2011).
- Semi-Lagrangian spectral element (SL-SE) algorithm using optimization to enforce mass conservation. Optimization is efficient and works for large time steps on unstructured grids (Peterson and Taylor, 2014).

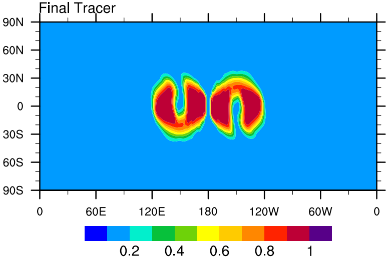
Deformational flow, 1.5° resolution



Mass error $= -3.44 \times 10^{-3}$

Min value $= -0.1070$

Max value $= 1.1934$



Mass error $= 1.69 \times 10^{-11}$

Min value $= 0.1$

Max value $= 0.9979$

Argonne National Laboratory · Lawrence Livermore National Laboratory · Los Alamos National Laboratory · NCAR · Oak Ridge National Laboratory · Pacific Northwest National Laboratory · Sandia National Laboratories

Terrestrial Biogeochemistry: Carbonyl sulfide (COS)

Objective:

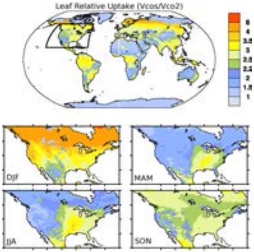
- Add important biogeochemical interactions between the biosphere and atmosphere to the Community Earth System Model (CESM).

New Science:

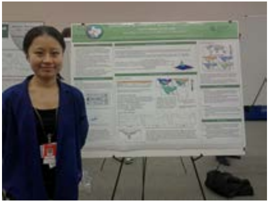
- Initial parameterization for carbonyl sulfide (COS) uptake by the biosphere tested in the Community Land Model (CLM4) by intern Wenting Fu.

Significance:

- COS provides a potentially powerful tracer for biosphere–atmosphere exchange of CO_2 and a constraint on global gross primary production.



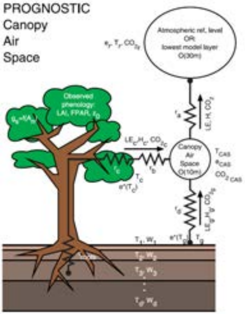
Leaf relative uptake from simulation testing COS uptake in CLM (Fu *et al.*, in prep.).



ORNL intern Wenting Fu presenting preliminary results from her initial implementation of COS uptake in CLM4.

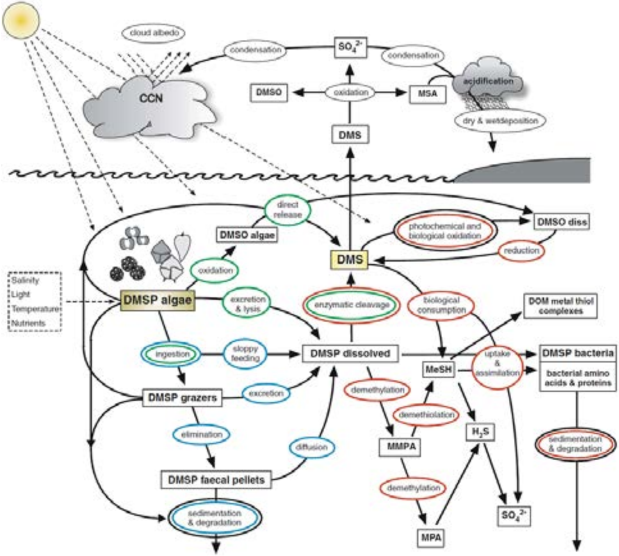
Terrestrial Biogeochemistry: Canopy Air Space

- Volatile Organic Compounds (VOCs)
  - Working with Alex Guenther at PNNL to improve VOC emissions factors (ε);
  - improving species-specific emissions in MEGAN2 model, initially by adding plant functional types (PFTs) in CLM4; and
  - will use GOAmazon2014 observations and UQ to improve model parameters in warm, moist tropical regions.
- Prognostic Canopy Air Space
  - Initially implementing single-layer canopy air space scheme of Vidale and Stöckli (2005) and
  - extending to a multiple-layer CAS to improve representation of canopy trace gas exchange.
- Ammonia
  - Focusing on canopy exchange and
  - adding soil and agricultural emissions (cattle, hogs, fertilizer).



