4) Airmass History Biases as Functions of Back Trajectory Model and Input Meteorological Data

4.1 Introduction

The goals of the evaluation of wind fields and back trajectory models within BRAVO were to determine, using this very data-rich project, 1) whether there are biases between the various available back trajectory models and available input fields; 2) whether past studies using the older Atmospheric Transport and Diffusion (ATAD) model with the usual raw rawinsonde data as input are valid in the light of new models and new gridded wind fields; 3) to determine the best model and wind fields to use for future routine back trajectory modeling both for Big Bend National Park and for other sites; 4) to determine whether the unavailability of the EDAS gridded wind field during October caused any biases during that month; and, 5) to determine whether any findings of biases in trajectory models and wind fields are valid for Big Bend during summer and fall 1999 only (the BRAVO study period) or whether similar biases exist for other regions of the country and for other seasons and other years in south Texas.

Models to be examined include ATAD, HYSPLIT ver. 4.5, and CAPITA Monte Carlo (CMC). Input meteorological data include North American rawinsonde data, these data plus information from the four BRAVO wind profilers, and three gridded wind fields, EDAS, FNL, and MM5. Details of the input meteorology are discussed in section 2.1.3 and details of the three back trajectory models are given in section 2.2. Table 4-1 briefly summarizes the major differences between the three back trajectory models. Appendix 4a has additional analyses and discussion of the MM5 and EDAS wind fields.

4.2. Evaluations to Determine Whether or Not There Is a Bias during the BRAVO Period

4.2.1 Simple Plotting of Trajectories Generated in Various Ways.

As a first step, we graphically examined the horizontal and vertical locations of back trajectories generated by each model/data set combination four times per day for the duration of the BRAVO study period (July 1–October 31, 1999). All of these trajectory plots are in Appendices 4b, 4c, and 4d. HYSPLIT trajectories were started at 10, 100, 200, 500, 1000, and 2000 m above ground. All trajectories were traced for 5 days backward in time from Big Bend National Park, TX (103.18 deg W, 29.30 deg N). In order to calculate a back trajectory, it is necessary to have meteorological data for a few days preceding and usually a few hours following the start time of the trajectory. Since the start and end dates of the MM5 data were July 1 and October 31, it was not possible to use the MM5 data for the first few days of the study, the last day, or part of the last day depending on the requirements of the back trajectory model. Several general observations resulted. Illustrations of these findings are discussed below.

Table 4-1. Summary of the major differences between the three trajectory models used for BRAVO.

Feature	ATAD	HySPLIT	CAPITA MC
Input Data Type	Designed for raw sounding data. Gridded data or supplemental profiler data can also be used	Gridded meteorological data. BRAVO study used EDAS, FNL, and MM5 36 km	Gridded meteorological data. BRAVO study used EDAS, FNL, and MM5 36 km
Available Input data	1946–present for raw soundings. 1997–present for EDAS and FNL, July–October 1999 for MM5 and BRAVO wind profilers	1997-present for EDAS and FNL July–October 1999 for MM5	1997-present for EDAS and FNL July–October 1999 for MM5
Horizontal Motion	Within the transport layer, average winds calculated from stations within a given radius.	Input horizontal winds	Input horizontal winds
Trajectory Start Height	As run for BRAVO, Depends on transport layer (see vertical motion).	Chosen by user, for BRAVO, 10, 100, 200, 500, 1000, 2000 m above ground.	Below mixing height
Vertical Motion	Trajectory stays in a variable calculated transport layer, approximately 300–3000 km above ground. Fixed vertical layers are also possible	For BRAVO, input vertical winds used to calculate transport. Maximum 10 km height. Other options are available.	Random vertical motion in mixed layer, vertical motion by input data above
Trajectory Start Frequency	Maximum of 4 times per day at 0, 6, 12, and 18 GMT.	Hourly at multiple heights	Unlimited particles per hour (used 20–50 for BRAVO applications)
Frequency of Output	Every 3 hours	Hourly * number of start heights	Hourly * number of particles per start time
Trajectory Length	Maximum 5 days	For BRAVO 5, 7, and 10 days	For BRAVO 5, 7, and 10 days

On some days all trajectories are similar, thus the choice of model, input data, and even starting height is unimportant. This is more likely to occur during the fall than during summer. An example for trajectories started on September 15, hour 18, is shown in Figure 4-1. Although there certainly are subtle differences in trajectory placement depending on start height, data, and model, all trajectories are arriving from east Texas. The CMC trajectories started at this time (shown on the right side) are horizontally similar to those generated by HYSPLIT and ATAD (on the left side), though those generated using MM5 input extend farther into the southern tier states to the east of Texas than do any of the others, including the other two models with the same input wind field. For this start time, the heights of the CMC trajectories generated with EDAS are much lower than those generated with MM5 and seem to be more similar to the heights of the HYSPLIT trajectories.

A common pattern, especially in the summer, is that there is a directional bias between trajectories generated with different input data. ATAD with raw rawinsonde data as input typically has airmasses arriving from the most southerly direction; HYSPLIT and CMC with EDAS input have airmasses arriving from the most easterly; while either HYSPLIT or CMC with FNL input generate trajectories between the extremes. Figure 4-2 (July 19, hour 6) shows an example of the typical type of directional bias that was observed frequently during July and

August. The ATAD trajectory with rawinsonde input (black) is arriving from the most southerly direction. Trajectories generated by HYSPLIT using EDAS (solid colored lines) generally arrive from the most northeasterly directions, while those generated by HYSPLIT with FNL (dotted colored lines) or MM5 (dashed colored lines) fall between the extremes. Note that ATAD trajectories generated with a gridded wind field (brown lines) are more similar to HYSPLIT trajectories than ATAD trajectories generated with raw sounding data. The CMC trajectories, shown in the right panel, also show a tendency for those generated using EDAS input to be more northeasterly than those generated with MM5 input for this start time.

An example of a day when choice of model and wind field made a large difference in the resultant back trajectories is shown in Figure 4-3 (September 20, hour 18). ATAD with rawinsonde data (black line) indicates airmasses arrived from deep within Mexico after possibly crossing the Gulf of Mexico. However, note trajectories generated with HYSPLIT and started at 200 m. The EDAS input (solid blue) creates a trajectory from the north, FNL input (dotted blue) creates a trajectory arriving from across the Gulf of Mexico, and MM5 input (dashed blue) generates a trajectory arriving from the west. The CMC trajectories, shown on the right, arrive mostly from the east with MM5 input, but mostly from the west with EDAS input. The differences are related to the height of the particles, with higher level airmasses arriving from the west. In this case, the CMC trajectories generated using EDAS input are higher than those generated with MM5 input. This is the reverse of the example shown in Figure 4-1.

For HYSPLIT, which requires specification of a start height, height differences between trajectories started at the same height, but generated with EDAS vs. FNL can be as large as several thousand meters after 5 days. In the example, shown in Figure 4-4 (July 24, hour 12), this is particularly evident in the trajectories started at 1000 m (green lines) and 2000 m (orange lines).

Directional differences between trajectories generated by HYSPLIT using EDAS vs. FNL can occasionally be as much as 180 degrees and, as expected, are largest when trajectory height differences are large. Again, in the example shown in Figure 4-5 (July 18, hour 6), this is particularly evident in the trajectories started at 1000 m (green lines) and 2000 m (orange lines). The CMC trajectories started at this same time are similar to the HYSPLIT trajectories started at the lower start heights. On this day, while the CMC heights are similar for both EDAS and MM5 inputs, the EDAS trajectories arrive from a more northeasterly direction.

For HYSPLIT, directional differences between input wind fields can exist even when the trajectory heights are similar. For example, see the HYSPLIT trajectories started at 500 m (cyan lines) in the example shown in Figure 4-6 (August 6, hour 18). This day also illustrates the occasional differences between outputs from the CMC model vs. HYSPLIT. Notice that except for the trajectory started at 2000 m and using EDAS input (dashed orange), all the HYSPLIT and ATAD trajectories arrive from the southeast. However, roughly half of the CMC trajectories, including those generated with both EDAS and MM5 input, arrive from a much more northeasterly direction.
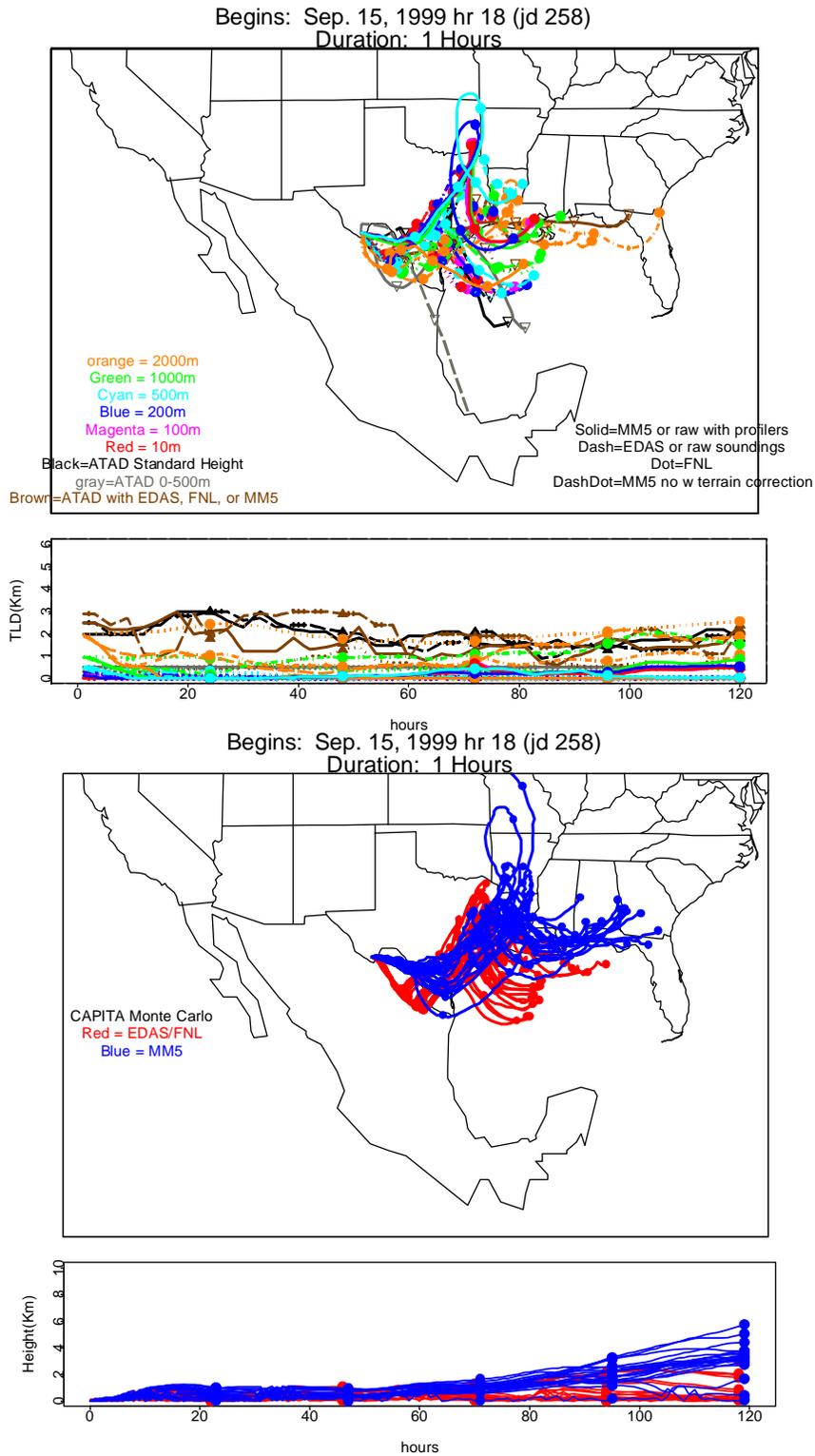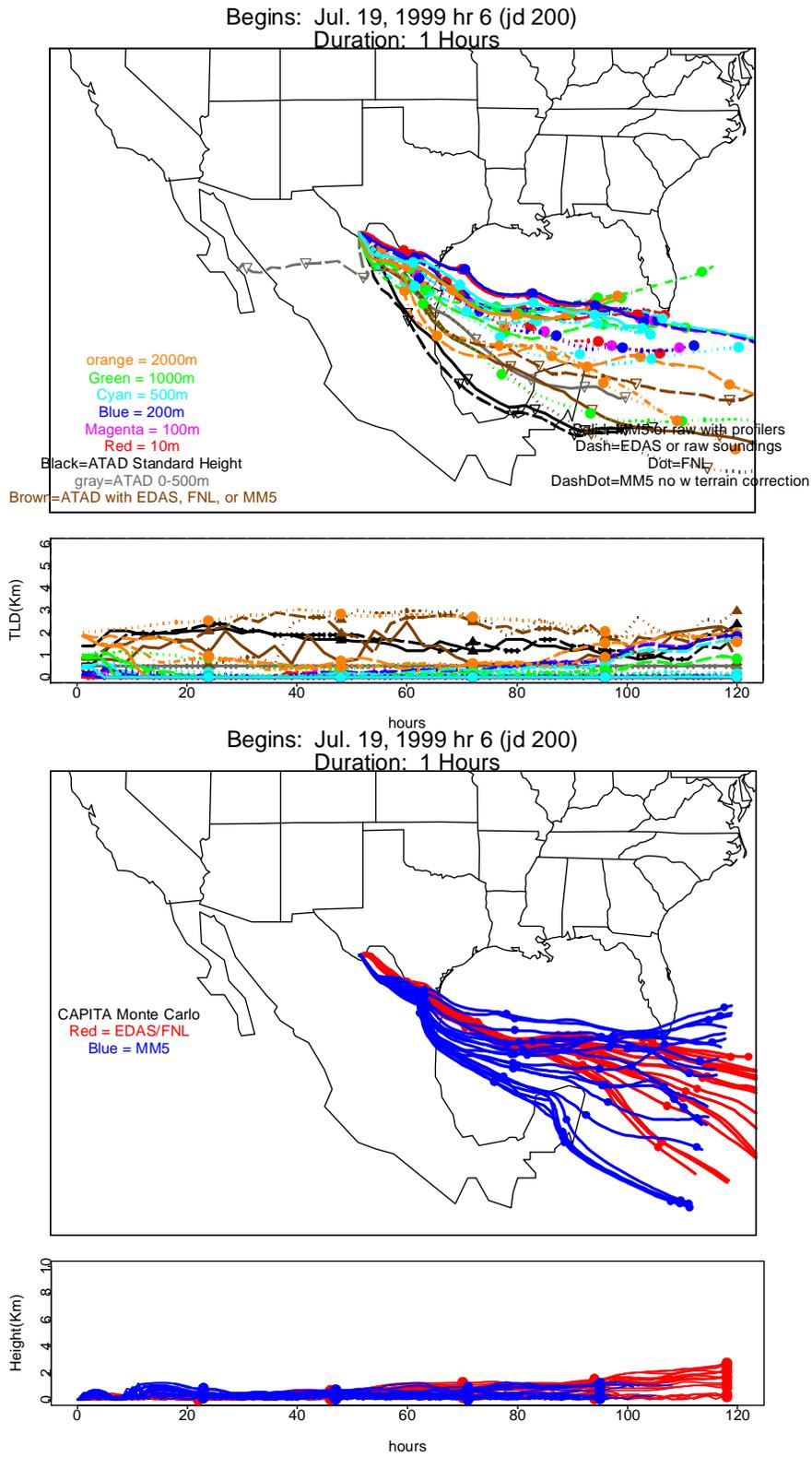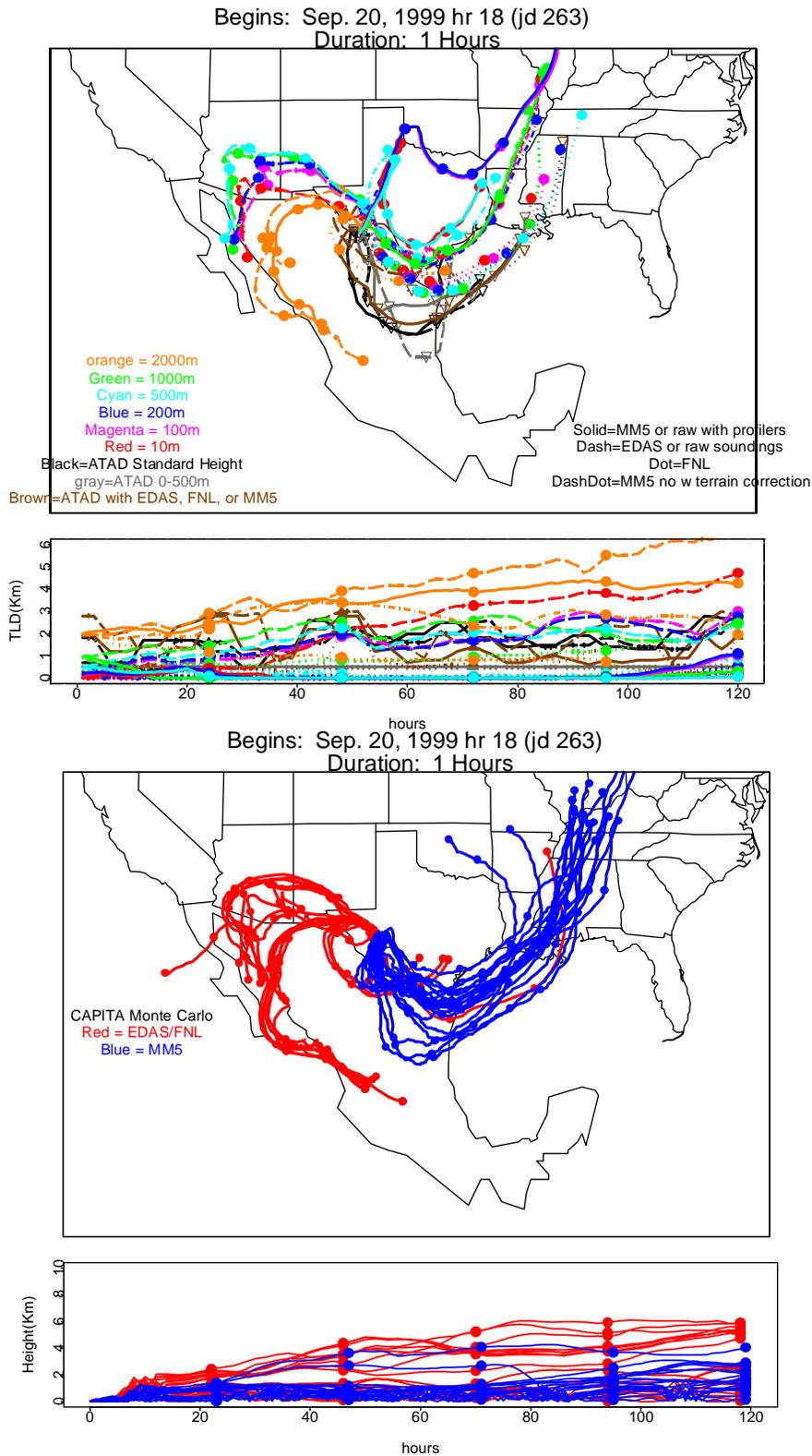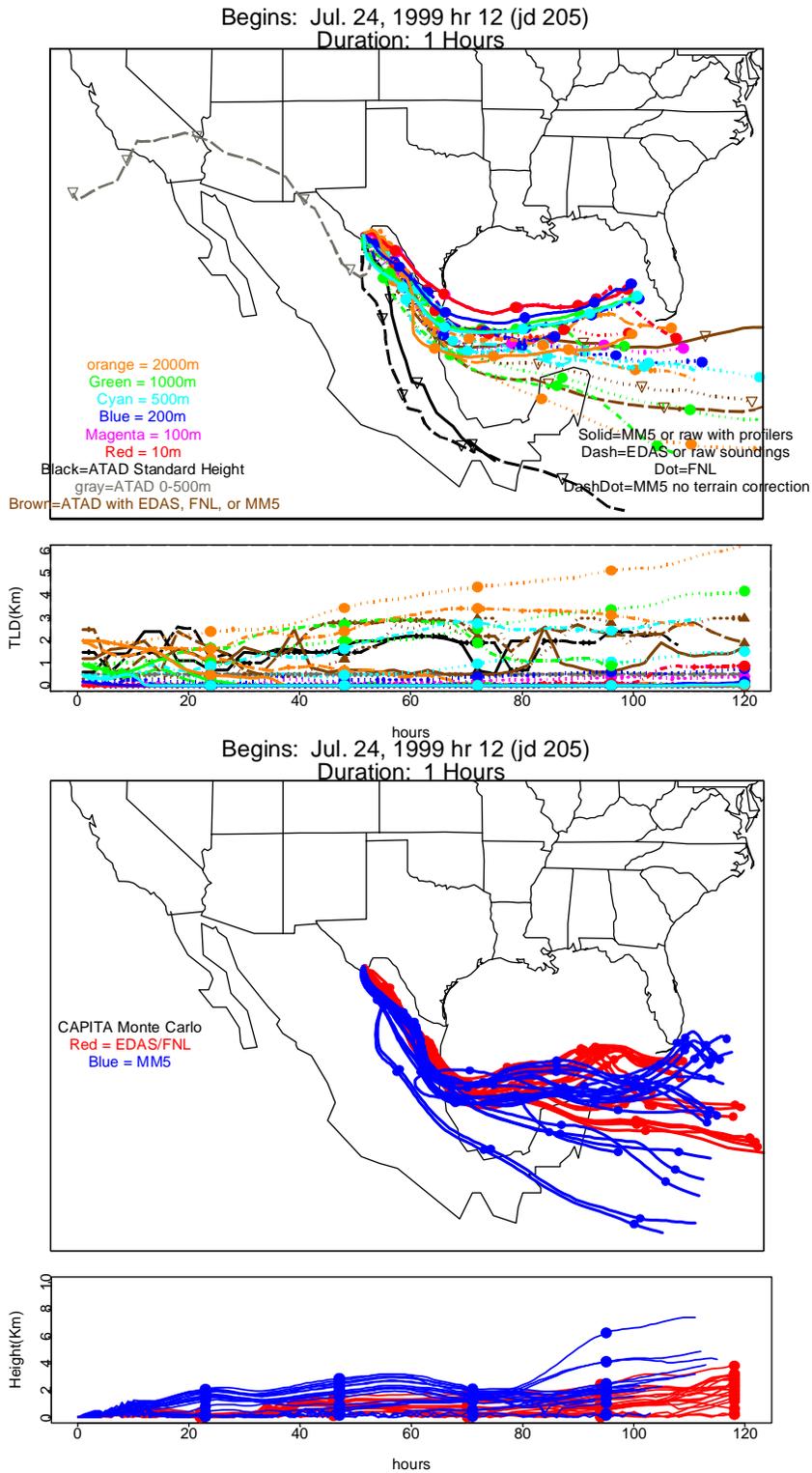
Figure 4-1. Five-day back trajectories generated by all model/input combinations beginning on 9/15/1999, 18:00 CDT. Top are HYSPLIT and ATAD outputs, bottom is CMC.

