Verification of the NOAA Smoke Forecasting System: Model Sensitivity to the Injection Height

ARIEL F. STEIN*

Earth Resources and Technology, Annapolis, Maryland

GLENN D. ROLPH, ROLAND R. DRAXLER, AND BARBARA STUNDER

National Oceanic and Atmospheric Administration/Air Resources Laboratory, Silver Spring, Maryland

MARK RUMINSKI

National Oceanic and Atmospheric Administration/National Environmental Satellite, Data, and Information Service, Camp Springs, Maryland

(Manuscript received 23 May 2008, in final form 20 October 2008)

ABSTRACT

A detailed evaluation of NOAA's Smoke Forecasting System (SFS) is a fundamental part of its development and further refinement. In this work, particulate matter with a diameter less than or equal to 2.5-μm ($PM_{2.5}$) concentration levels, simulated by the SFS, have been evaluated against satellite and surface measurements. Four multiday forest fire case studies, one covering the continental United States, two in California, and one near the Georgia–Florida border, have been analyzed. The column-integrated $PM_{2.5}$ concentrations for these cases compared to the satellite measurements showed a similar or better statistical performance than the mean performance of the SFS for the period covering 1 September 2006–1 November 2007. However, near the surface, the model shows a tendency to overpredict the measured $PM_{2.5}$ concentrations in the western United States and underpredict concentrations for the Georgia–Florida case. Furthermore, a sensitivity analysis of the model response to changes in the smoke release height shows that the simulated surface and column-integrated $PM_{2.5}$ concentrations are very sensitive to variations in this parameter. Indeed, the model capability to represent the measured values is highly dependent on the accuracy of the determination of the actual injection height and in particular to whether the smoke injection actually occurred below or above the planetary boundary layer.

1. Introduction

Rolph et al. (2009) introduced the National Oceanic and Atmospheric Administration's (NOAA) Smoke Forecasting System (SFS), a smoke forecasting tool within the National Weather Service's (NWS) National Air Quality Forecast Capability (NAQFC) (Stockwell et al. 2002). The SFS, composed of the Hybrid Single-Particle Lagrangian Integrated Trajectory (HYSPLIT) dispersion model (Draxler and Hess 1997, 1998), the BlueSky framework emission processing (Larkin et al. 2009), and the Hazard Mapping System (HMS; Ruminski et al. 2006) satellite analysis, was designed to provide air quality forecasters and the public with a daily forecast of the smoke concentration over North America from fires large enough to produce visible smoke from satellite analysis.

Rolph et al. (2009) conducted a model evaluation by comparing predicted smoke levels with actual smoke detected from satellites by the HMS and the Geostationary Operational Environmental Satellite (GOES) aerosol/smoke product. A figure of merit in space (FMS), defined as the intersection over the union of the observed and calculated smoke plumes, was calculated for the 2007 fire season (September 2006–November 2007). While the highest FMS scores for individual plumes

*Current affiliation: NOAA/Air Resources Laboratory, Silver Spring, Maryland.

Corresponding author address: Ariel F. Stein, NOAA/Air Resources Laboratory, R/ARL, 1315 East–West Highway, Silver Spring, MD 20910.
E-mail: ariel.stein@noaa.gov

DOI: 10.1175/2008WAF2222166.1

approached 60%, mean values over the United States for the 1 and 5 μg m^{-3} $PM_{2.5}$ concentration contours ranged between 6% and 12% using the HMS analysis and between 8% and 15% using the GOES analysis.

In this paper, as a complement to the standard evaluation, the concentration predictions at levels near the ground are compared to the surface measurement data where available. Comparison with $PM_{2.5}$ observations across the United States is complicated by the sparse data network and the fact that the anthropogenic contribution is not modeled by this system. In this work, we present a detailed analysis of four case studies involving large forest fire events including both satellite verification and a comparison of the modeled surface levels with measured $PM_{2.5}$ concentrations at surface observing stations where the smoke made a substantial contribution to the measured concentration.

An important part in the development of the SFS involves the inclusion of the latest scientific knowledge into the system while keeping the computational burden to a minimum in order to produce a forecast in a timely manner. Therefore, there is a trade-off between the choice of model settings currently implemented in the operational SFS and those settings needed to reflect the best research and development for improving the model's performance. Determining the optimal model settings is achieved through model sensitivity analysis and comparison with surface and satellite observations. One of the key aspects controlling the model performance is its sensitivity to changes in the smoke release or injection height. Many previous studies have shown that the geographical distribution of the smoke plumes as transported by the wind strongly depends on the choice of the initial vertical distribution of the smoke source (Turquety et al. 2007; Morris et al. 2006; Wotawa and Trainer, 2000). For instance, Colarco et al. (2004) tested the sensitivity of their model to smoke injection altitude by running simulations with releases at different altitudes for a Canadian forest fire that occurred in July 2002 in which smoke was observed and modeled to pass over the Washington, D.C., area. They found that the three-dimensional distribution of smoke was very sensitive to the assumed emissions height. Leung et al. (2007) performed a sensitivity analysis of surface carbon monoxide (CO) levels to the injection altitude of boreal fire emissions and showed that the simulated concentrations at near-field sites are especially sensitive to the injection height of the emissions.

This work presents a sensitivity analysis of the model response to variations in the methodology to estimate the smoke injection height. Section 2 presents a description of the four case studies chosen to compare the model output against the measured data. In section 3 we describe the model settings used to analyze the case studies along with the surface measurements used to compare the simulations. In section 4 a sensitivity analysis of the model to the release height estimation is presented and also the model predictions are compared with satellite and surface measurements of $PM_{2.5}$. Conclusions are presented in section 5.

2. Case studies

The period 1–15 September 2006 has been chosen to analyze the HYSPLIT model's performance against satellite and surface $PM_{2.5}$ measurements and to test the simulation's sensitivity to changes in the smoke release height (case 1). During this period large fires were distributed along the northwest U.S. coast; these fires produced smoke that traveled long distances (>1000 km) and covered a very large portion of the United States and Canada. However, the large geographic extent of the smoke makes it difficult to identify individual smoke sources and to perform a more detailed analysis of their behavior. Consequently, three additional cases have been analyzed to test the model performance and sensitivity in geographical areas where individual smoke plumes are evident. The first of these three cases extends over the period 21–27 October 2007 and involves forest fires that occurred in California where the smoke produced traveled across the state and into the Pacific Ocean (case 2). The second case deals with large fires that occurred at the border between the states of Georgia and Florida during 17–21 April 2007 (case 3). Finally, the last case took place in southwestern California during the period 23–24 September 2006 (case 4).

a. Case 1: 1–15 September 2006

Fires in northwestern California, northeastern Oregon, southeastern and central Washington, Idaho, and western and southern Montana produced large amounts of smoke that started spreading locally from 1 to 4 September 2006, but merged into a larger smoke cloud that traveled eastward reaching most of the U.S. Plains, the Mississippi River valley, the Great Lakes region, and the Ohio Valley from 5 to 8 September (Fig. 1a). On 9 September the smoke traveled north from the sources and covered most of Alberta and the central Northwest Territories of Canada (Fig. 1a). From 10 to 15 September, the smoke traveled eastward passing over all of southern Canada and the north-central United States. Also, fires in Manitoba and Ontario, Canada, contributed to an increase in the amount of smoke during the last 5 days of the period under study (Fig. 1a).

b. Case 2: 21–27 October 2007

Wildfires exploded on 21 October 2007 over southwestern California under the influence of the Santa Ana winds and produced a widespread area of smoke (Fig. 1b). Fires developed and spread rapidly in Santa Barbara, Ventura, Los Angeles, Orange, southwest San Bernardino, western Riverside, and San Diego Counties, as well as across the border into northern Baja California. The smoke moved rapidly to the west and southwest and by sunset had reached about 400 km into the Pacific west of San Diego. Also, on 21 October, several areas of blowing dust were noted across the southwestern United States. A large area of dust was emanating from the region around the Salton Sea in southern California and blowing south into the Gulf of California. By sunset, this area of blowing dust covered the northern three-fourths of the Gulf of California. Areas of blowing dust also originated from the west coast of Baja California. This area of dust moved southwest into the Pacific Ocean and also mixed with smoke from the California fires.