(Vidale and Stöckli, 2005)

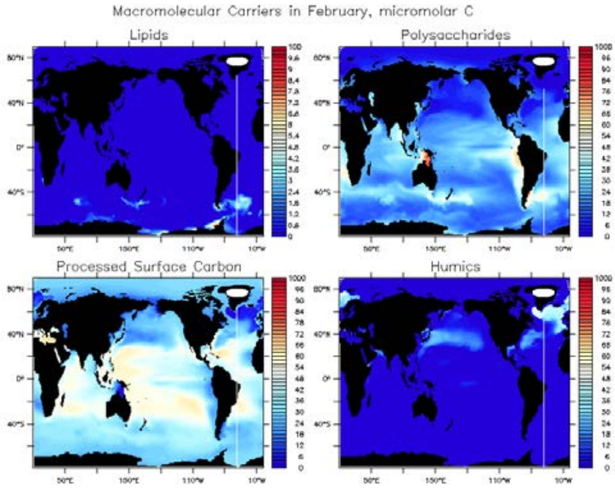
Marine Biogeochemistry



(Stefels et al., 2007)

Argonne National Laboratory · Lawrence Livermore National Laboratory · Los Alamos National Laboratory · NCAR · Oak Ridge National Laboratory · Pacific Northwest National Laboratory · Sandia National Laboratories

Marine Biogeochemistry

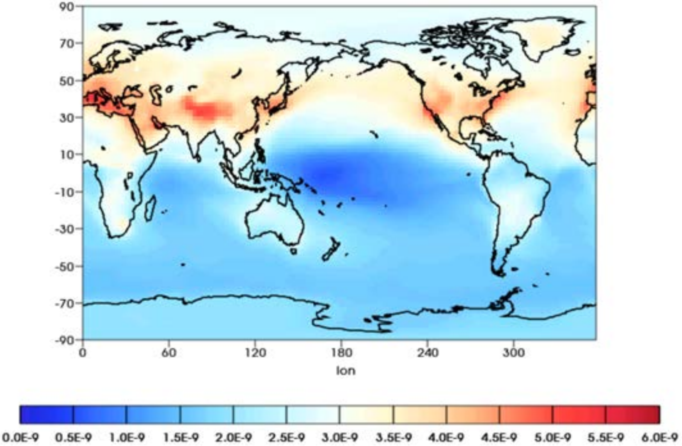


Scott Elliott (LANL)

Organic ocean emissions are sources of primary organic aerosols.

Macromolecules

Marine Biogeochemistry

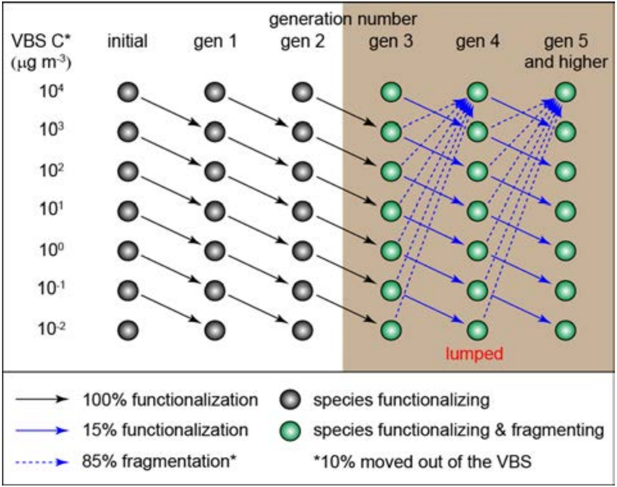


Scott Elliott (LANL)

Ozone concentration differences for simulations of enhanced Arctic Ocean methane released due to clathrate destabilization on continental slopes → complex HO_x/NO_x chemistry.

Argonne National Laboratory · Lawrence Livermore National Laboratory · Los Alamos National Laboratory · NCAR · Oak Ridge National Laboratory · Pacific Northwest National Laboratory · Sandia National Laboratories

