

**CRC Report No. A-55**

**IMPLEMENTATION OF THE  
TRAJECTORY-GRID ADVECTION  
ALGORITHM IN CMAQ**

**Final Report**

**June 30, 2006**



**COORDINATING RESEARCH COUNCIL, INC.**  
3650 MANSELL ROAD·SUITE 140·ALPHARETTA, GA 30022

## **Final Report**

### **CRC Project A-55**

#### **Implementation of the Trajectory-Grid Advection Algorithm in CMAQ**

June 30, 2006

by

Daewon W. Byun and Peter B. Percell

Institute for Multidimensional Air Quality Studies (IMAQS)  
312 Science and Research Building 1, University of Houston  
Houston, Texas, TX 77204-5007

#### **Abstract**

Numerical diffusion is a serious problem for Eulerian advection algorithms typically used in chemical transport simulators for air quality modeling. This extra diffusion damps local extremes and smears steep gradients in trace species concentrations, for example, those due to localized emissions of primary pollutants such as NO<sub>x</sub> and VOC's. This, in turn, compromises the model's chemistry calculations of the production of secondary pollutants, such as ozone, which depend nonlinearly on the relative concentrations of the precursors. In order to remove this source of inaccuracy, a Lagrangian advection algorithm called Trajectory-Grid (TG) has been implemented in the EPA's Community Multiscale Air Quality (CMAQ) model. The TG algorithm follows representative air "packets" along trajectories determined by the wind field. The air composition in a packet is unchanged by advection. The resulting prototype simulation model is called CMAQ-TG.

CMAQ-TG was compared to standard CMAQ for two-dimensional rotating, shearing, and stretching flows and a realistic three-dimensional wind field. Compared with the standard CMAQ advection algorithm, i.e., the piecewise parabolic method (PPM), CMAQ-TG advection produces a non-diffusive solution and maintains linearity of the advection process. Initial implementation has been limited to the simulation of photochemical air quality with an emphasis on the transport processes in CMAQ.

## 1. Introduction

Air quality models contain many components: (1) model assumptions, which are contained in the equations to be solved and the accompanying formulas and parameters, (2) input data, which include initial and boundary conditions, emissions inventory, meteorological information, etc., and (3) numerical methods for solving the equations using the input information. Because many of the model assumptions cannot be readily verified, large uncertainties are inherent in the input information and one encounters great numerical difficulties in accurately describing the advection of trace species in the air, it is difficult to be certain that the model predictions are indeed reliable, even when the model sometimes yields reasonable results compared to observation. This is because the model may yield the right result for the wrong reasons. A logical first step in improving the reliability of model predictions is to remove the potential errors inherent in the numerical methods. Once this has been done, errors in model assumptions can be better exposed and, potentially, corrected with greater confidence.

The most serious numerical problem encountered in air quality modeling is the solution of the advection equation, which tracks species concentrations as air is transported by the wind from one location to another. The typical numerical methods are based on an Eulerian frame of reference. Because of the need for discretization, substantial numerical errors occur due to constructive and destructive interferences as the different Fourier modes of the concentration profile move at different phase speeds. Attempts to retain certain key characteristics of the profile - such as non-negativity, mass conservation, etc. - lead to unpredictable distortions of the profile. This error alone can represent a substantial portion of errors in model predictions.

Chock et al. (1996, 2005) introduced a Lagrangian or trajectory approach, called Trajectory-Grid (TG) that transports points on a concentration profile along their trajectories and also automatically conserves mass, sign, and shape of the profile accurately. As demonstrated by the earlier investigations, Lagrangian advection methods need to be tested under a variety of flow conditions to demonstrate their numerical characteristics. Because the methods follow the motion of individual “representative” fluid particles (packets), they may also be called quasi-particle methods. Oran and Boris (2001) provide an overview of such methods as the *particle-in-cell* (PIC) method (Evans and Harlow, 1957; Harlow, 1964), the *marker-and-cell* (MAC) method (Harlow and Welch, 1965) and the *grid-and-particle* method (GAP) (Marder, 1975) which have been applied to related problems, and are conceptual cousins of TG. Because the trace species of interest for air quality studies do not significantly affect dynamic flow characteristics, TG takes input of meteorological and emission data from a regular Eulerian mesh as the basis of the computation, and the packets do not have to represent the dynamic changes in the flow characteristics.

In TG, Eulerian grids are used to solve the diffusion equation and other science processes such as cloud mixing and aqueous-phase chemistry. However, in addition to advection, other processes such as deposition, gas chemistry and aerosols can be calculated for individual packets and this is done whenever possible. The approach has other advantages. Since it transports all species at once, it saves execution time for the advection process, and the saving increases relative to Eulerian approaches as the number of species increases (Chock and Winkler, 2000; Chock et al., 2002). This advantage has been exploited in the application of the TG approach to solving the

condensation/evaporation equations in the aerosol module developed by Spyros Pandis of Carnegie Mellon University (Gaydos et al., 2003). Since the partial differential equations for advection are no longer solved, the approach does not have the constraint of the Courant number for numerical stability. Since one can modify packet density at will, a straightforward approach to grid nesting can be designed. However, the approach has problems of its own, which arise because each packet in a profile has infinitesimal spatial resolution, whereas a gridded model has a finite spatial resolution and, in theory, represents the whole computational domain, rather than a collection of samples.

Because convergence and divergence occur in air flow, initially evenly distributed packets can quickly become unevenly distributed, necessitating packet management, which requires packet density higher than that of the grid resolution in regions of interest. However, all these issues can be resolved without incurring excessive computation time and interpolation errors. The TG approach has been incorporated into the Comprehensive Air Quality Model with Extensions (CAMx) (Environ, 2004) and has been used for actual modeling applications. A similar implementation of the TG approach in the Community Multiscale Air Quality (CMAQ) modeling system (Byun and Ching, 1999) can similarly help to improve the predictive reliability of that model.

The purpose of the current project has been to incorporate the TG approach into CMAQ, test the resulting model, study actual applications, and possibly identify and remove errors in the model. Initial implementation has been limited to the simulation of photochemical air quality emphasizing the transport processes in CMAQ. Additional modification of aqueous-phase and aerosol process codes is needed, but deferred for future work. It is hoped that ultimately the process begun with this project will be a much improved CMAQ for scientific studies and regulatory applications. Although we verified the newly implemented code by comparing its results with the standard CMAQ here, further characterization of its numerical behavior and evaluation of the system will obviously be needed in the longer term.

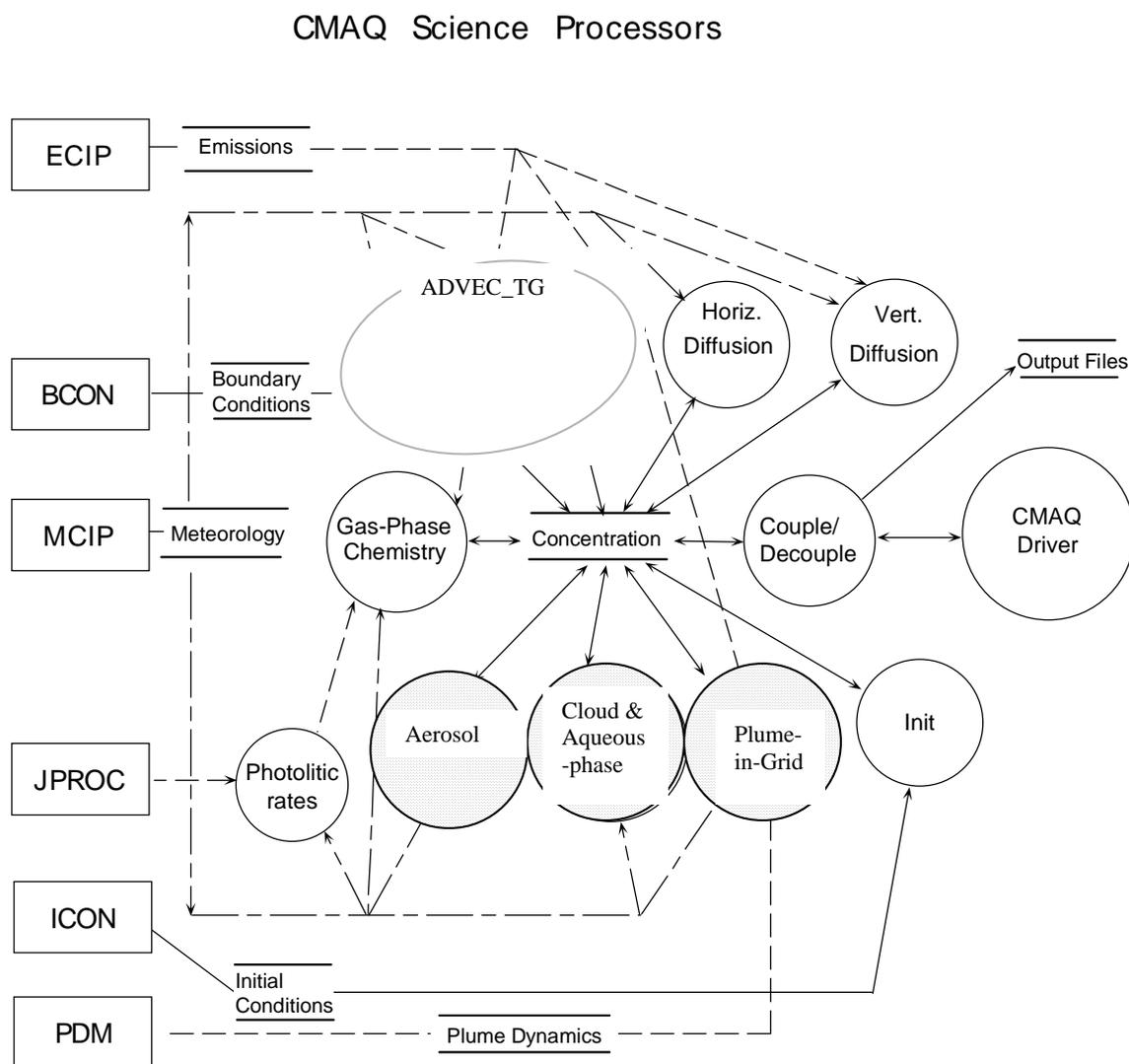
For the purposes of this report, “standard” CMAQ means CMAQ v4.4 with

- PPM (Piecewise Parabolic Method) used for horizontal and vertical advection
- Eddy diffusion (K-theory) used for horizontal and vertical diffusion
- CB4 gas chemistry mechanism and EBI-CB4 solver

## 2. Implementation of the Chemical Transport Module (CTM) in CMAQ-TG

In this section, we give a brief description of the key features of the TG implementation in CMAQ code (CMAQ-TG).

Figure 1 shows the suite of science modules in CMAQ. None of the modules that process and provide input to the CTM has been modified for TG, so CTM-TG uses the same meteorology, emissions, initial conditions, boundary conditions, and photolytic rates input as CMAQ's CTM, making "apples-to-apples" comparison of CMAQ and CMAQ-TG possible.



**Figure 1 Science process modules in CMAQ. The present version of the CMAQ-TG does not include the aerosol, cloud/aqueous-phase, and plume-in-grid modules (shaded circles).**

Replacing the separate horizontal advection (HADV) and vertical advection (VADV) modules with a 3-D Lagrangian TG advection module called ADVEC\_TG is the primary modification for CMAQ-TG. Because the TG approach transports packets, each with a set of mixing ratios, it conserves mass as long as the input meteorology data are mass consistent (see Chock et al., 2005). Therefore, there is no need to invoke a mass correction routine as in the Eulerian CMAQ. Also, CMAQ-TG does not use Couple/Decouple to convert between mixing ratios and concentrations, but instead maintains gas phase composition solely in the form of mixing ratios. The present prototype version of CMAQ-TG just supports advection, diffusion, gas chemistry, emissions, and dry deposition.

The following is a brief description of the core physical processes currently simulated over a common time step of CMAQ-TG's Chemical Transport Model.

### **2.1. The role of the Eulerian grid in CMAQ-TG**

Input to CMAQ, such as meteorological data, emissions data, initial conditions, and boundary conditions, is all associated with the cells of an Eulerian grid on the model domain. Also, output from CMAQ is most easily saved as grid data using netCDF (network Common Data Form), an interface for array-oriented data access (see, <http://www.unidata.ucar.edu/software/netcdf/>), and viewed using the Package for Analysis and Visualization of Environmental data (PAVE) available from <http://www.cmascenter.org>. For these two reasons alone, it is necessary to easily relate packet position to grid cell indices. Furthermore, some processes that are potentially global in nature, such as eddy diffusion, do not have satisfactory Lagrangian models at this time and are therefore modeled on the Eulerian grid. Finally, grouping packets in the same grid cell together is a convenient way to keep track of which packets are close to each other. Thus, the primary data structure for listing packets in CMAQ-TG is an array of lists of the packets in each grid cell, indexed by the column, row, and layer indices of the CMAQ Eulerian grid. The main data used to characterize each packet are simply the (x, y, z) coordinates of its location in the model domain and an array of mixing ratios for the species of interest.

Although an Eulerian grid is necessary in the current CMAQ-TG environment, whenever possible, it plays no role once input data have been interpolated to packet locations. For example, this is the case for the advection, dry deposition, emissions, and gas chemistry processes.

### **2.2. Time step selection**

The common time step for the CMAQ-TG science process modules, called the synchronization time step, is selected to satisfy the condition that a packet may not travel more than  $\frac{3}{4}$  of the distance between opposite faces of an Eulerian grid cell in any of the 3 coordinate directions in a single time step. This Courant-like condition is not based on a stability requirement of the advection solution, but rather on accuracy and a condition that a packet not "jump over" any grid cell that receives injections of emissions. Since TG packets move in all 3 dimensions simultaneously and can move freely between grid cells, the same time step must be used for packets in cells at all layers. Note that CMAQ's algorithms for simulating the individual physical processes have the option of subdividing the synchronization time step, as appropriate.

### 2.3. External sources and sinks (e.g., emissions and deposition)

Emissions input, given as a moles/sec rate of injection into grid cells, is converted to a change in mixing ratios based on cell volume. Once this has been done, the same change in mixing ratios is applied to each packet in a cell. Dry deposition is calculated packet-by-packet. In standard CMAQ, these processes are treated as sub-processes of vertical diffusion, which allows their deltas to be applied to cells at the frequency of subdivided vertical diffusion time steps. This practice may increase accuracy, but it defeats modularity. In CMAQ-TG, these processes have been broken out as independent modules, allowing greater flexibility in managing the sequence in which the CTM science processes are performed. Source and sink processes are performed at the start of each synchronization time step, before the transport processes of advection and diffusion.

### 2.4. Three dimensional advection and packet management

A packet is a small volume of air that moves along a trajectory determined by the wind field. Thus, the trajectory of a packet evolves in time following an ordinary differential equation

$$\frac{d\mathbf{y}(t)}{dt} = \mathbf{V}(\mathbf{y}(t), t) \quad (1)$$

where  $\mathbf{V}$  is the three-dimensional wind velocity and  $\mathbf{y}(t)$  is the position vector at time  $t$ .

