Case Study: An Integrated Approach for Steering, Visualization, and Analysis of Atmospheric Simulations

Yves Jean,[1] Thomas Kindler,[2] William Ribarsky,[3] Weiming Gu,[1] Gregory Eisenhauer,[1] Karsten Schwan,[1] and Fred Alyea[2]
[1]College of Computing
[2]School of Earth and Atmospheric Sciences
[3]Office of Information Technology
Georgia Institute of Technology

Abstract

In the research described here, we have constructed at tightly coupled set of methods for monitoring, steering, and applying visual analysis to large scale simulations. This work shows how a collaborative, interdisciplinary process that teams application and computer scientists can result in a powerful integrated approach. The integrated design allows great flexibility in the development and use of analysis tools. This work also shows that visual analysis is a necessary component for full understanding of spatially complex, time-dependent atmospheric processes.

Overview

For large scale atmospheric simulations one would like a tight coupling between the simulation, the observational database on which it is based, and the visualization/ analysis process by which it is understood. In fact there should be feedback, in the form of steering, between the latter and the simulation, since this will yield much more accurate representations of atmospheric processes and a significantly more focused investigation of behavior relevant to questions being asked. Since the data have complicated 3D structures and are highly time-dependent, the visualization approach must handle this dynamic data in a highly interactive fashion.

In the research described here, we have combined all these aspects into a single, integrated approach. This has required a collaborative, interdisciplinary process involving atmospheric scientists, experts in parallel high performance computing, visualization specialists, and experts in user interfaces. In particular, we find that it is important to have the scientists involved from the beginning in defining the steps of the project and evaluating its results. This constant evaluation allows an iterative refinement of the approach and aids everybody (including the scientists) in discovering new aspects of the problem that they did not foresee. We think that the process used here could serve as a template for building highly effective and powerful applications (and tools supporting them), a process where the developer comes away with a deeper understanding of user needs.

The Scientific Problem

The ultimate goal in climate modeling is the simultaneous simulation on a global scale of physical and chemical interactions in the ocean and atmosphere. This goal is still far from reach since, in addition to the problem's enormous complexity, parameters must be chosen to simulate processes that are not well understood or whose influence can only be approximated at the scale of current models.

Earth and atmospheric scientists at Georgia Tech have developed a global chemical transport model (GCTM) [1] that uses assimilated windfields for the transport calculations. These models are important tools to answer scientific questions about the stratospheric-tropospheric exchange mechanism or the distribution of species such as chlorofluorocarbons (CFC's), hydrochlorofluorocarbon (HCFC's), and ozone. This model uses a spectral approach, which is common to global models [2], to solve the transport equation for each species. In a spectral model, all variables are expanded into a set of orthogonal spherical basis functions. In a typical run our model contains 37 layers, which represent segments of the earth's atmosphere from the surface to approximately 50 km, with a horizontal resolution of 42 waves or 946 spectral values. When one transforms to a grid system, this corresponds to a resolution of about 2.8 degrees by 2.8 degrees per grid cell. Thus in each layer 8192 gridpoints have to be updated every time step. A typical time step increment is 15 simulated minutes, and for the usual annual run the number of grid values generated is over 10 billion. Of course, several variables may be evaluated at each grid point, and one might need many runs at different parameter settings to accurately simulate observed phenomena.

An Integrated Computational Approach

In order to enable our integrated approach, we have developed Falcon [3], a toolkit that collectively supports the on-line monitoring, steering, visualization, and analysis of parallel and distributed simulations. Falcon tools include:

- sensors, probes, and steering objects generated from monitoring and steering specifications. The partially analyzed monitoring information is sent to the graphical and visualization displays. Once steering decisions are made by the user or a steering algorithm, changes to the application's parameters and states are made by Falcon's steering mechanisms by invoking the steering objects embedded into the application code.
- An on-line display system to provide a flexible mechanism for attaching different types of graphical and visualization displays to an application's execution. On-line displays can be dynamically attached to and detached from the display system.
- A self-describing library, called Portable Binary I/O (PBIO), for communication of information among components of Falcon. PBIO allows binary information to be easily exchanged between heterogeneous machines, hiding differences in data representation, data sizes or data structure layout. The self-describing nature of PBIO data streams allows data exchange between applications with no compile-time knowledge of the form or content of the handled data.

