The ANU Translator: Facilitating Computer Visualization and Data Analysis of Climate Model Outputs

D. E. HYMAN, D. R. WHITEHOUSE*, J. A. TAYLOR, J. W. LARSON, J. A. LINDESAY**

Centre for Resource and Environmental Studies, Australian National University, Canberra
*ANU Supercomputer Facility, Australian National University, Canberra
**Department of Geography, Australian National University, Canberra
e-mail: Dena.Hyman@anu.edu.au

Abstract Computer visualization and data analysis tools are an essential component in the study and evaluation of climate model output. A number of useful visualization and analysis tools already exist, each with its own set of unique strengths. A translation tool is needed to convert data from the large number of model output formats into formats suitable for use with visualization and analysis tools. Such a translator would allow a researcher using a climate model to easily exploit a large number of visualization and analysis tools. Our goal is to develop a translation tool for UNIX computer systems that will be highly portable, very simple to use, and easily expandable. The translation program, The ANU Translator, will feature a graphical user interface when run on any system with an X11 interface. It will work with a number of climate model formats, including: CCM1, CCM2, RegCM2, MM5, CSIRO9, and ANU-CTM. It will produce output in a range of data file formats, including Vis5D, GrADS, and netCDF. NetCDF allows access to a range of analysis and visualization tools, including AVS, HDF, and GMT. Vis5D, GrADS, and the tools compatible with netCDF will enable the user to view data statically or in animation, including the following critical options: 1) geographic distribution plots, 2) two dimensional graphics, and 3) fully three dimensional renderings. It is intended that this translation package be made widely available for use by the climate modeling community as a research and teaching aid.

1. INTRODUCTION

The climate modeling community relies heavily upon data visualization for the evaluation and analysis of climate model results. Because there are no closed-form solutions for climate models, the analysis of model results must be achieved through the integration of vast amounts of data. Statistical methods are applied to some extent for the purposes of evaluating model performance, but it is often unclear whether or not such methods are strictly applicable to model results. Many statistical methods assume a random variable and/or a normal distribution; this poses a problem when analyzing output from General Circulation Models (GCMs) because they are deterministic systems. In addition to statistical analysis, numerical data analysis is used by climate modelers for preparing data for visualization, e.g. to time average the data before plotting.

Scientific visualization is of value in the following ways:

- it serves as a diagnostic for model performance with respect to observations;
- it provides the modeler with a quick method of checking that a model run is proceeding in a plausible manner. This assists the modeler to prevent the squandering of valuable computer resources when, due to human or computer error, a run is not proceeding as intended;
- it helps the modeler compare the output derived from sensitivity experiments with the output derived from control runs;
- it provides illustrations for scientific articles and conference discussions;
- it provides visual aids to help scientists meet their obligations to educate the public about the importance and implications of their work;
- it allows scientists to examine processes, which tend to incorporate the time-dependent behavior of two or more variables in concert; and
- it aids scientists in locating displaced climatic features that may be more difficult to detect through statistical methods alone, e.g. a somewhat displaced major circulation cell like the Icelandic Low.

One of the most extensive and widely used analysis and visualization packages in climate modeling is the National

Center for Atmospheric Research (NCAR) Modular Processor which is compatible with the NCAR Community Climate Models (CCMs) [Buja, 1992; Buja, 1993]. The NCAR Modular Processor also reads observational data written in CCM format, such as the European Centre for Medium Range Weather Forecasts (ECMWF) data. The NCAR Modular Processor was designed for use on the NCAR Crays and is highly dependent on Cray architecture. It has only been successfully ported to other Crays and to the IBM RS/6000. In Cray binary representation, the integer, character, and floating point words are all equal sized, 64 bit words. The Processor takes advantage of this memory structure to manage whole blocks of mixed mode (integer/real/character) memory as a large group of integers or as a large group of reals depending upon the task. The IBM RS/6000 has the rare property of being able to emulate this equal-size memory construct, but other high-end processors do not.