Begins: Jul. 19, 1999 hr 6 (jd 200)
Duration: 1 Hours

orange = 2000m
Green = 1000m
Cyan = 500m
Blue = 200m
Magenta = 100m
Red = 10m
Black=ATAD Standard Height
gray=ATAD 0-500m
Brown=ATAD with EDAS, FNL, or MM5

Solid=MM5 or raw with profilers
Dash=EDAS or raw soundings
Dot=FNL
DashDot=MM5 no w terrain correction

Begins: Jul. 19, 1999 hr 6 (jd 200)
Duration: 1 Hours

CAPITA Monte Carlo
Red = EDAS/FNL
Blue = MM5

Figure 4-2. Five-day back trajectories generated by all model/input combinations beginning on 7/19/1999, 6:00 CDT. Top are HYSPLIT and ATAD outputs, bottom is CMC.

87

Figure 4-3. Five-day back trajectories generated by all model/input combinations beginning on 9/20/1999, 18:00 CDT. Top are HYSPLIT and ATAD output, bottom is CMC.

Figure 4-4. Five-day back trajectories generated by all model/input combinations beginning on 7/24/1999, 12:00 CDT. Top are HYSPLIT and ATAD output, bottom is CMC.
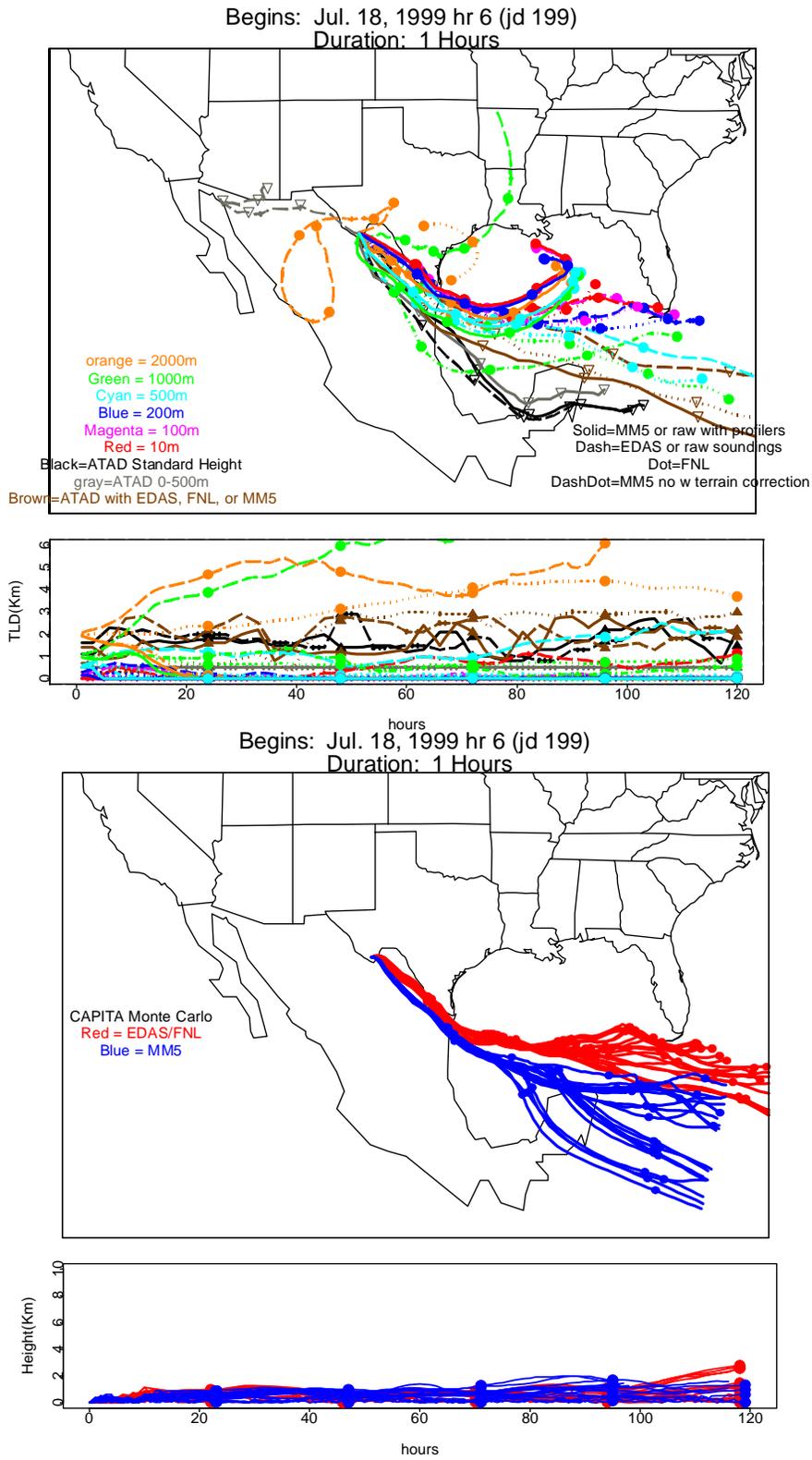
Figure 4-5. Five-day back trajectories generated by all model/input combinations beginning on 7/18/1999, 6:00 CDT. Top are HYSPLIT and ATAD output, bottom is CMC.
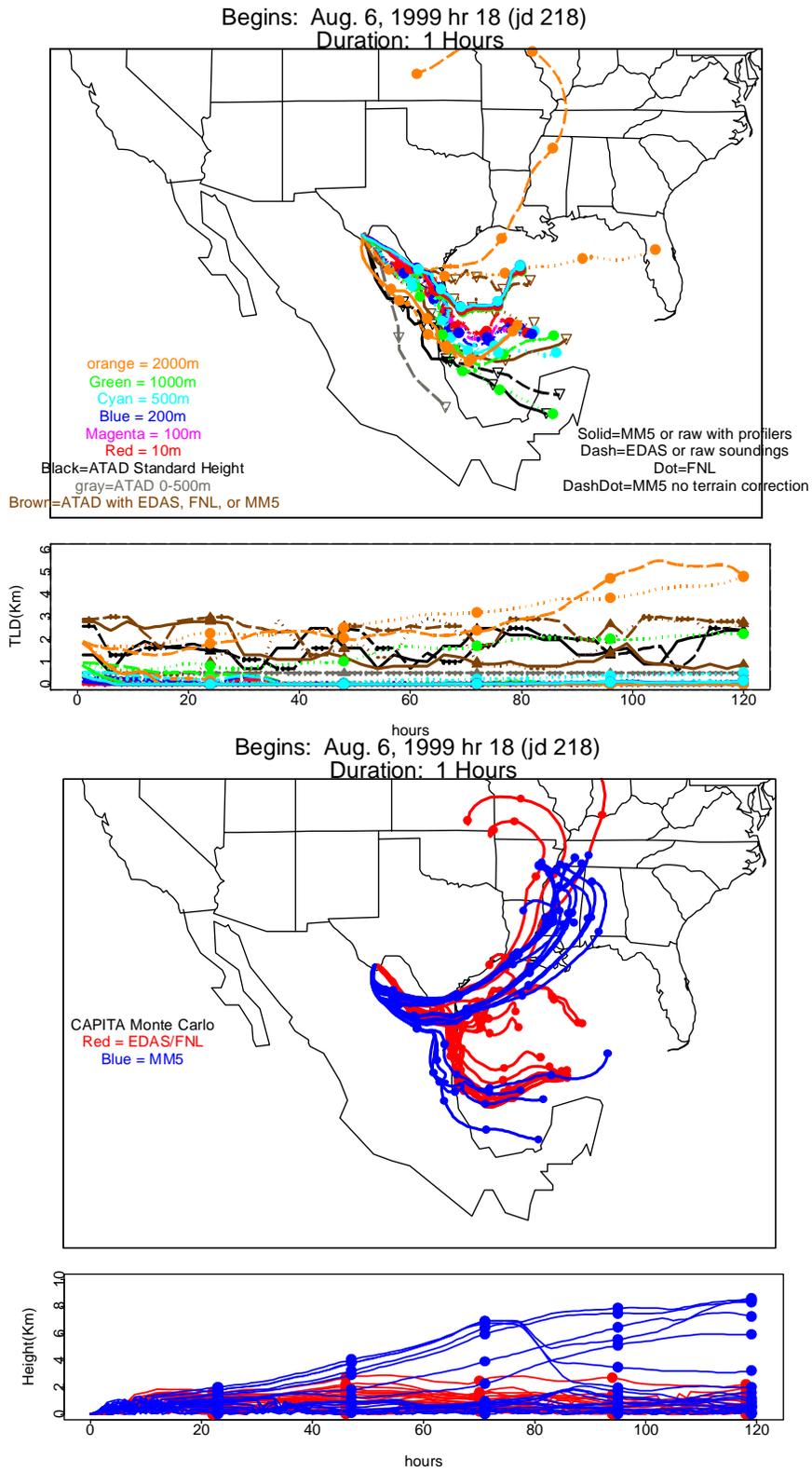
Figure 4-6. Five-day back trajectories generated by all model/input combinations beginning on 8/6/1999, 18:00 CDT. Top are HYSPLIT and ATAD output, bottom is CMC.

Trajectories generated with either HYSPLIT or CMC can reach heights as high as 10 km, clearly above the mixed layer that is likely to be carrying most air pollutants. An advantage of CMC over HYSPLIT is that the height of the mixed layer is known and available as output. Both CMC and ATAD provide estimates of transport integrated throughout the mixing layer. Due to its different design, ATAD does not generate trajectories above its "transport layer" which has a maximum of about 3 km above ground.

From these simple graphical analyses, it is apparent that there can be significant episodic differences between the modeled back trajectories depending on the model and input data used to generate the trajectories. These observations are examined in a more systematic manner in the following sections.