We solve this equation using a simple predictor-corrector scheme. The predictor step gives an initial estimate of the new position ( $\mathbf{y}^i$ ) at the end of a time step as

$$\mathbf{y}^i(t + \Delta t) = \mathbf{y}(t) + \mathbf{V}(\mathbf{y}(t), t)\Delta t \quad (2)$$

The final position ( $\mathbf{y}^f$ ), computed by the corrector step, is given as

$$\mathbf{y}^f(t + \Delta t) = \mathbf{y}(t) + 0.5[\mathbf{V}(\mathbf{y}(t), t) + \mathbf{V}(\mathbf{y}^i(t + \Delta t), t + \Delta t)]\Delta t \quad (3)$$

A packet carries with it a set of mixing ratios for the species of interest. These mixing ratios are unchanged by the advection process. Therefore, after updating boundary packets with current boundary conditions, all that needs to be done for an advection step is to move each packet one time step along its trajectory. This is done by interpolating wind velocity to the appropriate times and positions and using the simple predictor-corrector method just described to update packet positions. Once the new packet positions are known, the grid cell association of packets is updated and packets that have left the solution domain are removed.

#### Packet Management

Packet management is done to balance the density of packets in cells between the competing goals of covering the solution domain - to ensure completeness and accuracy - and limiting

computational expense. The packet management tasks of “spawning” and “pruning” are done as part of the advection process. The first step in packet management is setting the initial number of packets in each cell.

The following is an overview of packet management as currently implemented in CMAQ-TG.

### High Resolution (HR) Cells

- Defined by a multiplier, N, and a range of columns, rows, and layers. (Defaults: N=2, all columns and rows, first 2 layers)
- HR cells, including boundary cells, are initialized with  $N^2$  uniformly spaced packets at the vertical center of the cell. Normal resolution cells are initialized with just one packet at the cell center.
- High resolution cells are never allowed to be empty, except when spawning is turned off.

### Spawning

Create new packets in cells with too few packets. Spawning is done every time step by creating a new packet at the center of a cell, when needed. The initial composition in a spawned packet is estimated from nearby packets. HR cells are never allowed to be empty, except when spawning is turned off. The options are:

- No Fill - never add packets to interior cells; i.e., turn off spawning
- Fill All - never allow any cell to be empty (default)
- Sparse Fill - spawn in an interior normal resolution cell if it is empty and is surrounded, in its layer, by empty cells

When a boundary cell becomes empty, it is refilled with packets the same way it was initially filled.

### Pruning

Remove packets from cells that have become overpopulated. The options are:

- No pruning
- Keep packets that are closest to the cell center (default)
- Keep packets that are oldest
- Parameters
  - Pruning frequency (number of time steps) (default: 5)
  - Number of packets to keep after pruning (default:  $N^2$  in HR cells; 2 in other cells)
  - Tolerance above number to keep before pruning is triggered (default:  $N^2$  in HR cells; 2 in other cells)

Although packet management has strong benefits, it has some negative aspects as well.

- Spawning requires initializing new packet composition using some type of interpolation which is likely to be somewhat inaccurate and diffusive.
- Aggressive spawning, e.g., never allowing a cell to be empty, may be unnecessary if a nearby packet will soon move into the cell.
- Pruning can throw away good composition information which could be used later in the simulation if the need for spawning a replacement is discovered.

The options for packet management that are currently available are described in detail in Appendix A.

## 2.5. Diffusion

The underlying algorithms for both vertical and horizontal diffusion are very close to those in standard CMAQ (see Byun and Ching, 1999, for details). The discussion here focuses on the changes made to accommodate packets.

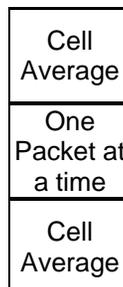
The requirements we set for the CMAQ-TG diffusion algorithms are:

1. For each species and cell, the average of the packet mixing ratios after a diffusion step should be the same as the cell mixing ratio obtained from Eulerian diffusion of the previous averages.
2. For each species, the maximum and minimum packet mixing ratios after a diffusion step should be within the range of values at the beginning of the step; in particular, this ensures that positive mixing ratios remain positive and that the maximum does not increase.
3. Linearity.

Initially we thought that horizontal and vertical diffusion could be adequately handled by doing Eulerian diffusion of cell averages and then adding cell deltas to the individual packets in cells. However, testing showed that this fast and simple procedure fails to satisfy our second requirement (see Figure 22). Then, an initial method we tried for fixing this problem failed to satisfy the third requirement, linearity (see Figure 23). These problems have been fixed by the algorithms described here (see Figure 24 and Table 3).

### Vertical Diffusion

- Set cell mixing ratios to the average over the cell's packets.
- In a loop over vertical diffusion sub-steps of the synchronization time step:
  - Solve the implicit Eulerian equations to find the cell averages at the end of the time step.
  - Find the mid-step average mixing ratios in cells.
  - With a packet representing its own cell and cell averages representing the neighboring cells, solve the Eulerian equations for the packet mixing ratios at the end of the time step using a predictor-corrector method.



**Note:** Solving the implicit Eulerian equations for each packet, rather than using a predictor-corrector method, does better at satisfying the first requirement for CMAQ-TG diffusion, but testing indicated that the slight improvement obtained is not worth the large increase in execution time.

### Horizontal Diffusion

- Set cell mixing ratios to the average over the cell's packets
- In a loop over horizontal diffusion sub-steps of the synchronization time step
  - Update each packet's mixing ratios with an explicit step of Eulerian diffusion between the packet representing it's own cell and cell averages representing the neighboring cells

	Cell Average	
Cell Average	One Packet at a time	Cell Average
	Cell Average	

- Update cell mixing ratios with an explicit step of Eulerian diffusion
- Do sub-grid diffusion
  - Pair-wise diffusion between packets in a cell is approximated by diffusion between each packet and an imaginary packet with the cell's average mixing ratios
  - The distance between each packet and the imaginary average packet is taken to be a little less than half the cell width

Note that sub-grid diffusion leaves the average over the packets unchanged

The algorithm for sub-grid diffusion in CMAQ-TG is very similar to that described in Chock et al., 2005. As was noted in that paper, further investigation of this algorithm is clearly needed.

Indeed, one of the major reasons for considering the TG approach is to remove the numerical diffusion characteristic of Eulerian advection algorithms, which typically dominates horizontal physical diffusion, so that the methods for simulating physical diffusion can be studied and improved. We hope to pursue this issue in future work.

## **2.6. Chemistry and other processes**

Chemistry in CMAQ-TG is processed for packet mixing ratios instead of grid cell mixing ratios. Otherwise, it is unchanged from standard CMAQ. In the future, support for all CMAQ processes, including aerosols and cloud/aqueous phase, will be added.

## 2.7. Output

Several representations of cell composition are available, each in its own file, including:

- AVG\_MIX: species-wise average mixing ratios over the packets in the cell
- CLS\_MIX: mixing ratios for the packet horizontally closest to the cell center
- MAX\_MIX: species-wise maximum mixing ratios over the packets in the cell
- MIN\_MIX: species-wise minimum mixing ratios over the packets in the cell
- OLD\_MIX: mixing ratios for the oldest packet in the cell

The PACKET file contains the following information for each cell

- COUNT: Number of packets
- NEW\_PACKETS: Number of new packets added during output step
- AVG\_AGE: Average age of the packets
- MAX\_AGE: Age of the oldest packet

At this time, output is limited to grid cell quantities and we depend on PAVE to view results. The ability to use a more flexible display tool, that can display packet quantities directly, is needed.

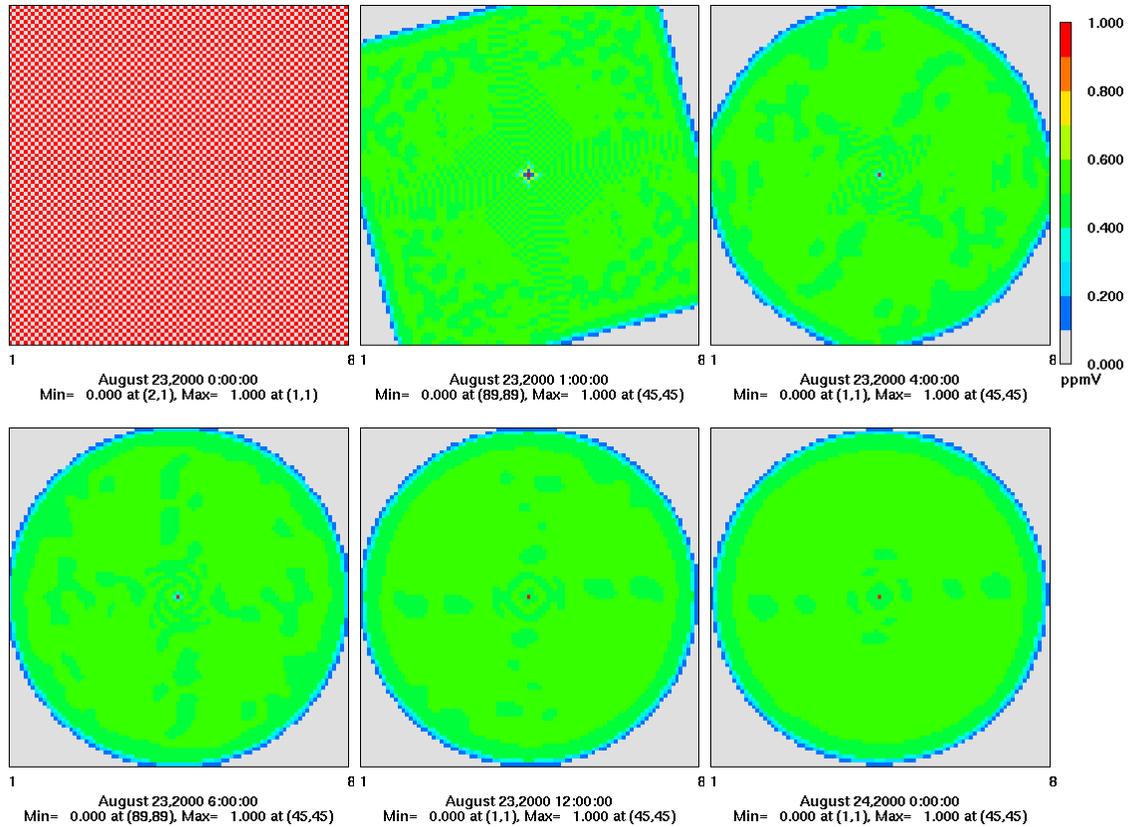
## 3. Evaluation of CMAQ-TG with Idealized Two-dimensional Wind Fields

In any numerical model development effort, new code must be thoroughly tested to verify that the design is valid and the implementation is correct. In this section, we evaluate the horizontal transport processes for properties such as preservation of peak values. First we look at advection alone, and then at advection with horizontal diffusion. We also examine the effectiveness and qualitative behavior of packet management strategies.

### 3.1. Test of TG advection with a two-dimensional rotational wind field

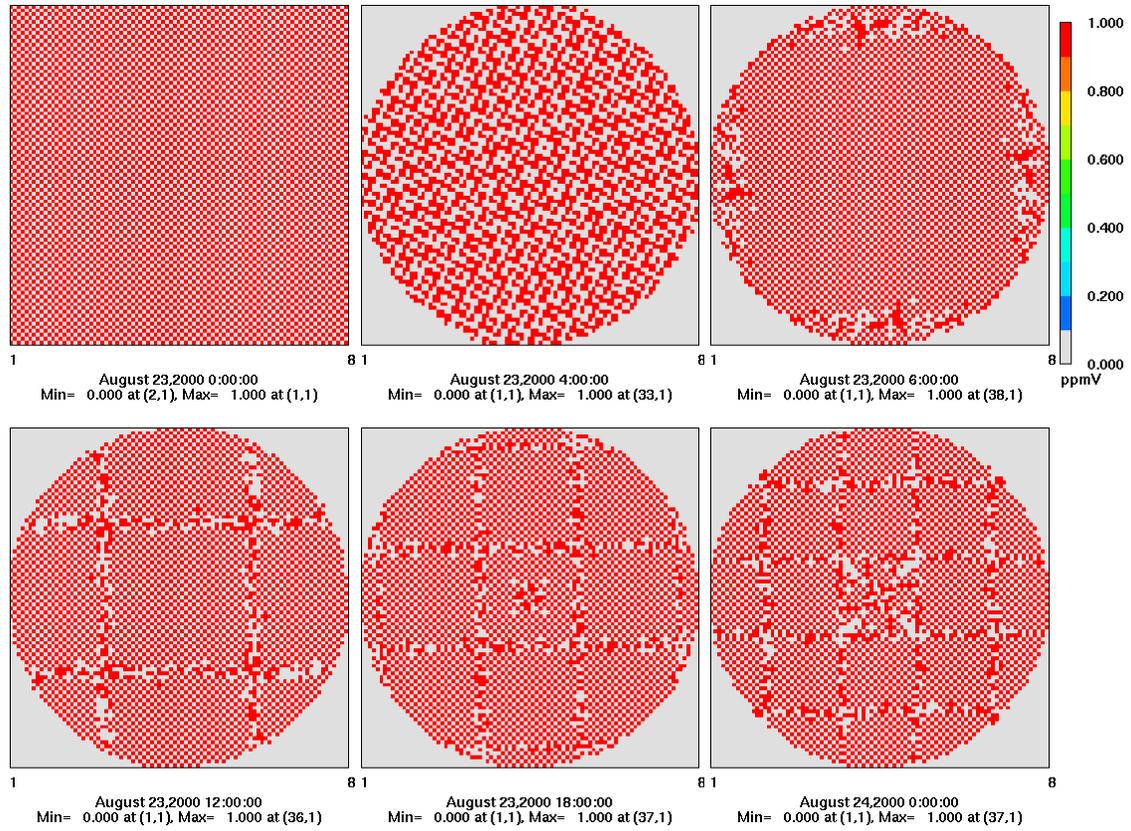
The wind field used in this section is a steady, two-dimensional rotation through  $360^\circ$  every 24 hours.

One of the critical shortcomings of Eulerian advection algorithms is numerical diffusion. As demonstrated by Figure 2, a high frequency  $2\Delta x$  signal, the checkerboard pattern, is lost when advection is simulated using the Eulerian piecewise parabolic method (PPM) in standard CMAQ. Note, however, that the PPM scheme does a much better job of preserving lower frequency signals, such as  $8\Delta x$  and  $16\Delta x$  (see Byun and Lee, 2002).