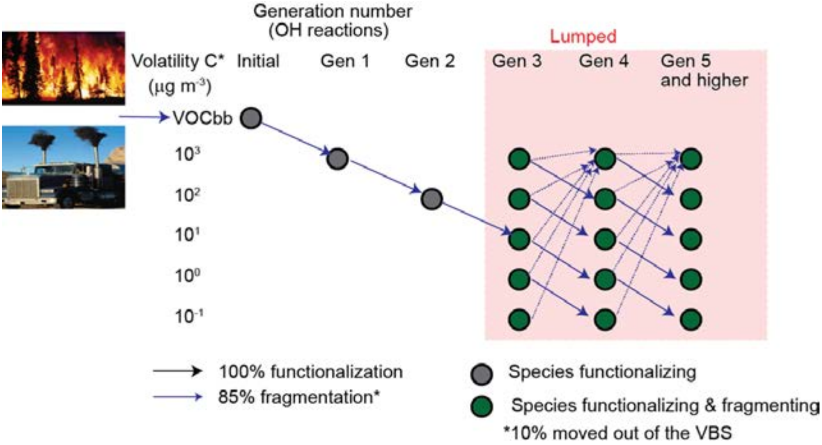
ACES4BGC Aerosol Research



(Shrivastava et al., 2013)

Volatility basis set (VBS) representation for secondary organic aerosol (SOA) precursors

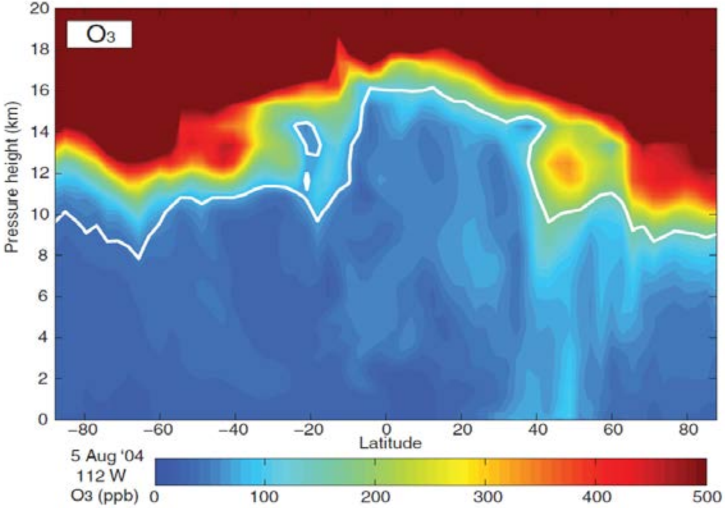
ACES4BGC Aerosol Research



(Shrivastava et al., 2013)

A very simplified volatility basis set (VBS) is used for current climate models, with only 8 tracers.

Atmospheric Chemistry

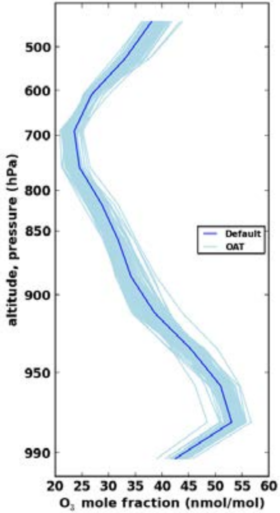


(Prather et al., 2011)

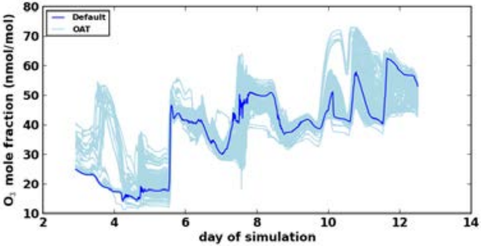
Ozone tracer emitted uniformly at surface with 90-day exponential decay.

ACES4BGC Uncertainty Quantification

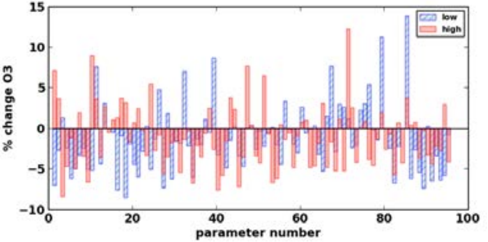
Vertical profile (days 5–10)
191 simulations



Time evolution (level 887 hPa)



Effects of parameter changes (days 5–10, lev 887)

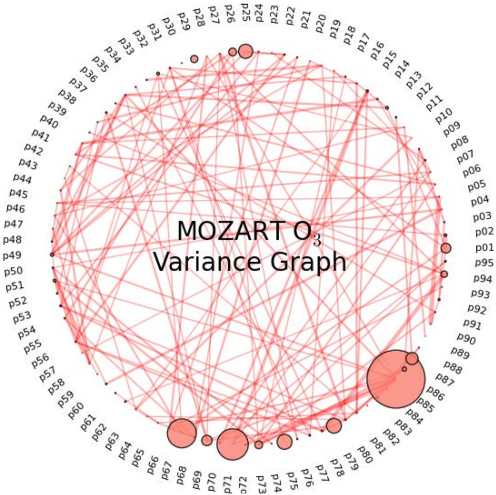


Don Lucas (LLNL)