4.2.2 Calculation of BRAVO Overall Residence Times and the Differences between Them

Overall residence times (ORT) were calculated for Big Bend for July–October, for July–September, for each of the four months individually, and for each of seven high sulfur episodes using 5-day and 10-day back trajectories generated with each model/data set. Evaluating periods without October in some cases allowed a comparison between results using EDAS vs. FNL inputs. Briefly, ORT is a count of the number of back trajectory endpoints or particles in each grid cell and shows the predominant direction(s) from which airmasses arrived as predicted by the model. The methodology is described in detail in section 2.3.1.3. While the graphical analysis described in the previous section made it apparent that there were occasional differences between the model/input combinations, ORT allows a more systematic evaluation of the average differences between the modeled back trajectories. Specifically, we would like to 1) examine the consequences of using ATAD with raw sounding data in previous studies, 2) examine the consequences of using FNL rather than EDAS data during October, 3) determine a reasonable start height or heights for HYSPLIT, and 4) determine whether there are systematic differences between the results when using a single model but different input meteorology or when using a fixed input data set in different models.

4.2.2.1 ATAD Model with Rawinsonde Data Compared to Other Combinations; Implications for Previous Studies

ORTs for July–October 1999 all generated with ATAD, but using four different input data sets, rawinsonde data only, rawinsonde data plus data from the four BRAVO wind profilers, MM5 36 km gridded wind field, and the EDAS/FNL gridded wind field, are shown in Figure 4-7. Moving from a) through d), ATAD predicts that airmasses arrive predominantly from progressively less southerly and more easterly directions. Use of the raw rawinsonde data in ATAD (Figure 4-7a) results in the most southerly predominant wind direction. The addition of even a small quantity of additional data, such as adding data from the four BRAVO wind profilers (Figure 4-7b), causes ATAD to predict a slightly more easterly average transport direction. ATAD with MM5 input (Figure 4-7c) yields yet a more easterly mean inflow direction, while ATAD with EDAS/FNL input (Figure 4-7d) is the most easterly of the four sets.
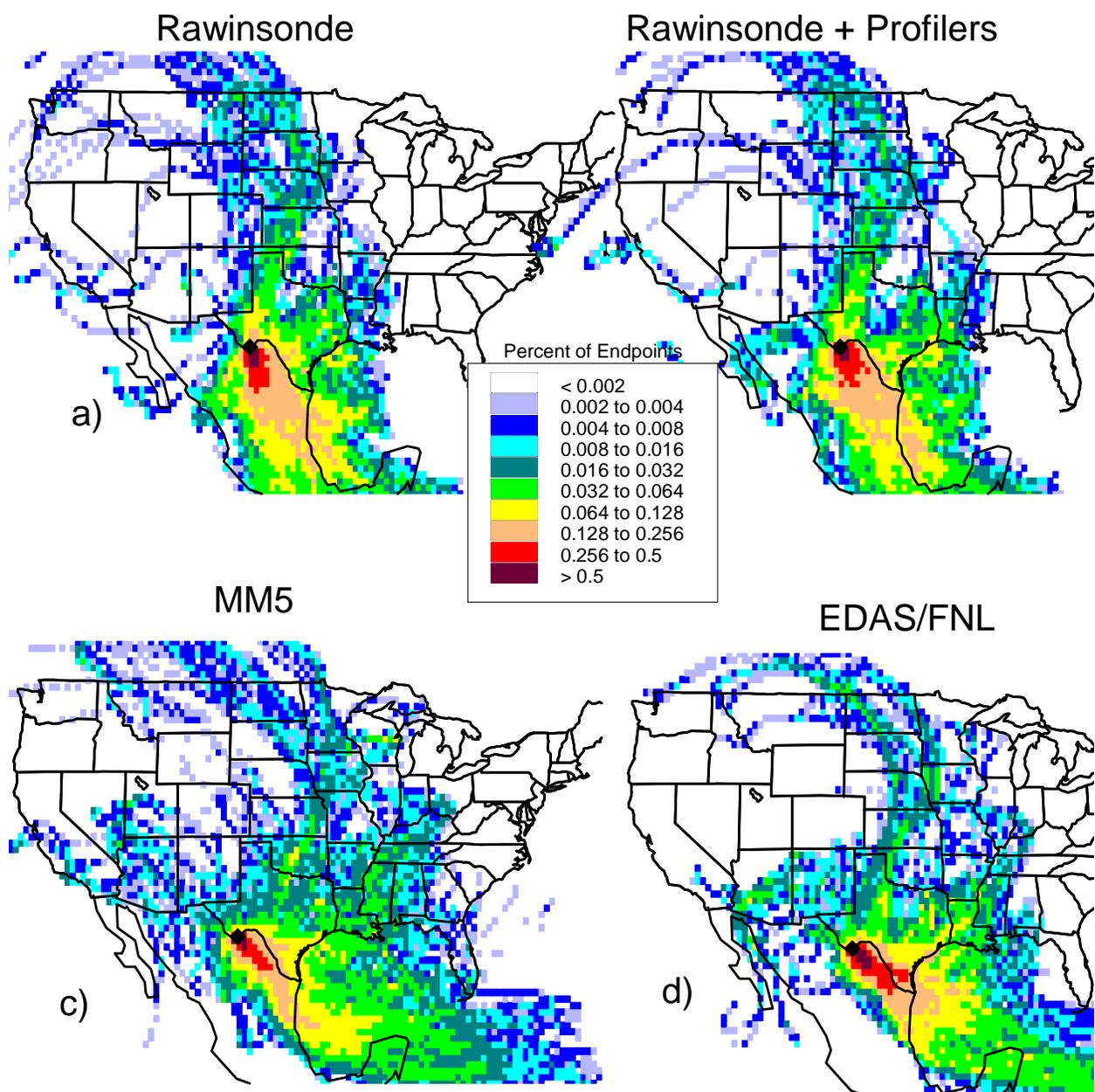
Figure 4-7. Overall Residence Times all generated with ATAD for July–October 1999, grid size ½ deg lat by ½ deg lon. a) using NOAA rawinsonde data b) using rawinsonde data plus data from four BRAVO wind profilers c) using MM5 36 km gridded meteorology d) using EDAS when available and FNL for October.

Big Bend is a particularly challenging receptor site to model using only rawinsonde data due to its proximity to both complex terrain and to data-sparse areas. In Mexico there are fewer balloon launch sites than in the U.S. and often there are data for these sites only once per day rather twice per day as is usual in the U.S. Also there are no rawinsonde data over the Gulf of Mexico. If the reason for the southerly bias in ATAD when the rawinsonde data are used alone as input is due to the sparseness of the input data, then trajectories generated by this combination of model/input data should not be used for further source attribution at Big Bend. This

conclusion does not necessarily extend to other receptor sites located in areas with less complex terrain and more rawinsonde data. Also, it's unclear from this limited analysis whether a similar problem exists during other seasons when the predominant wind directions in the region are different and more likely to arrive from areas in the U.S. where there are more data than in Mexico and over the Gulf of Mexico.

Another issue is whether ATAD trajectories are similar to or different from trajectories from the other two trajectory models when identical data sets are used as input. Figure 4-8 shows the four-month average ORT generated by ATAD, CAPITA MC, and HYSPLIT, all using MM5 data as input. Figure 4-9 is the same except all three models are using the EDAS/FNL combination input. For HYSPLIT, results shown are for an aggregation of start heights including 100, 200, 500, and 1000 m AGL.

For both MM5 and EDAS/FNL input, results from ATAD and CMC appear the most similar while the predominant wind direction calculated by HYSPLIT using the same input data is less southerly and more along the Texas-Mexico border. An obvious difference is in the smoothness of the contours. ATAD has orders of magnitude fewer endpoints, generating only 40 during a single 5-day trajectory compared to 120 by HYSPLIT and a variable number by CMC depending on how many particles were released per hour and how often the particle position was saved. While ATAD can have only 4 incoming pathways per day, HYSPLIT can have 24 per start height (four heights in this case) and CMC has 480 for this case of 20 particles per hour. The total number of endpoints used for July–October ORT statistics was approximately 20,000 for ATAD, $1.3\text{-}1.4 \times 10^6$ for HYSPLIT depending on the input data, 1.7×10^6 for CMC with EDAS/FNL, and 3.1×10^6 for CMC with MM5 input.

As we already know from examining the individual start times in the previous section, the height of the back trajectory can have a large influence on the horizontal placement of the trajectory; in general, the higher the mean height, the faster the mean speed of the trajectory. This has implications for source attribution because higher wind speeds will tend to implicate more distant sources at the expense of closer sources, with lower speeds preferentially implicating nearer source areas. Table 4-2 shows the mean endpoint heights and wind speeds for all model and input data combinations used for BRAVO. When EDAS/FNL data are used as input, HYSPLIT generated trajectories were higher on average than when MM5 data were used at the same start height. However, ATAD and CMC generated trajectories with higher mean heights when MM5 data were used as input. Also, there is a large difference in mean heights between HYSPLIT and the other two models. This is a function of the start heights chosen for HYSPLIT. A HYSPLIT start height of 1000 m most closely matches the mean heights generated by ATAD and CMC which do more averaging within the mixed layer. For HYSPLIT with start heights of 500 m and greater, the mean trajectory height is approximately equal to the start height. For lower start heights, the mean trajectory height was never lower than 200 m, no matter what the start height.
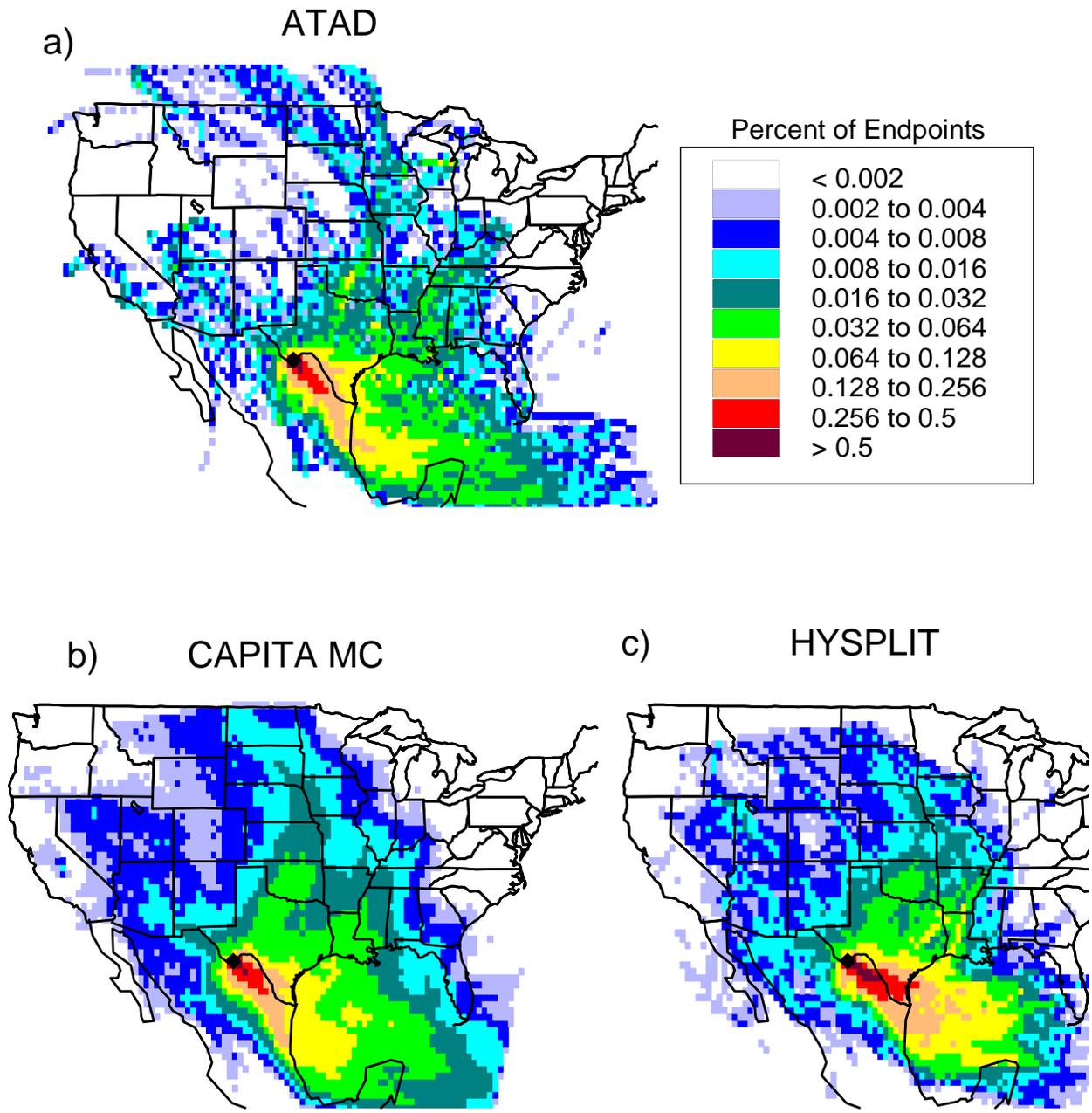
Figure 4-8. Overall Residence Times for July–October 1999, grid size ½ deg lat by ½ deg lon. all with 36 km MM5 input a) ATAD b) CAPITA MC c) HYSPLIT started at 100, 200, 500, 1000 m AGL.
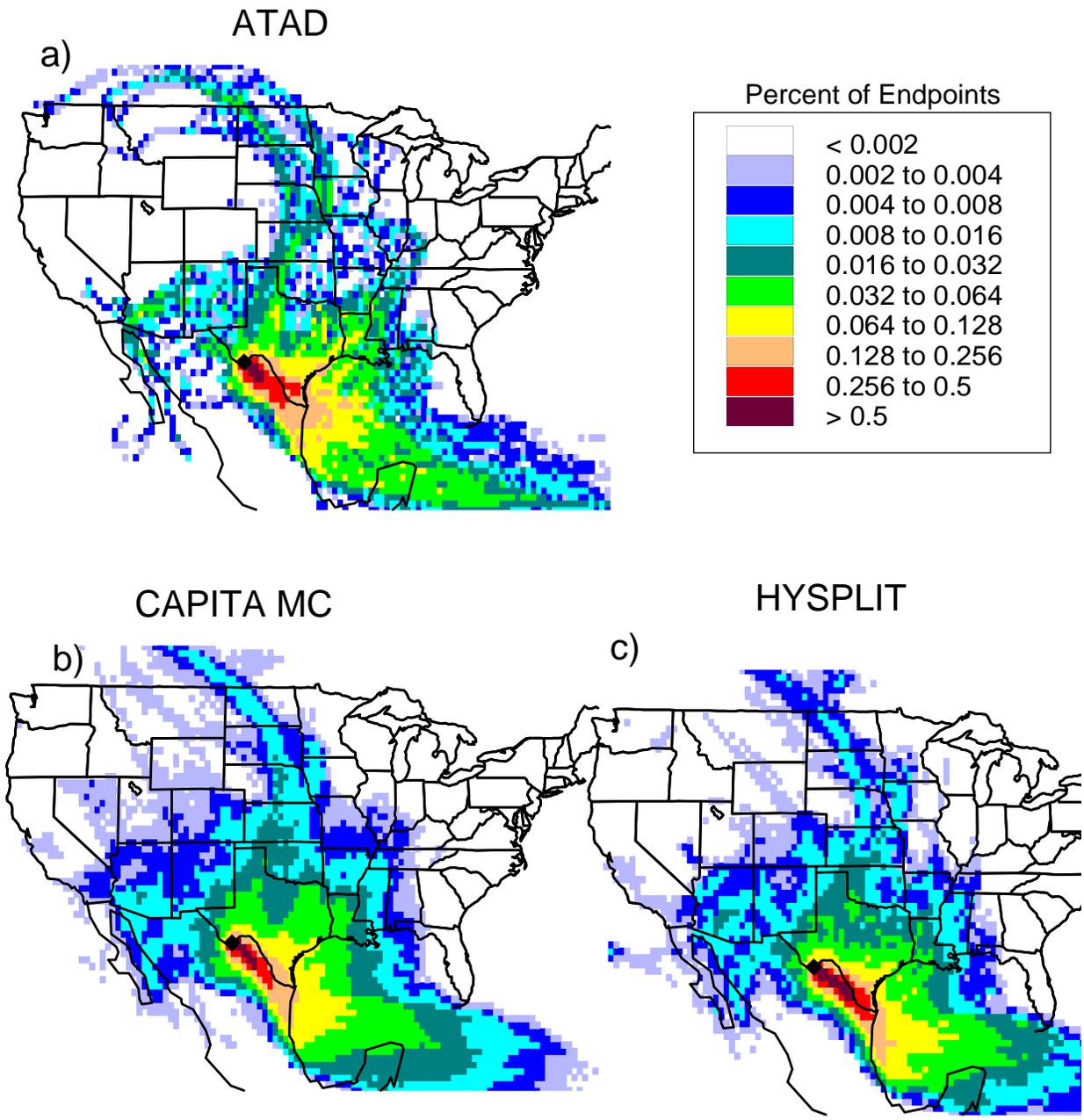
ATAD

a)

Percent of Endpoints

- < 0.002
- 0.002 to 0.004
- 0.004 to 0.008
- 0.008 to 0.016
- 0.016 to 0.032
- 0.032 to 0.064
- 0.064 to 0.128
- 0.128 to 0.256
- 0.256 to 0.5
- > 0.5

CAPITA MC

b)

HYSPLIT

c)

Figure 4-9. Overall Residence Times for July–October 1999, grid size ½ deg lat by ½ deg lon. All with EDAS/FNL input a) ATAD b) CAPITA MC c) HYSPLIT started at 100, 200, 500, and 1000 m AGL.