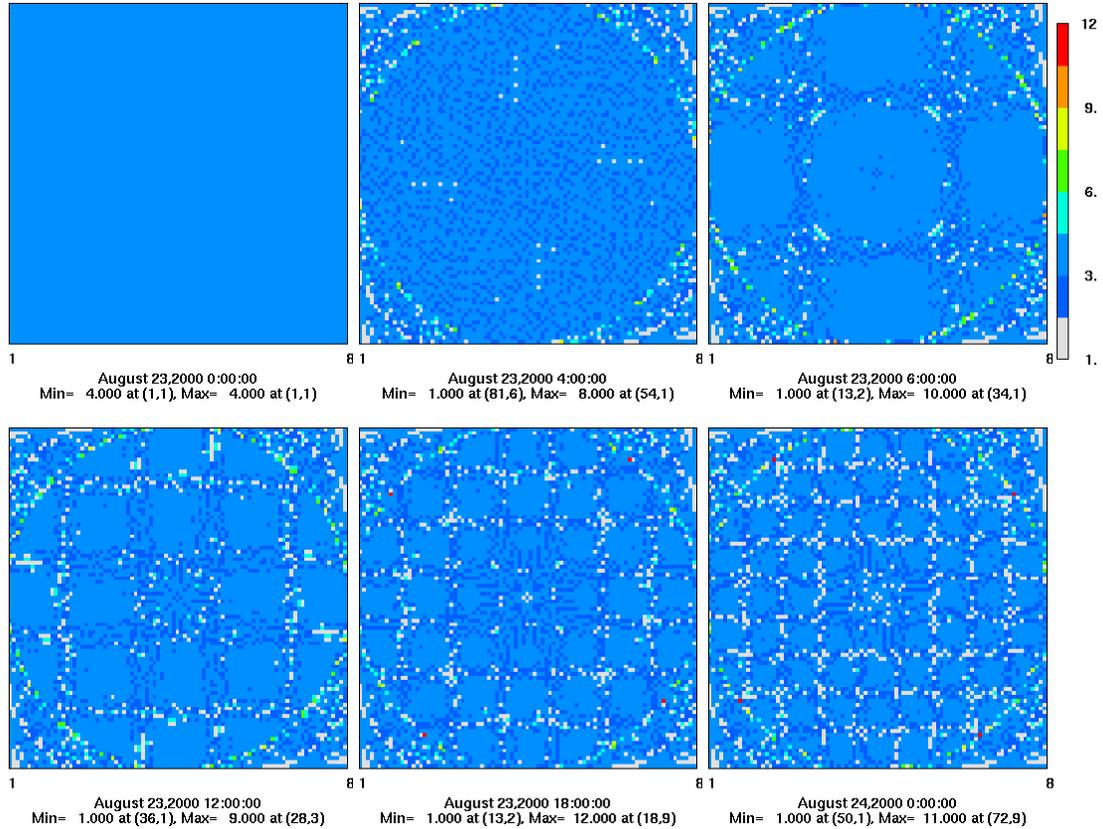


**Figure 2 Demonstration of the loss of high frequency signals, caused by numerical diffusion, when advection is simulated using the Eulerian piecewise parabolic method (PPM).**

The very first algorithm we tried for calculating packet trajectories was the first order forward Euler method. This turned out to be too crude for the time steps chosen by the time step selection algorithm. When cells were initialized with 4 uniformly distributed packets, the algorithm yielded the interesting, but erroneous, geometric patterns seen in Figure 3. The checkerboard pattern should be recovered every 6 hours, during which the packets go through a  $90^\circ$  rotation. The erroneous results were caused by trajectories that strayed too far from being perfect circles. After a  $90^\circ$  rotation, packets that started in the same cell should again be together in the same cell. Figure 4 shows that this was failing to happen. In particular, every 6 hours there should be 4 packets in each cell that is inside the largest disk contained in the domain.



**Figure 3** Result of the same test as in Figure 2, with the first order forward Euler algorithm used to calculate trajectories for the TG method. This algorithm is too crude.



**Figure 4** Number of packets in grid cells at different integration times. The problem with the low order trajectory algorithm, demonstrated in Figure 3, can be attributed to trajectories that stray too far from being circles.

However, the problem was easily fixed by switching to the more accurate, second order, predictor-corrector algorithm, as is seen in Figure 5. For the time being, we feel that this choice is sufficiently accurate. But, there are many more accurate algorithms available, and we may decide to switch again, if we feel that improved accuracy is desirable.

Advection with the TG approach does not have any numerical diffusion and the original checkerboard pattern is recovered every 6 hours, as expected. However, the choice of the AVG\_MIX representation of cell composition exhibits what we call “display diffusion” at hr=4. This is due to the averaging of the packets in each cell for reporting. The packet mixing ratios in the checkerboard pattern are actually 1 and 0, as is seen in Figure 6, where the mixing ratio of the packet closest to the cell center (CLS\_MIX) is displayed instead of the average.

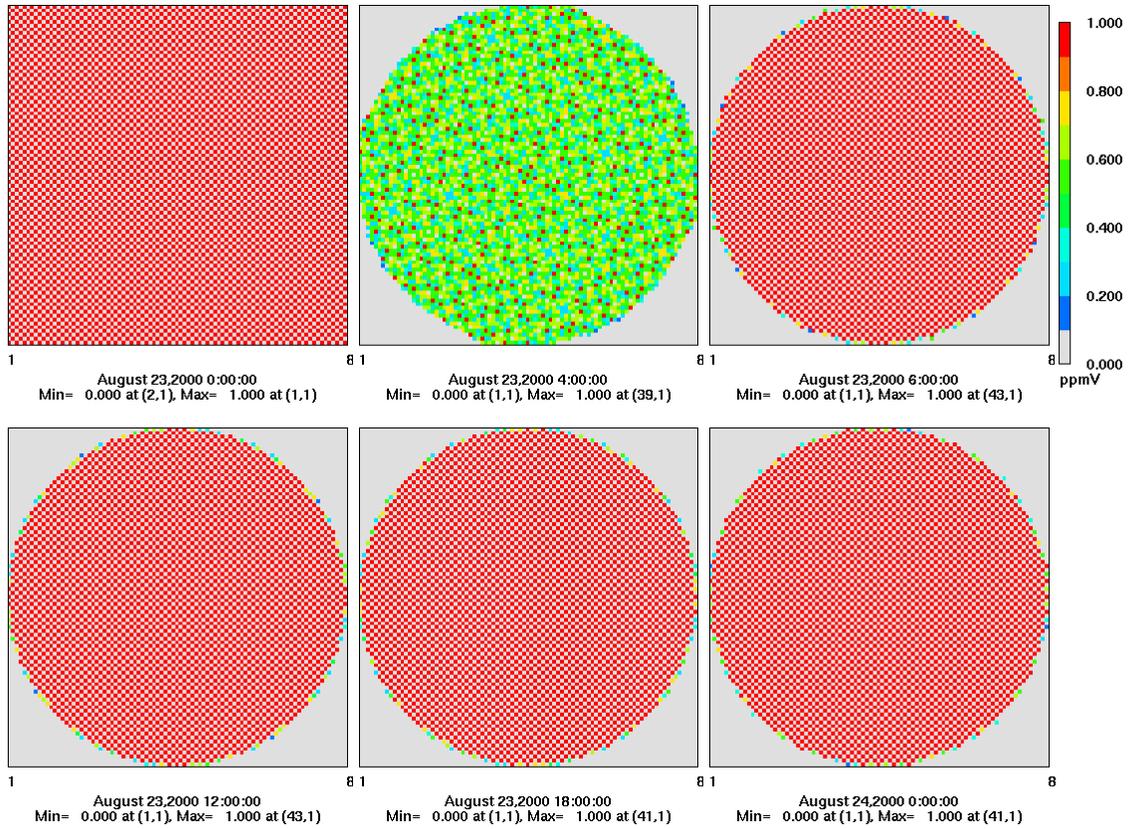


Figure 5 The problem was solved by switching to the predictor-corrector trajectory method. Displaying average mixing ratios in cells shows “display diffusion” at hr=4, but the signal is recovered almost perfectly every 6 hours.

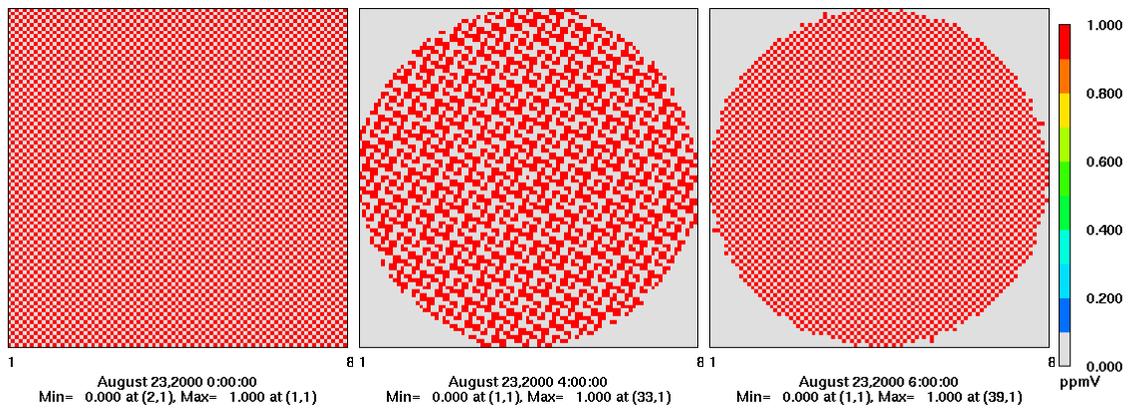


Figure 6 Displaying the mixing ratio of the packet closest to a cell center shows no “display diffusion”, but the checkerboard pattern is still temporarily lost, at hr=4, and then restored, at hr=6.

## The SPOS signals

Important properties of a numerical advection scheme can be studied by advecting signals of different shapes and monitoring the interrelationship of the signals. The three signals  $h_A$  (SPOS\_A),  $h_B$  (SPOS\_B), and  $h_C$  (SPOS\_C) have been designed for superposition tests of transport algorithms. The base shape of the concentration mounds is described by the Witch of Agnesi surface

$$h_s = \frac{a^2}{a^2 + (x - x_r)^2 + (y - y_r)^2} \quad (4)$$

where  $a$  is the radius of the mound,  $(x_r, y_r)$  is the position of the peak of the mound, and  $h_s$  is the height of the mound. The parameters have been set so that the mound is centered at cell (29,29) in the modeling domain, and  $a$  is 10 cell widths. The three tracer signals  $h_A$ ,  $h_B$ , and  $h_C$  are then defined as follows:

$$h_A = q_{\max}(1 + h_s) + q_{\min}, \quad (5)$$

$$h_B = 2q_{\max}(1 + h_s) - q_{\min}, \quad (6)$$

$$h_C = -q_{\max}(1 + h_s) + 2q_{\min}, \quad (7)$$

where  $q_{\max}$  and  $q_{\min}$  determine the amplitude and background values of the signals. Here, both are set to 50. Therefore, the SPOS\_A signal, for example, has a peak height of 150 with a background value of 100.

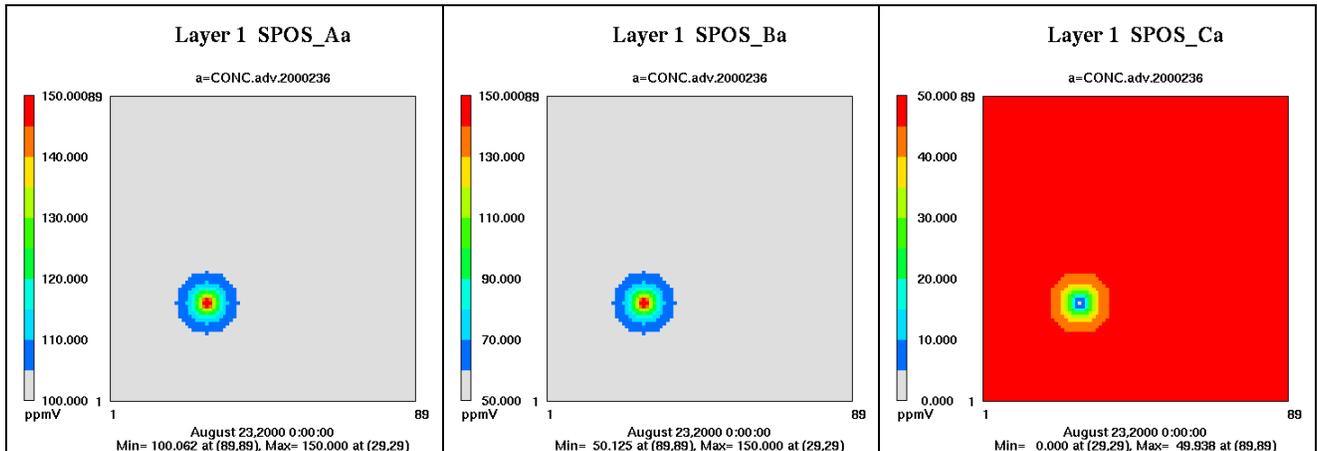


Figure 7 Initial profiles of the three SPOS signals.

We now compare TG advection of the SPOS\_A mound with advection using the Eulerian PPM scheme in standard CMAQ, see Figure 8. TG advection is expected to move initial profiles at wind speed along trajectories, and to preserve amplitude. The comparison shows that TG advection has these expected properties, while the numerical diffusion of PPM advection significantly dampens the peak of the mound, bringing it down by over 16 ppm from the initial height of 50 ppm above the background value. PPM's numerical diffusion also distorts the shape of the mound, making the PPM solution as much as 5 ppm higher than the TG solution on the trailing side of the mound. The ripples seen in the display of the difference (TG – PPM) are entirely due to the behavior of the PPM scheme with the rotational wind field.