In Falcon the original simulation code has probes and sensors placed with an instrumentation tool and then is run in instrumented form (as a highly parallel or distributed application). A steering server interprets commands to the running application from the steering interface and passes information back to the interface. Monitors control what is displayed in the visualization/analysis (VA) interface. These data can either be shipped directly to the VA interface or filtered and stored for later investigation. The Falcon sensor, probe, and steering objects permit us to build flexible and adjustable steering and visualization interfaces across multiple platforms.

The Steering Interface

Fig. 1 shows a simple steering interface built on Falcon; it presents a screen display of the distribution of C^{14} at a latitude of 70^{o} N. The display has two parts: one for showing both the computed and the observed concentration values of the atomic species, and the other for accepting steering requests. The computed vertical C^{14} distribution at the chosen lattitude and longitude is represented by a plotted curve from atmospheric layer 0 to 37, and it is updated for every model time step. The observed C^{14} concentration at a number of atmospheric layers is represented by discrete triangular points, and is used to judge whether the current computation is "correct" or "wrong." The user can dynamically modify the application execution to conform more closely to the observation. For example, the curve shown in Fig. 1 demonstrates that the computed concentration of C^{14} is consistently higher than the observed values from layer 10 to 15, but it is lower from layer 16 to about 23. The user can alter the vertical wind velocity at these atmospheric layers to remove these discrepancies in future time steps. The user can also stop the application's execution, change parameters, and restart the execution at another checkpointed time step. Using Falcon tools it is straightforward to modify the steering interface to display different variables or control different parameters.

The Visualization/Analysis Module

For simulations such as the atmospheric model described here, we must go well beyond 1D or even 2D displays of averaged variables. Without full 3D visualization it is hard to see, for example, important correlative effects between vertical wind velocities and transport of atmospheric species between layers. There are also localized, time-dependent effects in temperature, species concentration and other variables that are not contained in one layer or at one latitude and thus would be hard to see fully without interactive visualization.

The Falcon system allows us to build independent VA tools in a dataflow environment. We have used the SGI Explorer environment and provided a module to read the sockets from the monitoring controller in Falcon and then convert the data from the PBIO format to our own data representation. For the interactive VA tasks we have used a specially organized and constructed set of modules in our GlyphMaker system [4] (built on top of the Explorer environment). Among the modules are ones for direct conversion of selected data from spectral to gridded form for visualization and a reader for converting the data to PV-Wave format. Thus we can run, as separate processes on the same workstation, the steering interface, the GlyphMaker system, and the PV-Wave program--providing maximum feedback, control of the simulation, and analysis.

GlyphMaker helps the user explore relations between spatially complex and time-dependent data [4]. It allows the user to customize visual representations by defining and then binding data to 0-, 1-, 2-, or 3-D glyphs--graphical objects whose elements (e.g., position, size, shape, color, transparency, texture, orientation, etc.) are bound to data. In addition to the 0D glyphs shown here, we now have the capability for binding to 1D (streamlines) and 2D (surface) glyphs to provide rich and flexible visual representations. The streamlines are useful since they can, for example, show in a continuous fashion the flow of the wind vector field and its relation to the other variables in the simulation. Our GlyphMaker structure allows the user needed flexibility to select and focus on important regions in dense and complex data. It has a calculator module for computing new variables or setting limits on the range of a variable to be viewed. It also has a Conditional Box for choosing a spatial region by direct manipulation; the data inside that region can then be bound to special glyphs or made to appear alone. For time sequences from the atmospheric data, we have built an animation interface that allows one to choose time steps to be visualized one at a time or in sets (with skipping between time steps, if desired). Finally we have an interface for selecting beginning dates and variables to be viewed from the many generated by the model. The information for this interface is gleaned from the PBIO file and thus provides the user with an up-to-date description of the model, variables, and time ranges being run.

Data Analysis. To show our VA tools in action, we choose a simulation of N_2O run within the Falcon system. The distribution of N_2O in the atmosphere has a rich structure and is significantly affected by horizontal and vertical wind fields. In addition to GlyphMaker, we have added Explorer modules constructed by Morin et. al. [5] for depicting slices of data at various longitudes, latitudes, and altitudes and in various projections. We focus on the correlations between horizontal and vertical wind fields (taken from satellite observations) and the changes in N_2O distribution. These correlations are hard to see using traditional visualization methods, but they can be critical in assessing the accuracy of the model and in understanding the processes by which species spread through the atmosphere. Since the correlation with horizontal windfield is easier to see, we first looked at the horizontal stream function variable (Psi) which represents the rotational part of the horizontal windfields. We wanted to see high roatational movement, so we chose Psi in the upper fifth of its range and bound it to red spheres where the magnitude of Psi was mapped to the size of the sphere. We then stepped through several time steps to get an overview of the wind distribution and displayed N_2O concentration layers at altitudes that coincided with high rotational winds. Shown in Figs. 2, 3, 4 are successive time steps (over a six hour period in simulation time) for this mapping taken from the early Spring part of the simulation. The high rotational winds are localized and, as expected for this time of year, have a strong equatorial concentration at higher altitudes. The wavelike nature of the winds is clearly seen in these figures as is the correlated movement of higher concentrations of N_2O in the equatorial region.