On 22 October, a multitude of fires continued to rage across southern California and produced a large area of smoke that extended over 1200 km into the Pacific Ocean (Fig. 1b). Moreover, on 23 October, these wildfires continued to burn and the smoke associated with them could be seen moving up the coast over the Pacific Ocean (Fig. 1b). On 24 October, smoke initially spread locally near the fire sources and then eventually moved to the west out over the Pacific Ocean. Indeed, Fig. 1b shows that visible smoke could be seen extending 1600 km off the California coastline. On the other hand, on 25 October, the smoke was not nearly as dense as it had been over the previous days and the smoke began moving back inland over the entire southern half of California, into southern Nevada and western Arizona. By 26 October, many large fire hotspots could still be detected with satellite imagery (Fig. 1b) and these fires were producing very dense smoke plumes that extended due north covering all of the Santa Ana range up to south San Bernardino County. A second pocket of dense smoke was farther downstream over the central Colorado River, northwestern Arizona, the southern tip of Nevada, and the southern counties of Utah. Finally, on 27 October, although the fires were still burning, incoming clouds covered the fire source regions and, therefore, smoke could not be observed from satellite imagery (Fig. 1b).

c. Case 3: 17–21 April 2007

On 17 April 2007, two very large brush fires in northern Ware County in southeastern Georgia produced very dense smoke that moved southeast across southeastern Georgia and northeastern Florida (Fig. 1c). The smoke associated with these fires continued to move southeast across the northern sections of the western–central Bahamas. The smoke moved more south-southeast to southeast earlier in the day, affecting eastern Florida and central and eastern portions of Cuba and points farther south. On 18 April, clouds obscured the satellite image of the fire for much of the day, but smoke was noted over the Atlantic off the Georgia coast in the late afternoon. During 19 April, the smoke initially moved to the southwest, through the Florida panhandle, south along Florida's Gulf coast, and out over the Gulf of Mexico. The densest smoke was found relatively close to the fire, although dense smoke could also be seen moving to the southwest. Later in the evening of 19 April, the winds changed direction causing the smoke to shift to a more easterly direction, moving out over Georgia's Atlantic coast and over the Atlantic Ocean. Figure 1c shows that on 20 April the enormous fire in Georgia continued to produce massive amounts of very dense smoke that was fanning out to the south, covering nearly all of Florida, the Bahamas, and circulating through the Gulf of Mexico. On 21 April, the smoke extended to the southwest over the Florida panhandle and thinned as it moved along the Gulf coast toward southern Louisiana and the northwestern Gulf of Mexico (Fig. 1c).

d. Case 4: 23–24 September 2006

On 23 September, a large wildfire burned throughout the Los Padres National Forest in Ventura County in southwestern California, producing dense smoke that moved mainly south, but also extended nearly 250 km to both the east and the west of the fire location (Fig. 1d). On 24 September, the smoke moved off to the north–northwest. Some of this dense smoke spread in a thin line that stretched eastward across central California, through Arizona, and into New Mexico (Fig. 1d). This line of smoke was nearly 50 km thick and spread nearly 1600 km inland.

3. Methodology

a. HYSPLIT

To compare the model output with measurements for the four case studies, the HYSPLIT model was configured with the same settings currently used in the operational SFS run (defined as the base case or BC) and described in Rolph et al. (2009). Only the settings relevant to the sensitivity analysis, namely the release height estimation methods, are described here. The initial

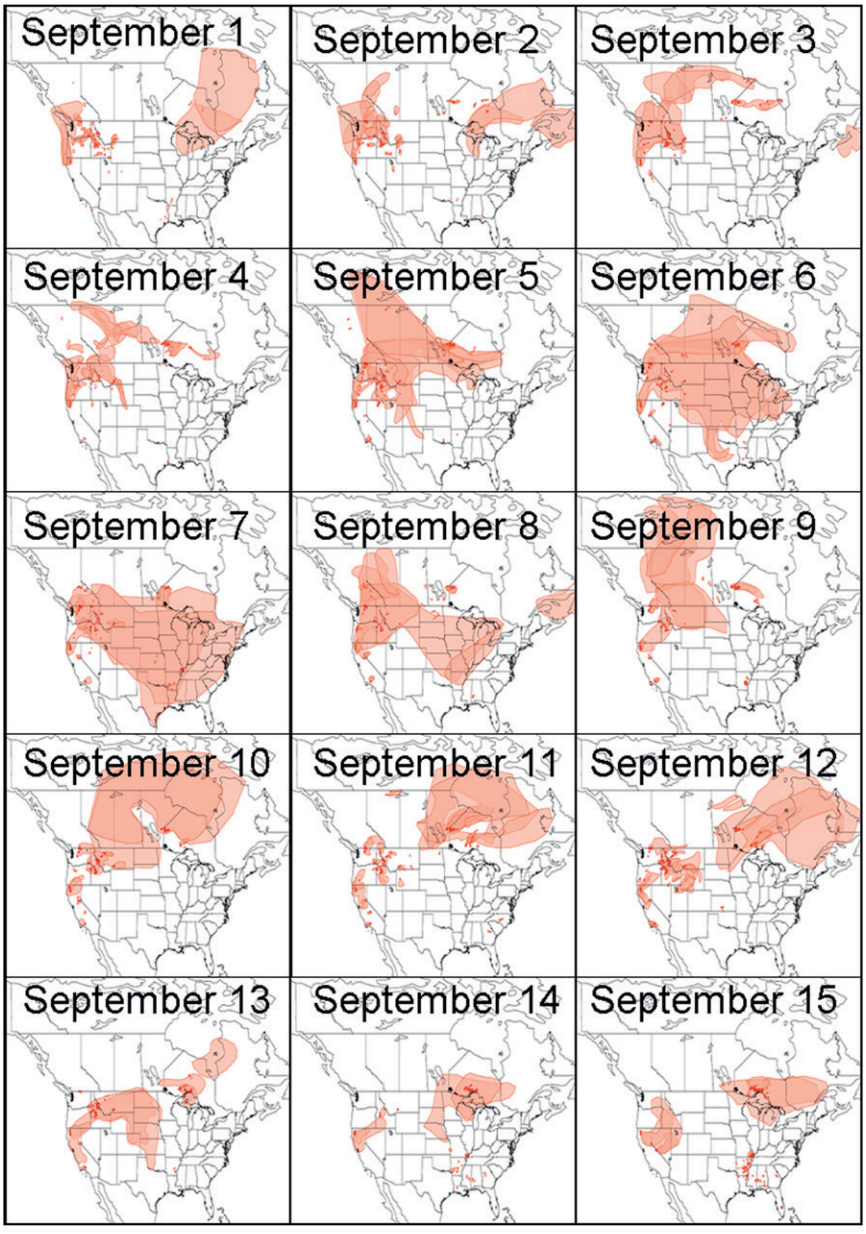


FIG. 1. Composite of HMS images corresponding to the four cases under study: (a) 1–15 Sep 2006, (b) 21–27 Oct 2007, (c) 17–21 Apr 2007, and (d) 23–24 Sep 2006.

particle height has been assumed equal to the final buoyant rise height as computed using the method of Briggs (1969) with the fire heat release estimated by the $PM_{2.5}$ emissions model (Larkin et al. 2008). In the BC run, the release height has been set with an upper limit not to exceed 2 times the mixing depth under stable or neutral conditions and up to 0.75 of the mixing depth under unstable conditions. Moreover, like in the operational SFS, two air concentration grids have been defined for each simulation. One grid consists of hourly averaged air concentrations of primary $PM_{2.5}$ from the surface up to 5 km (i.e., column integrated) for verification with satellite smoke plume observations, while the second one defines the layer in the lowest 100 m for comparison with surface measurements.

HYSPLIT only predicts the contribution of primary $PM_{2.5}$ as emitted from wildfires. Therefore, when comparing measured and modeled $PM_{2.5}$ concentrations, it

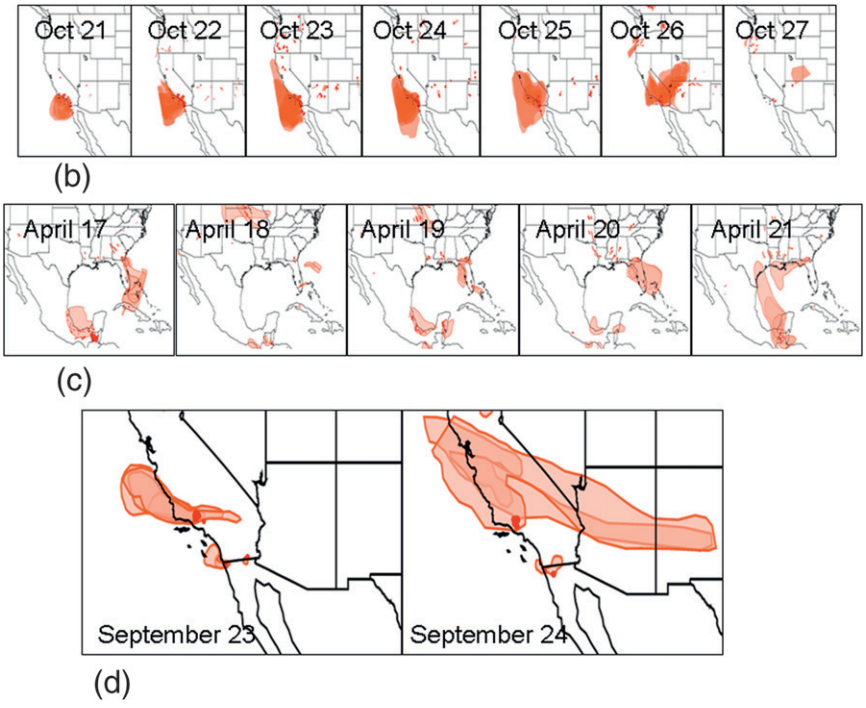


FIG. 1. (*Continued*)

is necessary to add the contributions of the secondary and anthropogenic sources when their contribution is substantial. Although the smoke from the forest fires corresponding to case 1 made an important contribution to the $PM_{2.5}$ measured over the western continental United States, the magnitude of the anthropogenic contribution to the $PM_{2.5}$ formation cannot easily be determined. Therefore, we have added to the HYSPLIT output in case 1 the $PM_{2.5}$ concentrations calculated by the Community Multiscale Air Quality (CMAQ) model (Byun and Ching 1999; Byun and Schere 2006) to represent the contribution from secondary sources of $PM_{2.5}$. The anthropogenic contribution to the other three cases was determined to be minor compared to the contribution from the wildfires and therefore no additional CMAQ simulations were considered to be necessary.

b. Measurement data

The HMS satellite-based visible smoke plumes (Ruminski et al. 2006) were compared to the model-predicted column-integrated $PM_{2.5}$ concentrations. A detailed description of the HMS can be found elsewhere (Rolph et al. 2009). For case 1, surface measurements of $PM_{2.5}$ at 164 stations west of 95°W were obtained from the AIRNow database (information online at http://www.epa.gov/airnow/) for comparison with the model calculations. For the remaining cases, four stations in the state of Florida (Fig. 2a) were chosen for case 3 and 21 surface measurement sites located within California were selected for cases 2 and 4. All of these $PM_{2.5}$ measurements were obtained from the AIRNow-Tech database (information online at http://www.airnowtech.org).