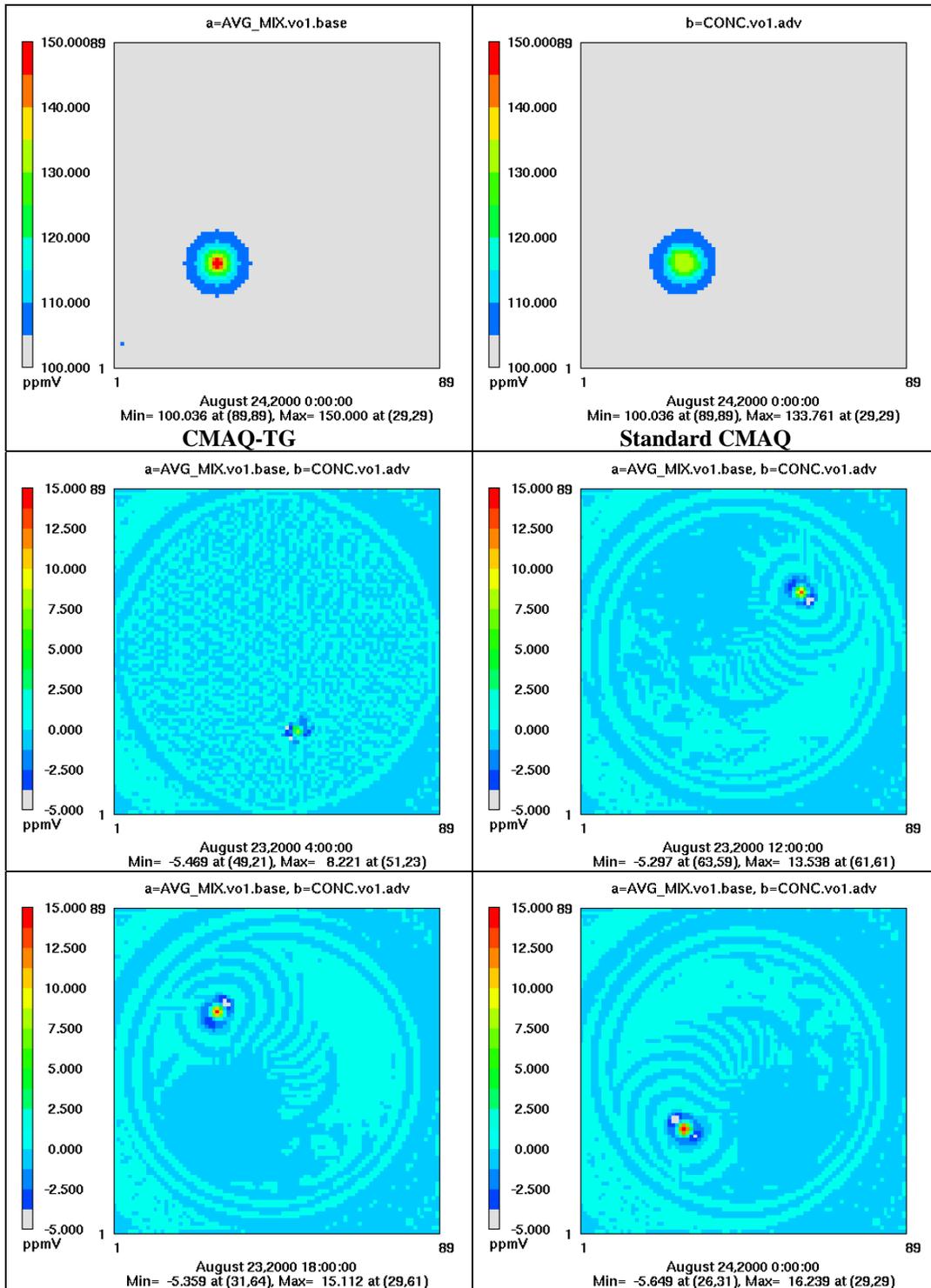
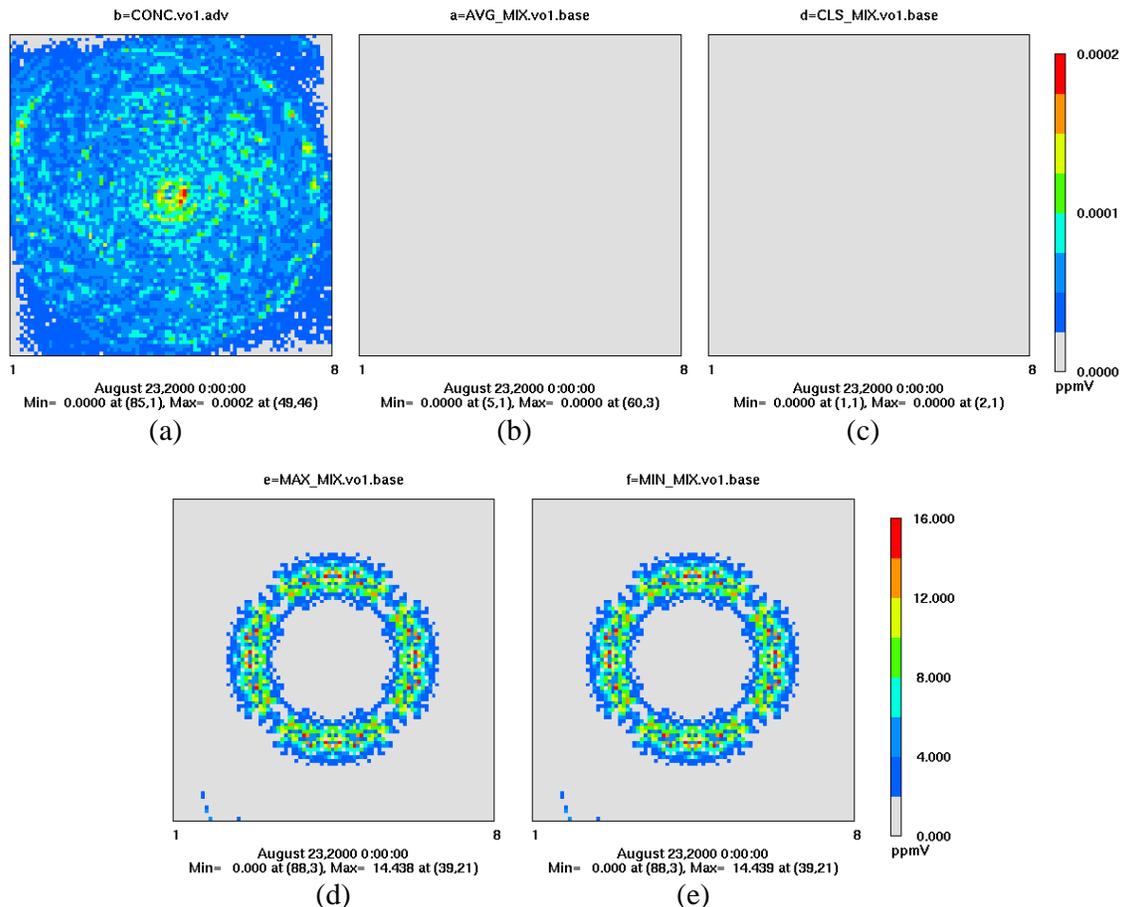


Figure 8 Comparison of TG advection of the SPOS\_A mound with PPM advection. The top two panels show the profiles at the end of 24 hours. The lower four panels show the evolution in time of the difference, TG - PPM. TG's peak value remains at 150 and its profile remains symmetric around the peak. Numerical diffusion causes PPM's peak value to steadily decrease, losing about 1/3 of its height over a period of 24 hours, and its profile to become asymmetric.

### 3.2. Linearity of TG advection algorithm

One way to evaluate possible nonlinearity of an advection algorithm is to examine the extent to which  $\max|h_A - (h_B + h_C)|$  differs from zero. Byun and Lee, 2002, showed that the PPM algorithm exhibits a very minor degree of nonlinearity, see panel (a) of Figure 9. We expect the TG advection algorithm to be perfectly linear because it does not change packet composition. Panels (b) and (c) of Figure 9 support this expectation. Note that the nonlinearity indicated in panels (d) and (e) is actually a property of the display method, not the advection algorithm itself. The reason is that selecting the species-wise maximum or minimum mixing ratio over the packets in a cell is likely to pick different packets for different species, which means that the advected profiles  $h_A$ ,  $h_B$ , and  $h_C$  are evaluated at different points in space. Furthermore, averaging and selecting one packet, e.g., the one closest to a cell center, are linear functions, while max and min are not. This must be taken into account when maximum concentration values are used to assess source apportionment for air quality problems.



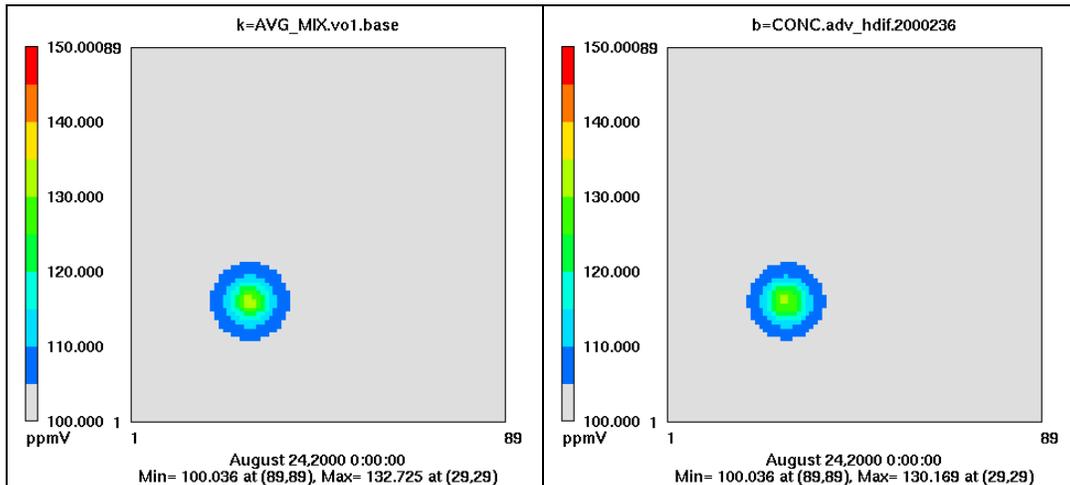
**Figure 9** A test of linearity of the PPM and TG advection algorithms using the three SPOS signals in a rotational wind field. Departure of  $\max|h_A - (h_B + h_C)|$  from 0 shows nonlinearity of the solution. (a) PPM, (b) AVG\_MIX for TG, (c) CLS\_MIX for TG, (d) MAX\_MIX for TG, (e) MIN\_MIX for TG. Panels (b) and (c) show that the TG advection algorithm is linear to machine precision. The nonlinearity shown in panels (d) and (e) is due to the output algorithm, not the advection algorithm.

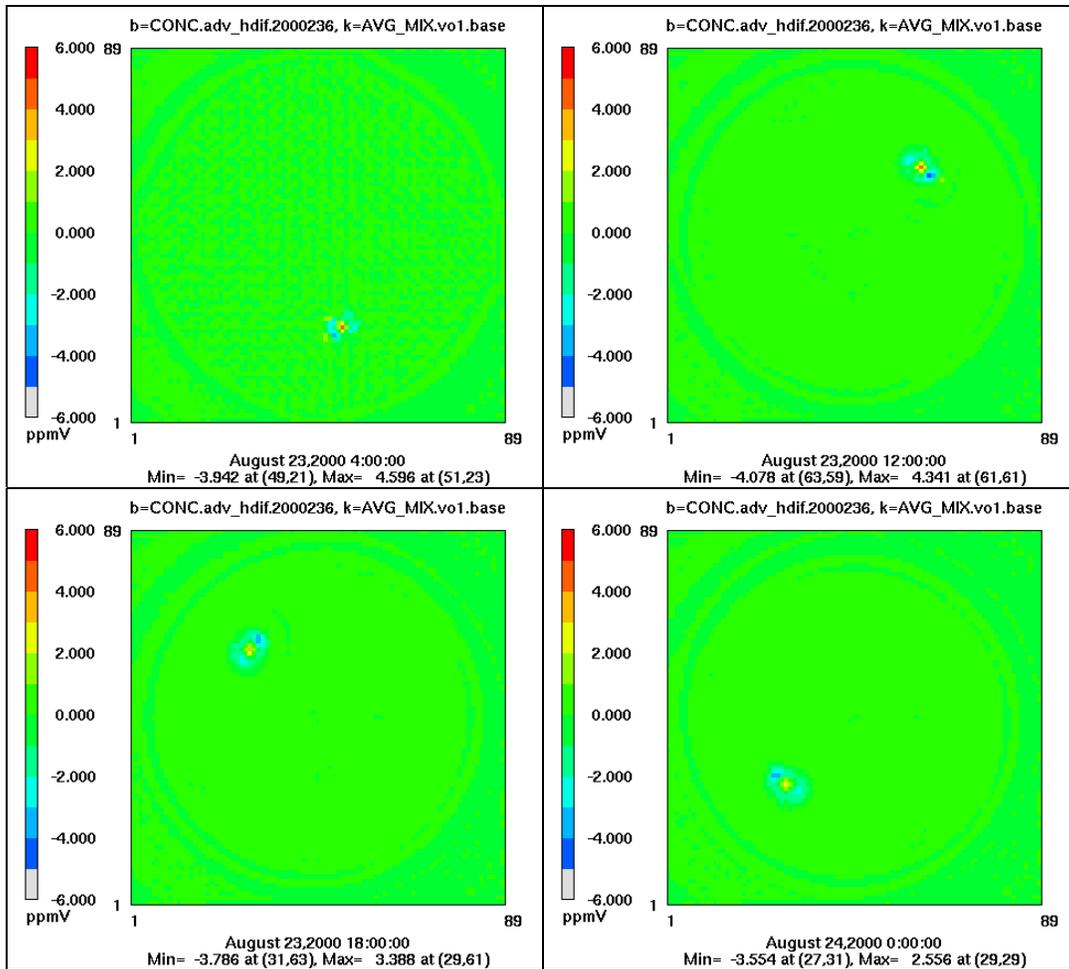
### 3.3. Testing of TG advection with horizontal diffusion in a two-dimensional rotational wind field

We now add horizontal diffusion to advection in our comparison of CMAQ-TG and standard CMAQ. As described above, horizontal diffusion in CMAQ-TG includes a subgrid diffusion step that may optionally be turned off. Subgrid diffusion is intended to simulate diffusion between the packets inside cells.

Because of the numerical diffusion in the current Eulerian advection algorithm, PPM, the horizontal diffusion process has been ill-defined in CMAQ. In particular, the current horizontal diffusion module in standard CMAQ handles physical diffusion and tries to compensate for the non-physical numerical diffusion. Now, because TG advection is not numerically diffusive, the horizontal diffusion module needs to be re-assessed, if the physical diffusion phenomena are to be properly captured.

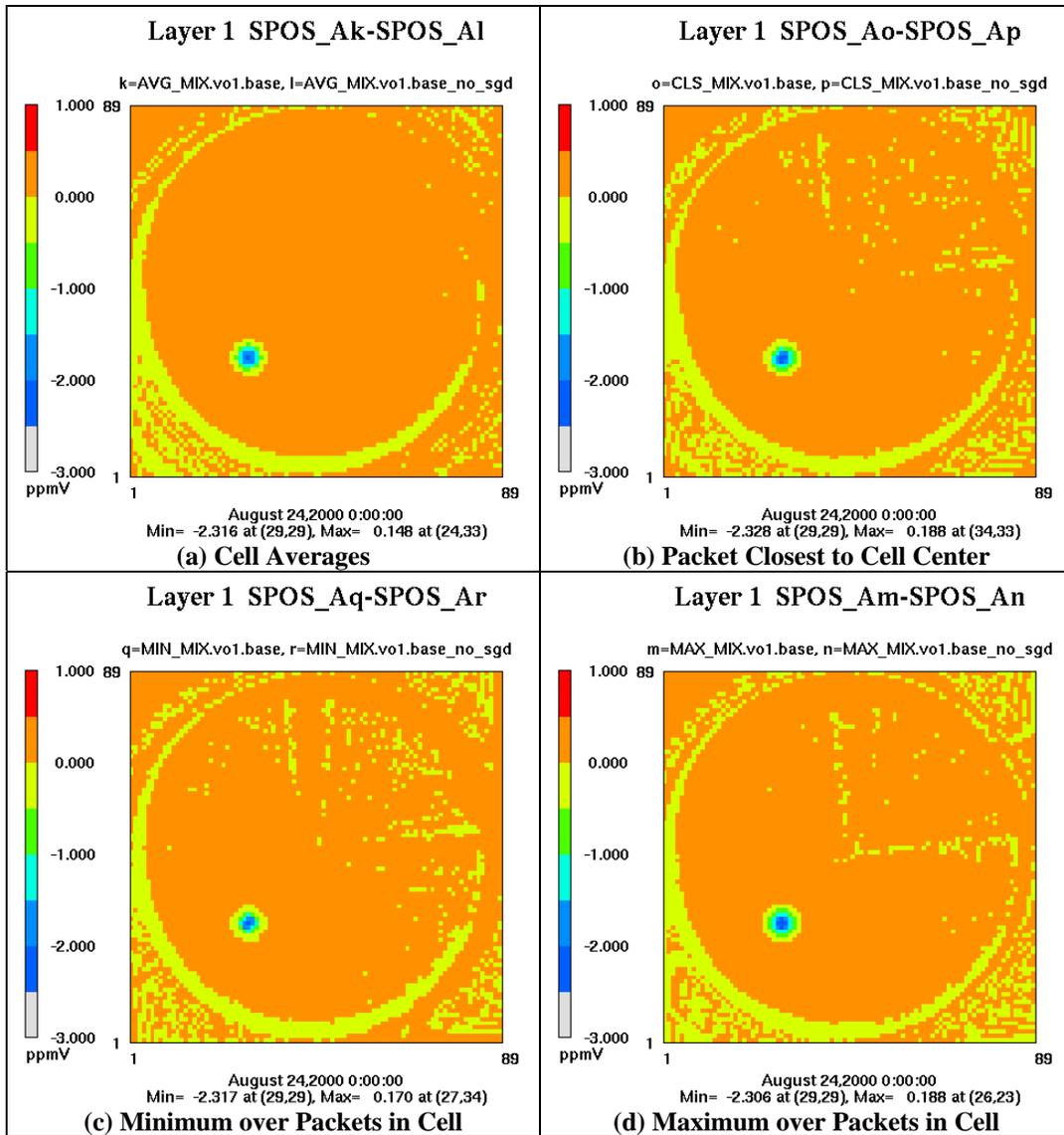
Figure 10 shows a comparison between CMAQ-TG and standard CMAQ for advection with horizontal diffusion, like the comparison in Figure 8 for advection alone. The magnitude of the differences here is significantly smaller that in Figure 8. This suggests that as more processes are incorporated, the effects of using different advection algorithms, Lagrangian or Eulerian, become less apparent.





