Final Report

For

BAAQMD Contract Number: 2009-078

2009 – 2011

Application of the Weather Research and Forecasting Model for Air Quality Modeling in the San Francisco Bay Area

From

Aijun Deng, P.I.

Department of Meteorology
The Pennsylvania State University
University Park, Pennsylvania 16802
deng@meteo.psu.edu
(814)863-8253

And

David R. Stauffer

Department of Meteorology
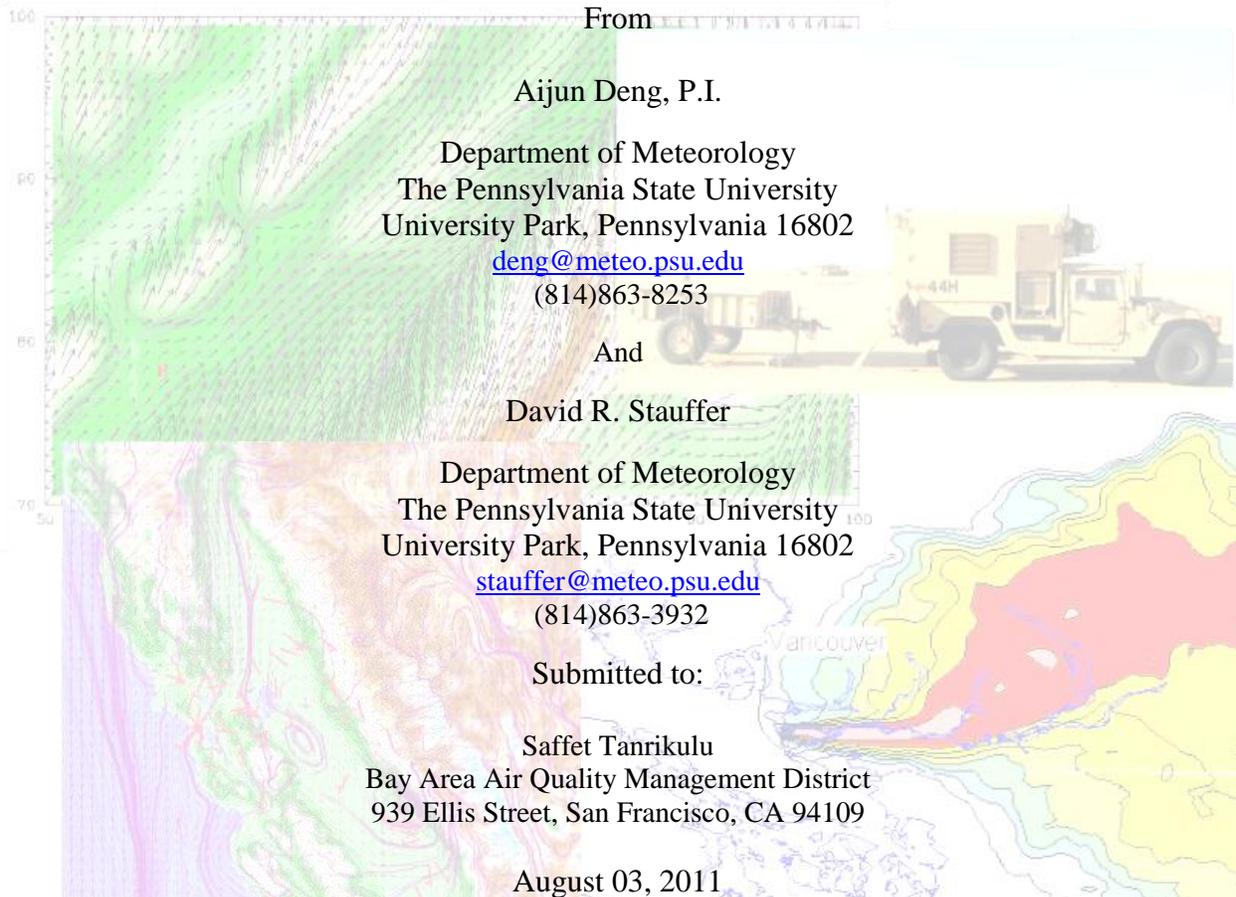The Pennsylvania State University
University Park, Pennsylvania 16802
stauffer@meteo.psu.edu
(814)863-3932

Submitted to:

Saffet Tanrikulu
Bay Area Air Quality Management District
939 Ellis Street, San Francisco, CA 94109

August 03, 2011

Application of the Weather Research and Forecasting Model for Air Quality

Modeling in the San Francisco Bay Area

Final Report

For

BAAQMD Contract Number: 2009-078

Submitted to:

Dr. Saffet Tanrikulu

Bay Area Air Quality Management District

939 Ellis Street, San Francisco, CA 94109

By

Aijun Deng and David R. Stauffer

Department of Meteorology
The Pennsylvania State University
University Park, Pennsylvania 16802

Technical Contact: Aijun Deng, (814)863-8253, deng@meteo.psu.edu, FAX (814)865-3663