4. Results and discussion

a. HMS versus HYSPLIT

Modeled column-integrated $PM_{2.5}$ concentrations were compared against HMS data using the methodology described in Rolph et al. (2009). The daily mean FMS (Mosca et al. 1989; Boybeyi et al. 2001) scores corresponding to the 5 $\mu g\ m^{-3}$ smoke concentration contour was chosen as the statistical measure of model performance because it gave the best overall FMS score of any contour, suggesting that it most closely represented the value at the edge of the HMS plume extent visible to the satellite analyst.

Figure 3a shows the daily mean FMS scores for case 1. Daily mean FMS scores generally exceed 5% with some days reaching levels above the long-term average of 11.6% calculated for the period 1 September 2006–1 November 2007 [the full study period in Rolph et al. (2009)]. Good model performance was observed on 2–4, 8–9, and 13–15 September 2006. During these days the model was able to capture the geographical location and extent of the 5 $\mu g\ m^{-3}$ concentration contour.

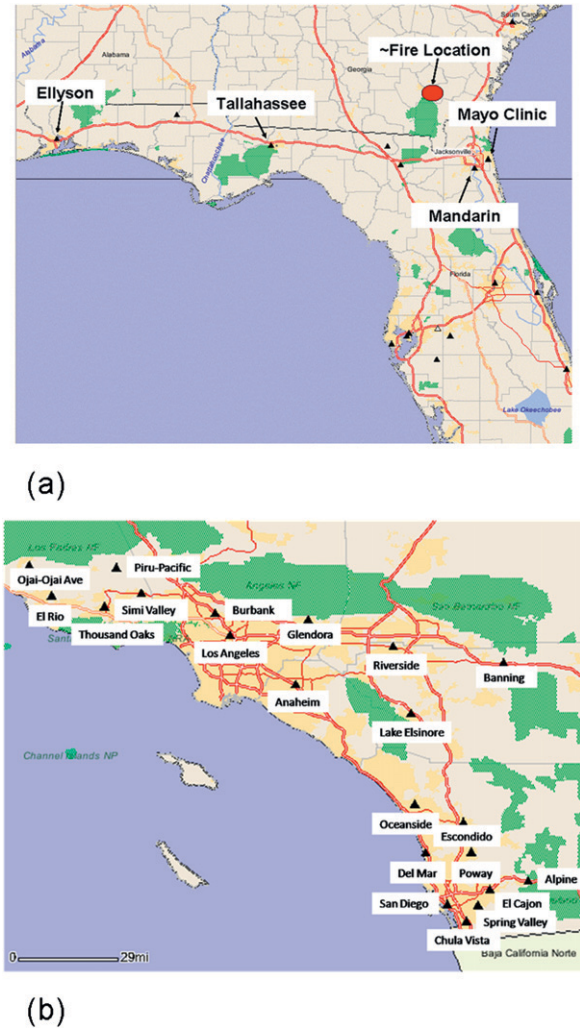


FIG. 2. Geographical distribution of surface measurement stations in (a) FL and (b) southern CA.

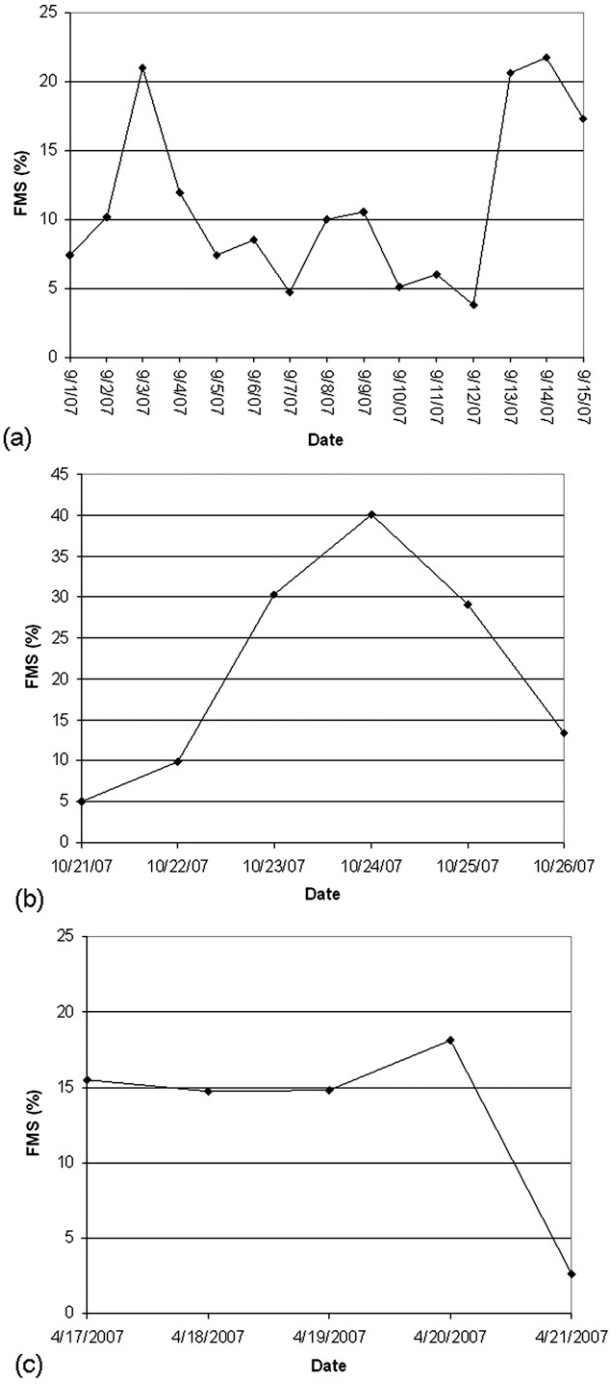


FIG. 3. Daily mean FMS scores corresponding to the 5 μg m^{-3} smoke concentration contours for the periods (a) 1–15 Sep 2006, (b) 21–27 Oct 2007, and (c) 17–21 Apr 2007.

Nevertheless, on 1 September 2006 the low FMS score is mainly attributed to the inability of the simulation to capture the smoke across parts of British Columbia and over the Ontario and Quebec provinces of Canada. From 5 to 7 September the model did not capture the smoke that was observed to be transported through the central and eastern United States. Also, during this period the simulated smoke column showed high concentrations over the Pacific Ocean west of California where no smoke was detected by satellite. Another degradation of the model performance was observed from 10 to 12 September. During this period HYSPLIT was unable to capture the smoke located over Hudson Bay in Canada. In addition, the simulation for the 5 μg m^{-3} concentration contour over the northwestern United States covered a larger area than the one reported in the HMS analysis.

To determine if these results are statistically different from those observed during the full study period, we applied the Kolmogorov–Smirnov statistical test (Wilks 1995, p. 133) between the FMS values for case 1 and the FMS scores calculated for the full study period. This

nonparametric test seeks differences between two datasets independent of the nature of the underlying distribution. We reject the null hypothesis of no difference between the two datasets if the p value is lower than 0.05. The calculated p value of 0.57 indicates that the dataset corresponding to case 1 is not significantly different from the one covering the full study period. Therefore, the 15-day episode under study in case 1 can be considered representative of the full study period.

On the other hand, during the California fires of 22–27 October 2007 (case 2) the FMS values reached as high as 40.1% (Fig. 3b). This corresponds to the second highest value calculated over the entire study period. On 22 October, however, the extent of the simulated smoke contour covered only a portion of that analyzed by inspection of the satellite imagery, explaining the poor comparison for that particular day. Nevertheless, for the rest of the case study the model exhibited excellent agreement with the HMS analyses.

Moreover, Fig. 3c shows that, for the most part, the April 2007 case features FMS scores higher than the 11.6% average calculated for the full study period, with the exception of 21 April 2007, when the predicted smoke contour extended over a wider area than the one derived from satellite imagery, explaining the low FMS on that day.