In addition to the NCAR Modular Processor, a number of other highly useful scientific visualization and analysis tools exist, each with its own set of unique strengths. The most inhibiting factor for climate modelers wishing to access these tools is the problem of converting vast quantities of model output into formats suitable for visualization and analysis tools. Our goal is to develop a computer program that will convert large amounts of data from a number of model output formats into formats suitable for inputting into a variety of visualization and analysis tools. We aim to create a program that will be highly portable, easily extendible, and very simple to use. We believe that such a data translation tool will be of considerable value to the wider climate modeling community.

The conversion program, The ANU Translator, is being written by Dena E. Hyman and Drew Whitehouse at the Australian National University (ANU). The code is written in ANSI C with a Tcl interpreter. It is designed to be highly portable, capable of running on any UNIX platform. The Translator is also designed to be simple to use and easily expandable. It interacts with the user through a Tcl/Tk user-friendly, graphical interface, and can also be run in non-graphical mode if an X11 interface or Tk are unavailable. Our goal is to incorporate the following models into The ANU Translator in the specified order: CCM1, CCM2, RegCM2, MM5, CSIRO9, and ANU-CTM. A brief description of each of these models follows below in section 1.1. Our goal is also to incorporate the following output formats/tools into The ANU Translator in the specified order: Vis5D, GrADS, and netCDF. NetCDF allows access to AVS, GMT, HDF, and other tools. A brief description of the output formats and tools mentioned here follows in section 1.2.

1.1 Model Descriptions

a. CCM1. The NCAR Community Climate Model Version 1 (CCM1) is a twelve-vertical-layer pseudo-spectral atmospheric global GCM with spectral resolution R15, corresponding to a grid resolution of 4.44° latitude × 7.5° longitude. CCM1 can be run with an optional interactive surface hydrology based on a simple bucket model. For a complete description of CCM1, see Williamson et al. [1987].

b. CCM2. The NCAR Community Climate Model Version 2 (CCM2) is an eighteen-vertical-layer pseudo-spectral atmospheric global GCM with spectral resolution T42, corresponding to a grid resolution of 2.8° latitude × 2.8° longitude. Compared to CCM1, the physical parameterizations are improved upon for a number of key climate processes, including: the planetary boundary layer, moist convection, transport, and clouds and radiation. A diurnal cycle is added. CCM2 can be run coupled to the Biosphere-Atmosphere Transfer Scheme (BATS) for more advanced treatment of land surface processes. For a complete description of CCM2, see Hack et al. [1993].

c. BATS. We plan to give The ANU Translator the capability of reading BATS output files. The purposes of BATS as coupled to the CCM are [Dickinson et al., 1993]:

- to calculate the net exchange of thermal infrared radiation of different surfaces and the fraction of incident solar radiation absorbed by them;
- to compute temperature and moisture values at land and sea ice surfaces;
- to determine the transfers of moisture, momentum, and sensible heat between the atmosphere and the surface; and
- to compute values for wind, temperature, and moisture at the level of surface observations, within vegetation canopies, and in the atmosphere.

d. RegCM2. RegCM2 is an updated version of the Penn State/NCAR Mesoscale Model Version 4 (MM4). Compared to a GCM, a mesoscale model has a limited spatial domain (usually continental scale or smaller) and relatively high spatial and temporal resolution. MM4 is a compressible, hydrostatic, primitive-equation atmospheric model with a horizontal Cartesian grid and vertical terrain-following sigma coordinates. The grid structure incorporates a staggering of model variables in the horizontal and the vertical. For a complete description of MM4 see Anthes et al. [1987]. The standard Kuo version of MM4 was enhanced to produce RegCM2 by Giorgi et al. [1993] who made the model more physically complete and more computationally efficient. RegCM2 includes the following improvements over MM4:

- use of a split-explicit time integration technique;
- incorporation of the radiative transfer scheme of CCM2;
- use of BATS to parameterize the surface physics;
- utilization of an improved formulation of cumulus convection; and
- inclusion of a lake model;

e. MM5. MM5 is the latest version of the Penn State/ NCAR mesoscale model. It incorporates a number of improvements and increased options compared to the previous version. It includes the option of running the model in non-hydrostatic mode. For a complete description of MM5 see Grell et al. [1994], and for information about running MM5, see the following URL: `http://www.mmm.ucar.edu/mm5/mm5-home.html`.