**Figure 10** The same as Figure 8, but showing the difference between CMAQ-TG and standard CMAQ for advection with horizontal diffusion.

Figure 11 shows several views of the difference between the TG transport with and without subgrid scale diffusion after one full rotation of the SPOS\_A hill signal. Without subgrid diffusion, peak values, at the center of the hill increase, as expected, but the base of the hill shows lower values, because high values diffuse down the hill more slowly. After a full rotation, the peak of the hill is back at its starting position in cell (29, 29), where the difference “with” minus “without” is at its minimum. Note that the location (which depends on the display type) where the difference is at its maximum is always within a radius of 6 cells from the peak, in other words, roughly at the base of the hill. Since the maximum has a positive value, this is a point on the hill where including subgrid diffusion gives a higher value than not doing subgrid diffusion.



**Figure 11** Difference between CMAQ-TG advection with horizontal diffusion with and without subgrid scale diffusion after one full rotation of the cone signal (24 hours). The difference shown is “with” minus “without”.

### 3.4. Packet management for some idealized, linear, two-dimensional wind fields

In this section we look at the challenge of the packet management issue for the TG approach. In order for a CMAQ-TG simulation to be considered to have a spatial resolution comparable to a corresponding Eulerian simulation, it seems reasonable to require that each Eulerian cell contain at least one packet, but not many more than one packet. This is the goal for the packet management strategies that have been implemented in CMAQ-TG. However, as has been noted above, spawning and pruning to control spatial density of packets has undesirable side effects. An obvious way to avoid spawning and pruning would be to initialize the domain with enough packets so that nothing further is needed to maintain a minimal packet density. The examples in this section will show that, for now, spawning and pruning are a necessity.

In addition to the rotational wind field used so far, we introduce two more wind fields, stretching and shearing, to test with. All three are non-divergent.

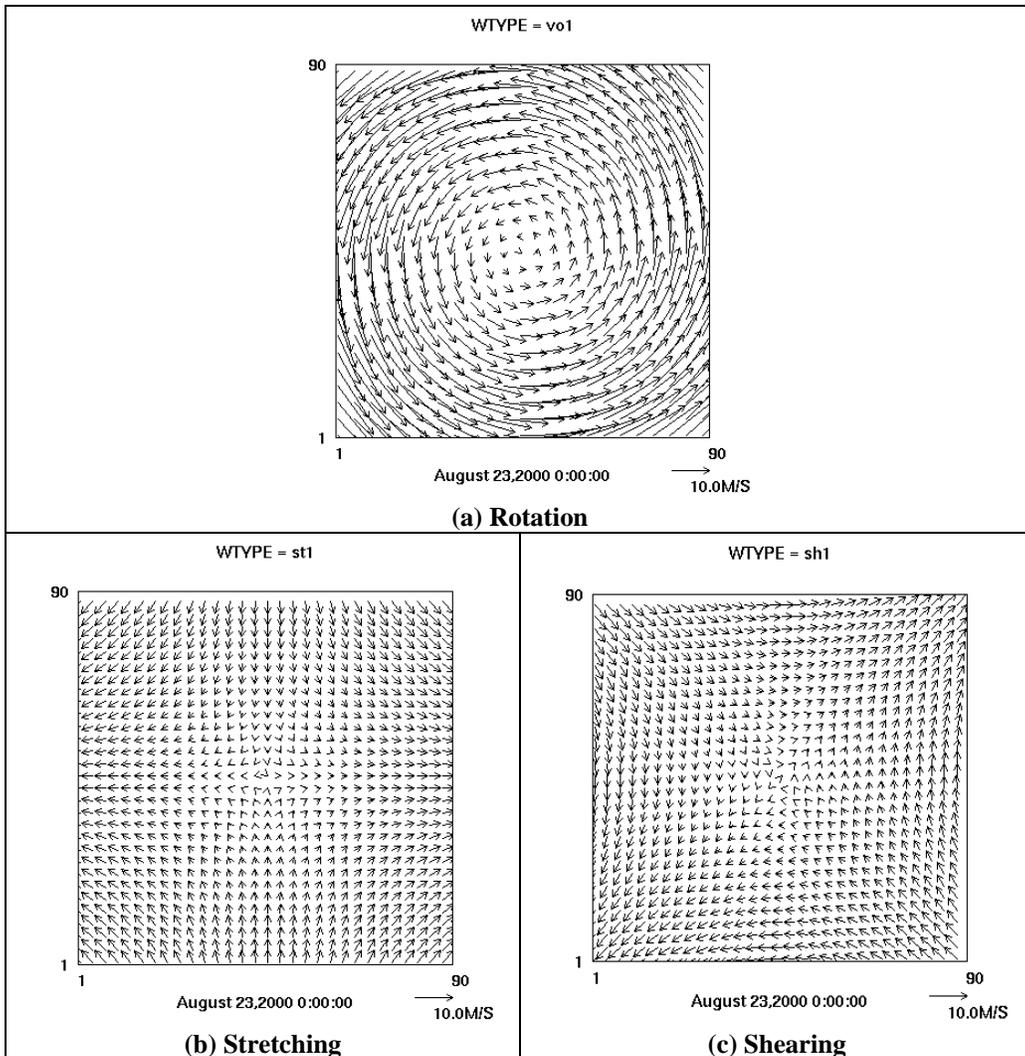


Figure 12 Some idealized, linear, two-dimensional wind fields.

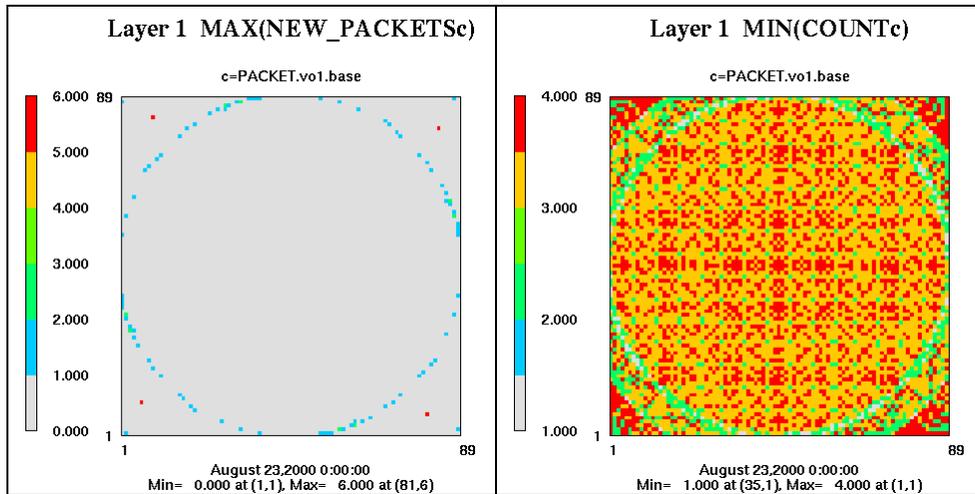


Figure 13 Rotation, default strategy: few cells in layer 1 ever become empty and only four cells need as many as 6 packets spawned during an hour.

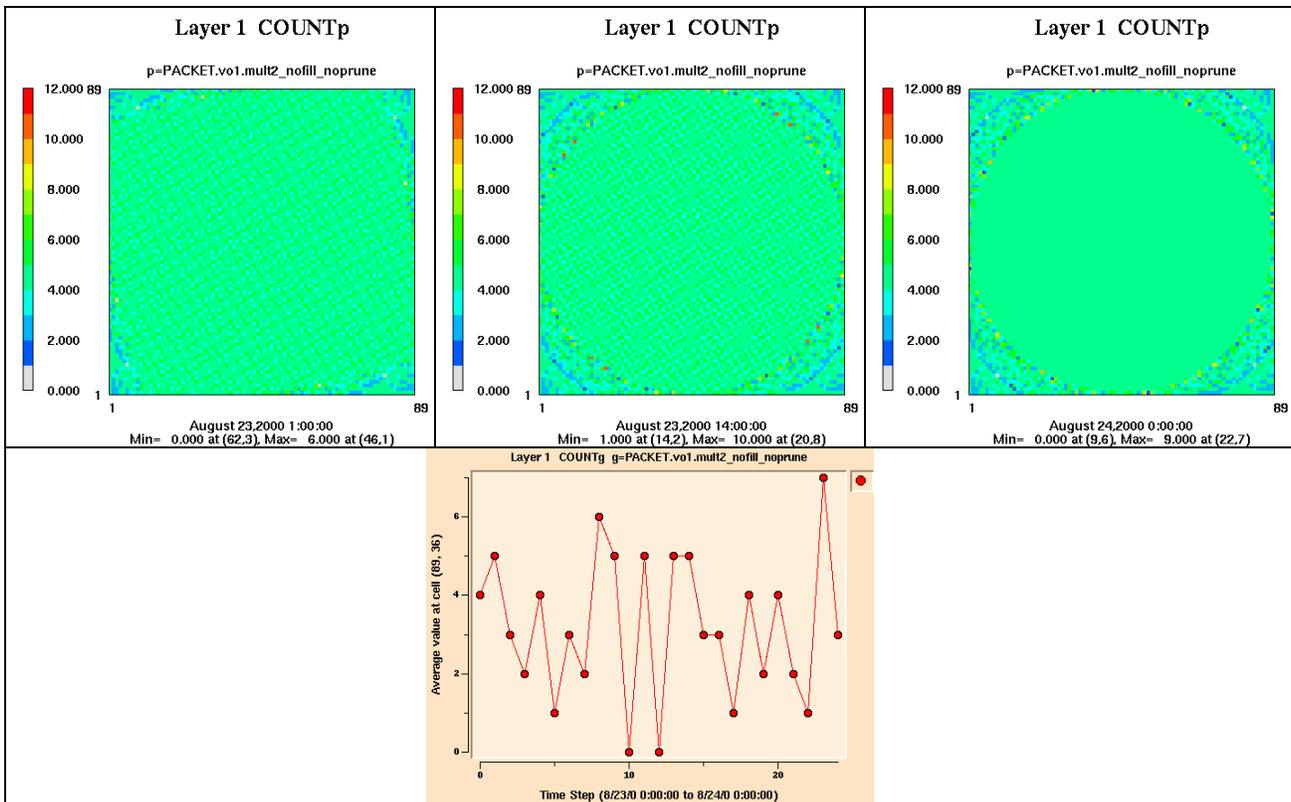


Figure 14 Rotation, cells initialized with 4 packets, no spawning or pruning: even the few cells that do become temporarily empty, don't stay empty.

Figure 13 and Figure 14 show that packet management is not a significant issue for the rotational wind field.

However, Figure 15 indicates that the situation might be quite different for the stretching wind field. Uniform cell initialization, with no spawning, would require starting with 25 packets in each cell to ensure that no cell ever became empty. In this case, the minimum number of packets in a cell would never drop below 11 and some cells would accumulate up to 50 packets. Assuming that for an air quality study each packet has roughly the same computational cost as a cell in an Eulerian approach, the computational cost of the TG approach with this kind of packet management would be enormous. Figure 16 indicates a somewhat better potential situation for the shearing wind field.

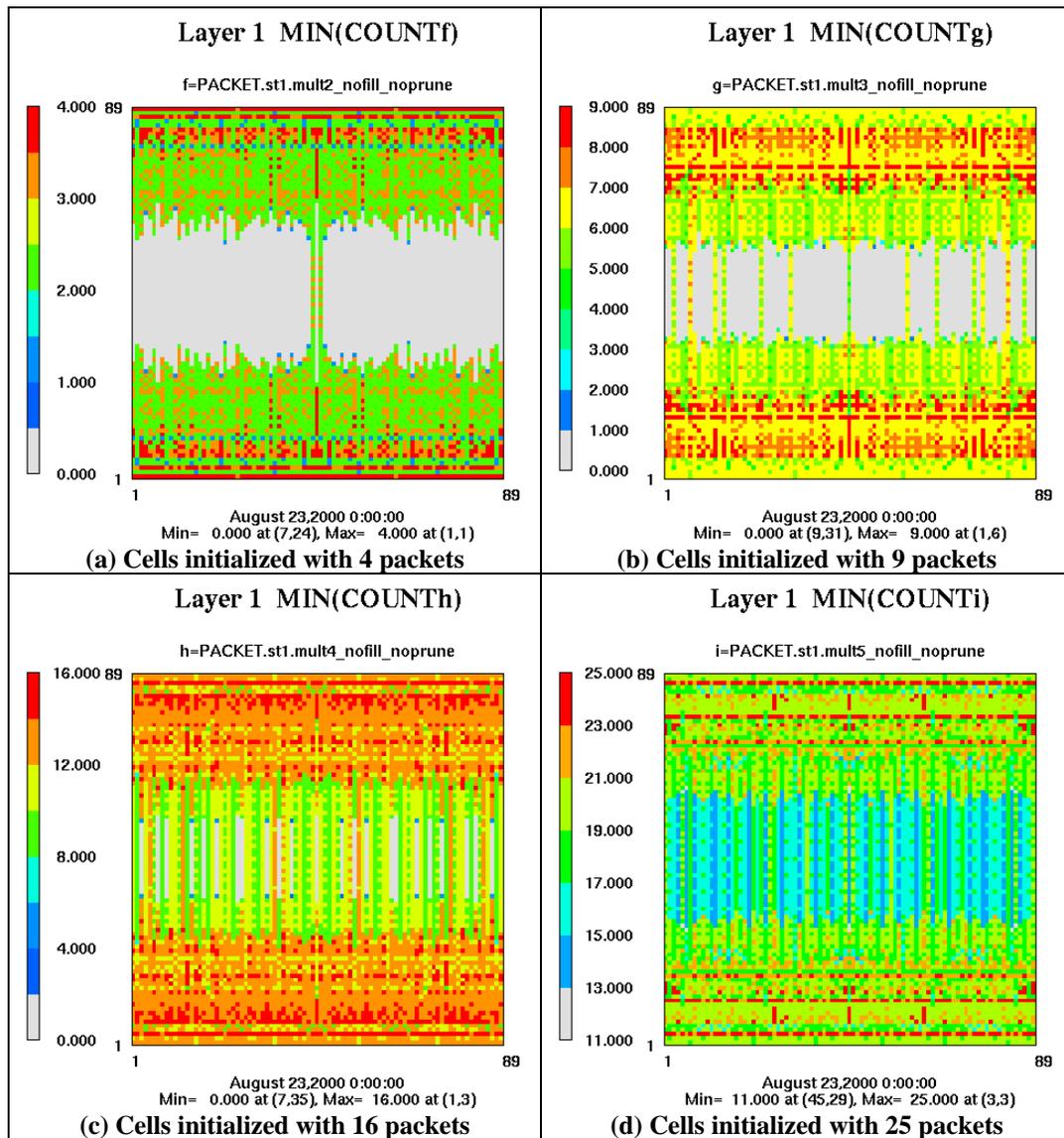
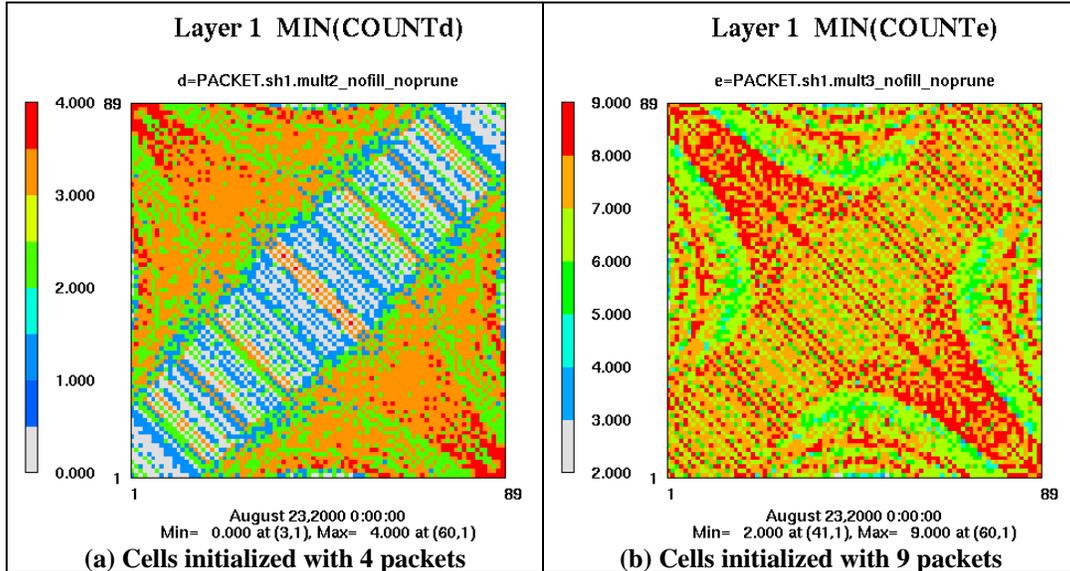


Figure 15 Stretching wind field. Without spawning, for a uniform initialization number to ensure that no cell ever became empty, it would be necessary to initialize every cell with 25 packets.