Table 4-2. Mean endpoint heights and wind speeds for 5-day back trajectories for July–October 1999 generated by all the model and input data combinations including all the start heights for HYSPLIT. A HYSPLIT start height of 100–1000 refers to an average of the results from 100, 200, 500, and 1000 m start heights.

Model	Input Data	Start Height (m AGL)	Mean Endpoint Height (m AGL)	Mean Wind Speed (m/sec)
ATAD	Rawinsonde	NA	1721	5.3
	Rawinsonde + BRAVO profilers	NA	1726	5.4
	EDAS/FNL	NA	1718	6.2
	FNL	NA	1181	5.9
	MM5	NA	1255	6.9
HYSPLIT ver. 4.5	EDAS/FNL	10	238	4.6
		100	364	5.0
		200	447	5.3
		500	706	5.8
		1000	1240	6.2
		2000	2617	6.9
		100–1000	686	5.6
	MM5	10	238	5.0
		100	252	5.1
		200	288	5.2
		500	462	5.3
		1000	950	5.4
		2000	2009	5.5
		100–1000	487	5.3
CMC	EDAS/FNL	NA	1156	6.2
	MM5	NA	1164	7.0

Figures 4-10 and 4-11 show typical representations of the heights of the layers of the MM5, EDAS, and FNL meteorological fields calculated using a U.S. Standard Atmosphere to convert pressures to heights. Because sigma height coordinates are used in all of the gridded data, (see section 2.1.3) the actual heights of the layers vary in time and space, but these figures illustrate typical differences between the three sets of data. Figure 4-10 shows all layers and 4-11 shows the lowest 5000 m so that the lower levels can be seen and include dotted red lines to show the HYSPLIT start heights. For BRAVO, HYSPLIT was run with a maximum height of 10,000 m and ATAD does not use data above 3000 m. It is evident from the thickness of the layers as shown in these figures that HYSPLIT trajectories will be least sensitive to start height when FNL data are used and most sensitive when MM5 data are input.
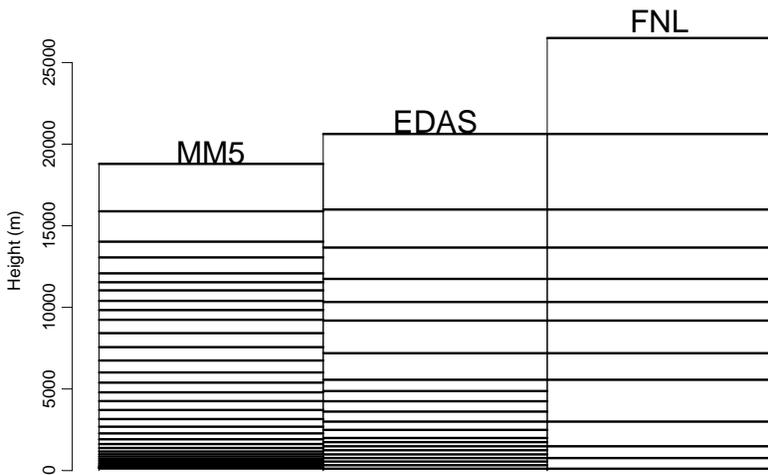
Figure 4-10. Typical heights of all vertical layers for each of the three gridded wind fields. See Figure 4-11 for detail of the lowest layers.
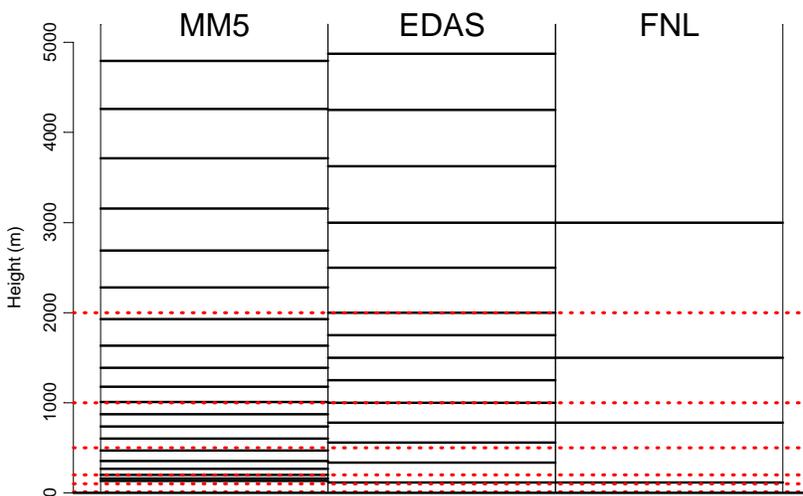
Figure 4-11. Typical heights of the vertical layers for each of the three gridded wind fields for the lowest 5000 m (black solid lines) and five HYSPLIT start heights of 10, 100, 200, 500, 1000, and 2000 m, shown as dotted red lines.

The fastest mean wind speeds are generated with the CMC model, with MM5 input being faster on average than EDAS/FNL input, possibly due to the mean height also being somewhat higher with MM5 input. Fastest mean wind speeds from ATAD are also with MM5 input, followed by EDAS/FNL input.

Figures 4-12 and 4-13 show maps of the differences between pairs of overall residence time output fields. Figure 4-12a shows the differences between ATAD with MM5 and CMC with MM5, while Figure 4-12b shows the differences between ATAD with MM5 and HYSPLIT with MM5. Figure 4-13 is the same except all results are using EDAS/FNL input data rather than MM5. In all cases the warmer colors, yellow through dark red, indicate a positive difference, or areas where the first model of the pair was more likely to place endpoints, while the cooler blue and green colors indicate a negative difference, or areas where the second model was more likely to place endpoints. For MM5 input, both CMC and HYSPLIT are somewhat more likely than ATAD to place endpoints along the Texas-Mexico border, while ATAD is somewhat more likely to place endpoints to the south of this pathway. This is also true for HYSPLIT with EDAS/FNL input. However, CMC with EDAS/FNL input is more likely than ATAD with EDAS/FNL to have endpoints along the border. In all cases, the ATAD model is somewhat less likely than either of the other two models to place endpoints in north and northwest Texas.

### 4.2.2.2 Consequences of Using FNL Data Rather Than EDAS during October

It is assumed that the EDAS wind field is superior to the FNL field due to its finer horizontal and vertical grid resolution and better resolution in time. One method to determine possible differences caused by use of FNL vs. EDAS data during October 1999 is to examine the differences between the ORTs generated with these two input fields by a single model. ORT difference results can be examined for July–September during BRAVO and also for October during other years. Both EDAS and FNL data are currently available for five Octobers during 1997–1998 and 2000–2002. Figure 4-14 shows these ORT difference fields during July–September 1999 and Figure 4-15 shows the same for the five Octobers for which data are available. CMC was not run with FNL input, so only results from ATAD and HYSPLIT can be compared. In both figures, the areas in the warmer colors, yellow through red-black, are where the given model with EDAS input was more likely to place endpoints while the cooler colors, blue through green, show areas where FNL input resulted in more endpoints.

During the first three months of the BRAVO study, July–September 1999, with both the ATAD and HYSPLIT models EDAS input was more likely to result in endpoints along the Texas-Mexico border, while FNL input resulted in more endpoints following the same general direction, but more likely to be south of the border.

Unfortunately, for ease of interpretation, the differences are not exactly the same for the five Octobers. During the Octobers of 1997–98 and 2000–02, EDAS input with either model was more likely than FNL input to result in endpoints to the southwest of Big Bend. When the HYSPLIT model is used, endpoints were also more likely to arrive from the northwest of Big Bend, especially Arizona. This is also seen to a lesser extent in the July–September 1999 difference fields for both ATAD and HYSPLIT. FNL input in either HYSPLIT or ATAD during these five Octobers, similar to FNL input during July–September 1999, was more likely to result in endpoints arriving from the northeast area of Mexico just south of the border. During the five Octobers it was also more likely for FNL input in either model to result in more endpoints in east Texas and more endpoints in northwest Mexico than predicted using EDAS input.
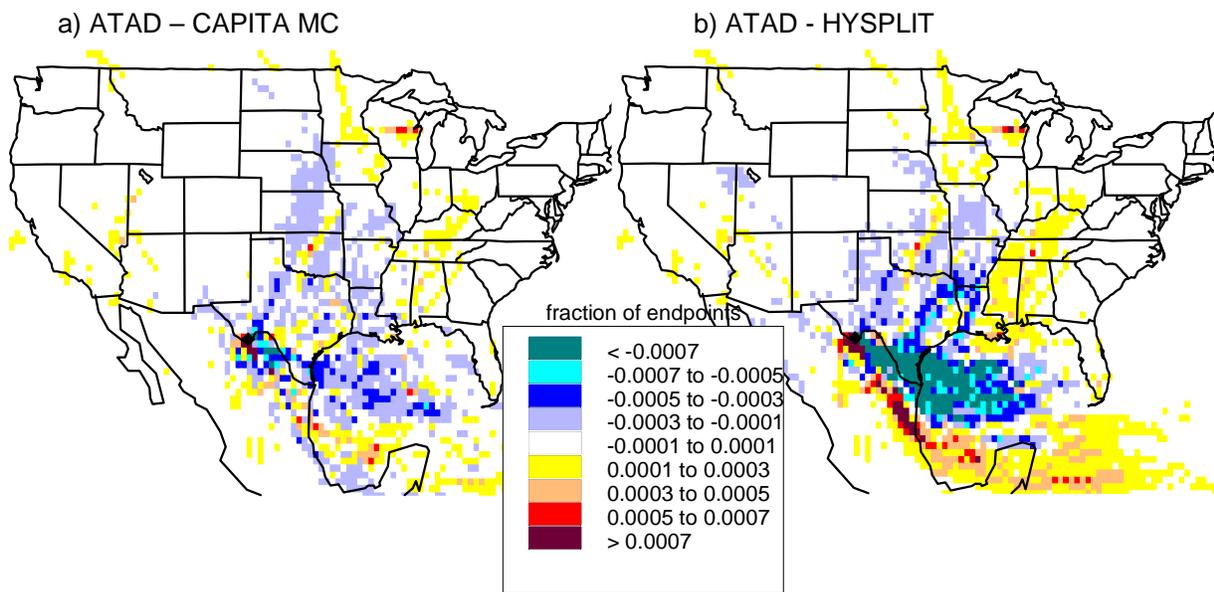
Figure 4-12. Overall Residence Time Differences for July–October 1999, grid size ½ deg lat by ½ deg lon. All with 36 km MM5 input a) ATAD-CAPITAMC b) ATAD-HYSPLIT.
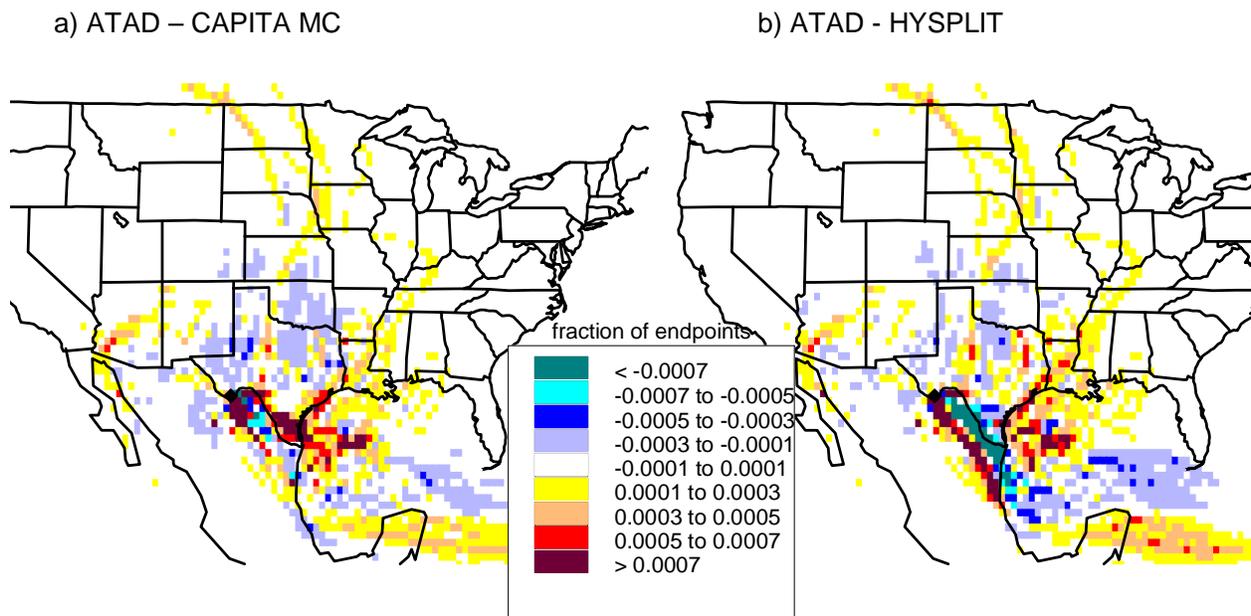
Figure 4-13. Overall Residence Time Differences for July–October 1999, grid size ½ deg lat by ½ deg lon. All with 36 km EDAS/FNL input a) ATAD-CAPITAMC b) ATAD-HYSPLIT.

Table 4-3 shows the mean endpoint heights and mean wind speeds for the combinations tested during the five Octobers. The most obvious difference is that the EDAS trajectories are substantially higher, sometimes twice as high on average, as those generated using FNL data. Like the analyses discussed in the previous section for the BRAVO period, the lower heights associated with FNL result in lower average wind speeds and so preferentially place endpoints

closer to the receptor rather than in more distant source areas. Also, consistent with the previous analyses, is that lower heights are associated with fewer endpoints in the western U.S.

Table 4-3. Mean endpoint heights and wind speeds for 5-day back trajectories for Octobers of 1997–1998 and 2000–2002 generated by all the model and input data combinations including all the start heights for HYSPLIT. A HYSPLIT start height of 100–1000 refers to an average of the results from 100, 200, 500, and 1000 m start heights.

Model	Input Data	Start Height (m AGL)	Mean Endpoint Height (m AGL)	Mean wind speed (m/sec)
ATAD	Rawinsonde	NA	1500	6.0
	EDAS	NA	1598	6.1
	FNL	NA	717	6.0
HYSPLIT	EDAS	100	651	5.6
		200	726	5.8
		500	945	6.3
		1000	1299	7.1
		100–1000	899	6.2
	FNL	100	236	5.2
		200	377	5.8
		500	742	6.9
		1000	1286	7.5
		100–1000	660	6.3

It seems likely that, on average, the meteorology during October 1999 was more similar to the average meteorological conditions during five other Octobers than during July–September of 1999. Transport directions are highly seasonal at Big Bend with airmasses tending to arrive with greater frequency from the south and southeast during spring and summer, from the southeast and northeast during the fall, and from more westerly directions during the winter. If this is true, then the results shown in Figure 4-15 are probably more similar to October 1999 than the results shown in Figure 4-14, though this is speculation and some inter-annual variability is expected.