Using the same methods, we then looked at the vertical windfields by binding the vertical velocity variable to blue cubes. Again we looked at higher velocities (the upper fourth of the range, showing only upward pointing winds) and bound the magnitude to the size of the cube. Our visualizations show that the vertical velocities are small, localized, and tend to jump from region to region. In Fig. 5 we display a slice through the N_2O data at constant longitude positioned at a particular cluster of upward winds. We see that there is a bulge (in both the blue and yellow parts of the surface) that shows mixing from lower to higher layers and that is associated with the upward winds. Since this is the main mechanism for mixing between layers in the atmosphere, it is important to understand in detail the effects of these velocity changes and species concentration responses. Finally we show in Fig. 6 how the correlations between several variables could be studied. We have built up a visualization of vertical and horizontal wind velocities and of N_2O concentration for two slices using the same bindings as before. We can clearly see the effects of the different variables in combination and singly and can follow them through time using our animation capability.

Conclusions and Future Work

Our steering and VA systems have proved quite useful in pinpointing behavior and following effects in atmospheric phenomena that are still not well understood. Our integrated approach gives us great flexibility in setting up and monitoring calculations and considerable power to move to larger simulations. The interplay between atmospheric and computation scientists has resulted in rapid prototyping of the steering interface and in considerable development and refinement in the VA tools. All too often visualization tools are developed based on what is possible in display algorithms or what has been done in the past without sufficient regard for application scientists' needs. Joint efforts plus the flexibility of our visualization system aid us in avoiding this pitfall.

We are now embarked on several improvements to our system. Prime among these will be to combine our steering and VA tools to provide a visual interface for steering. Thus, for example, the user will be able to start a simulation by placing a particular species concentration distribution in a particular spatial region. This approach will be critical to understanding atmospheric chemistry and modeling species profiles that are closer to observation than is possible now.

Acknowledgements

This work is supported in part by the NASA AISRP Program under contract number NAGW-3886 and by NSF under grant number NCR-90000460.

References

1. T. Kindler, K. Schwan, D. Silva, M. Trauner, F. Alyea. A parallel spectral model for atmospheric transport processes. Submitted to *Concurrency: Practice and Experience*, 1994.
2. W.M. Washington and C.L. Parkinson. An introduction to three-dimensional climate modeling. Oxford University Press, 1986.
3. Weiming Gu, Greg Eisenhauer, Eileen Kraemer, Karsten Schwan, John Stasko, Jeffrey Vetter, and Niru Mallavarupu. Falcon: On-line Monitoring and Steering of Large-Scale Parallel Programs. *Proceedings of The Fifth Symposium of The Frontiers of Massively Parallel Computation*, McLean, VA, February, 1995.
4. William Ribarsky, Eric Ayers, John Eble, and Sougata Mukherjea. Using Glyphmaker to Create Customized Visualizations of Complex Data. *IEEE Computer*, July, 1994.
5. P.J. Morin, T. Tanimoto, D.A. Yuen, and Y. Zhang *Pixel*, Vol.3, no. 3, pp.20-26 (March 1992)

Figure Captions

1. The steering interface for the atmospheric simulation. Shown are results and parameters for C^{14} distribution.

2,3,4. A time sequence from the simulation over a 2 hour period from Feb. 1, 1993. A horizontal cross-section of N_2O at about 5 mbars is displayed with concentration mapped to color and with red spherical glyphs representing high rotational winds.

5. Cross-sections of N_2O at fixed longitude from Feb. 1, 1993 showing mixing between layers at different altitudes. N_2O concentration is again mapped to color and the blue cuboid glyphs represent high vertical winds.

6. A combined mapping of N_2O, high rotational winds, and high vertical winds for one time step on Feb. 1, 1993. The glyph and color mappings are the same as above.