ACES4BGC Uncertainty Quantification

MOZART Ozone Example

- Implemented PDFs for 100 photochemical parameters in the MOZART mechanism.
- Ran $\sim 10^4$ ensemble SCAM simulations using Latin hypercube sampling.
- Analyzing ensemble variance using new UQ methods in collaboration with **QUEST Institute**.



Don Lucas (LLNL)

Second order decomposition of the variance of daily mean O_3 concentration in the middle troposphere is dominated by 10 laboratory rate constant parameters.

OLCF Director's Discretion Project

How do we pull together all these individual science components?

- Combine code onto a single ACES4BGC branch of ACMEv0.
- Perform new science simulations, turning on new biogeochemical and aerosol processes incrementally.
- Develop a science and computational performance baseline before switching to CSLAM on unstructured grids.
- Validate and use a GPU-enabled version of the existing finite volume tracer advection scheme with large numbers of tracers.
- Enabled by **SUPER Institute** collaboration and the Oak Ridge Leadership Computing Facility (OLCF) Director's Discretion Project: **ACES4BGC SciDAC-3 Partnership Project Gen2ESM Foundry**
  - Titan: 6M core-hours; Rhea: 6K core-hours

ACES4BGC Benefitting from SciDAC Institute Partnerships

- **FASTMath Institute** – Development and implementation of MOAB by Iulian Grindeanu and Vijay Mahadevan (ANL) will provide a significant advance to enable new science.
- **SUPER Institute** – Critically important performance tracking and optimization by Pat Worley (ORNL) continues to enable more science in less time and fewer computing resources and supporting performance portability across new architectures.
- **QUEST Institute** – Uncertainty quantification tools and methods being used by Don Lucas (LLNL) to develop framework for biogeochemistry sensitivity studies.
- **SDAV Institute** – Began after our project, but we are starting a conversation with Rob Ross (ANL) about *in situ* analysis, parallel I/O, and tracer visualization.

Poster #15

Intersection of Distributed Meshes for Multi-Tracer Transport Schemes

Iulian Grindeanu[†], **Kara Peterson**[β], **Timothy J. Tautges**[†], **Mark A. Taylor**[β], **and Patrick H. Worley**[§]

[†]Argonne National Laboratory, [§]Oak Ridge National Laboratory, and [β]Sandia National Laboratories

Introduction

- One of the important goals of ACES4BGC project is to implement and optimize new computationally efficient advection algorithms for large number of tracer species
- The resulting upgrades to the Community Earth System Model (CESM) will deliver new scientific capabilities, offer unprecedented accuracy in representing biogeochemical interactions, and yield *improved predictive skill and computational performance.*
- Significant computational challenges must be overcome to meet the scientific requirements for supporting large numbers of reactive tracers in CESM. The most significant challenge is presented by the advection of many tracers in the atmosphere and ocean component models.

Advection / Transport Problem in Climate Codes

- The transport problem can be cast in Eulerian, Lagrangian or Arbitrary Lagrangian-Eulerian form. One popular variant is semi-Lagrangian method, run in Lagrangian mode for one time step and then interpolate back to the Eulerian mesh (Staniforth and Cote, 1991)
- At every time-step, the semi-Lagrangian approach involves interpolating quantities from a distorted Lagrangian mesh to a regular Eulerian mesh or vice versa, depending on the trajectories. The transport problem is reduced to a regridding problem if the Lagrangian mesh movement is prescribed (Lauritzen et al., 2010)
- Conservative regridding (remapping) has been made practical by the application of Gauss-Green's theorem over the intersection mesh, which converts area-integrals into line-integrals. This approach has been applied for up to second-order static grid-to-grid remapping in (Jones, 1999) and the method was extended for regular lat-lon grids and cubed-sphere grids (Ulrich et al., 2009)
- For unstructured meshes that arise in either CAM-SE or MPAS formulations (figure 1) a more general framework for advection/transport is needed.