With these caveats in mind, then it seems probable that the result of using FNL data rather than EDAS for trajectory modeling during October of the BRAVO study is that airmasses arriving from west and southwest Mexico and the southwestern U.S. are underestimated during this month, while airmasses arriving from east Texas and from Mexico just south of the Texas border are overestimated. However, it is impossible to quantify the actual differences during October 1999. Also, since most forward and back trajectory modeling was used for qualitative and quantitative source attribution for the average of the four-month study period and not for October alone, the difference in mean results for July–October 1999 due to the FNL data substitution will be smaller than for only October 1999.
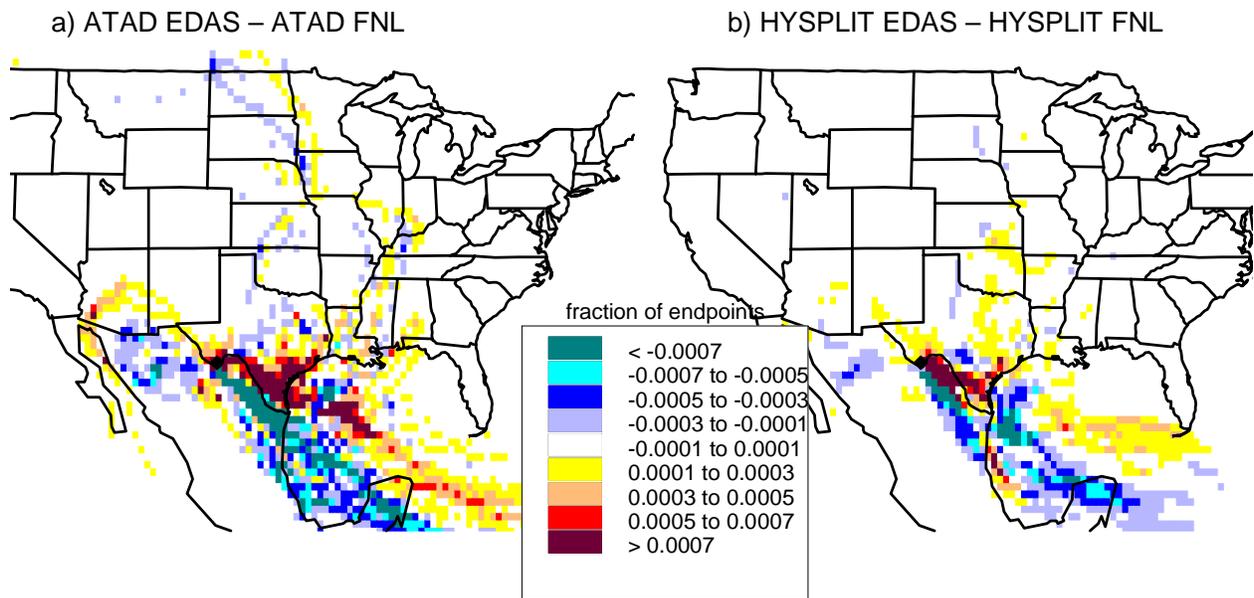
a) ATAD EDAS – ATAD FNL

b) HYSPLIT EDAS – HYSPLIT FNL

fraction of endpoints

	< -0.0007
	-0.0007 to -0.0005
	-0.0005 to -0.0003
	-0.0003 to -0.0001
	-0.0001 to 0.0001
	0.0001 to 0.0003
	0.0003 to 0.0005
	0.0005 to 0.0007
	> 0.0007

Figure 4-14. Overall Residence Time differences during July–September 1999 a) ATAD/EDAS-ATADFNL, b) HYSPLIT EDAS-HYSPLIT FNL.
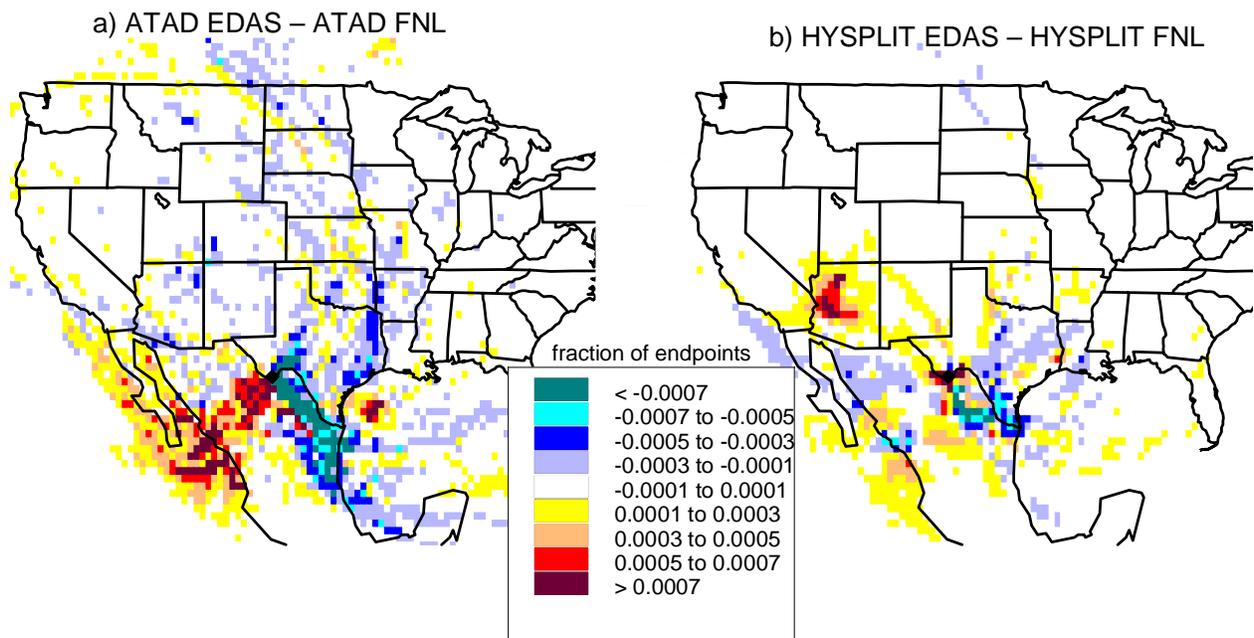
a) ATAD EDAS – ATAD FNL

b) HYSPLIT EDAS – HYSPLIT FNL

fraction of endpoints

	< -0.0007
	-0.0007 to -0.0005
	-0.0005 to -0.0003
	-0.0003 to -0.0001
	-0.0001 to 0.0001
	0.0001 to 0.0003
	0.0003 to 0.0005
	0.0005 to 0.0007
	> 0.0007

Figure 4-15. Overall Residence Time differences during Octobers 1997, 1998, 2000–2002 a) ATAD/EDAS-ATAD/FNL, b) HYSPLIT EDAS-HYSPLIT FNL.

4.2.2.3 Biases between HYSPLIT and CMC Models

To examine differences in trajectory placement due to the trajectory model rather than the input meteorological data, the differences in ORT for a single input data set with two different models are examined. Figure 4-16a shows the differences between CMC and HYSPLIT when both used EDAS/FNL data as input and Figure 4-16b shows the differences between CMC and

102

HYSPLIT when both have MM5 data as input. The patterns are similar. With both sets of input data, HYSPLIT is more likely than CMC to place endpoints along the Texas-Mexico border, while CMC is more likely to place them both southeast of this pathway in Mexico and northeast of this pathway in either east Texas (for EDAS/FNL data) or in the eastern U.S. (for MM5 data).

a) CAPITA MC EDAS/FNL – HYSPLIT EDAS/FNL   b) CAPITA MC MM5 – HYSPLIT MM5
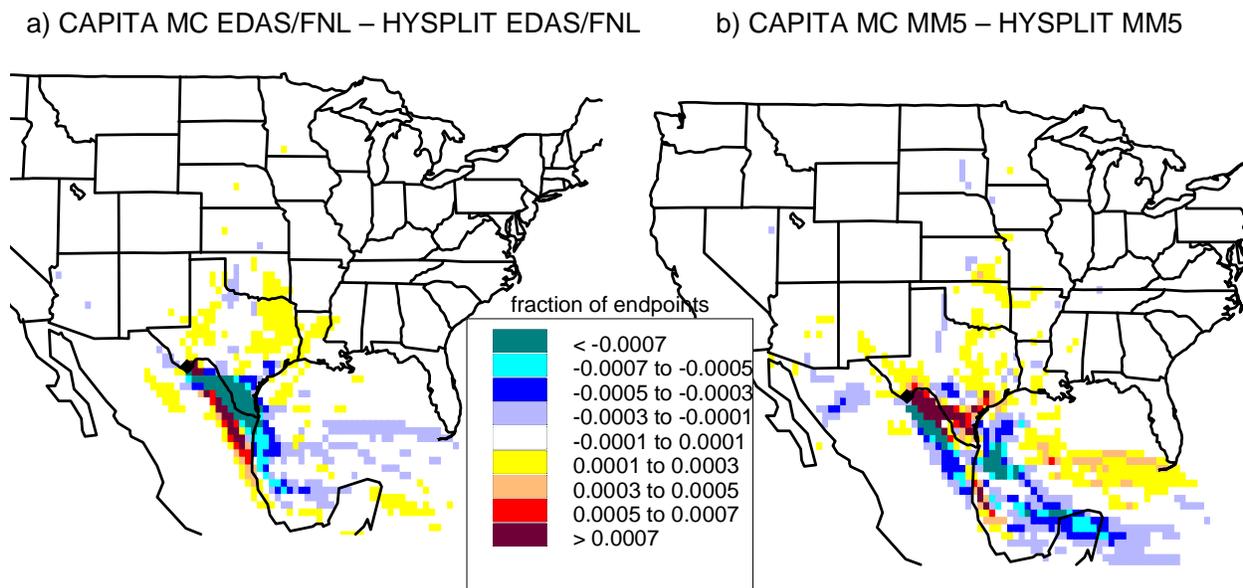
Figure 4-16. Overall Residence Time differences during July–October 1999 a) CMC with EDAS/FNL-HYSPLIT with EDAS/FNL, b) CMC with MM5-HYSPLIT with MM5.

4.2.2.4 Further Evaluation of HYSPLIT Start Heights

An issue when using the HYSPLIT model is the choice of trajectory start-height or heights at the receptor. This is not an issue for either ATAD or CMC. By default, ATAD calculates a transport layer depth and averages the horizontal winds within this transport layer (see section 2.2.2 for details of how the transport layer depth is determined). ATAD can also generate trajectories within fixed layers, but this was not how the model was run for BRAVO. In CMC, particles below the mixed layer are allowed random vertical movements within this layer, so most particles started below the mixing height will tend to follow a similar pathway (see section 2.2.4). However, HYSPLIT, which does less vertical averaging than ATAD and does not use the random vertical movements of CMC, is more sensitive to start height. This is especially true if different start heights fall within different vertical layers of the input meteorological data.

Figures 4-17 and 4-18 show the resultant five-day ORTs for each of several start heights for EDAS/FNL and MM5 inputs, respectively. The start height needs to be low enough to capture the transport of pollutants within the mixed layer, but high enough to avoid the difficulties of measuring and modeling near-surface wind flow. In each figure, the top left corner has a start height of 10 m AGL, top right is 100 m AGL, middle row 200 and 500 m AGL, and bottom row is 1000 and 2000 m AGL. The average ORTs for the four-month study period have subtle differences for subtle changes in start heights. As shown in Table 4-1, as the start heights increase the mean wind speeds increase, thus allowing endpoints to get farther from the receptor. The lowest start heights a have very small fraction of endpoints in the western U.S. and for the EDAS/FNL input, there are very few endpoints in Mexico west and southwest of Big Bend. The

103

fractions of endpoints in these areas increase as the start heights and, consequently, the average trajectory heights increase. The most different pattern is for the 2000 m start height, which for both HYSPLIT and MM5 results in the most endpoints west of Big Bend both in Mexico and in the U.S.

Because the 2000 m start height is, on average, above the mixed layer where most pollutants are transported, it was not used for source attribution of sulfate. Also, because the trajectories started at 10 m and 100 m were very similar, they were deemed redundant, and the trajectories started at 10 m were also not used for source attribution. The remaining start heights all generate trajectories that are on average within the mixed layer and it was decided to aggregate the endpoints from start heights of 100–1000 m when doing source attribution for sulfate.
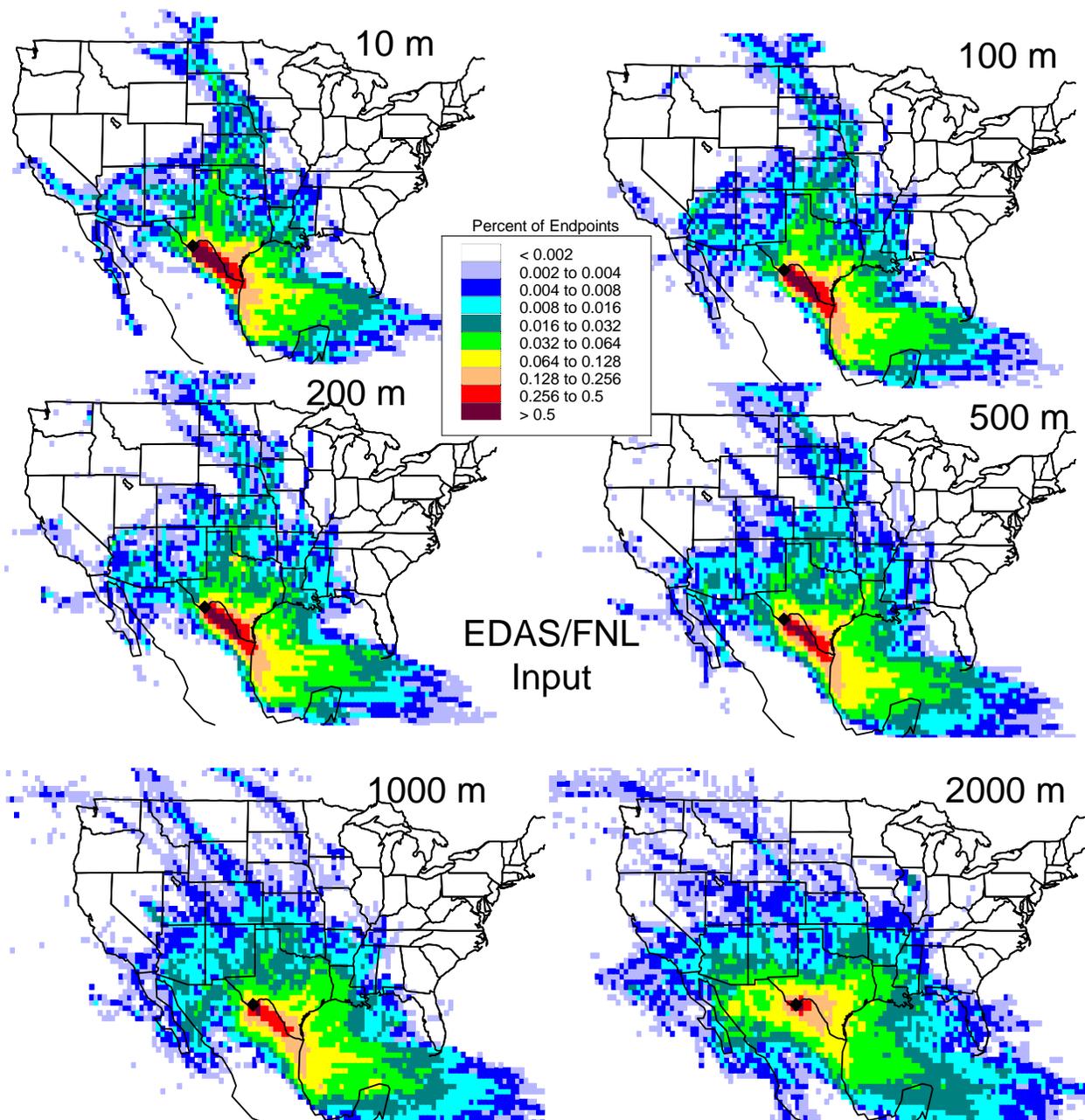
Figure 4-17. Overall Residence Times calculated with HYSPLIT and EDAS/FNL data during July–October 1999 with start heights of 10, 100, 200, 500, 1000, and 2000 m AGL.
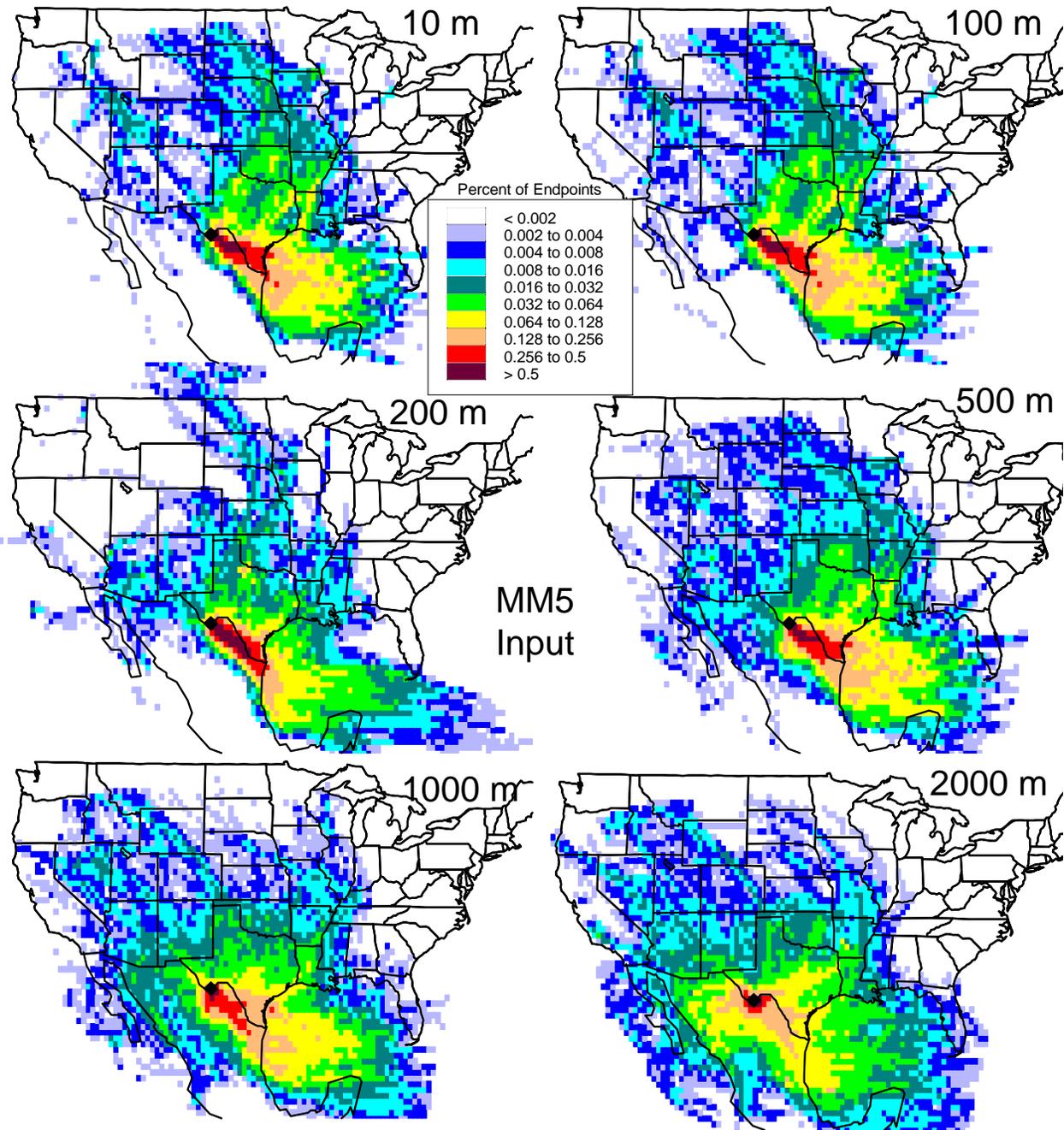
Figure 4-18. Overall Residence Times calculated with HYSPLIT and MM5 data during July–October 1999 with start heights of 10, 100, 200, 500, 1000, and 2000 m AGL.