f. CSIRO9. CSIRO9 is a nine-vertical-level atmospheric GCM developed by the Australian Commonwealth Scientific and Industrial Research Organisation (CSIRO) Division of Atmospheric Research (DAR). It's resolution in spectral space is R21; in grid space it has 64 evenly spaced longitudes on the globe and 28 unevenly spaced latitudes in each hemisphere. CSIRO9 includes a diurnal cycle and does not include any biosphere modeling. Three layers are used for the computation of soil temperatures, with the upper layer being the surface; soil moisture is computed in the reservoir method. Ocean grid points adjacent to sea ice are modeled as 50 m deep mixed layer ocean (MLO) points; the location of MLO points changes as sea ice extent waxes and wanes. Expansion of the MLO points to include the entire ocean is in progress. CSIRO9 computes screen temperatures, facilitating diagnostic comparisons with observational data. The model includes a particle trajectory facility for analyzing flow patterns and estimating final positions of particles introduced into the atmosphere (e.g. volcanos, 1991 Gulf war oil fires). For a complete description of the model see McGregor et al. [1993].

g. ANU-CTM. The ANU Chemical Transport Model (ANU-CTM) is a three-dimensional Lagrangian model designed to advect air parcels in the atmosphere. Each air parcel represents a known mass of a tracer gas in air. The tracers are advected according to a stochastic scheme based upon ECMWF observational wind data on a $2.5° \times 2.5°$ grid with seven vertical levels. A mean and a time-varying component are included in the wind field. The flux of a trace gas is added or removed from air parcels at the lowest vertical level. Mixing ratios are computed by translating Lagrangian parcel coordinates to Eulerian grid coordinates. The model incorporates diffusive mixing between air parcels [Taylor, 1989; Taylor et al., 1991]. ANU-CTM is currently being extended using fourteen and fifteen level T42 ECMWF re-analyses which cover the period 1985-1994 [Trenberth, 1992]. The enhanced model will include a boundary layer and stratosphere and the capability of simulating inter-annual variability.

1.2 Descriptions of Visualization and Analysis Tools

a. Vis5D. Vis5D is a public-domain program for UNIX workstations designed for the interactive visualization of large 5-D gridded data sets. It was fashioned with numerical weather models in mind. The authors of the most recent version are Bill Hibbard and Brian Paul of the University of Wisconsin Space Science and Engineering Center. Vis5D is largely controlled through a simple-to-use graphical user interface (GUI). It can be used to make full color two dimensional and three dimensional static or animated renderings, and it has features for making contour line slices, colored slices, and volume renderings. Vis5D has special provisions for wind vectors and trajectories and for constructing isosurfaces. A number of variables can be viewed and animated at once. Several different map projections are available including the ability to view data on a spherical globe, and images can be easily rotated with the mouse. Graphical output may be saved and printed. For more information on Vis5D's requirements and capabilities, sample images of Vis5D output, and information on obtaining Vis5D, please see the following URL: `file://iris.ssec.wisc.edu/pub/www/vis5d.html`.

b. GrADS. The Grid Analysis and Display System (GrADS) is a public-domain program developed by Brian E. Doty at the Center for Ocean-Land-Atmosphere Interactions at the University of Maryland. It is an interactive tool for the analysis, manipulation, and display of earth science data on UNIX workstations and DOS based PC's. GrADS expects input data sets to be described by four dimensions, usually: latitude, longitude, level, and time. GrADS includes the following features [`http://grads.iges.org/grads/gradover.html`]:

- it can handle station data and gridded data, including non-linearly spaced grids and variable resolution data;
- different data sets can be both compared analytically and graphically overlaid;
- user-defined or built-in functions can be performed on the data interactively;
- graphics capabilities include: bar graphs, line graphs, scatter plots, contour lines and shaded contours, wind vectors, streamline plots, grid and shaded grid boxes, and station model plots. The user can control many aspects of the graphical output; the default settings are geophysically intuitive. Graphics may be output in Postscript for color or monochrome printing;
- GrADS includes an interpreted scripting language. Scripts can be written to include widgets and accept mouse input from the user and to automate complex calculations or displays; and
- GrADS may be run in batch mode.