**Figure 16 Shearing wind field. Without spawning, for a uniform initialization number to ensure that no cell ever became empty, it would be necessary to initialize every cell with 9 packets.**

Figure 15 and Figure 16 indicate that, even for simple non-divergent wind fields, spawning is necessary in order to avoid empty cells. Fortunately, by just looking at all the cells that would ever become empty during a one day simulation, they paint a picture of likely spawning requirements for these wind fields that is far too pessimistic.

By looking at snapshots at the end of 24 hours and at time series, Figure 17 shows that for the stretching wind field, with spawning and pruning, even one initial packet per cell, might be feasible.

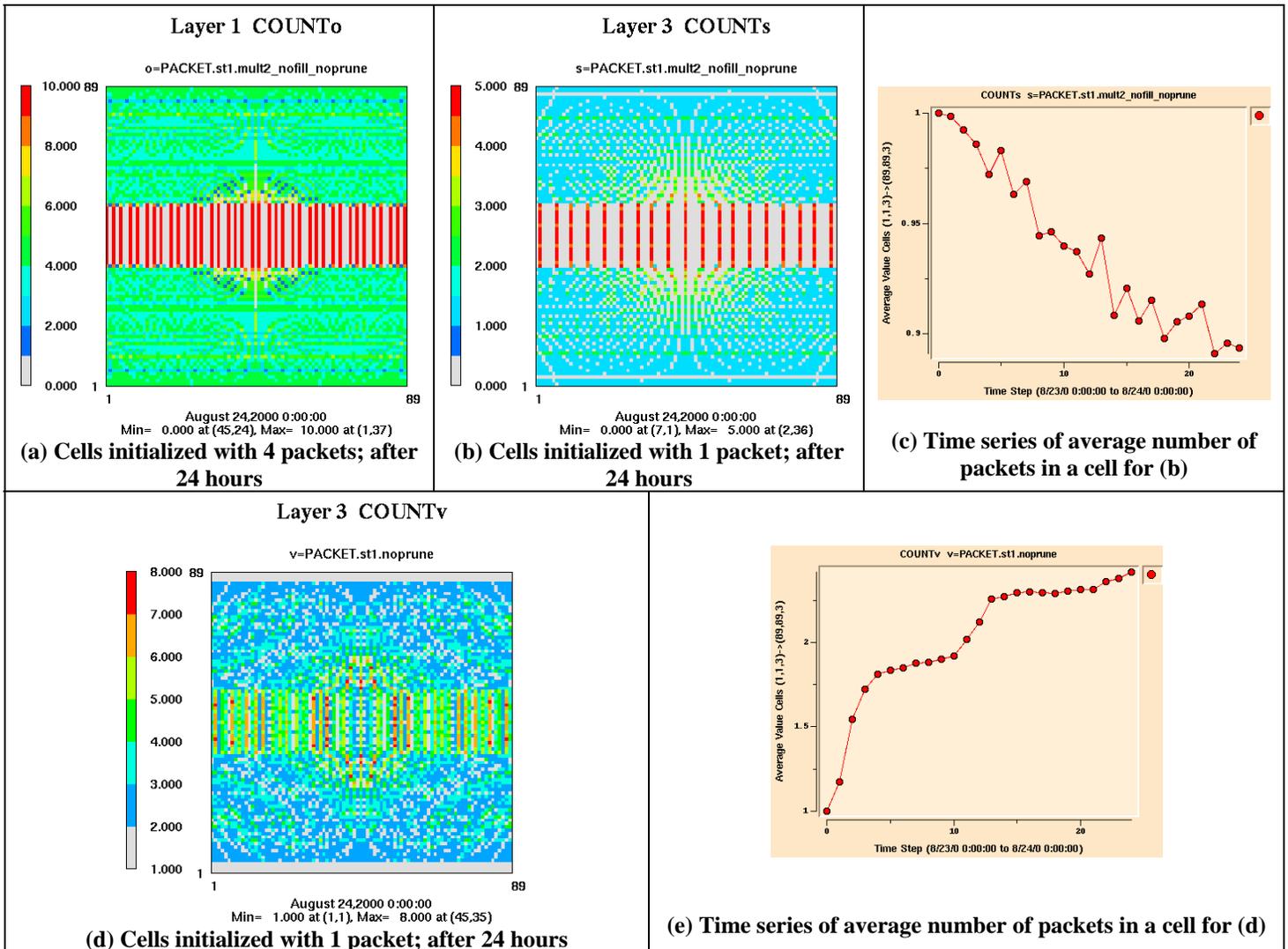


Figure 17 Stretching wind field. (a), (b), (c) no spawning or pruning; (d), (e) spawning, but no pruning

This has just been a preliminary look at the challenge of packet management to give some idea of the problem. While the testing done in other sections shows that the current default strategy is not bad, a good deal of experience with real wind fields will be needed before it will be possible to confidently identify a reliable default strategy for packet management.

#### 4. Evaluation of CMAQ-TG with a Realistic Three-dimensional Wind Field

Although a two-dimensional rotational field provides an opportunity to study various transport characteristics, the TG algorithm must be tested with realistic meteorological fields. We use realistic 3-D output from the Mesoscale Model version 5 (MM5) for August 25, 2000 for testing CMAQ-TG with advection and diffusion. CMAQ was run with 23 vertical layers. Table 1 summarizes the 3-D wind test cases for standard CMAQ and Table 2 for CMAQ-TG.

**Table 1 Summary of 3-D wind test cases for standard CMAQ (with concentration adjustment)**

Case	Characteristics
CTM/ADV	8 tracer species; 3-D advection
CTM/TRANS	8 tracer species; 3-D advection and 3-D diffusion
CTM/CB4	CB4 species; 3-D advection, 3-D diffusion (including dry deposition and emissions), CB4 chemistry
CTM/CB4_NO_HDIF	same as CTM/CB4 but without horizontal diffusion

**Table 2 Summary of 3-D wind test cases for CMAQ-TG**

Case	Characteristics
Ta: TG/ADV	8 tracer species; 3-D advection; default TG control options
Tb: TG/ADV_FILL_ALL_NO_PRUNE	same as TG/ADV except that pruning is turned off
Tc: TG/ADV_SP_FILL_NO_PRUNE	same as TG/ADV except that the SPARSE_FILL option is used for spawning and pruning is turned off
Td: TG/ADV_HR_ALL	same as TG/ADV except that all layers are made high resolution, i.e., cells are initialized with 4 packets and empty boundary cells are filled with 4 packets
Te: TG/TRANS	8 tracer species; 3-D advection and 3-D diffusion; default TG control options
Tf: TG/CB4	CB4 species; dry deposition, emissions, 3-D advection, 3-D diffusion, CB4 chemistry; default TG control options
Tg: TG/CB4_NO_HDIF	same as TG/CB4 but without horizontal diffusion

### 4.1. Comparison of 3-D advection and transport between PPM and TG

The first test we performed with a 3-D wind field from MM5 was to see how the concentration fields are affected by a complicated wind field. Figure 18 compares transport of the checkerboard and IC1\_BC0 (initial conditions = 1, boundary conditions = 0) patterns by the TG and PPM advection algorithms after 24-hours. The results show that 3-D movement of the checkerboard pattern by PPM and TG (with display diffusion) produces qualitatively similar results. However, the PPM results show a smoothed concentration field, resulting from numerical diffusion, while the graininess of the TG results indicates that the checkerboard pattern is being preserved. Figure 19 shows a vertical cross section of the vertical velocity field from MM5 and a snapshot of the number of packets being used in the TG. It shows that packets are moved by the updrafts and downdrafts and therefore the signals are affected by the transport from upper layers.

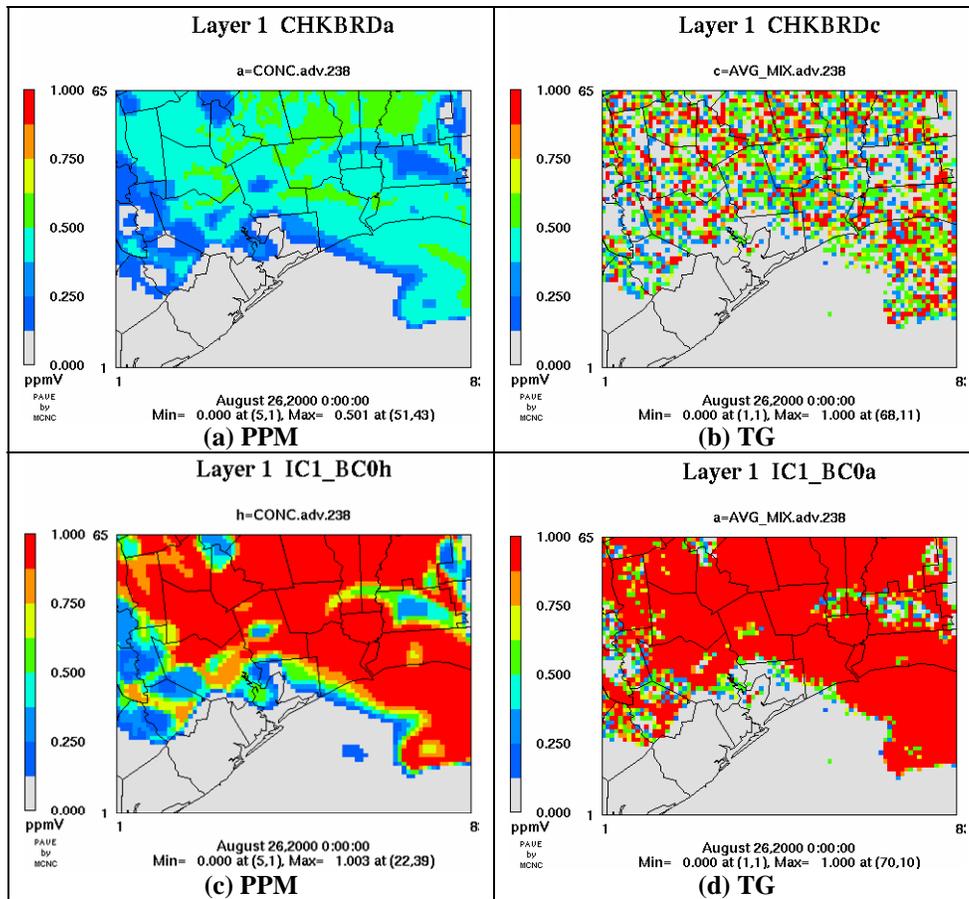
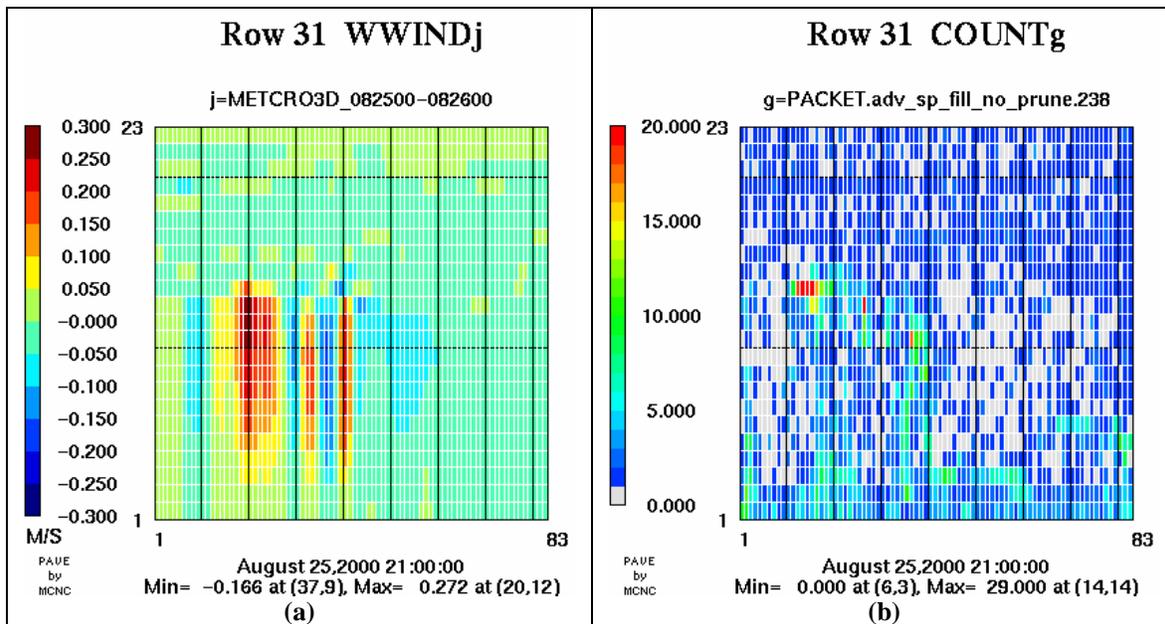


Figure 18 Advection of the checkerboard and IC1\_BC0 signals by a three-dimensional wind field from MM5 after 24 hours. The signals are affected by the transport of packets from the upper layers. The TG results show display diffusion.



**Figure 19 (a) Vertical cross section of the vertical velocity field across row 31, and (b) number of packets used in the TG calculation with a three-dimensional wind field from MM5.**

Depending on subtle differences in the packet management scheme used, final cell concentrations can be different (Figure 20).

When the horizontal and vertical diffusion processes are added to advection, the difference between standard CMAQ and CMAQ-TG transport is significantly reduced (Figure 21).

Testing of our first versions of horizontal and vertical diffusion revealed serious problems with methods we tried for distributing cell deltas to packets. Diffusion is a smoothing process and should not yield a packet mixing ratio that is outside the range of values found in packets before a diffusion step. Figure 22 shows violations of this. Furthermore, Figure 23 demonstrates that CMAQ-TG with diffusion became nonlinear for some of the methods we tried. However, Figure 24 and Table 3 show that both of these problems have been fixed by the methods, described in Section 2.5, that are now used in CMAQ-TG.