(a) (b)

Figure 1: *Unstructured meshes of interest*

Transport Equations

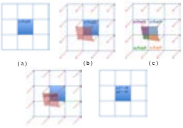
A tracer, represented by its mixing ratio q and mass ρq, is transported in the flow with velocity $\mathbf{u}$

$$\left.\begin{aligned}\frac{\partial\rho}{\partial t} + \nabla\cdot\rho\mathbf{u} = 0\\ \frac{\partial\rho q}{\partial t} + \nabla\cdot\rho q\mathbf{u} = 0\end{aligned}\right\} \rightarrow \frac{Dq}{Dt} = 0$$

New methods should satisfy

- Local conservation of ρq
- Monotonicity or shape preservation of q
- Free stream preserving (consistency) when $q = 1$

Semi-Lagrangian Remap Algorithm for Transport



(a) (b) (c)

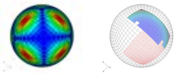
(d) (e)

- Consider a cell c_i with cell-averaged density ρ_i and tracer q_i at time t
- Compute Lagrangian departure grid
- Compute tracer reconstructions on Eulerian (arrival) grid
- Find intersections between departure grid and arrival grid
- Integrate tracer reconstructions over cell intersections
- Update tracer on arrival grid $q_i(t+\Delta t) = q_i^L(t)$

By carefully choosing reconstruction function the method is guaranteed to be tracer mass conserving and free stream preserving. The method is stable for large time steps and in practice can allow for CFL ~ 5 in the atmosphere. It also is efficient for large numbers of tracers because the cell intersections are computed once and used for all tracers.

Coupling with CAM-SE

The MOAB grid intersection algorithm is now being coupled with CAM-SE. The algorithm for tracer transport begins with the initialization of the CAM-SE Eulerian tracer grid in MOAB, distributed on each processor (CAM-SE coarse grid). The CSLAM fine grid is instantiated on each processor, and the parallel infrastructure of MOAB is leveraged, so the ghosting and communication with neighboring processors is relatively easy to do.



(a) (b)

Figure 3: *Initial CAM-SE mesh (a) and CSLAM mesh (refined) on 2 out of 16 processors (b)*

Figure 3 shows the initial CAM-SE coarse mesh, for ne=15, and the distribution of the refined CSLAM mesh on the first 2 processors, for a small run, on 16 processors. The initial partition of mesh is done using an HSFC method implemented in HOMME, and the instantiation of the MOAB mesh data structures is done locally, on each task/processor. Global identificators are used in MOAB to resolve sharing and ghosting of local data with the neighbors.

For each time step, departure point locations computed in CAM-SE are passed to MOAB. MOAB will then compute intersections between departure (Lagrangian) grid and arrival (Eulerian) grids and return the intersection polygons to HOMME, which will continue with remapping.

Figure 4: *Software link between HOMME/CESM and MOAB*

Currently, the interface between MOAB and HOMME is done using a new module in HOMME, data is flowing between HOMME and MOAB for initialization stage. A special library is developed in MOAB (mbcslam) , which contains all intersection related code and initialization.

Linear Complexity Algorithm for Intersection of Overlapping Meshes

- The intersection algorithm used in this paper follows the ideas from (Gander and Japhet, 2008), in which two meshes are covering the same domain. At the core is an advancing front method; For simplicity, label the two meshes red and blue, as they are covering about the same domain.
- First, 2 convex cells from the red and blue meshes that are intersecting are identified, by using a search tree. This will constitute the seed of the front. Advancing in both meshes, using adjacency information, we incrementally compute all possible intersections. Important for the algorithm is also a robust scheme, in which 2 intersecting cells from the different meshes are overlapped and resolved. Every edge on one of the meshes maintains a list of intersection points, and geometric tolerances are used to eliminate duplicates and ensure consistency.
- The intersection of cells from different meshes is simplified and more robust if they are convex. If in the initial meshes there exist concave polygons, they are decomposed in simpler convex polygons, by simply splitting along interior diagonals. The intersection between 2 convex cells will then be a convex polygon. The final result of meshes intersection will be a set of convex polygons, that cover the common domain.
- All edges of cells on the sphere are considered to be great circle arcs. Using gnomonic projection, on one of the six planes, each arc becomes a straight line in the gnomonic plane, and a robust intersection is much easier to achieve.