Finally, during case 4, the FMS on 23 September 2006 was 22.5%, but dropped significantly to 1.7% on 24 September 2006. This degradation in the comparison can be attributed to the model failing to predict the line of smoke extending from southwestern California to the east as far as Arizona and New Mexico.

b. Surface measurements versus HYSPLIT

Figure 4a compares measured $PM_{2.5}$ concentrations with those calculated by HYSPLIT and CMAQ averaged over all of the 164 monitoring stations west of 95°W obtained from the AIRNow database for each hour during case 1. The models tend to capture the gross features of the day-to-day relative variation in the $PM_{2.5}$ levels, with increasing concentrations from 1 to 8 September, a sharp decrease from 9 to 10 September, a new buildup on 11 September with a peak on 12 September, and a decrease toward the end of the period. However, the simulations tend to overpredict the modest measured diurnal variation of about 5 $\mu g\ m^{-3}$: nighttime $PM_{2.5}$ surface concentrations are overpredicted and daytime levels are underpredicted as compared with surface monitoring measurements. This is readily apparent in Fig. 4b, which shows the average difference between the measurements and the model results. Moreover, HYSPLIT, as well as CMAQ, predict that $PM_{2.5}$ concentrations peak generally around 1300 UTC, while the measurements display two maximums around 0300 and 1700 UTC. Several modeling studies attempting to estimate $PM_{2.5}$ concentrations have shown this nighttime overprediction and the misplacement of the concentration maximum (Mathur et al. 2008; McKeen et al. 2007). These discrepancies have been attributed to a variety of causes including errors in emissions inventories, dry deposition, and biases in the estimation of the planetary boundary layer behavior during sunrise and sunset.

In case 2, the HYSPLIT model also showed a tendency to overpredict the $PM_{2.5}$ concentrations when compared with surface measurements. Figures 4c and 4d show that HYSPLIT grossly overpredicted concentrations from 23 to 25 October, but with a noticeable improvement toward the end of the period. Curiously, the model underpredicted the $PM_{2.5}$ concentrations from 21 to 22 October. This underprediction was also associated with high PM_{10} (particulate matter with diameters $\leq 10\ \mu m$) levels at 3 out of the 21 stations. Indeed, satellite measurements have indicated that these high measurements were associated with active dust sources during this period. Smoke and blowing dust have a very similar appearance in GOES satellite visible imagery, which is the primary tool used for smoke identification (Ruminski et al. 2007). There are three primary means that have been employed to distinguish between the two. First, the 3.9-μm imagery has been probed for a hotspot at the source of the plume. Therefore, if one is present, it would be an indication that the plume corresponds to smoke. Second, the land type at the source of the plume is also analyzed. If the plume is emanating from a desert or scrub-type vegetation, it is likely indicative of blowing dust in the absence of a hotspot. Third, blowing dust has a higher spectral response than smoke (which typically has no response) in the 11-μm imagery.

To test the hypothesis that dust sources contributed to the enhancement of the PM levels, HYSPLIT was setup to run using its dust module (Escudero et al. 2006; Draxler et al. 2001). Dust source locations were obtained from satellite observations and a 20 km^2 emission area was assumed for each source. HYSPLIT calculates PM_{10} emissions based on a threshold friction velocity (see Escudero et al. 2006 for further details). In only one of the three stations where the underestimation occurred did the model predict significant levels of PM due to dust sources. Figure 5 shows a comparison of the observed PM concentrations and model estimations at Lake Elsinore, California. Figure 5 illustrates that HYSPLIT captures the largest increase in PM on 22 October associated with dust sources and the PM associated with smoke for the rest of the period under study.

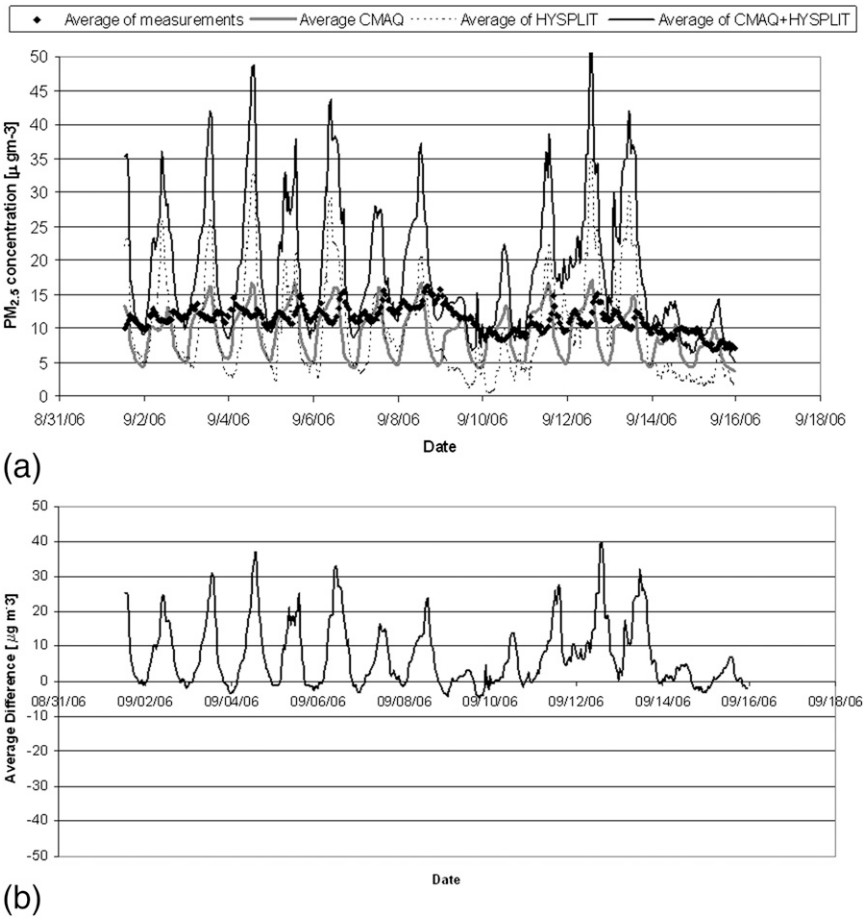


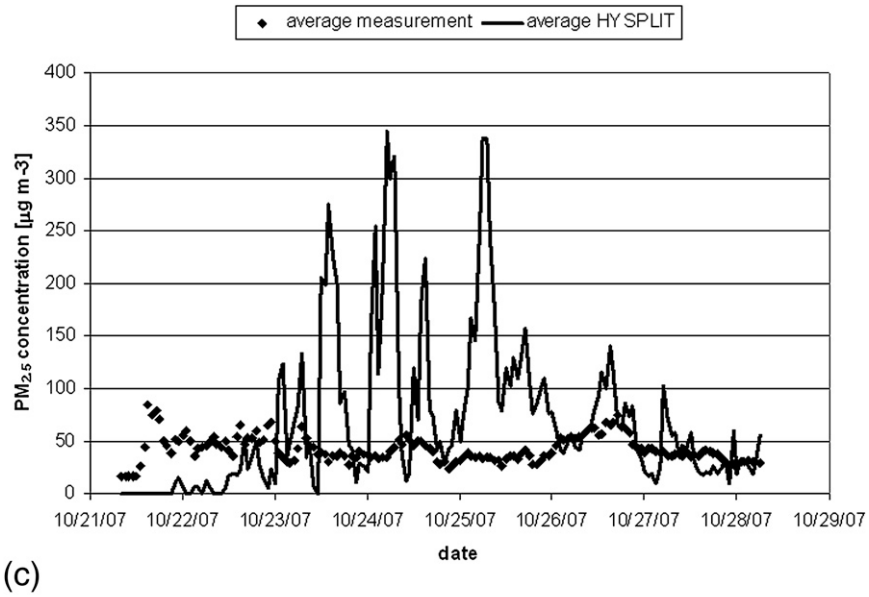
FIG. 4. Comparison of modeled and measured hourly averaged $PM_{2.5}$ surface concentrations for all stations: (a) 1–15 Sep 2006, (c) 21–27 Oct 2007, and (e) 17–21 Apr 2007. (b),(d),(f) As in (a),(c),(e) but for the average of the difference between modeled and measured $PM_{2.5}$ surface levels. All hours correspond to UTC.

Contrary to the former two case studies, the model tended to underpredict the $PM_{2.5}$ concentrations for the Georgia–Florida fires in case 3 (Figs. 4e and 4f). However, the magnitude of this underprediction was not as significant as that corresponding to the overprediction for cases 1 and 2. Furthermore, the model captured the onset and relative magnitude of the $PM_{2.5}$ levels during the entire case study. For this case, the discussion of the transport of the smoke plume helps in interpreting the measurements and the model results (see section 2c). For instance, on 17–18 April smoke was transported to the southeast of the fire to stations at Mandarin and Mayo Clinic, near Jacksonville, Florida. The remaining two stations, namely Tallahassee and Ellyson, Florida, were influenced by this fire on 21 April when the smoke plume traveled directly toward the southwest.