For more information on GrADS, including sample images and instructions on how to obtain the software, please see the following URL:
`http://grads.iges.org/grads/head.html`.

c. netCDF. The network Common Data Form (netCDF) is a widely used, freely available format, interface, and library for the creation, access, and sharing of scientific data. It was created by Russ Rew, Glenn Davis, and Steve Emmerson at the Unidata Program Center. NetCDF provides a machine-independent format for the representation of scientific data, an interface for accessing the data, and a library that specifies implementation of the interface. A large number of commercial and free software packages can read netCDF data, including but not limited to the following: AVS, GMT, HDF, IDL, NCAR Graphics, and Spyglass Dicer and Slicer. CSIRO stores GCM results in netCDF form. Many more software packages are currently being expanded to include netCDF capabilities [`http://www.unidata.ucar.edu/packages/netcdf/`].

d. AVS. Advanced Visual System's Application Visualization System (AVS) is a commercially available, highly generic and customizable, powerful visualization environment. For more information on AVS, see the following:

1. `http://www.mcnc.org/HTML/ITD/IAC/IAC.html`,
2. `http://www.avs.com/products/avs.html`,
3. the file `/avs_readme/WHAT_IS_AVS` available from the anonymous ftp site `avs.ncsc.org`.

e. HDF. The National Center for Supercomputing Applications' (NCSA) Hierarchical Data Format (HDF) is a public-domain multi-object file format that expedites the transfer of various data types between different machines and operating systems. HDF data sets can be n-Dimensional and are self-defining and easily extendible. HDF includes utilities to create images [`http://hdf.ncsa.uiuc.edu:8001`].

f. GMT. Generic Mapping Tools (GMT) is a package created and maintained by Paul Wessel and Walter H. F. Smith. It is a collection of about 50 UNIX tools for manipulating 2D and 3D data and producing illustrations. GMT is freely available to government agencies and nonprofit educational organizations worldwide [`http://www.soest.hawaii.edu/soest/gmt.html`].

2. METHODS

2.1 Structure of The ANU Translator

The ANU Translator is being designed with the following three goals in mind: 1) portability, 2) ease of use, and 3) extensibility. Portability is achieved by the avoidance of implementation-specific coding: ANSI C is used with a standard Tcl/Tk interpreter [Ousterhout, 1994]. Tcl and Tk encapsulate the graphical and the non-graphical user interfaces. The GUI is designed so that a user can run the Translator without relying significantly on documentation (Figure 1). The ease of use should assist researchers and greatly facilitate the introduction of climate modeling to students.

The extensibility of The ANU Translator is being achieved primarily through the modular nature of the C code and the use of an internal representation type common to all input and output types (Figure 2). The C code has four basic components: 1) Introduction, 2) Interrogation, 3) Extraction, and 4) Output. The "Introduction" module receives the user's request for file interrogation from

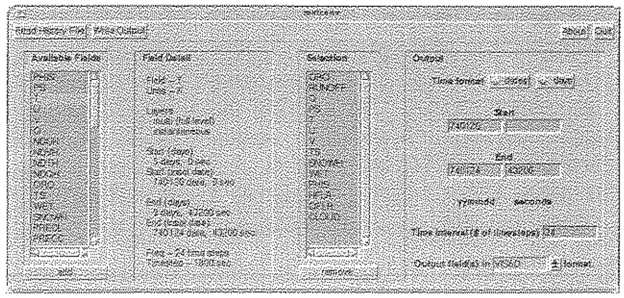

Figure 1: The graphical user interface (GUI) of The ANU Translator.