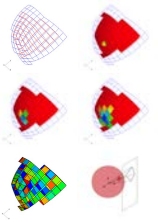
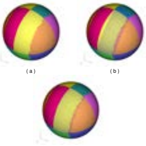


Figure 5: *Advancing front for local intersection*

Figure 5 shows how the algorithm advances; Each red cell is resolved, by building a local queue of blue cells that intersect the red cell. Blue cells are added to the local queue incrementally, using adjacency. At the same time, a global queue with seeds is formed, and it contains pairs of red/blue cells that have the potential to intersect. When there are no blue cells in the local queue, algorithm advances to the next seed from the global queue

Parallel Implementation Considerations

- Use parallel infrastructure from MOAB (Tautges et al., 2004)
- The two meshes to be intersected will be distributed; the arrival mesh is fixed, and can be distributed in advance (use Zoltan (Devine et al., 2002))
- The algorithm works with any distribution of the arrival mesh, and CAM-SE code use space filling curve method.
- The departure mesh needs to be distributed at run time
  - compute boundary boxes for each arrival processor mesh
  - departure points position stored in a MOAB tag, on each arrival node
  - departure cell bounded box computed locally, for each arrival cell
  - sent to processors whose bounding boxes are within tolerance

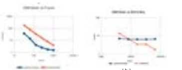


(a) (b)

(c)

Figure 6: *(a) initial distributed arrival grid. (b) departure grid on one of the processors (blue mesh) (c) covering departure mesh after communication*

Figure 6(a) shows initial arrival mesh, distributed over processors. Each vertex knows its departure position, and the algoritm computes departure cells, blue mesh (b). Departure cells that intersect bounding boxes of other processors are sent there. Each arrival grid will be covered by departure mesh (c).



(a) (b)

Figure 7: *(a) results on fusion.lcrc.anl.gov (b) results on mira.alcf.anl.gov*

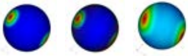
Figure 7 shows the scalability of the algorithm up to 64k tasks, on an MPAS model with 65M cells. Arrival mesh is distributed using Zoltan when loading the model; Recursive Coordinate Bisection is used, which ensures a balanced distribution, and that the bounding boxes of partitions have a low overlap. Intersection is almost embarassingly parallel, as long as the mesh is communicated between processors, and it scales well. It is important to note that this intersection needs to be carried out at each time step and at each level.

At high processor count we notice that the communication cost for departure mesh is not decreasing anymore. Also, these two machines have different networks, and this explains the differences in results. We may need to change the algorithm, or the methodology. One method would be to ghost sufficient layers in the beginning, on arrival mesh. Considering that the target CFL number would be at most 5, 5 ghost layers would be enough. The departure mesh would lay then in arrival mesh and the ghosted region on each processor, so no communication would be needed for the covering mesh.

Tracer Transport Example in MOAB

A simplified transport/advection method is implemented in MOAB, and it is exercising the intersection algorithms explained above. Quantities of interest are discretized on the arrival mesh at each time step.

Higher order approximations should be employed, but in this study, first order approximations were used, i.e., quantities of interest were averaged over each cell. *The original CSLAM type implementation uses higher order approximations.* So for this MOAB implementation, only the area of the intersection polygons determines the resulting tracers concentrations. Still, as some departure polygons will end up on different processors, their tracer mass concentrations need to be communicated to the corresponding arrival cell. So for each time step in the approximation, two communication steps are involved, one to distribute the departure mesh to processors, and second to communicate back the tracer concentrations from the intersection polygons. For a higher order approximations of the tracers on the arrival/departure mesh, complete information about polygons will be needed.



(a) (b) (c)

Figure 8: *(a) initial tracer field (b) velocity field (c) final tracer field*

Figure 8 shows the results of a transport simulation as in (Nair and Lauritzen, 2010), the cosine bell example and non-divergent flow. The transported scalar field follows complex trajectories and undergoes severe deformation during the simulation, but it reverses its course halfway through and returns to the initial position by the end of the simulation. If there would be no errors in the simulation, the initial and final plot should be the same. The pictures show that the constant tracer approximation is not sufficient, higher polynomial order is mandatory.

Conclusions

An efficient intersection algorithm for distributed unstructured meshes has been developed and is currently being integrated into CAM-SE. This intersection algorithm will provide CAM-SE with the ability to perform efficient tracer transport on fully unstructured quadrilateral meshes on the sphere. The infrastructure and methodology will be leveraged for MPAS framework, too.

References

K. Devine, E. Boman, R. Heaphy, B. Hendrickson, and C. Vaughan. Zoltan data management services for parallel dynamic applications. *Computing in Science and Engineering*, 4(2):90–97, 2002.