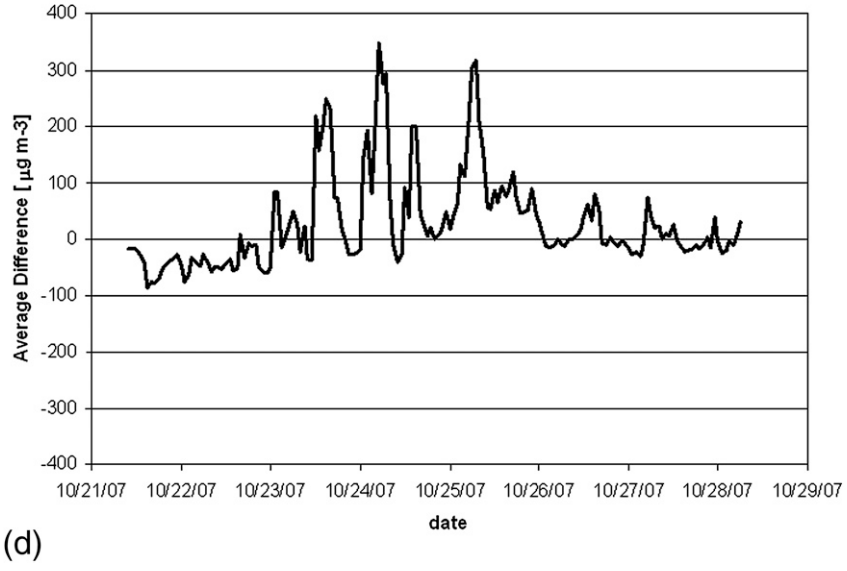
In case 4, no apparent effect from the large wildfires was noted in the surface station measurements. However, the HYSPLIT Base case run predicted a noticeable smoke contribution in 3 out of the 21 stations (i.e., El Rio, Piru, and Simi Valley, California). This discrepancy can be attributed to errors in the model estimation of the smoke release height. Some of the simulated smoke was emitted below the boundary layer, thus making an important contribution to the surface concentrations. A more comprehensive explanation is presented in the next section where the effect of the injection height of the fires is analyzed.

c. *Sensitivity to release height*

Case 1 has been chosen for further testing of a wider range of sensitivity tests of the model release height over the other three cases due to its representativeness of the full study period (September 2006–October 2007) used for the recent evaluation of the operational SFS (Rolph et al. 2009). Because the objective of this work is to evaluate the sensitivity of the current operational SFS



(c)



(d)

FIG. 4. (*Continued*)

to changes in the injection height, we will focus our analysis of the results on the FMS resulting from the comparison of the model output and the HMS product, which is used for verification of the operational product. A comparison of model-predicted $PM_{2.5}$ levels with measurements at surface stations has been performed only for those sensitivity cases where noticeable statistical improvements in the FMS with respect to the BC have been found. Accordingly, the model–measurements comparisons for the other three case studies have also been focused only on sensitivity cases where the model showed better performance when compared to the operational SFS.

1) FIXED HEIGHT

A series of five simulations were performed to study the model sensitivity to changes in the injection altitude. In each run, Lagrangian particles were released at an initial height of 100, 750, 1500, 3000, and 5000 m, respectively. Figure 6 shows differences larger than 5% in the daily mean FMS scores observed between each sensitivity case and the BC run with no particular sensitivity run consistently outperforming the rest for the entire period 1–15 September 2006. However, for the period covering 1–5 September, the BC and the 100-, 750-, and 1500-m runs show somewhat better daily

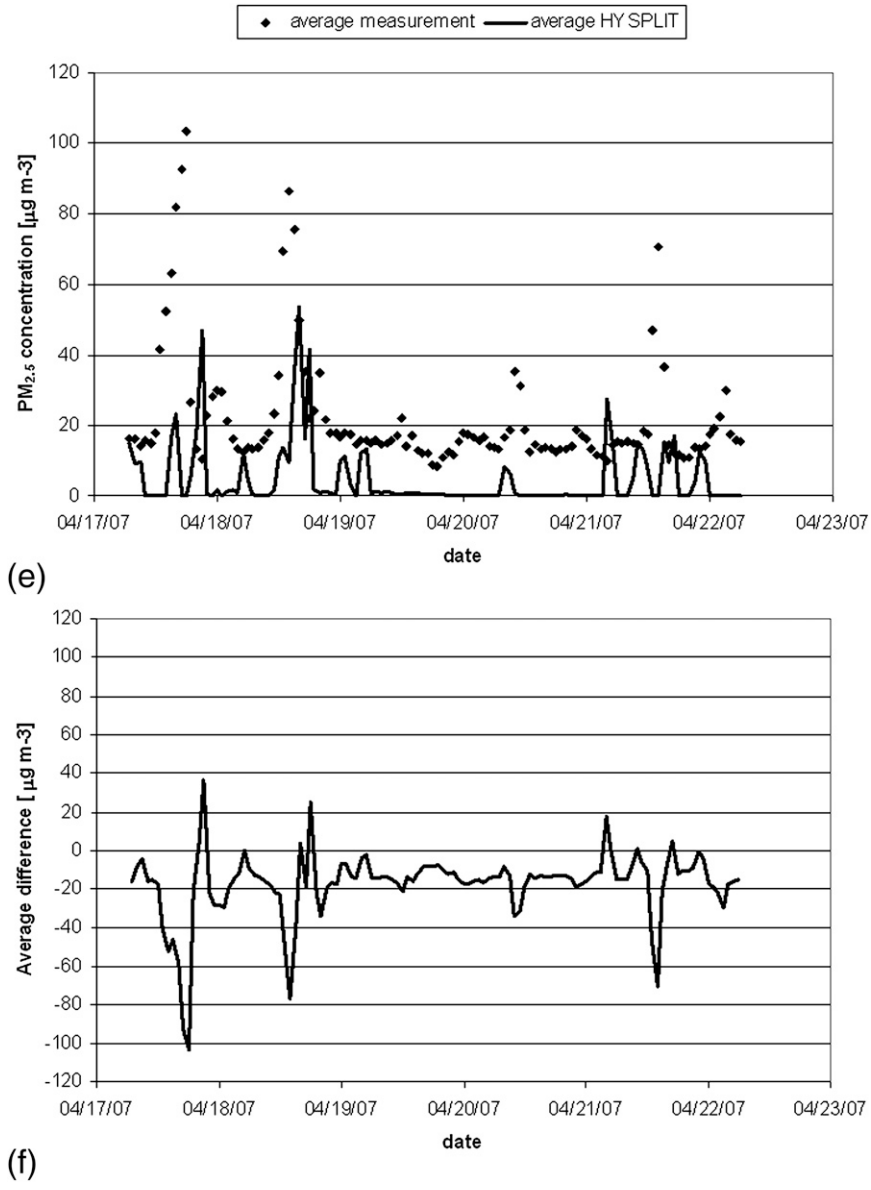


FIG. 4. (*Continued*)

mean FMS scores than the 3000- and 5000-m runs, except for 3 September, in which all runs with release heights less than or equal to 3000 m outperformed the BC run. On 6–7 September, the 5000-m run outperformed the rest of the simulations. Between 8 and 12 September, no important differences were observed among the runs. However, between 13 and 15 September, the 3000-m run showed the best performance.

2) VARIABLE RELEASE HEIGHT

Additional simulations were performed to test the sensitivity of the modeled $PM_{2.5}$ levels to a variable smoke injection height. As described in section 3a, the smoke release height is assumed to be equal to the final buoyant plume rise height as computed using Briggs (1969), implying that the final rise is a function of the estimated fire heat release rate, the atmospheric stability, and the wind speed. In the BC run, the injection height is set not to exceed 2 times the planetary boundary layer (PBL) depth under stable or neutral conditions and up to 0.75 of the PBL height under unstable conditions. With these limits, the Lagrangian particles are always released below the PBL under unstable conditions. Consequently, three sensitivity runs were defined to allow the injection of particles above the PBL if the estimated final plume rise was higher

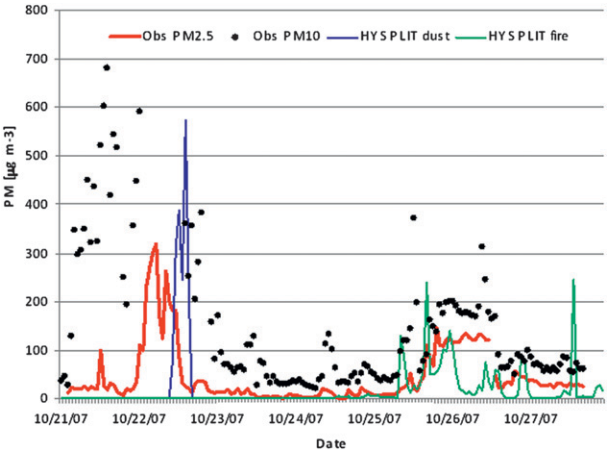


FIG. 5. Comparison of observed and modeled surface PM_{10} and $PM_{2.5}$ concentrations at Lake Elsinore, CA.