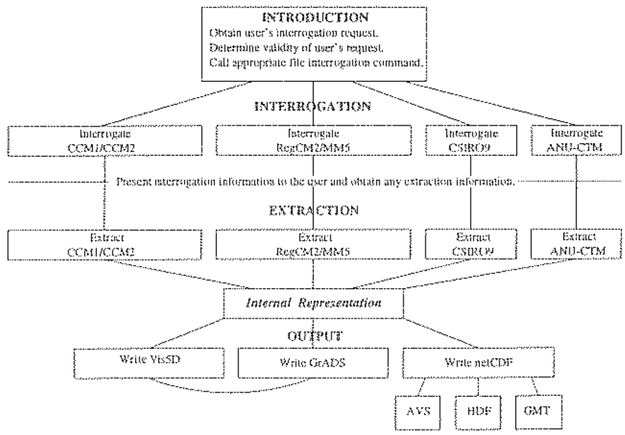

Figure 2: The structure of The ANU Translator program.

Tcl. It determines whether the user's requested model type is within the Translator's capabilities and calls the appropriate "Interrogation" module. There is one "Interrogation" module for each general model type, (for coding purposes CCM1 and CCM2 are considered one general model type, as are RegCM2 and MM5). The "Interrogation" modules read the model output files requested by the user and indicate what information is present in the files. This information is presented graphically when the program is being run in graphical mode. The Translator then receives any extraction requests from Tcl. An "Extraction" module exists for each general model type. Data is extracted from each model into a common internal representation. The "Output" stage proceeds from this internal representation format, with each different output type having its own "Output" module. We hope to include the ability to share output tools, especially where a visualization tool does not include adequate numerical analysis capabilities, e.g. data could be averaged in GrADS and then funneled into Vis5D.

Rather than having one reading and writing routine for each input-to-output conversion type, the Translator is designed so that each input type requires only one reading routine, and each output type requires only one writing routine. A researcher wanting to add an input model type to the Translator should be able to do so by adding new "Interrogation" and "Extraction" modules and making small adjustments to the "Introduction" section and the GUI Tcl/Tk code. Similarly, a new output type can be added easily by writing one "Output" module.

2.2 Hardware and Software Requirements

The ANU Translator should run on any UNIX-based system. Because climate model output data are usually extensive, the Translator is being designed in such a way that all of the data being analyzed together will not have to be in memory at once. Actual memory limitations and disk space requirements are not known at this time.

Compilation of The ANU Translator requires an ANSI C compiler and a Tcl installation. To run the graphical version of the Translator, Tk and an X interface are also necessary. We expect to make the Translator compatible with the most recent non-beta versions of Tcl, Tk, Vis5D, GrADS, and netCDF available at the time of release. Input to the Translator will be in the standard format for the particular model type being input. An installation of one or more of the packages compatible with the Translator's output will be needed for the actual visualization and/or analysis. The C compiler, Tcl, Tk, most of the climate models, and most of the analysis and visualization tools discussed in this report may all be obtained free of charge via the Internet.

2.3 Availability

We plan to make The ANU Translator freely and publicly available via anonymous ftp and the World Wide Web, complete with easy-to-follow installation instructions. We hope that other researchers will add modules to the code to increase its range of model input types and its analysis and visualization output types. We believe that The ANU Translator will assist the research and teaching purposes of the climate modeling community.

3. DATA AND RESULTS

The RegCM2 output shown (Figure 3) is from a run conducted by Jay Larson on the Fujitsu VP 2200 at ANU. The model run incorporates a seventeen-vertical-level grid with a horizontal resolution of 16 km, and it begins on radiation day 1 December 1991. The grid is centered over Japan at 36°N, 139°E, and is subdivided into 40 latitudes and 41 longitudes. The output shown in Vis5D is for the last time step of the second day, i.e. 2 December 1991, 18Z. The topography and the map boundaries are taken from the Vis5D datasets.

The CSIRO9 output shown (Figure 4) is from a $2 \times CO_2$ Greenhouse run (660 ppmv) conducted by Martin Dix and Ian Watterson of CSIRO DAR on the CSIRO Cray Y-MP in August-October 1991. The $2 \times CO_2$ run was initialized from the state at the end of the ten-year climatology run described in McGregor et. al [1993]. The ten-year climatology run relied on prescribed sea surface temperatures