M. J. Gander and C. Japhet. An algorithm with optimal complexity for non-matching grid projections. In *Proceedings of 18th International Conference on Domain Decomposition Methods*, Springer Lecture Notes in Computer Science, 2008.

P. Jones. First- and second-order conservative remapping schemes for grids in spherical coordinates. *Monthly Weather Review*, 127:2204 – 2210, 1999.

P. H. Lauritzen, R. D. Nair, and P. A. Ullrich. A conservative semi-Lagrangian multi-tracer transport scheme (CSLAM) on the cubed-sphere grid. *J. Comput. Phys.*, 229(5):1401–1424, Mar. 2010. doi: 10.1016/j.jcp.2009.10.036.

R. D. Nair and P. H. Lauritzen. A class of deformational flow test cases for linear transport problems on the sphere. *Journal of Computational Physics*, 229(23):8868 – 8887, 2010. ISSN 0021-9991. doi: http://dx.doi.org/10.1016/j.jcp.2010.08.014. URL `http://www.sciencedirect.com/science/article/pii/S0021999110004511`.

A. Staniforth and J. Cote. Semi-lagrangian integration schemes for atmospheric models a review. *Monthly Weather Review*, 119(9):2206 – 2223, 1991. doi: DOI: 10.1175/1520-0493.

T. J. Tautges, R. Meyers, K. Merkley, C. Stimpson, and C. Ernst. MOAB: a mesh-oriented database. SAND2004-1592, Sandia National Laboratories, Apr. 2004. Report.

P. A. Ulrich, P. H. Lauritzen, and C. Jablonowski. Geometrically exact conservative remapping (gecore): Regular latitude longitude and cubed - sphere grids. *Monthly Weather Review*, 137:1721 – 1741, 2009. doi: DOI: 10.1175/2008MWR2817.1.

ACKNOWLEDGEMENTS: Funding provided through ACES4BGC, FASTMath and SUPER. This research used resources of the ALCF at ANL, which is supported by the Office of Science of the U.S. Department of Energy under contract DE-AC02-06CH11357

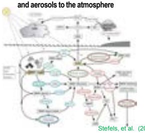
Poster #36 → #15

Computational Challenges Posed by Simulating Complex Ecosystems and Large Numbers of Tracers in Earth System Models

Philip Cameron-Smith (pjc@llnl.gov), ACES4BGC team (ANL, LANL, LLNL, ORNL, PNNL)

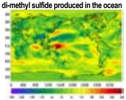
Biological emissions affect climate and air-quality

Ocean ecosystem is the source of many gases and aerosols to the atmosphere



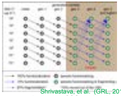
Stefels, et al. (2007)

Decrease in soar surface heating due to di-methyl sulfide produced in the ocean



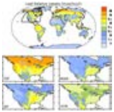
Global mean = 3.5 W/m^2 W/m^2

Organic chemicals in atmosphere oxidize into aerosols



Shrivastava, et al. (GRL, 2013)

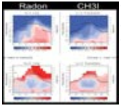
Land ecosystem is source of CO_2 and hydrocarbons



Carbonyl sulfide (OCS) is a natural tracer that is similar to CO2, and can be used to quantify the difference between photosynthesis and respiration.

Tracers can answer: who, what, where, why?

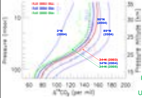
Natural tracers can validate unmeasurable processes



- ^{222}Radon affected by convective updrafts.
- CH_3I tracks uplift from oceans.
- SF_6 quantifies Interhemispheric exchange rate.

Rosa, et al. (2012)

Natural tracers can constrain stratospheric motion and fossil CO_2

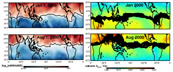


- ^{14}C tracks strat motion & fossil CO_2.
- $^7Be/^{10}B$ tracks stratospheric air in troposphere.

Kanu et al. (LLNL-UCBerkeley)

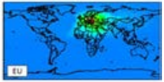
Idealized tracers can unambiguously identify tropical mixing barrier

Atmospheric equator / Variation over month



Tracers emitted north and south with 90-day decay.
Holmes & Prather (UC Irvine)

Tracers can be tagged in model to quantify contribution from source region or plant type



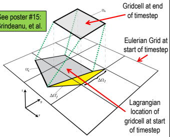
Effect of PM2.5 from European emissions
Anenberg, et al (2014)

Improving tracer performance with FASTMath & SUPER

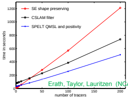
Existing tracer advection schemes usually solve the advection equations separately for each tracer, so computational cost scales linearly.