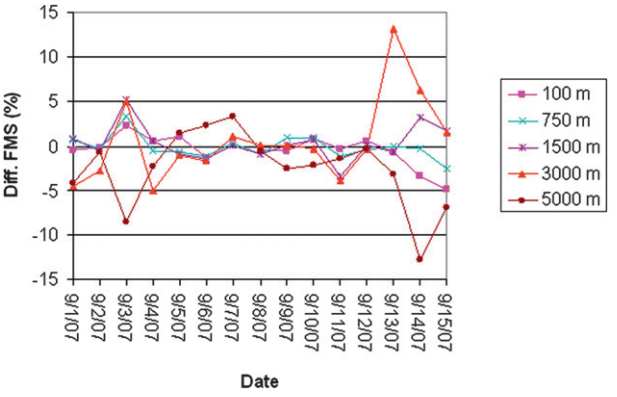


FIG. 6. Difference in daily average FMS scores between the sensitivity case and BC runs corresponding to the fixed injection height sensitivity analysis. Positive values indicate that the FMS sensitivity case performed better than the FMS BC.

than the PBL height, with an upper limit of 3000 m above ground level. This upper limit has been found to be a typical estimate of the top of the smoke release height for most of the low-altitude injections found in Canada and the United States (Colarco et al. 2004; Wotawa and Trainer 2000). In the first sensitivity run (VAR1) the heat release from the fire was used to estimate the release height (i.e., same settings as the BC without the PBL limits) and the particles could be released up to 3000 m. Run VAR2 featured two release heights: one fixed at 100 m and the second one variable up to 3000 m as in VAR1. The third run (VAR3) had up to four release heights: 100, 750, 1500, and 3000 m, respectively, with the upper limit defined by the fire heat and not to exceed 3000 m.

Figure 7 shows the difference in daily FMS scores between the sensitivity runs and the BC. In general, there was an overall improvement in the statistical performance in all runs with respect to the BC. However, the VAR1 run tended to outperform the other two. Accordingly, we will focus the rest of this analysis on the VAR1 run. Figure 8 shows a comparison of modeled (HYSPLIT and CMAQ) and measured surface $PM_{2.5}$ concentrations corresponding to the VAR1 run, and there was a considerable reduction (>10 μg m^{-3}) in the overprediction of the $PM_{2.5}$ concentrations compared to the BC shown in Fig. 4a.

The comparison between the two datasets of model results (namely, the BC and VAR1 runs) with the measurements can be further explored by using a slightly modified measure of effectiveness (MOE) from the one described in Rolph et al. (2009). A two-dimensional MOE can be computed based on marginal over- and underpredictions of the concentration as described in Warner et al. (2004). A region of false positive (A_{fp}) is calculated by summing the differences between the predicted and observed $PM_{2.5}$ values at all the measurement sites where the prediction is greater than the observation. Similarly, the region of false negative (A_{fn}) can be estimated based on the samplers that have measured values higher that the ones predicted by the model. The MOE components can be calculated as follows:

$$\text{MOE} = (x, y) = [1 - (A_{fn}/A_{ob}),\ 1 - (A_{fp}/A_{pr})], \quad (1)$$

where A_{ob} and A_{pr} are the observed and predicted values, respectively. Figure 9 shows a comparison of the hourly MOE for the base case and the VAR1 run. The MOE values calculated from the VAR1 run show a noticeable improvement over the ones estimated from the base case that can be attributed to the injection of smoke at higher levels in the atmosphere with a

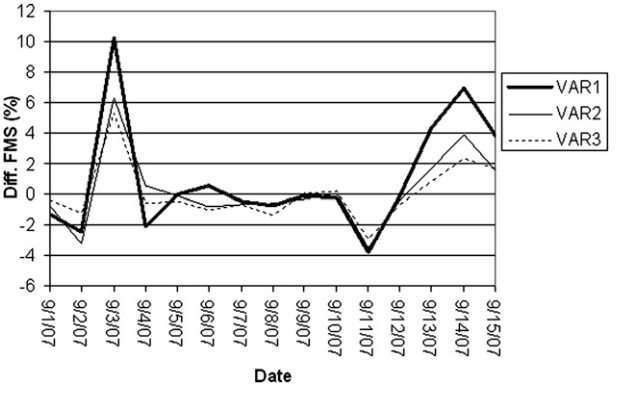


FIG. 7. Difference in daily mean FMS scores between the sensitivity case and BC runs corresponding to the variable release height sensitivity analysis. Positive values indicate that the FMS sensitivity case performed better than the FMS BC. See text for a description of the sensitivity runs.

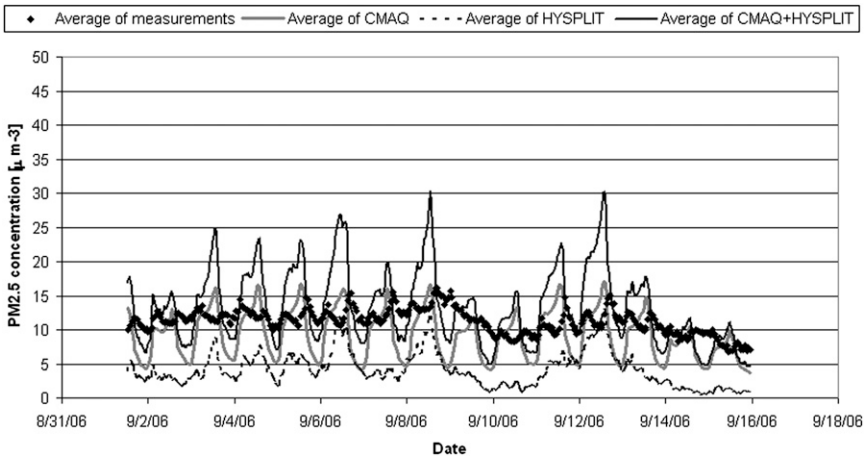


FIG. 8. Comparison of modeled (HYSPLIT and CMAQ) and measured hourly averaged surface $PM_{2.5}$ concentrations corresponding to the VAR1 run.

subsequent reduction in the surface $PM_{2.5}$ concentrations as estimated by HYSPLIT.

However, due to the wide geographical distribution of wildfires, as well as the extensive spatial coverage of smoke during the 1–15 September 2006 period in case 1, it was very difficult to analyze in detail the underlying causes of the model response to the changes introduced in the VAR1 simulation. Therefore, cases 2–4 were rerun using the VAR1 sensitivity settings in order to perform an in-depth analysis of the behavior of individualized smoke plumes.

The use of the VAR1 settings in the October 2007 California (case 2) study did not produce any noticeable change in the FMS scores with respect to the BC. Indeed, changes of less than 3% in FMS scores were observed. The surface $PM_{2.5}$ concentrations comparison showed only a very small improvement with respect to the BC run in both the difference in concentrations and the MOE. Therefore, for this particular period the model did not show a noticeable sensitivity to variations in the release height settings as represented by the VAR1 run.

In contrast to the previous cases, the VAR1 simulation for case 3 presented an evident degradation in the FMS scores and MOEs for surface $PM_{2.5}$ concentrations when compared to the BC simulation (not shown). This apparent decline in the model performance can be directly associated with the injection of smoke above the PBL occurring as a consequence of relaxing the restrictions in the estimation of the release height in the VAR1 simulation (i.e., 3000-m upper limit). For instance, during the daytime hours on 19 April 2007 the modeled particles were mainly released at 3000 m. The vertical profile data (Table 1) corresponding to that date at 1500 UTC near the location of the fire suggest that the PBL height lies at about 750 m. Therefore, the particles simulated by the VAR1 run injected at 3000 m were transported by winds blowing from a different direction than in the BC run in which the particles were released below the PBL height. Indeed, particles released below 1500 m, as in the BC run, were influenced by northeasterly winds whereas those released above that height, as in the VAR1 simulation, experienced a different transport pattern featuring north-northwesterly winds (Table 1). Figures 10a and 10b show the HYSPLIT 5 $\mu g\ m^{-3}$ $PM_{2.5}$ concentration contour compared with HMS for the BC and VAR1 runs for 19 April 2007 at 1500 UTC. The patterns in Fig. 10 confirm that the influence of northeasterly winds is the key factor

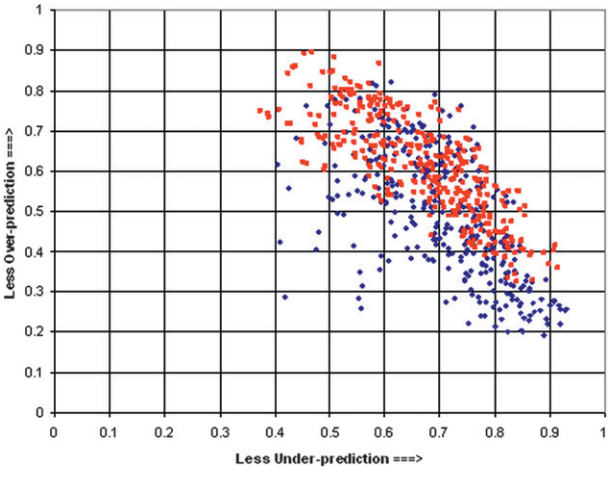


FIG. 9. Hourly MOE corresponding to the period 1–15 Sep 2006. MOE values for the BC run are indicated by blue diamonds and those for the VAR1 run by red squares.

TABLE 1. Meteorological profile corresponding to the fire located at 31.1°N, 82.5°W at 1500 UTC 19 Apr 2007 (case 3).