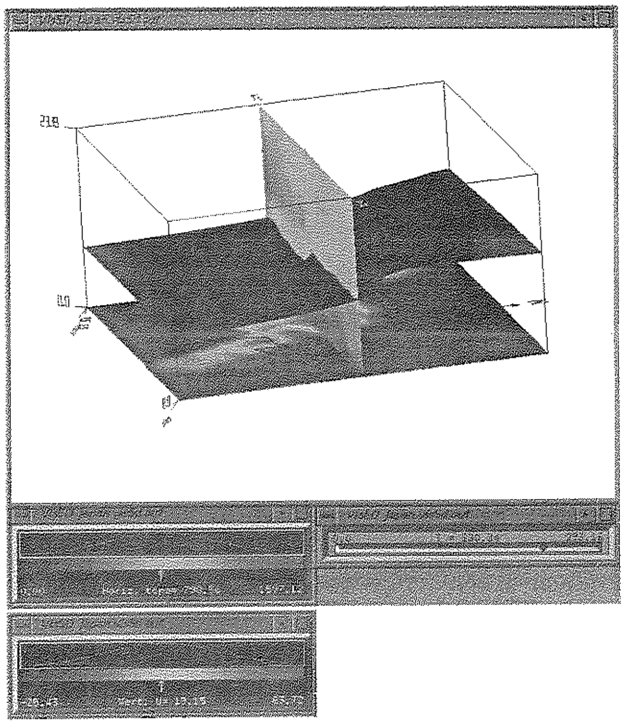

Figure 3: Output from RegCM2 run over Japan displayed in Vis5D using a rectilinear projection. The time step is 18Z on day 2 of the run, corresponding to radiation day 2 December 1991. The horizontal surface slice shows model topography (topo) in m. The lightly shaded vertical slice shows the zonal wind (u), in m s^{-1}. The dark 3-D surface is the temperature (T) isosurface at 230 K. The black lines on the surface indicate political boundaries.

(SSTs). Surface heat fluxes from the ten-year climatology run were used to generate implied ocean heat fluxes (q-flux); these were supplied to the 50 m mixed layer ocean model included in the Greenhouse run. The $2 \times CO_2$ run was allowed to equilibrate for 27 years and then continued for 30 additional years to form a climatology. The data shown are the 30-year July monthly averages.

Several features of Vis5D are illustrated by the RegCM2 and CSIRO9 output displayed here, including:

- horizontal and vertical colored slices;
- two-dimensional wind vector maps;
- three-dimensional isosurfaces;
- ability of the user to choose among three-dimensional projection types;
- display of the map and topography datasets built into Vis5D; and
- simultaneous display of multiple variables.

Static, black and white, two-dimensional images do not adequately convey the power and flexibility of computer software, especially GUIs and visualization packages. At MODSIM '95 we will present a computer display of the ANU Translator and sample model output animated within Vis5D.

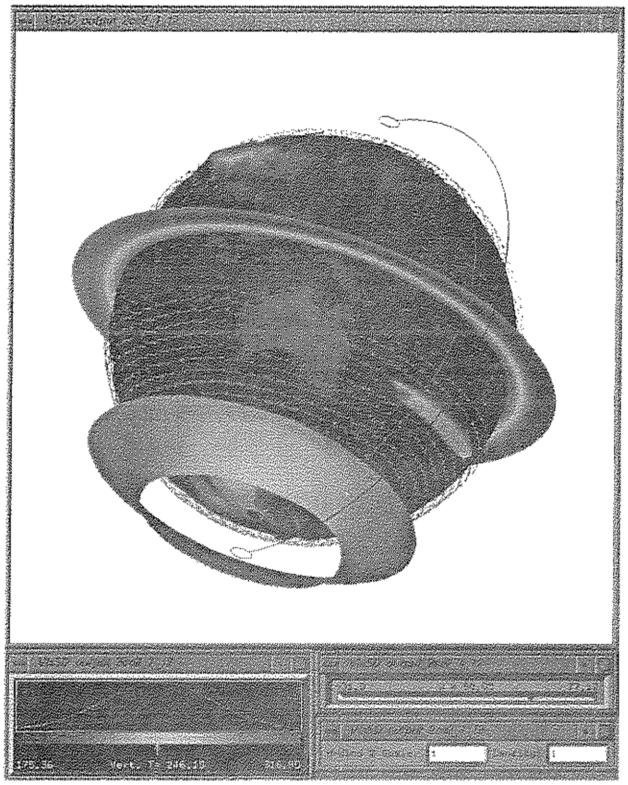

Figure 4: Output from the CSIRO9 2 × CO_2 (660 ppmv) displayed spherically in Vis5D. Data are July 30-year averages. A vertical slice of temperature (T) in K is shown over the equator. The horizontal uv wind vector slice is in m s^{-1}. The isosurface shows the zonal wind (u) at 50 m s^{-1}.