4.3 Qualitative Evaluations Using Tracer Data

Several analyses can take advantage of the BRAVO tracer data to help determine the accuracy of the various back trajectories. Among the simplest qualitative analyses are those to determine whether the back trajectories correctly indicate a transport pathway between the known tracer-release sites and the receptor, especially for time periods when the tracer concentrations are high. An example is shown in Figure 4-19. Each row of the graph is for the day with the maximum 24-hour average measured tracer concentration for one of the four tracers

at Big Bend (K-Bar). Each column shows the trajectories generated by one of the back trajectory models. The colors indicate the input wind field: blue is MM5, red is EDAS (EDAS/FNL for CMC), green is FNL, black is rawinsonde data only, and brown is the rawinsonde data plus the BRAVO wind profilers. The black dots on each map show the tracer release site for the tracer which had the maximum concentration on that day.

Results are mixed and inconclusive. All models and wind field combinations are able to correctly locate the proper tracer release sites on the days of the highest Eagle Pass (August 22) and San Antonio (October 5) tracer concentrations. None of the combinations does well with the maximum Big Brown tracer (October 23), although CMC and HYSPLIT with MM5 come closest. For the day with the maximum Houston tracer (October 12), CMC with either input and ATAD with all inputs except MM5 correctly pass over the release site, while HYSPLIT with either FNL or MM5 input and for all start heights produces trajectories that are too far north.

A more efficient method of determining which model/wind field is best at locating the correct tracer emission sites is to examine the differences between the high concentration residence time and the overall residence times for each of the tracers. In this case, a high concentration was defined to be a 24-hour tracer concentration at the 80th percentile or higher at Big Bend (K-Bar). Peak (HRT-ORT) differences should occur over the tracer release site or at least along the transport pathway between the tracer release site and the receptor.

As was the case with simply plotting trajectories for individual high tracer days, there was not a single back trajectory model or a single wind field input that uniformly performed better for all four tracers. All combinations did reasonably well at correctly finding the Eagle Pass and San Antonio release sites and, conversely, most combinations were unable to generate peak residence time differences in the vicinity of Big Brown for the iPPCH tracer. The two model/wind field combinations which did best for Big Brown were ATAD with EDAS/FNL and ATAD with MM5 data. For the Houston tracer, any model with EDAS/FNL input did reasonably well and ATAD with MM5 performed moderately well. In general, when model/wind field combinations missed the correct tracer release site, all combinations tended to miss in the same direction. For ocPDCH (Eagle Pass) those combinations that did less well missed the correct location by being too southerly (or too clockwise). For PDCB and PTCH from San Antonio and Houston, respectively, the trajectories tended to either be correct or be too far to the northwest (or too counterclockwise.) These results are illustrated in Figures 4-20 through 4-23.

Other, more quantitative analyses involving back trajectories and tracer concentrations are discussed in section 7.1 and analyses using forward trajectories from the CMC model are discussed in section 5.2.
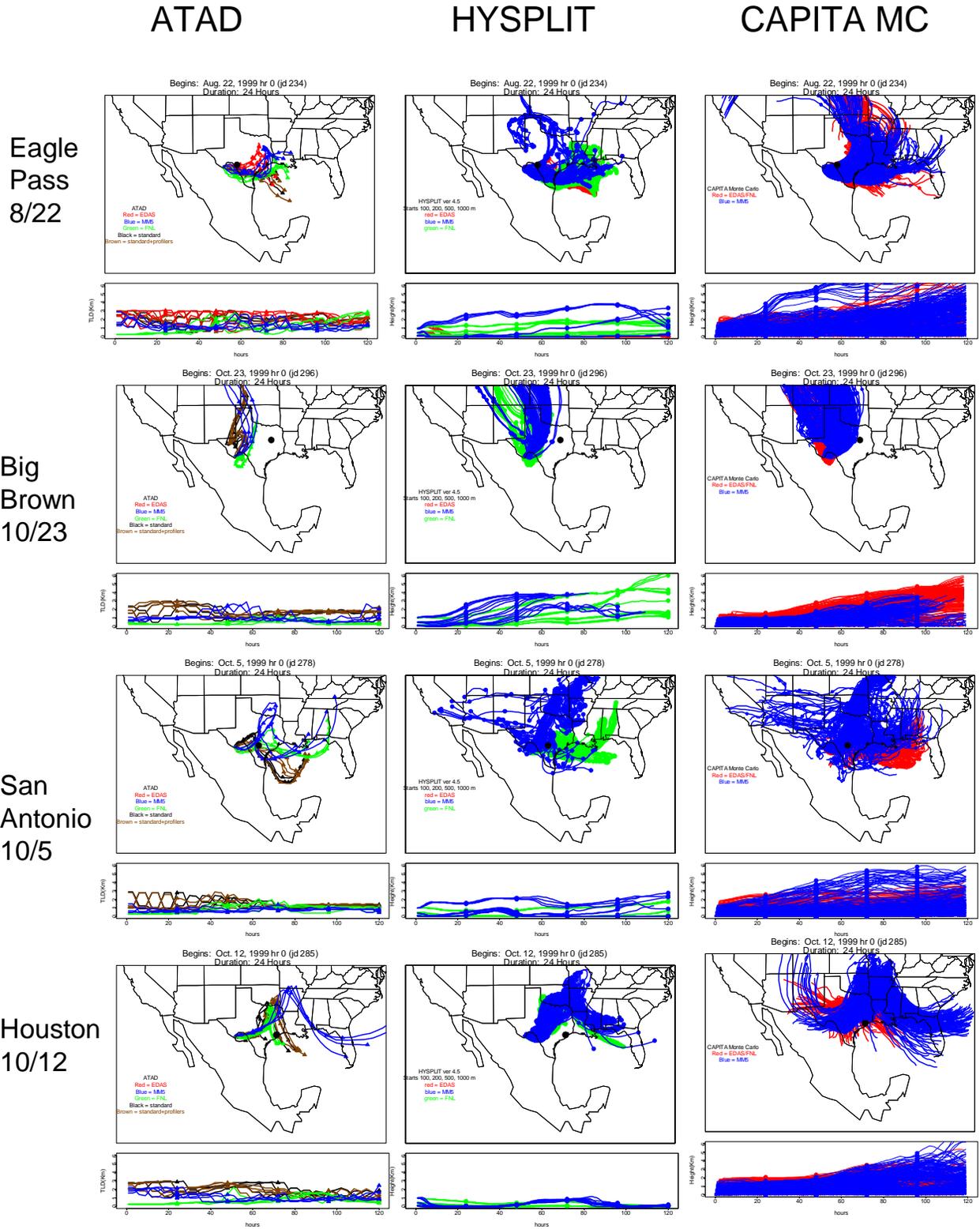
Figure 4-19. Back trajectories generated by all model and wind field combinations for the day of maximum tracer concentration at Big Bend for each tracer. The black dots indicate the tracer release sites for the tracer that had the high concentration on that day.
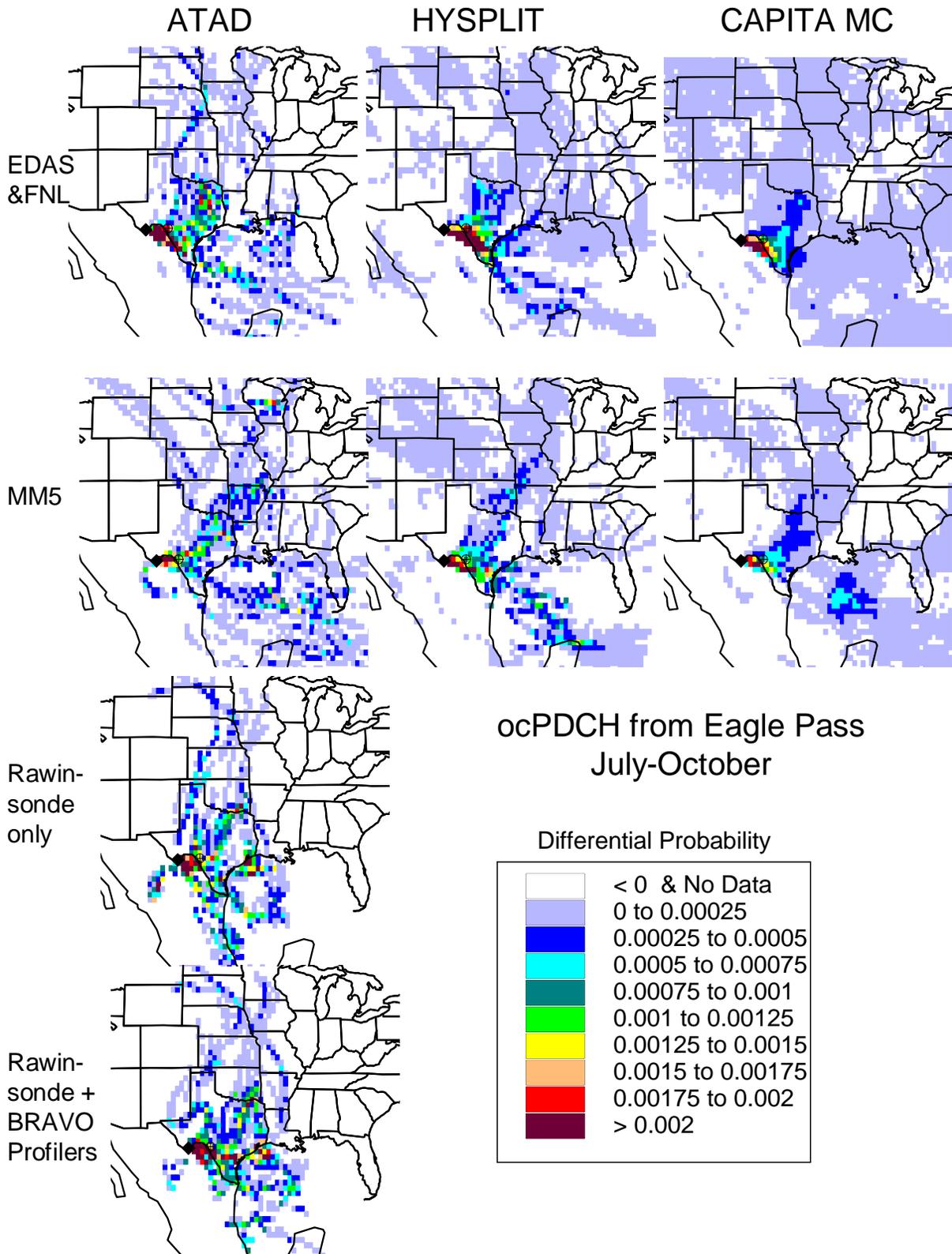
Figure 4-20. Differential Probability (HRT-ORT) for the Eagle Pass Tracer (ocPDCH). A high concentration is defined as 80th percentile or greater.
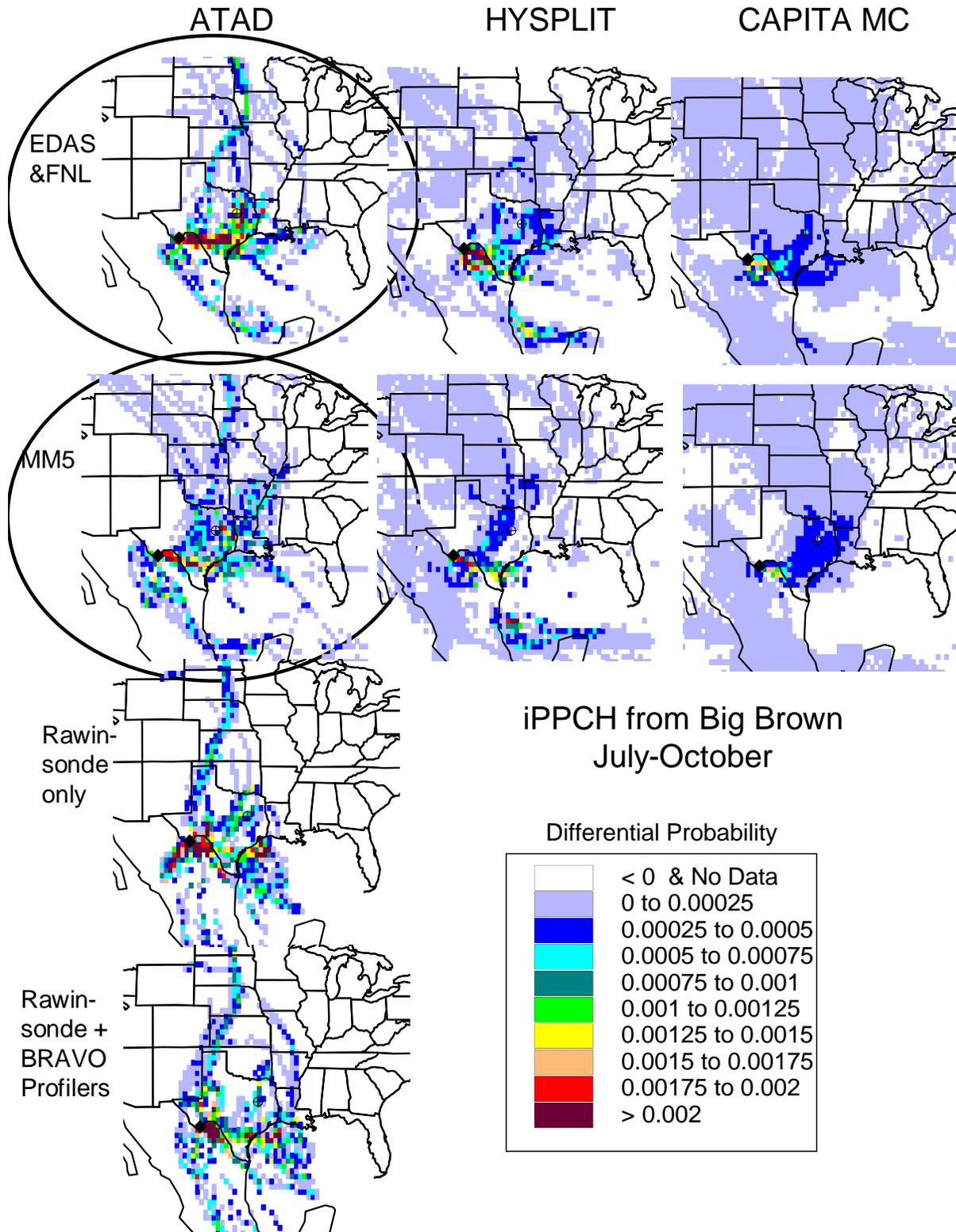
Figure 4-21. Differential Probability (HRT-ORT) for the Big Brown (iPPCH). A high concentration is defined as 80th percentile or greater. The two combinations that were subjectively determined to do the best are circled.
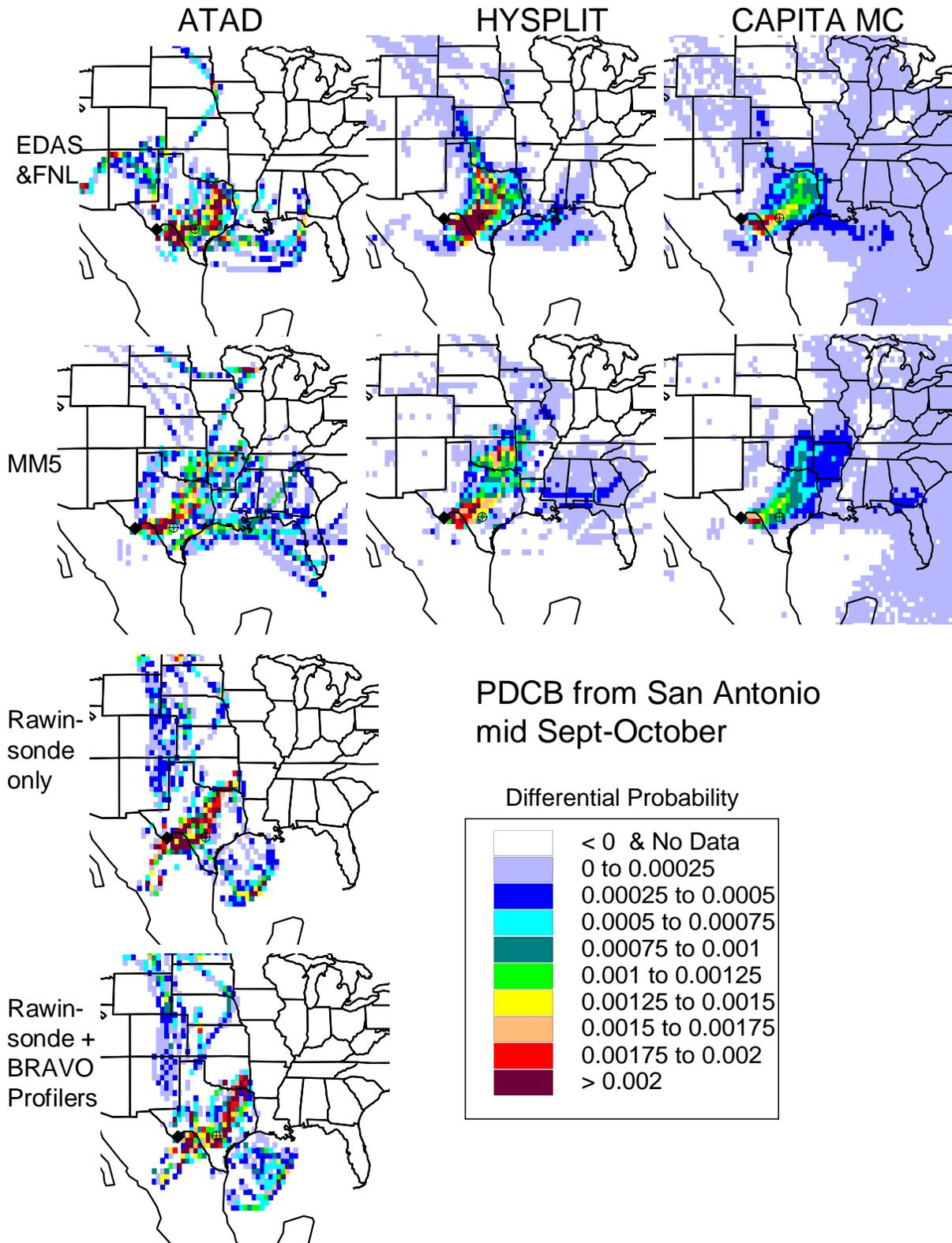
Figure 4-22. Differential Probability (HRT-ORT) for the San Antonio Tracer (PDCB). A high concentration is defined as 80th percentile or greater.
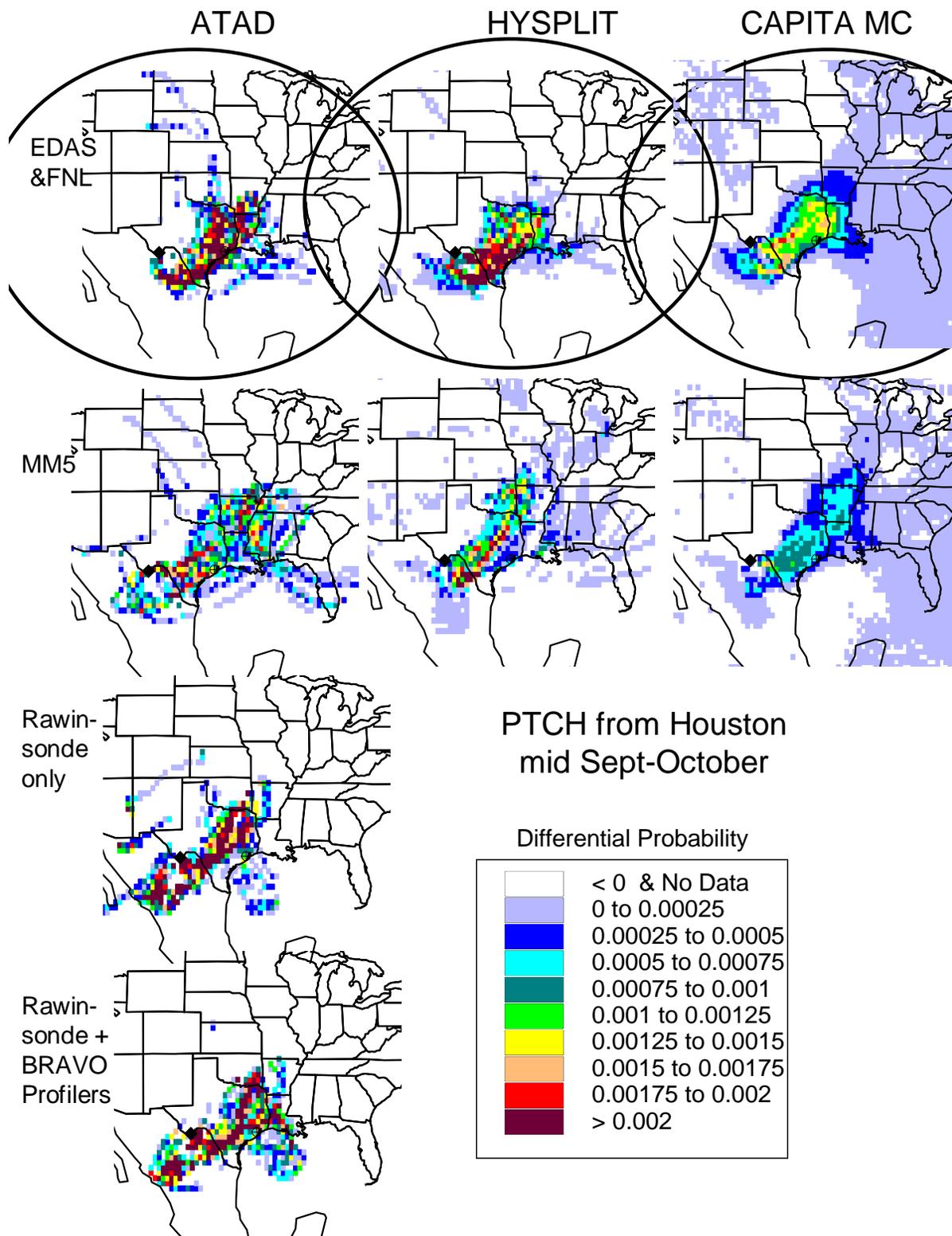
Figure 4-23. Differential Probability (HRT-ORT) for the Houston Tracer (PTCH). A high concentration is defined as 80th percentile or greater. The combinations that were subjectively determined to be doing better than the others are circled.