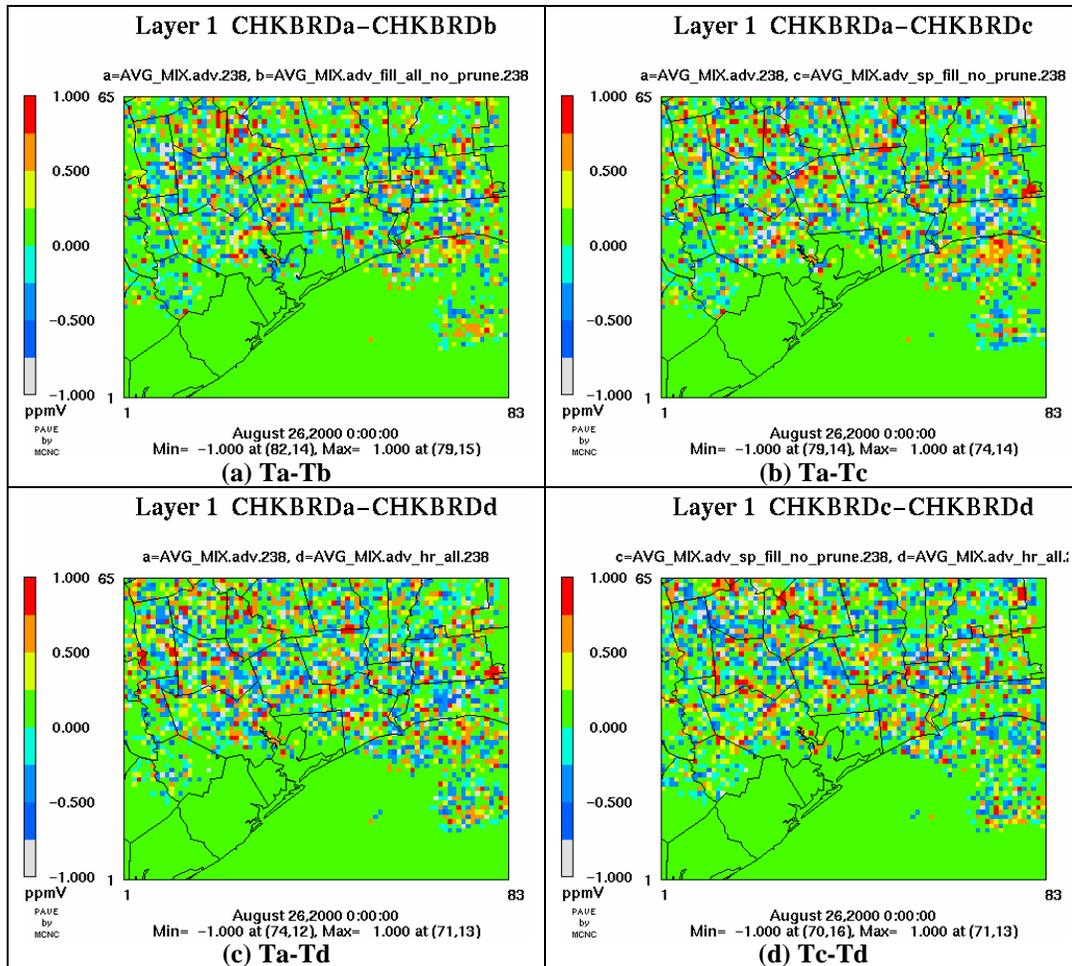


Figure 20 Differences in TG advection of the checkerboard pattern with different packet management methods after 24 hours.

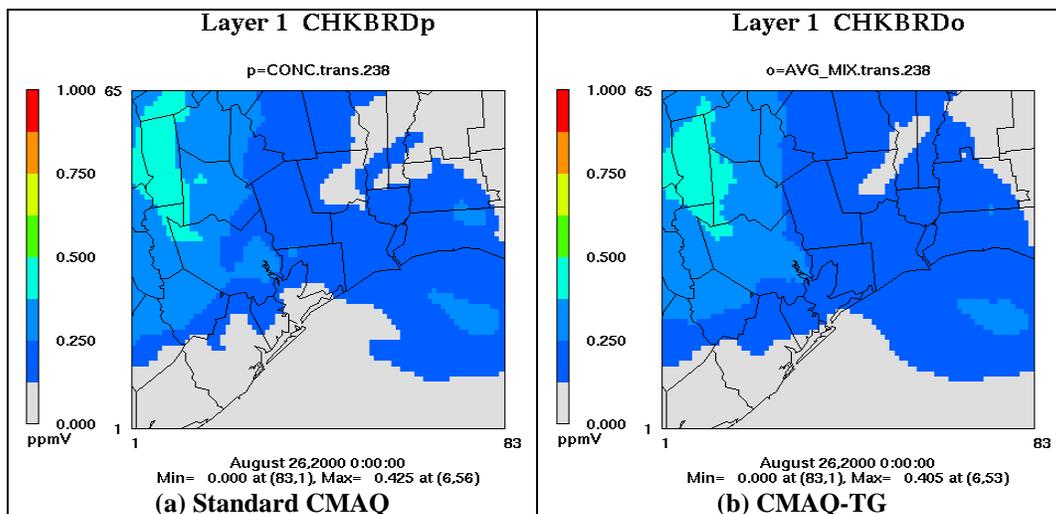


Figure 21 Results of TG advection with horizontal and vertical diffusion of the checkerboard pattern after 24 hours.

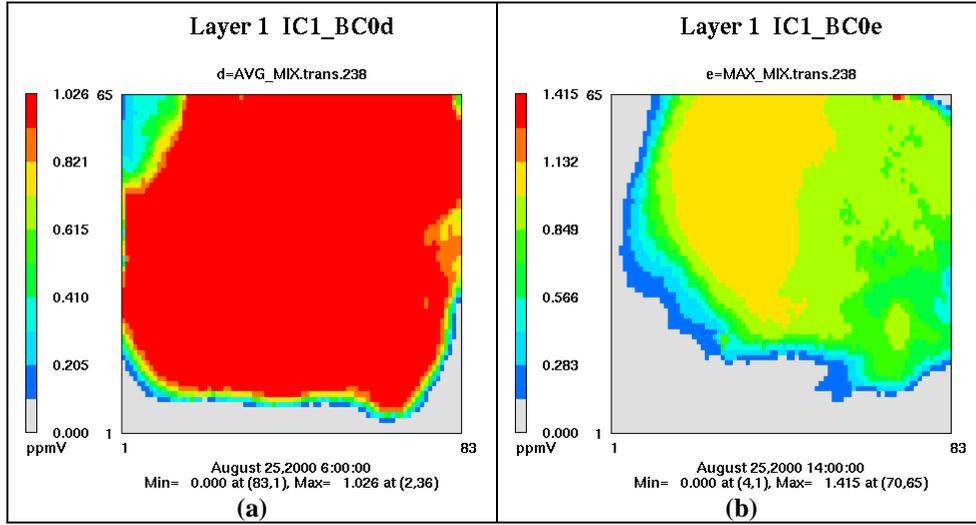


Figure 22 Bad results from one of the early diffusion algorithms that was tried in CMAQ-TG. The simulation is for a species, IC1\_BC0, with initial condition 1.0 and boundary condition 0.0. Both the average value in a cell (a) and the maximum value in a packet (b) become significantly greater than 1.0 during the simulation (in different places, at different times).

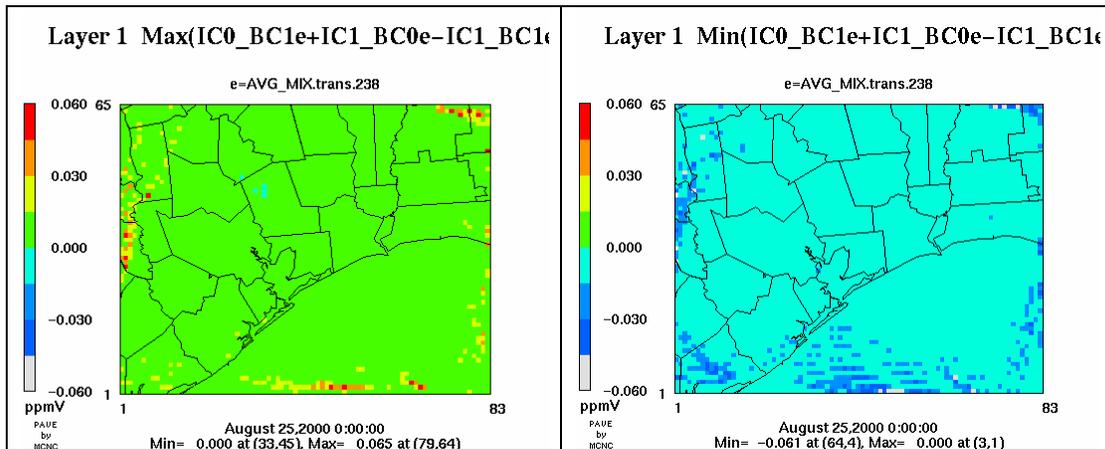


Figure 23 Bad results from an early diffusion algorithm tried in CMAQ-TG that failed to be linear. Deviation of  $IC_0\_BC_1 + IC_1\_BC_0 - IC_1\_BC_1$  from zero demonstrates nonlinearity in the solution. Maximum of the cell averages (left) and minimum of the cell averages (right) over the 24-hour period are shown.

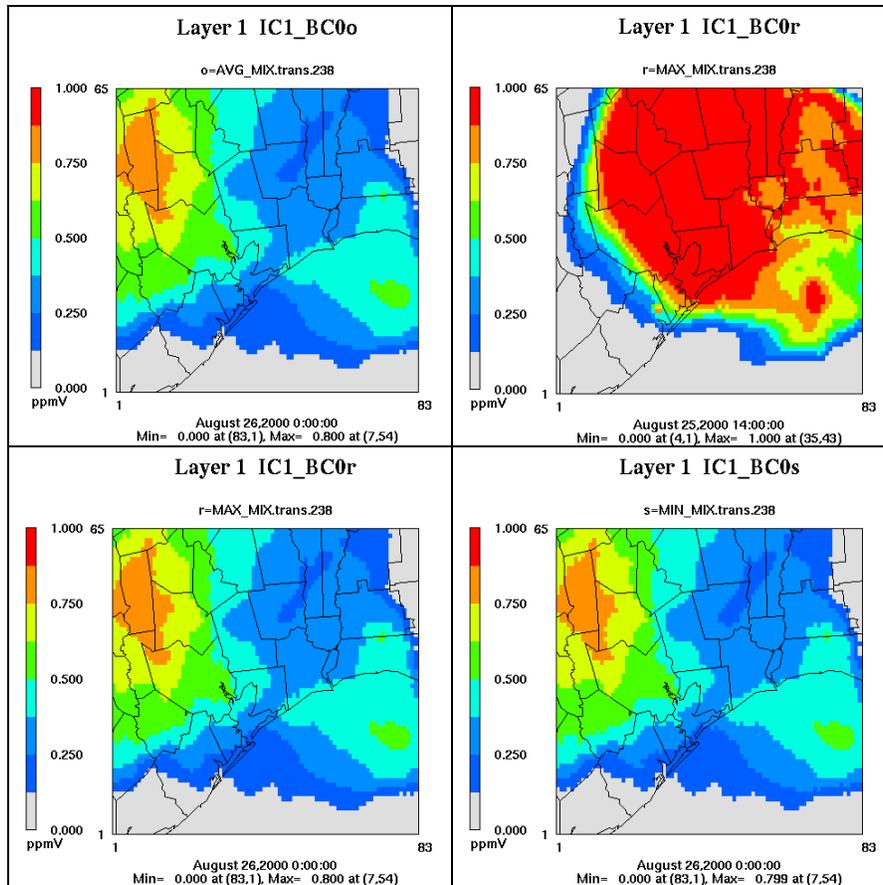


Figure 24 The problem shown in Figure 22 has been fixed by the diffusion algorithms now used in CMAQ-TG.

Table 3 Results of linearity tests of the transport algorithms (advection and diffusion) now in CMAQ-TG and in standard CMAQ with a realistic wind field. Linearity for CMAQ-TG is excellent.

	IC1_BC1 - IC1_BC0 - IC0_BC1		SPOS_A - SPOS_B - SPOS_C	
	CMAQ-TG	Standard CMAQ	CMAQ-TG	Standard CMAQ
Layer 1	(-0.0000, 0.0000)	(-0.0190, 0.0238)	(-0.0005, 0.0015)	(-0.3704, 0.3227)
Layer 13	(-0.0000, 0.0000)	(-0.2316, 0.1127)	(-0.0010, 0.0013)	(-2.0829, 1.4139)
Row 31	(-0.0000, 0.0000)	(-0.0867, 0.0468)	(-0.0028, 0.0021)	(-0.2743, 0.2470)

## 4.2. Preliminary testing of CMAQ-TG for ozone simulation

The next step in the CMAQ-TG code verification process was to study the interaction of the transport modules with the gas-phase chemistry module. Because of the high nonlinearity of the chemistry module, and due to the non-diffusive nature of Lagrangian advection, we expected CMAQ-TG to exhibit a somewhat wider range of chemical concentrations than standard CMAQ. The simulations in this section include CB4 gas phase chemistry, emissions, dry deposition, advection, and diffusion.

In Figure 25 we compare daily maximum ozone values simulated by CMAQ-TG and standard CMAQ. Note that, while standard CMAQ yields the highest daily maximum ozone value, by 8 ppb, CMAQ-TG produces individual cell values that differ from standard CMAQ by as much as -22 ppb below and 19 ppb above. CMAQ-TG shows the largest difference above standard CMAQ near the Houston Ship Channel, a site with high emissions.

Figure 26 examines the same simulations for the affect of horizontal diffusion on daily maximum ozone concentration. For both CMAQ-TG and standard CMAQ, not doing horizontal diffusion increases the predicted daily maximum ozone value by 2 ppb. For CMAQ-TG, both cell average output and cell maximum output are shown, with a small difference in the expected direction.

Some of the differences in simulated ozone concentrations can be related differences in the precursor, radicals, and secondary air pollutants. Figure 27 presents differences in mixing ratios between CMAQ-TG and standard CMAQ for CO, HO<sub>2</sub>, HNO<sub>3</sub>, and O<sub>3</sub>. It is evident that these patterns correspond to the distribution of emissions inputs from point, mobile, and marine sources.

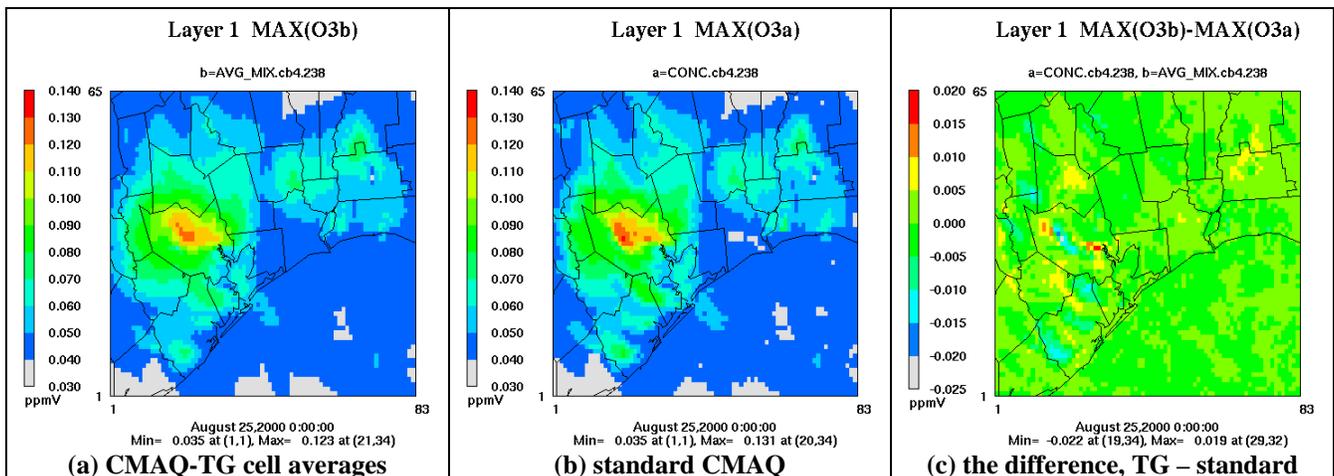


Figure 25 Comparison of daily maximum ozone values simulated by CMAQ-TG and standard CMAQ. CMAQ-TG shows the largest difference above standard CMAQ near the Houston Ship Channel, a site with high emissions.