Using MOAB to solve advection once and apply to all tracers

See poster #15: Grindeanu, et al.

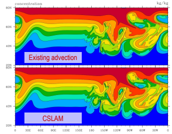


A simpler version (CSLAM) limited to neighboring gridcells (ie CFL=1), is working well.



Erath, Taylor, Lauritzen (NCAR & SNL)

Comparison between advection schemes shows no significant loss of accuracy

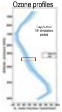


Erath, Taylor, Lauritzen (NCAR & SNL)

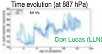
Uncertainty quantification, validation, & tuning with QUEST

Uncertainty in laboratory chemical rate constants results in significant differences in simulated ozone

Ozone profiles

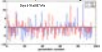


Time evolution (at 887 hPa)



Don Lucas (LLNL)

Effects of parameter changes



QUEST tools enable high dimensional UQ analysis of ozone sensitivity to laboratory parameters

2nd order decomposition of the variance of O_3 concentration.

- PDFs for 100 photochemical parameters.
- Ran ~10^4 ensemble column models using Latin hypercube sampling.
- Space complexity ~ $O(NC^2)$
  N = # training data
  C = # variance features
- MOZART graph
  N ~ 10^5 simulations
  C ~ 10^4 – 10^6
  (2^{nd} – 4^{th} order)
- Memory limits to 2^{nd} order variance graphs.

Strategies to circumvent memory limitations
These problem sizes require 10-1000GB for 2nd to 4th order analyses. Hence, many computers limit our UQ variance analyses to low orders (e.g. Titan has 32GB per node). "Big data" machines have more memory (e.g. LLNL's Catalyst has 800GB per node). Sparse UQ methods that retain only the relevant information (e.g. compressive sensing) are being implemented to scale to even higher dimensions and orders.

Acknowledgements

Support for this work was provided through Scientific Discovery through Advanced Computing (SciDAC) program funded by U.S. Department of Energy, Office of Science, Advanced Scientific Computing Research and Biological and Environmental Research.

This research used resources of the National Energy Research Scientific Computing Center, which is supported by the Office of Science (BER) of the U.S. Dept. of Energy under Contract No. DE-AC02-05CH11231, the National Center for Computational Sciences at Oak Ridge National Laboratory, which is supported by the Office of Science of the Department of Energy under Contract DE-AC05-00OR22725, and the Argonne Leadership Computing Facility at Argonne National Laboratory, which is supported by the Office of Science of the U.S. Department of Energy under contract DE-AC02-06CH11357.

This work is supported by the Earth System Modeling program of the Office of Science of the United States Department of Energy. This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under contract DE-AC52-07NA27344. LLNL-POST-

GC007: Earth System Modeling at the Extreme Scale

Co-conveners: William Collins (LBNL), Forrest M. Hoffman (ORNL), and Stephen F. Price (LANL)

Earth system modeling is entering a new era as the climate community transitions to extreme-scale computing and as the need for more capable models becomes increasingly evident in the face of "no-analogs" climate regimes emergent under global warming. We invite presentations at the frontiers of climate simulation that address the prospects, early development, and proof-of-principle experiments with the next generation of Earth System Models (ESMs). We seek talks and posters demonstrating how to best exploit recent advances in theory, applied mathematics, computational science, process-scale modeling, uncertainty quantification, and observational assimilation to make ESMs more accurate, robust, scalable, and extensible. These more robust ESMs are characterized by increasing reliance on more mechanistic, and more computationally intensive, treatments of core climate processes. Examples include global cloud resolving models, regional large eddy simulation models, full physics, high-resolution land ice models, and high-throughput treatments of chemical and biogeochemical transport across the climate system.

Invited speakers:

- Omar Ghattas, Univ. of Texas at Austin (next-generation ice sheet models)
- Peter Lauritzen, NCAR (atmospheric transport)
- Charlotte DeMott, Colorado State University (cloud/turbulence interactions)
- Hisashi Yashiro, RIKEN (non-hydrostatic global cloud-resolving models)

Abstracts are due August 6, 2014 at `http://fallmeeting.agu.org/2014/`

Acknowledgements

This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Advanced Scientific Computing Research, Scientific Discovery through Advanced Computing (SciDAC) program.