Pressure (hPa)	Height (MSL, m)	Temp (°C)	Dewpoint (°C)	Wind direction (m s^{-1})	Wind speed (m s^{-1})
1006	42	21.1	14.1	61.7	2.5
1000	92	19.6	12.4	32.1	1.8
975	309	16.6	11.7	32.9	2.2
950	529	14.4	11	35.7	2.3
925	754	13.5	5.9	49.1	3.3
900	985	12.7	2.2	40.5	3.9
875	1220	11.7	−3.9	30.9	4.6
850	1462	10.1	−7.4	21.5	3.9
825	1709	8.3	−9.4	352.5	3.2
800	1962	6.7	−9.8	317.4	4
775	2221	5.3	−9.1	303.1	5.3
750	2488	3.5	−8.8	304.7	6.9
725	2762	1.6	−10.9	311.8	8.5
700	3044	0.4	−16.8	318.8	11.3
650	3636	−0.7	−33.7	322.6	16.6
600	4270	−4.7	−41.5	322.3	19.5
550	4948	−9.2	−49.6	320	25.8
500	5679	−12.5	−49.8	305.2	28.6
450	6476	−16.6	−45	294.5	35.3
400	7352	−21.8	−57.3	288.9	40.7
350	8324	−26.8	−61.1	291.4	51.7
300	9421	−33.6	−56.4	295.7	63.1
250	10673	−43.7	−61.3	295	63.8
200	12131	−55	−70.8	296.5	70.9
150	13937	−61.9	−77.7	293.7	46.4
100	16397	−66.6	−76.8	292.7	25.2
50	20601	−62.8	−82.7	310.5	3.7

controlling the transport of smoke during this particular period and the resulting difference in the model-predicted plume overlap with the HMS pattern between the BC and VAR1 simulations. Considerably more overlap occurred for the BC run than for the VAR1 run.

The last case in this analysis is focused on understanding the failure of the operational SFS to produce smoke moving eastward from a fire source located in southwestern California as observed by the satellite images on 23 and 24 September 2006. The height of the PBL in the vicinity of Los Padres National Forest on 23 September at 2100 UTC was estimated (from Table 2) to be approximately 1200 m above ground level (or 2500 m above mean sea level) at the location of the fire. Consequently, the injection height calculated in the BC run was below 1200 m with the corresponding wind blowing from the east. The smoke release height corresponding to the VAR1 run reached 3000 m at 2100 UTC on 23 September. At that altitude, the winds were blowing from the west (Table 2), causing the smoke particles to be transported eastward, and reproducing the pattern observed in the satellite image (Figs. 11a–d). It is also worth mentioning that the overprediction at the El Rio, Piru, and Simi Valley stations for the BC run were not observed in the VAR1 run.

5. Summary and conclusions

The SFS has been evaluated in detail for four different case studies by comparing surface and column-integrated $PM_{2.5}$ concentrations with surface and satellite observations, respectively. The first case, covering the period 1–15 September 2006, featured the long-range transport of smoke that extended over a very large portion of the United States and Canada originating from large fires distributed along the northwestern United States. Daily mean FMS scores generally exceeded 5% with some days reaching levels above the average of 11.6% calculated for the full study period (September 2006–October 2007). The model statistical performance for this case study was not significantly different from that of the full study period. Comparison of the simulated $PM_{2.5}$ concentrations with surface measurements showed that the model was able to capture the gross features of the relative day-to-day variations in the hourly $PM_{2.5}$ levels; however, the model tended to overpredict the nighttime concentrations and underpredict the daytime concentrations. These discrepancies can be attributed to a variety of causes including errors in emissions inventories, dry deposition, and difficulties in estimating whether the wild fire plume rise stays within or exceeds the planetary boundary layer height.

The second case study covered the period 21–27 October 2007 where the smoke extended over parts of California and the Pacific Ocean. The daily mean FMS scores for this particular case showed excellent model performance with values well above the average for the full study period. For this case, the model showed a tendency to overpredict the hourly $PM_{2.5}$ levels when compared with surface measurements. HYSPLIT grossly overpredicted concentrations from 23 to 25 October; however, the results improved noticeably toward the end of the period. Moreover, during the first 2 days of the period, when the contributions of dust sources were included in the emission and transport of PM, the simulation results showed that the high measured PM values could be attributed, at least in part, to a contribution from dust sources.

The third case featured a large fire located near the border of Georgia and Florida, where the model showed a tendency to underpredict the hourly $PM_{2.5}$ surface levels. However, the model did capture the timing of the peaks as well as their relative magnitude with respect to the measured data. The comparison with satellite analyses showed that the model performance was above the average corresponding to the full study period.

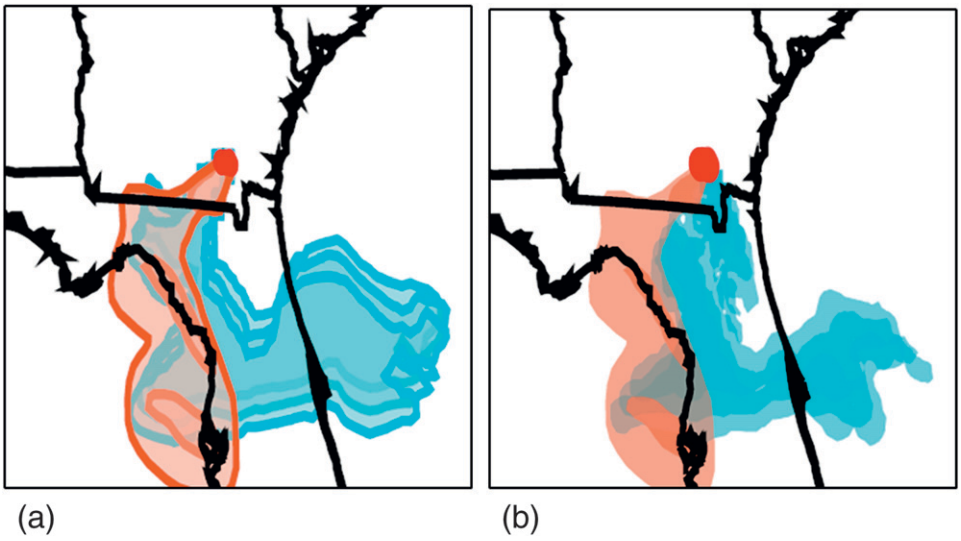


FIG. 10. Simulated 5 μg m^{-3} $PM_{2.5}$ concentration contour (light blue shading) compared to the HMS (orange shading) for the (a) BC and (b) VAR1 runs between 1545 and 1715 UTC 19 Apr 2007 (case 3). The red dot corresponds to the location of the fire.

Finally, the FMS score for the last case on 23 September 2006 reflected very good model performance. However, on 24 September, the score experienced a sharp deterioration. On this day the model was unable to capture the thin line of smoke extending from southwestern California to the east producing the observed degradation in the simulation performance. Although the sparse measurement data did not show any appreciable $PM_{2.5}$ attributable to fires, the model presented a noticeable amount of smoke at the surface. This disagreement can be associated with errors in the model estimation of the smoke injection height. Indeed, the simulation emitted smoke below the boundary layer, making an important contribution to the surface concentrations.

The sensitivity analysis to the injection height showed that the model response to a fixed smoke release height did not show a consistent improvement in the FMS scores with respect to the BC for the period 1–15 September 2006 among the five cases chosen, namely 100, 750, 1500, 3000, and 5000 m. However, the sensitivity to a variable release height produced an overall improvement in FMS scores for the three runs performed, with the VAR1 run, where the heat release from the fire was used to estimate the release height and the particles were released up to 3000 m, showing the best performance. Also, the settings used for the VAR1 run contributed to a noticeable improvement in the model performance when compared with the $PM_{2.5}$ surface concentrations for the 1–15 September 2006 period.

The VAR1 settings used in the other three study periods, where the identification of individual smoke plumes facilitated the analysis, showed mixed results that can be directly attributed to whether the smoke injection actually occurred below or above the PBL. For instance, when using the VAR1 settings for 19 April

TABLE 2. Meteorological profile corresponding to the fire located at 34.7°N, 119.1°W at 2100 UTC 23 Sep 2006 (case 4).