4.4 Biases at Big Bend during Other Seasons and Other Years

All of the analyses to this point are valid only for Big Bend during July–October 1999. Differences between the models and between the wind fields may be quite different during other years, during other seasons, and at other receptor sites. MM5 data and BRAVO wind profiler data are not readily available for other time periods. However, it is possible to easily examine differences between seasonal overall residence times as generated in previous studies with ATAD and using HYSPLIT with the most readily available wind fields, EDAS and FNL. All of these combinations can be run for six full years, 1997–2002.

Figure 4-24 shows the overall residence times for each season for these years and for each of these three combinations. Seasons are climatological, e.g., winter is December, January, and February.

For the fall season, 1999 is not included in any of the analyses to avoid having the HYSPLIT/EDAS analysis be for different years. For the summer season, the biases discussed for the BRAVO year appear to be consistent throughout the other years. ATAD is the most southerly, EDAS is the most easterly, and FNL falls between. During all other seasons, the major difference between ATAD and the HYSPLIT analyses is that HYSPLIT with either EDAS or FNL input has more trajectories arriving from the west. This indicates that previous analyses based only on ATAD trajectories may have underestimated the transport of airmasses, particularly from northwestern Mexico and Arizona.

Another simple test was to count the number of endpoints in each quadrant (NW, NE, SW, and SE) from the receptor and compare the results of each model/input data combination. Data for 1999 and 2000 were used, except fall included only 2000 to avoid the month of missing EDAS data.