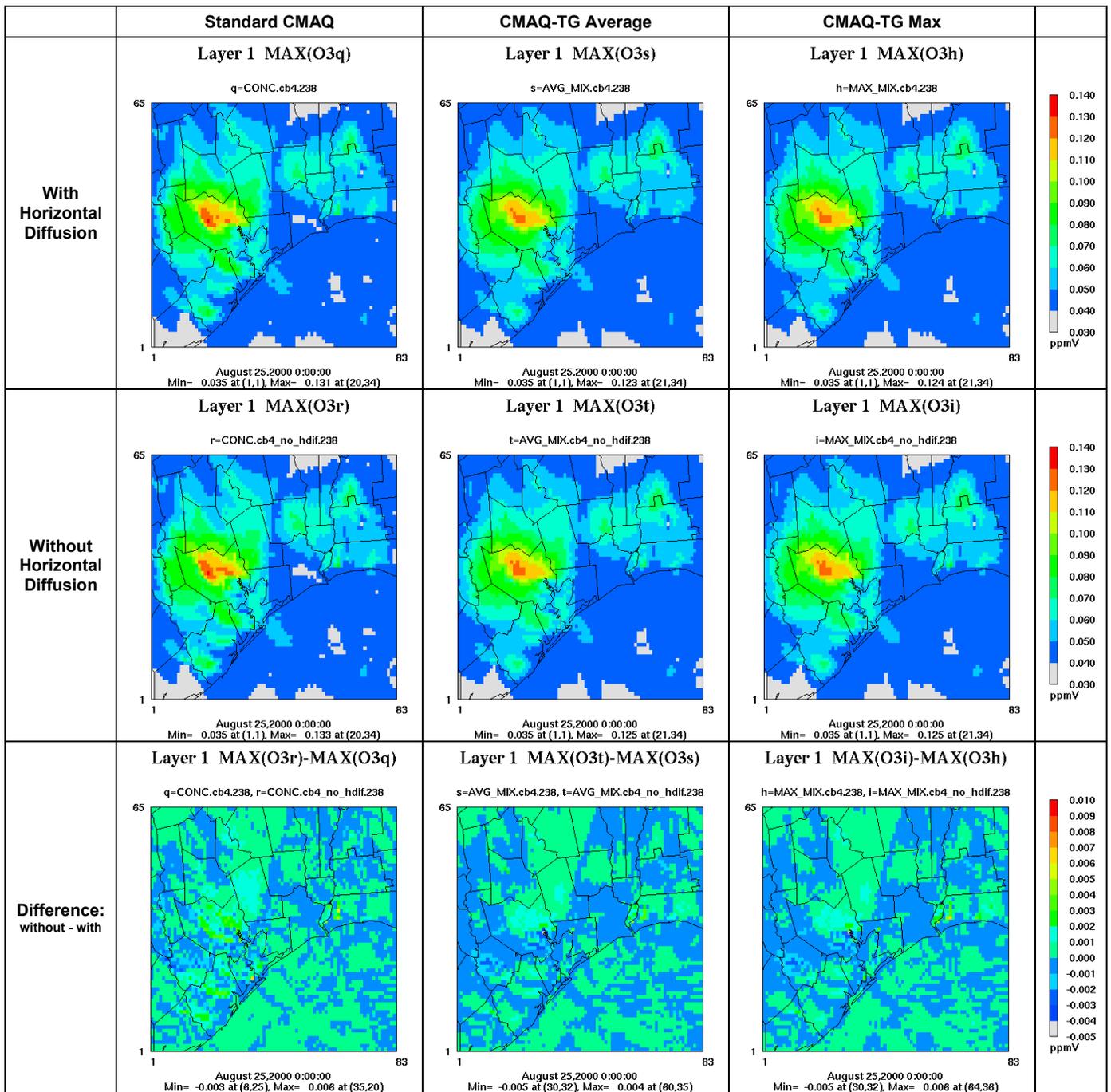


Figure 26 Comparison of the affect of horizontal diffusion on the same two simulations as in Figure 25. For both, not doing horizontal diffusion increases the daily maximum ozone value by 2 ppb. For CMAQ-TG, both cell average output and cell maximum output are shown, but are only slightly different.

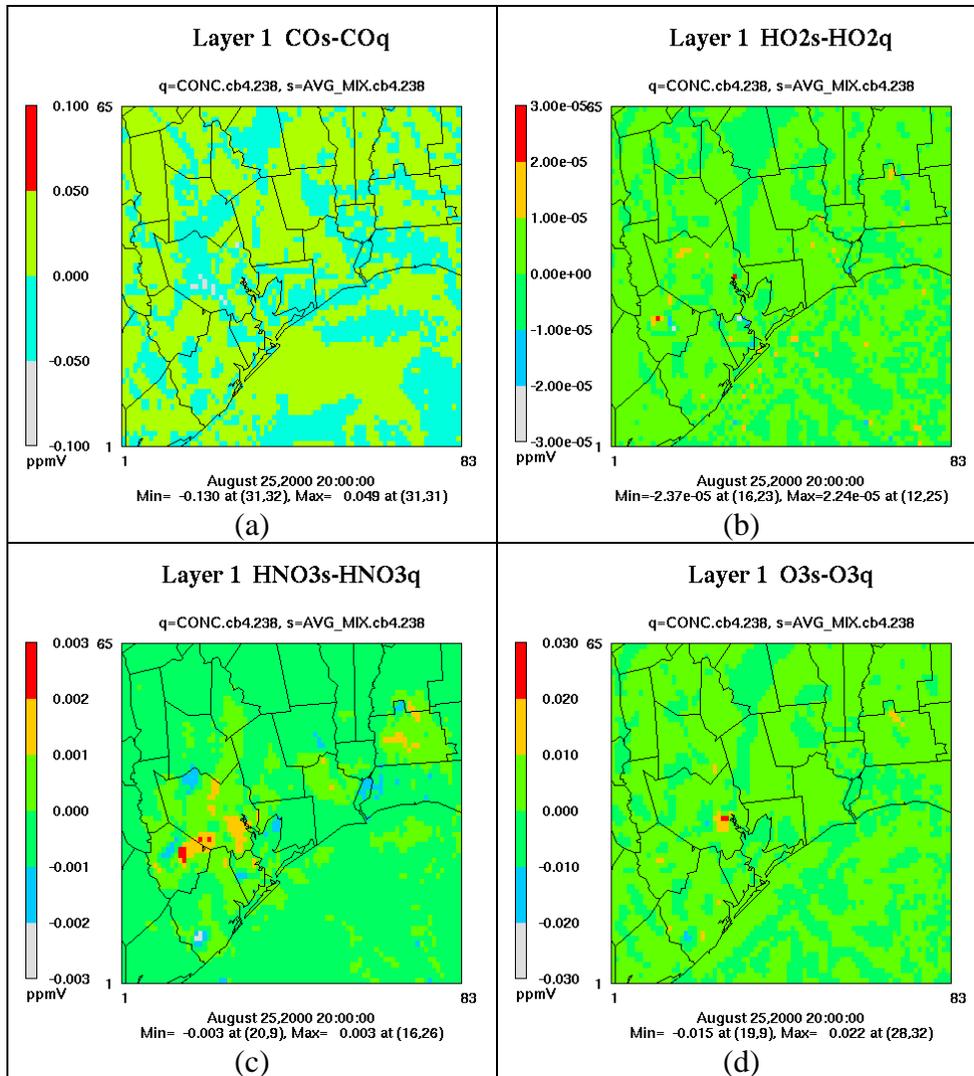


Figure 27 Differences in mixing ratios between CMAQ-TG and standard CMAQ for (a) CO, (b) HO<sub>2</sub>, (c) HNO<sub>3</sub>, and (d) O<sub>3</sub>.

## 5. Concluding Remarks and Future Directions

In its current form, CMAQ-TG supports many of the basic physical processes required for a CTM, including advection, diffusion, gas chemistry, emissions and dry deposition. The TG technique eliminates the serious spurious diffusion associated with Eulerian methods, without introducing other side effects such as false oscillations and negative concentrations. However, another form of numerical diffusion may be introduced by the packet spawning process that is part of the current packet management strategies. Basic versions of packet management tasks have been implemented, but effective packet management is probably the biggest obstacle to CMAQ-TG becoming competitive with standard CMAQ. Several representations of cell composition are available. However, displaying results using existing tools designed for uniform

grid structures causes information loss or "display diffusion", so improved display methods are needed.

Basic future development tasks for CMAQ-TG that we intend to complete include

- Implementing the remaining physical processes included in standard CMAQ, in particular, aqueous-phase and aerosol modeling
- Improving packet management
- Adding better display methods
- Supporting options for more accurate trajectory generation

Along with improvements to basic functions of CMAQ-TG, we hope to

- Add features so that CMAQ-TG can be used to identify source-receptor linkage, e.g., save full trajectories for packets and display trajectories of selected packets.
- Improve modeling of physical diffusion. Because TG advection does not have the side effect of numerical diffusion, current modeling of physical diffusion can be re-assessed, with the goal of improving it.

Further characterization of numerical properties and evaluation of the CMAQ-TG results against a comprehensive dataset from a field experiment is needed. It is hoped that ultimately a much improved CMAQ will be developed for scientific studies and regulatory applications.

## **6. Disclaimer and Acknowledgment**

The research described in this article has been funded mainly by the Research Agreement between the University of Houston and Coordinating Research Council (CRC) (Project No. A-55). Software developed or used in this project should be available to CRC without copyright or licensing restrictions. This work was also funded in part by the United State Environmental Protection Agency and Texas Air Research Center and their contributions are acknowledged. However, it has not been subjected to the peer and policy review of these institutions, and therefore does not necessarily reflect the views of the institutions and no official endorsement should be inferred.

## References

- Byun, D.W. and J.K.S. Ching, Ed., 1999: Science Algorithms of the EPA Models-3 Community Multiscale Air Quality (CMAQ) Modeling System, EPA Report, EPA/600/R-99/030, NERL, Research Triangle Park, NC.
- Byun, D. W., and S. Lee, 2002: Numerical solution of trace species advection under non-uniform density distribution: Experiment with two-dimensional linear flows. In IMA volume on Air Quality Modeling, Ed., D. Chock and G. Carmichael. 109-151
- Chock, D. P., Sun, P., and Winkler, S. L., 1996: Trajectory-grid: An accurate sign-preserving advection-diffusion approach for air quality modeling. *Atmospheric Environment*, **30**, 857-868.
- Chock, D. P and Winkler, S. L., 2000: A trajectory-grid approach for solving the condensation and evaporation equations of aerosols. *Atmospheric Environment*, **34**, 2957-2973.
- Chock, D. P, Winkler, S. L., and Sun, P., 2002: Effects of grid resolution and subgrid assumptions on the model prediction of non-homogeneous atmospheric chemistry. In IMA volume on Air Quality Modeling, Ed. D. Chock and G. Carmichael. 81-107.
- Chock, D. P., M. J. Whalen, S. L. Winkler, P. Sun, 2005: Implementing the trajectory-grid transport algorithm in an air quality model. *Atmospheric Environment*, **39**, 4015-4023.
- Environ, Novato, CA, 2004, <http://www.camx.com/overview.html>
- Evans, M.W., and Harlow, F.H., 1957; The particle-in-cell method for hydrodynamic calculations, LASL report no. LA-2139. Los Alamos, N. Mex.: Los Alamos National Laboratory
- Gaydos, T.M., Koo, B., Pandis, S.N., Chock, D.P., 2003: Development and application of an efficient moving sectional approach for the solution of the atmospheric aerosol condensation/evaporation equations. *Atmospheric Environment*, **37**, 3303-3316.
- Harlow, F.H., 1964; The particle-in-cell computing method for fluid dynamics. In *Methods in computational physics* 3, Eds. B. Adler, S. Fernback, and M. Rotenberg, 319-343. New York: Academic Press.
- Harlow, F.H., and Welch, J.F., 1965; Numerical calculation of time-dependent viscous incompressible flow of fluid with free surface. *Physics of Fluids*, **8**, 2182-2189.
- Marder, B.M., 1975; GAP – A PIC-type fluid code. *Mathematics of Computation*, **24**, 434-436.

Oran, E.S. and Boris, J.P., 2001; *Numerical Simulation of Reactive Flow*, 2<sup>nd</sup> Edition. Cambridge University Press, 529 pp.

## Appendix A - CMAQ-TG Controls

Options for packet management and for limiting sub-grid diffusion are controlled by environment variables that a user can set in a run script. These are the only additional controls on CMAQ-TG beyond those for the standard version of CMAQ. To get documentation for standard CMAQ, go to <http://www.cmascenter.org/>.

### High Resolution (HR) Box

Cells inside the HR box are initialized with  $HR\_MULT^2$  uniformly spaced packets at the vertical center of the cell; cells outside the HR box are initialized with just one packet at the center of the cell

- Parameters
  - `HR_MULT` - resolution multiplier for box [default: 2]
  - `HR_BCOL_ECOL` - column range for box [default: "0 NCOLS+1" (all columns)]
  - `HR_BROW_EROW` - row range for box [default: "0 NROWS+1" (all rows)]
  - `HR_TOP` - top layer for box [default: 2]

### Spawning

Create new packets in cells with too few packets; spawning is done every time step by creating a new packet at the center of a cell; HR cells are never allowed to be empty, except when spawning is turned off

- Available Methods (possible values for the environment variable `FILL_METHOD`)
  - `NO_FILL` - never add packets to interior cells; i.e., turn off spawning
  - `FILL_ALL` - never allow any cell to be empty  
[default `FILL_METHOD`]
  - `SPARSE_FILL` - spawn in an interior low resolution cell if it is empty and is surrounded, in its layer, by empty cells

### Pruning

Reduce the number of packets in a cell to the initial count for the cell

- Available Methods (possible values for the environment variable `PRUNING_METHOD`)
  - `NO_PRUNING` - no pruning
  - `KEEP_CLOSEST` - keep packets that are closest to center of cell  
[default `PRUNING_METHOD`]
  - `KEEP_OLDEST` - keep packets that are oldest
- Parameters
  - `PRUNING_FREQ` - pruning frequency (number of time steps) [default: 5]
  - `HR_KEEP_IN_CELL` - number of packets to keep after pruning HR cell  
[default:  $HR\_MULT^2$ ]
  - `HR_KEEP_TOL` - prune HR cell when  $NUM\_IN\_CELL > HR\_KEEP\_IN\_CELL + HR\_KEEP\_TOL$   
[default:  $HR\_KEEP\_IN\_CELL$ ]

- NR\_KEEP\_IN\_CELL - number of packets to keep after pruning normal resolution (NR) cell [default: 2]
- NR\_KEEP\_TOL - prune NR cell when  $\text{NUM\_IN\_CELL} > \text{NR\_KEEP\_IN\_CELL} + \text{NR\_KEEP\_TOL}$  [default: 2]

### Sub-grid Diffusion

- Parameters
  - MAX\_SGD\_FAC - maximum sub-grid diffusion factor allowed [default: 0.1] (choose 0.0 for no sub-grid diffusion)