Pressure (hPa)	Height (MSL, m)	Temp (°C)	Dewpoint (°C)	Wind direction (m s^{-1})	Wind speed (m s^{-1})
862	1387	20.9	−7.1	97.2	8.2
850	1508	17.3	−4.4	92.2	7.2
825	1761	14.2	−7.3	89	10.4
800	2019	11.9	−9	87.5	11.2
775	2283	8.9	−9.4	87.8	11.2
750	2553	7.6	−8.2	85	10.2
725	2831	5.9	−5.4	83.6	7.7
700	3118	3.7	−2.5	85.2	5.3
650	3715	0.4	−4.1	103.4	4.5
600	4360	2.2	−17.5	266.4	3.8
550	5058	−1.2	−18.2	290.2	7.5
500	5811	−5.3	−24.6	315.5	8.2
450	6628	−11.7	−34.5	316.8	8.1
400	7517	−18.9	−41.1	313.7	8.2
350	8493	−27.1	−48.9	297	8.5
300	9583	−36	−58.9	304.4	8.5
250	10823	−45.8	−63.5	269.5	12.3
200	12271	−55.8	−67.2	257.4	18.6
150	14069	−63.2	−77.2	257.4	23.6
100	16505	−69.5	−78.9	275	15
50	20717	−60.6	−86.6	344.2	0.8

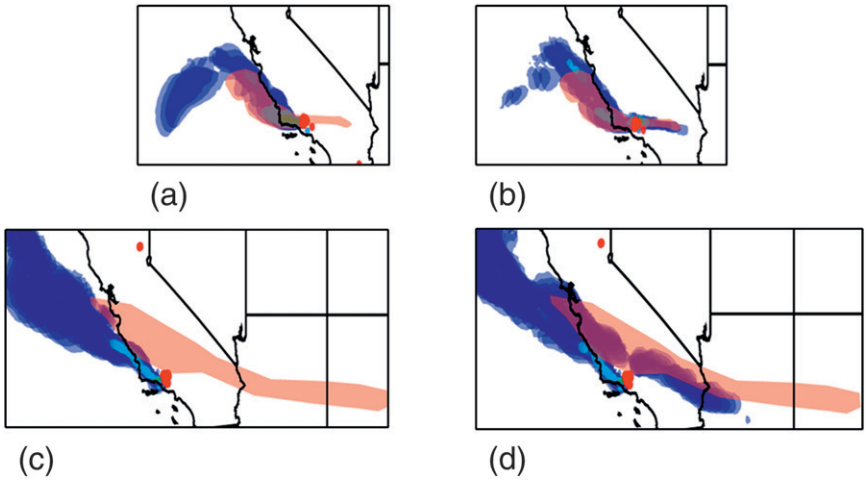


FIG. 11. Simulated 1 (purple shading) and 5 (blue shading) μg m^{-3} $PM_{2.5}$ concentration contours compared to the HMS (orange shading) for the (a) BC and (b) VAR1 runs at 2215–0000 UTC 23–24 Sep 2006. (c),(d) As in (a),(b) but at 1100–1630 UTC 24 Sep 2006 (case 4). The red dots correspond to the locations of the fires.

2007 the modeled injection height took place above the estimated PBL. At that height, the winds were blowing from a different direction than those compatible with the satellite image, which was consistent with emission heights below the PBL, further explaining the misplacement of the simulated smoke contour. The opposite is true for the 23–24 September 2006 case where the eastward transport of smoke can only be explained if the simulated injection height reached the height at which the wind was blowing from the west (approximately 3000 m).

In conclusion, it is clear from our analysis that an accurate estimation of the smoke release height is important when trying to calculate $PM_{2.5}$ concentrations within or about the PBL. This initial investigation suggested some simple changes to the current operational system to improve the emission height estimation. However, further investigation is required to determine alternative means to accurately estimate the smoke release height, which depends upon many factors that may never be known with sufficient accuracy to estimate the smoke plume heights using only numerical modeling techniques. Future methods may need to rely more heavily on satellite observations to initialize the forecast. However, in the meantime, the current operational SFS includes a 24-h analysis period computation to initialize the 48-h forecast. This analysis calculation is available online (http://www.arl.noaa.gov/ready-bin/smokevrf.pl) with each day's forecast and can be used by the forecaster to determine the relative uncertainty for each day's prediction. The real-time analysis calculation can be compared with the HMS analyses to deduce the potential uncertainty of the forecast in a manner similar to our case study evaluation.

Acknowledgments. This project was carried out under the auspices of NOAA's National Air Quality Forecast Capability and the views expressed in this paper are those of the authors and do not necessarily represent those of NOAA. The authors thank Dr. Rohit Mathur for providing the CMAQ model data derived from developmental PM forecast simulations. Useful comments from Paula Davidson are greatly appreciated.

REFERENCES

Boybeyi, Z. N., N. Ahmad, D. P. Bacon, T. J. Dunn, M. S. Hall, P. C. S. Lee, R. A. Sarma, and T. R. Wait, 2001: Evaluation of the Operational Multiscale Environment model with grid adaptivity against the European Tracer Experiment. *J. Appl. Meteor.,* **40,** 1541–1558.

Briggs, G. A., 1969: Plume rise. TID-25075, USAEC Critical Review Series, National Technical Information Service, Springfield, VA, 81 pp.

Byun, D. W., and J. K. S. Ching, Eds., 1999: Algorithms of the EPA MODELS-3 Community Multiscale Air Quality (CMAQ) modeling system. EPA/600/R-99/030, Office of Research and Development, Washington, DC, 574 pp.

——, and K. Schere, 2006: Review of the governing equations, computational algorithms, and other components of the Models-3 Community Multiscale Air Quality (CMAQ) modeling system. *Appl. Mech. Rev.,* **59,** 51–77.

Colarco, P. R., M. R. Schoeberl, B. G. Doddridge, L. T. Marufu, O. Torres, and E. J. Welton, 2004: Transport of smoke from Canadian forest fires to the surface near Washington, D.C.: Injection height, entrainment, and optical properties. *J. Geophys. Res.,* **109,** D06203, doi:10.1029/2003JD004248.

Draxler, R. R., and G. D. Hess, 1997: Description of the HYSPLIT_4 modeling system. NOAA Tech. Memo. ERL ARL-224, NOAA/Air Resources Laboratory, Silver Spring, MD, 24 pp.

——, and ——, 1998: An overview of the HYSPLIT_4 modeling system of trajectories, dispersion, and deposition. *Aust. Meteor. Mag.,* **47,** 295–308.

——, D. A. Gillette, J. S. Kirkpatrick, and J. Heller, 2001: Estimating PM10 air concentrations from dust storms in Iraq, Kuwait, and Saudi Arabia. *Atmos. Environ.,* **35,** 4315–4330.

Escudero, M., A. Stein, R. R. Draxler, X. Querol, A. Alastuey, S. Castillo, and A. Avila, 2006: Determination of the contribution of northern Africa dust source areas to PM10 concentrations over the central Iberian Peninsula using the Hybrid Single-Particle Lagrangian Integrated Trajectory model (HYSPLIT) model. *J. Geophys. Res.,* **111,** D06210, doi:10.1029/2005JD006395.

Larkin, N. K., S. M. O'Neill, R. Soloman, C. Krull, S. Raffuse, M. Rorig, J. Peterson, and S. Ferguson, 2009: The BlueSky Smoke Modeling Framework: Design, application, and performance. *Int. J. Wildland Fire,* in press.

Leung, F. T., and Coauthors, 2007: Impacts of enhanced biomass burning in the boreal forests in 1998 on tropospheric chemistry and the sensitivity of model results to the injection height of emissions. *J. Geophys. Res.,* **112,** D10313, doi:10.1029/2006JD008132.

Mathur, R., Y. Shaocai, D. Kang, and K. L. Schere, 2008: Assessment of the wintertime performance of developmental particulate matter forecasts with the Eta-Community Multiscale Air Quality modeling system. *J. Geophys. Res.,* **113,** D02303, doi:10.1029/2007JD008580.

McKeen, S., and Coauthors, 2007: Evaluation of several $PM_{2.5}$ forecast models using data collected during the ICARTT/NEAQS 2004 field study. *J. Geophys. Res.,* **112,** D10S20, doi:10.1029/2006JD007608.

Morris, G. A., and Coauthors, 2006: Alaskan and Canadian forest fires exacerbate ozone pollution over Houston, Texas, on 19 and 20 July 2004. *J. Geophys. Res.,* **111,** D24S03, doi:10.1029/2006JD007090.

Mosca, S. G., G. Graziani, W. Klug, R. Bellasio, and R. Bianconi, 1989: A statistical methodology for the evaluation of long-range dispersion models: An application to the ETEX exercise. *Atmos. Environ.,* **32,** 4302–4324.

Rolph, G. D., and Coauthors, 2009: Description and verification of the NOAA Smoke Forecasting System: The 2007 fire season. *Wea. Forecasting,* **24,** 361–378.

Ruminski, M., S. Kondragunta, R. Draxler, and J. Zeng, 2006: Recent changes to the Hazard Mapping System. *15th International Emission Inventory Conf.: Reinventing Inventories—New Ideas in New Orleans,* New Orleans, LA, EPA. [Available online at http://www.epa.gov/ttn/chief/conference/ei15/session10/ruminiski.pdf.]

——, ——, ——, and G. Rolph, 2007: Use of environmental satellite imagery for smoke depiction and transport model initialization. *16th Annual International Emission Inventory Conf.: Emission Inventories—Integration, Analysis, and Communications,* Raleigh, NC, EPA. [Available online at http://www.epa.gov/ttn/chief/conference/ei16/session10/ruminski.pdf.]

Stockwell, W. R., and Coauthors, 2002: The scientific basis of NOAA's Air Quality Forecasting Program. *EM*, December, Air and Waste Management Association, Pittsburgh, PA, 20–27.

Turquety, S., and Coauthors, 2007: Inventory of boreal fire emissions for North America in 2004: Importance of peat burning and pyroconvective injection. *J. Geophys. Res.,* **112,** D12S03, doi:10.1029/2006JD007281.

Warner, S., N. Platt, and J. F. Heagy, 2004: User-oriented two-dimensional measure of effectiveness for the evaluation of transport and dispersion models. *J. Appl. Meteor.,* **43,** 58–73.

Wilks, D. S., 1995: *Statistical Methods in the Atmospheric Sciences*. Academic Press, 467 pp.

Wotawa, G., and M. Trainer, 2000: The Influence of Canadian forests fires on pollutant concentrations in the United States. *Science,* **288,** 324–328.