Results for Big Bend are shown in Figure 4-25. ATAD in this figure refers to ATAD run with the typical rawinsonde data. EDAS and FNL refer to HYSPLIT output with these input data started at 1000 m. This is the start height that gives trajectories most similar to ATAD. HYSPLIT results in these analyses were limited to four start times per day and ATAD endpoints were interpolated to 1-hourly positions from the standard 3-hourly output. This gives the same total number of endpoints for each model/input combination. Days without full trajectories were eliminated. All models agree that the predominant wind direction is easterly during the summer (red). However, HYSPLIT/EDAS has nearly 5 times as many endpoints in the NE during the summer than ATAD, which has many more from the SE. This is consistent with all of the analyses for the BRAVO period. Predominant wind direction in the fall (yellow) is southerly with all model/input combinations showing little differences between SE and SW. However, HYSPLIT with either EDAS or FNL input has many more endpoints in the NW quadrant during the fall than does ATAD, which again shows more from the SE. During the winter (blue) and spring (green), ATAD is more likely to have trajectories arriving from the SE as opposed to HYPSLIT/EDAS with more from the SW.
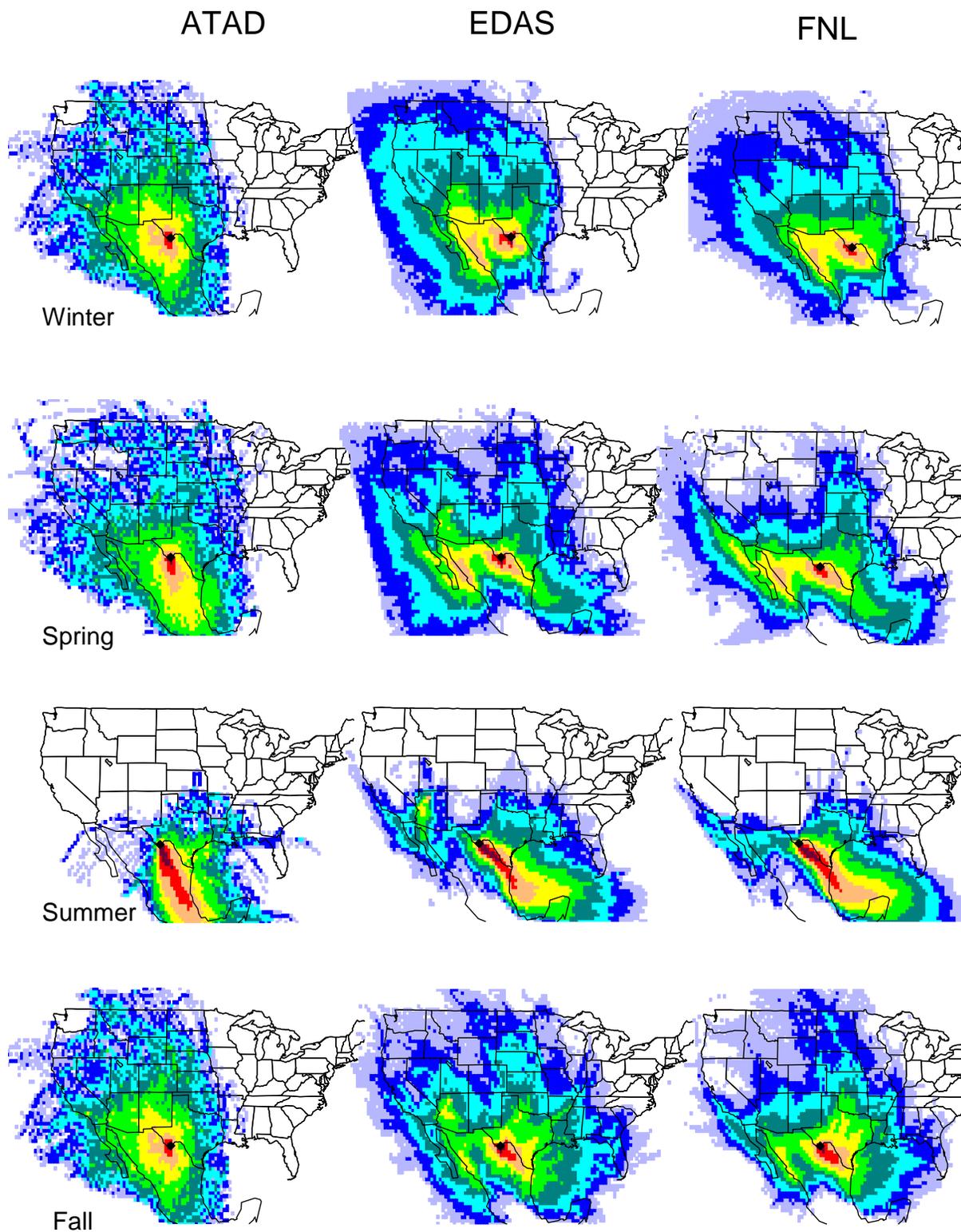
Figure 4-24. Overall Residence Times for 1997–2002 by season, and using ATAD with rawinsonde data, HYSPLIT with EDAS data, and HYSPLIT with FNL data. HYSPLIT results include start heights of 100, 200, 500, and 1000 m. Fall does not include 1999. Colors are the same as in Figure 4-18.
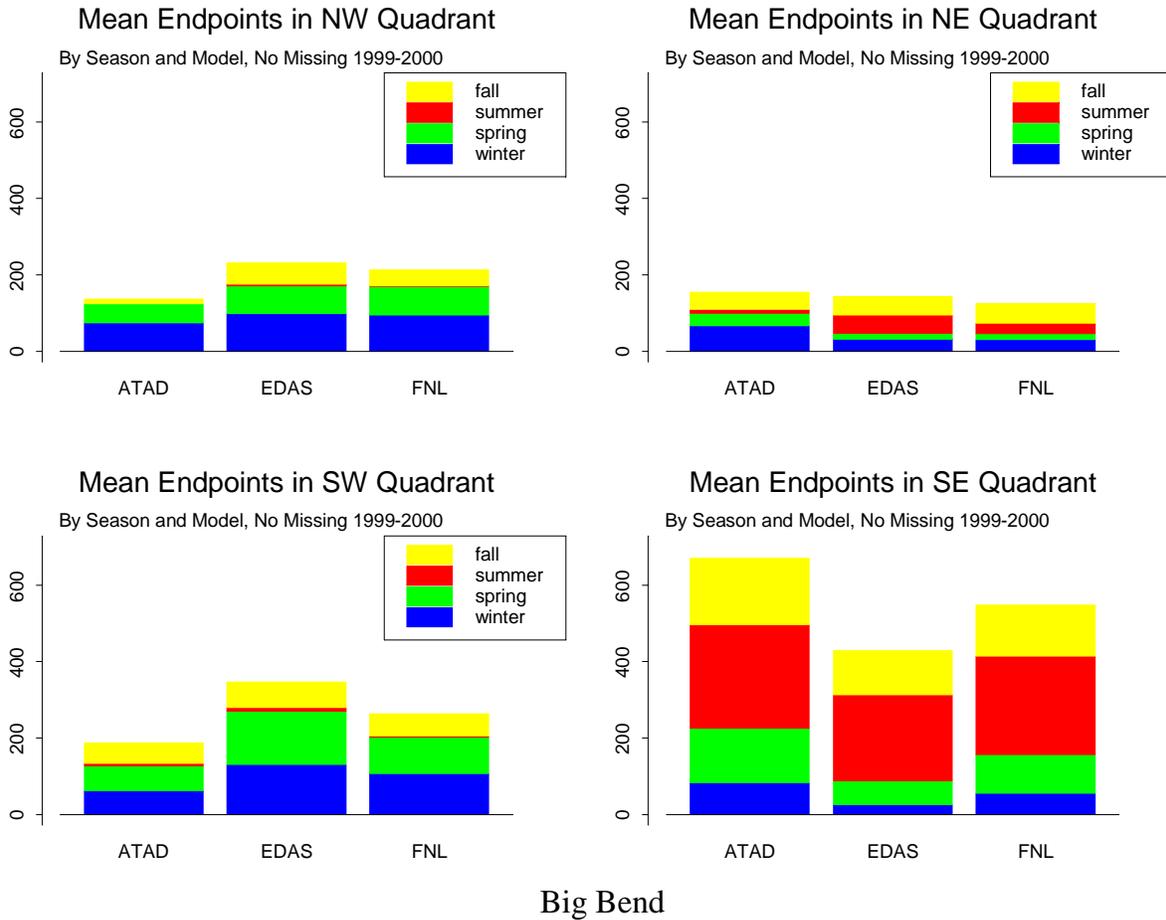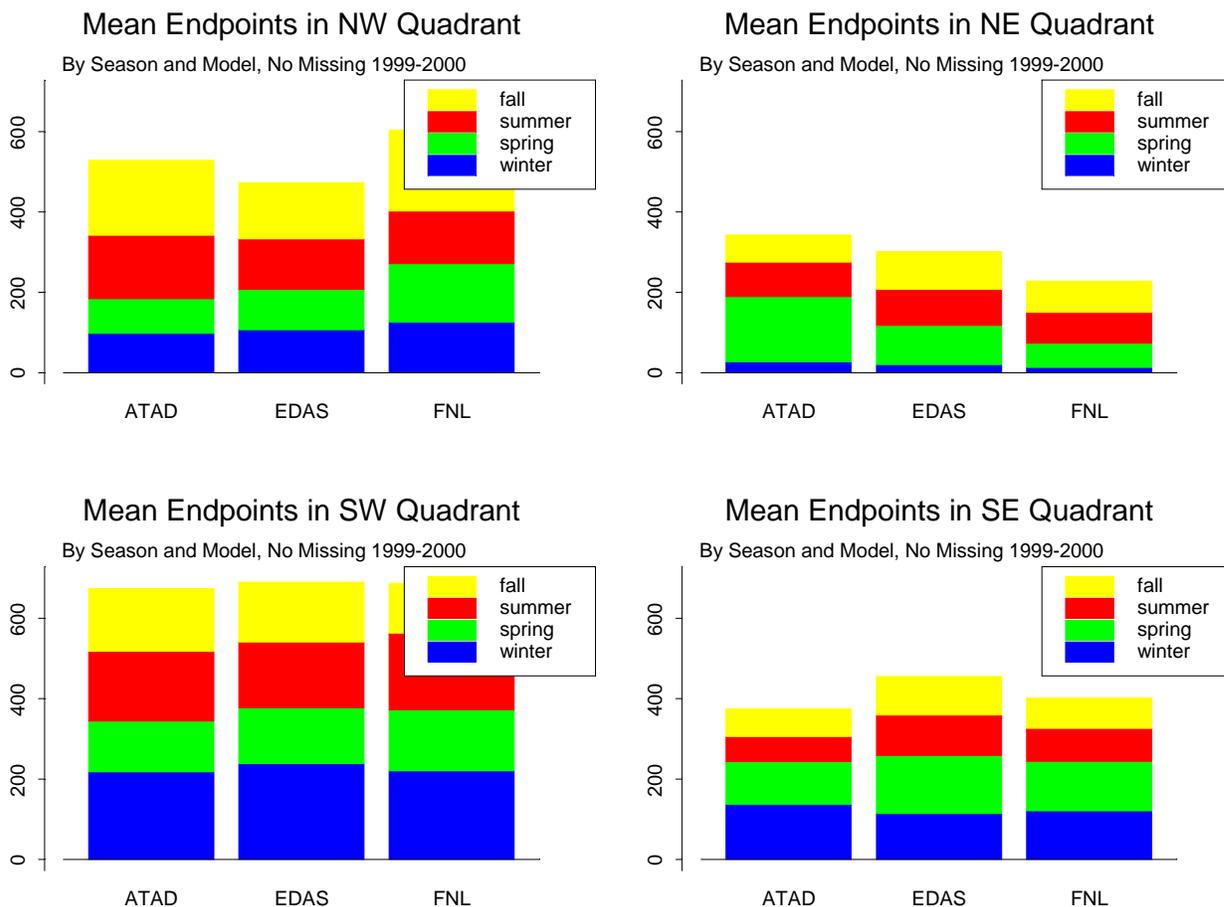
114

Figure 4-25. Mean number of back trajectory endpoints in each directional quadrant from Big Bend using ATAD with rawinsonde data and HYSPLIT with EDAS and FNL data started at 1000 m. Bar colors indicate seasons.

4.5 Biases in Other Regions during All Seasons

Comparisons between ATAD with rawinsonde input and results from HYSPLIT with both EDAS and FNL input, similar to those discussed in the previous section for Big Bend, were also compiled for Grand Canyon National Park, Arizona, and Great Smoky Mountains, Tennessee. This was to get a qualitative indication of whether similar directional differences in back trajectory output also exist for other parts of the country and whether conclusions drawn from BRAVO can be extrapolated to other sites and seasons.

Differences between model/input sets are smallest at Great Smoky Mountains, Tennessee. The largest differences at that site are in the spring (green bars) when ATAD has more trajectories arriving from the NE than does HYSPLIT/FNL, which has more from the NW. At Grand Canyon, Arizona, differences are small for all seasons for the predominant wind direction (airmasses from the SW). Largest differences are during the winter (blue bars) when ATAD is much more likely to have trajectories arriving from the NE and HYSPLIT/EDAS much more likely to show them arriving from the SE. On an annual average, HYSPLIT/EDAS has more trajectories arriving from the SE than either of the other combinations at both Grand

115

Canyon and Great Smoky Mountains. Interestingly, this is opposite of the directional bias at Big Bend where HYSPLIT/EDAS has the fewest endpoints arriving from the SE. These are shown in Figures 4-26 and 4-27. Results for these other sites will not be discussed in detail here, but in general, the results are different enough from the results for Big Bend that it seems likely that any biases between models and wind fields in other regions may be different from what was observed at Big Bend during BRAVO.

Figure 4-26. Mean number of back trajectory endpoints in each directional quadrant from Great Smoky Mountains National Park using ATAD with rawinsonde data and HYSPLIT with EDAS and FNL data started at 1000 m. Bar colors indicate seasons.
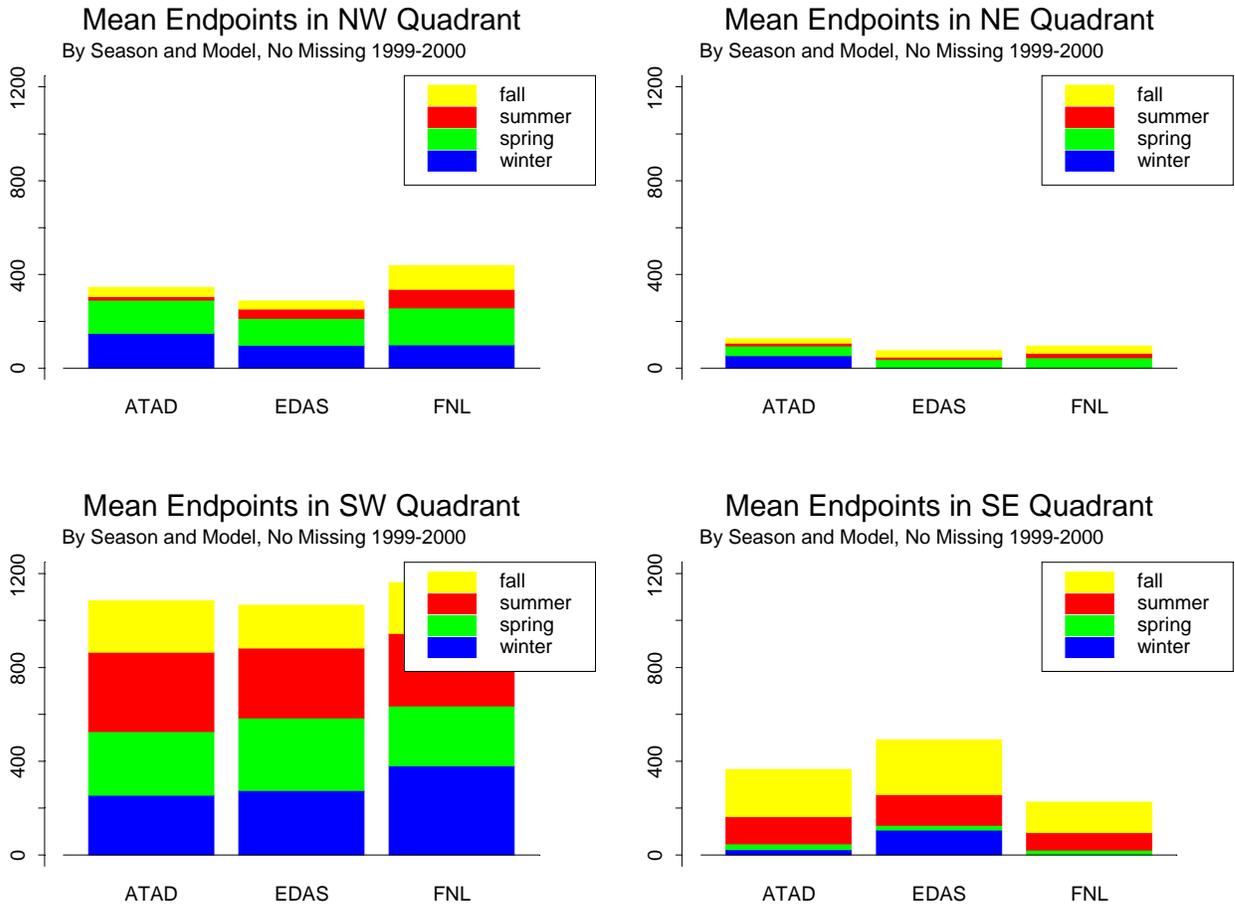
116

Figure 4-27. Mean number of back trajectory endpoints in each directional quadrant from Grand Canyon National Park using ATAD with rawinsonde data and HYSPLIT with EDAS and FNL data started at 1000 m. Bar colors indicate seasons.

4.6 Conclusions

Big Bend is a particularly challenging receptor site for which to generate back trajectories due to its proximity to both complex terrain and data-sparse areas. Analyses of back trajectories by several methods have shown that there are systematic differences between the results of different back trajectory model/input data combinations at Big Bend National Park during BRAVO.

Much of the difference appears to be due to the input wind field rather than the back trajectory model, although there are also some differences due to model alone. Using rawinsonde data alone (ATAD model only) gives back trajectories arriving from the most southerly direction, EDAS data results in the most easterly trajectories, with those generated using MM5 and FNL data falling between these extremes. These general directional biases are consistent across models and are more pronounced during summer than fall. The BRAVO tracer data clearly indicate that ATAD with rawinsonde data alone results in trajectories that are too southerly, especially during the summer. This is most likely due to the scarcity of rawinsonde data in Mexico and the absence of these data over the Gulf of Mexico. Even supplementing the

rawinsonde data with minimal additional data, e.g., information from the four BRAVO wind profilers, causes the ATAD/rawinsonde trajectories to be more similar to those generated using the gridded wind fields.

There are also some differences between results from different trajectory models even when the same input meteorological data are used. In general ATAD and CMC give results more similar to each other than to HYSPLIT. This is primarily because both ATAD and CMC, though different in mechanism, have the result of averaging the horizontal transport within the mixed layer. The predominant wind direction calculated by HYSPLIT is less southerly and more along the Texas-Mexico border than that calculated by ATAD and CMC using the same input wind fields. This is true for both EDAS/FNL and MM5 input.

HYSPLIT, which uses the gridded wind fields without the averaging of ATAD and without the random vertical movements of CMC, is more sensitive to start location, start height, and placement of individual endpoints. This is because slight changes in position can affect the grid cell for which data are used for the next time step. Slight differences in trajectory height can result in large differences in horizontal placement of trajectories in HYSPLIT during time periods when there is significant directional wind sheer with height, so that moving a trajectory endpoint from one horizontal level to the next drastically changes the wind direction. CMC can also be somewhat sensitive to height if a particle is moved from within the mixed layer to above. Differences caused by occasional bumping of a trajectory from one grid cell and/or horizontal layer to a neighboring one will tend to be minimized if trajectory data are aggregated over longer time periods, several start heights in HYSPLIT, more particles in CMC, and more start times per day. Both HYSPLIT and CMC allow trajectories to get as high as 10 km. Due to its different design, ATAD does not have trajectories above the mixed layer, with a maximum height of 3 km. An advantage of the CMC model over HYSPLIT is easy output of the height of the mixed layer.

Start heights in HYSPLIT affect the mean trajectory height and so the mean trajectory speed. Higher speeds and heights will tend to implicate more distant source areas, while lower speeds and heights give more weight to nearer sources. The lowest start heights result in the fewest endpoints in the western U.S. during BRAVO. Also for EDAS/FNL input, lower start heights have fewer endpoints in Mexico to the west and southwest of Big Bend. The fractions of endpoints in these areas increase as the start heights increase.

During October 1999, one quarter of the BRAVO time period, the EDAS winds were unavailable due to a lab fire. Evaluations of the possible consequences of having to substitute the lower-resolution FNL data during this month were investigated. The differences during October 1999 are not quantifiable, but during five other Octobers, using FNL data rather than EDAS resulted in fewer airmasses arriving from west and southwest Mexico and the southwestern United States, and more airmasses arriving from east Texas and from Mexico just south of the Texas border. During these five Octobers, HYSPLIT trajectories generated using FNL data were on average much lower than those generated with EDAS data and consequently had lower average wind speeds. The consequences of these biases will be smaller for the 4-month average of the BRAVO period than for October alone.

Cursory investigation of trajectory differences for other regions and other seasons indicate that the magnitude of the directional differences between the ATAD (rawinsonde), EDAS and FNL wind fields appears to be both geographically and seasonally dependent. So,

though there are also likely to be biases in other regions, the directions of these biases cannot be inferred from the BRAVO data.

The 36 km MM5 data and the EDAS/FNL data had approximately equal skill at predicting the tracer release locations and both data sets will be used for further back trajectory modeling. The MM5 data have a higher mean wind speed and will therefore be somewhat more likely to attribute concentrations to source areas farther away from Big Bend than will the EDAS/FNL wind fields. Further quantitative testing of the back trajectories against known attributions of both measured tracer concentrations and simulated sulfate concentrations were able to eliminate some combinations of trajectory models and input data. These tests are discussed in detail in chapter 7.