4. CONCLUSIONS

The ANU Translator is a work in progress. We believe that The Translator has a great deal to offer the climate modeling community as a research tool and a teaching aid. We hope that it will enjoy wide usage and will form the basis for a cooperative expansion of climate model visualization efforts.

Acknowledgments. A grant from the Quality Assurance Program of the Department of Employment Education and Training supported this work (DEET R30884). We thank CSIRO Division of Atmospheric Research for providing the CSIRO9 data.

5. REFERENCES

Anthes, R.A., E. Hsie, and Y. Kuo, *Description of the Penn State/NCAR Mesoscale Model Version 4 (MM4),* NCAR Tech. Note, NCAR/TN-282+STR, 66 pp., Boulder, Colorado, 1987.

Buja, L.E., *Introduction to the UNICOS CCM Processor,* NCAR Tech. Note NCAR/TN-383+IA, 62 pp., Boulder, Colorado, 1992.

Buja, L.E., *CCM Processor Users' Guide (UNICOS Version),* NCAR Tech. Note, NCAR/TN-384+IA, Boulder, Colorado, 1993.

Dickinson, R.E., A. Henderson-Sellers, and P.J. Kennedy, *Biosphere-Atmosphere Transfer Scheme (BATS) Version 1e as Coupled to the NCAR Community Climate Model,* NCAR Tech. Note, NCAR/TN-387+STR, 72 pp., Boulder, Colorado, 1993.

Giorgi, F., G.T. Bates, and S.J. Nieman, The multiyear surface climatology of a regional atmospheric model over the western United States, *J. Climate, 6,* 75-95, 1993.

Grell, G.A., J. Dudhia, and D.R. Stauffer, *A Description of the Fifth-Generation Penn State/NCAR Mesoscale Model (MM5),* NCAR Tech. Note, NCAR/TN-398+STR, 138 pp., Boulder, Colorado, 1994.

Hack, J.J., B.A. Boville, B.P. Briegleb, J.T. Kiehl, P.J. Rasch, and D.L. Williamson, *Description of the NCAR Community Climate Model (CCM2),* NCAR Tech. Note, NCAR/TN-382+STR, 108 pp., Boulder, Colorado, 1993.

McGregor, J.L., H.B. Gordon, I.G. Watterson, M.R. Dix, and L.D. Rotstayn, *The CSIRO 9-level Atmospheric General Circulation Model,* CSIRO Division of Atmospheric Research Tech. Paper No. 26, 89 pp., CSIRO, Australia, 1993.

Ousterhout, J.K., *Tcl and the Tk Toolkit,* Addison-Wesley, 458 pp., Reading, Massachusetts, 1994.

Taylor, J.A., A stochastic Lagrangian atmospheric transport model to determine global CO_2 sources and sinks - a preliminary discussion, *Tellus, 41B,* 272-285, 1989.

---, G.P. Brasseur, P.R. Zimmerman, and R.J. Cicerone, A study of the sources and sinks of methane and methyl chloroform using a global three-dimensional Lagrangian tropospheric tracer transport model, *J. Geophys. Res., 96*(D2), 3013-3044, 1991.

Trenberth, K.E., *Global analyses from ECMWF and atlas of 1000 to 10 mb of circulation statistics,* NCAR Tech. Note, NCAR/TN-373+STR, 191 pp., Boulder, Colorado, 1992.

Williamson, D.L., J.T. Kiehl, V. Ramanathan, R.E. Dickinson, and J.J. Hack, *Description of NCAR Community Climate Model (CCM1),* NCAR Tech. Note, NCAR/TN-285+STR, 112 pp., Boulder, Colorado, 1987.