August 03, 2011

1

ABSTRACT

The Bay Area Air Quality Management District (BAAQMD) has been using the Pennsylvania State University (PSU)/National Center for Atmospheric Research (NCAR) Mesoscale Model Version 5 (MM5) to create the meteorological datasets for use as inputs into photochemical models such as the Community Multiscale Air Quality (CMAQ) model and the Comprehensive Air Quality Model with Extensions (CAMx) model. However, development on MM5 has been discontinued, and the BAAQMD is interested in transitioning to the Weather Research and Forecasting (WRF) model when it can perform as well as the MM5. The objective of this research is to assist the BAAQMD to transition to the WRF modeling system by finding the optimal WRF model configuration for the Bay Area and Central Valley region based on our WRF simulations for both winter and summer seasons.

Two case studies were conducted: a winter particulate matter (PM) case and a summer ozone case. The investigation started by determining the optimal set of physics packages to use for the region before four dimensional data assimilation (FDDA) was applied. The FDDA was applied throughout the model integrations to produce dynamic analyses of the meteorology for use in the atmospheric chemistry models. The two atmospheric radiation schemes tested were the Rapid Radiative Transfer Model (RRTM) and the RRTM for general circulation models (RRTMG). It was found that the RRTM radiation scheme performed equal to or better than the RRTMG scheme. Using the RRTM radiation scheme, four land surface models (LSM) (the 5-layer thermal diffusion LSM, Noah LSM, Rapid Update Cycle (RUC) LSM, and the Pleim-Xiu (PX) LSM) were compared to determine the best LSM for use in the baseline configuration before FDDA was applied. The Pleim-Xiu land surface model was found to produce the smallest error for the winter PM case, and performed equally well to the other LSMs for the summer ozone case.

With the baseline configuration using the RRTM radiation and PX PBL/land surface physics, six experiments were conducted to compare different FDDA strategies for the winter case and three experiments were run for the summer case. The FDDA strategies used were analysis (3D and surface) nudging, observational nudging, and multiscale FDDA that was a combination of both analysis nudging and observational nudging. Statistical and subjective analyses were performed to compare the model output from the FDDA experiments with the observations and the best experiment was chosen. Similar verifications were also performed to compare the baseline and the best FDDA experiment for three subregions within the 4-km domain: the Bay Area, the Sacramento Valley in the northern Central Valley, and the San Joaquin Valley in the southern Central Valley. The incoming marine flow over the Bay Area was examined, as well as the wind flows in the Sacramento and San Joaquin Valleys.

The major conclusions from this research include: 1) WRF is able to simulate the major mesoscale features over the area including the Fresno eddy and Schultz eddy; 2) FDDA significantly reduces model errors both statistically and subjectively, with the multiscale FDDA strategy producing the lowest errors for all four fields; 3) there is added value to using the special surface wind observations taken by the BAAQMD observation network; and 4) over the Bay Area subregion, FDDA is not as effective as in the Sacramento Valley and San Joaquin Valley.

TABLE OF CONTENTS

LIST OF ACRONYMS

AIRS	Aerometric Information Retrieval System
AQS	Air Quality System
ARB	Air Resources Board
AUSPEX	Atmospheric Utility Signatures, Predictions, and Experiments
BAAQMD	Bay Area Air Quality Management District
CIMIS	California Irrigation Management Information System
CAMx	Comprehensive Air Quality Model with Extensions
CAPEX	Cross-Appalachian Tracer Experiment
CMAQ	Community Multiscale Air Quality
DTRA	Defense Threat Reduction Agency
EPA	Environmental Protection Agency
FDDA	Four Dimensional Data Assimilation
GCM	General Circulation Model
IOP	Intensive observation period
LSM	Land surface model
MAE	Mean absolute error
ME	Mean error
MM5	Penn State/NCAR fifth generation mesoscale model
NAAQS	National Ambient Air Quality Standards
NARSTO	North American Research Strategy for Tropospheric Ozone
NCAR	National Center for Atmospheric Research
NCDC	National Climatic Data Center
NCEP	National Centers for Environmental Prediction
NWS	National Weather Service
PBL	Planetary boundary layer
PM	Particular matter
PX	Pleim-Xiu
RAWS	Remote Automated Weather Stations
SARMAP	Regional Modeling Adaptation Project
SCIPUFF	Second-Order Closure Integrated Puff
SCOS	Southern California Ozone Study
SFBA	San Francisco Bay Area
SJVAQS	San Joaquin Valley Air Quality Study
WMO	World Meteorological Organization
WRF	Weather Research and Forecasting Model

LIST OF FIGURES

LIST OF TABLES

1   INTRODUCTION

Meteorology and its effects on air pollution has been the subject of modeling and observational studies for more than two decades in the San Francisco Bay Area (SFBA) by the Bay Area Air Quality Management District (BAAQMD). Understanding the process from the emission of pollutants to the formation of smog and haze and their transport is paramount to developing strategies to reduce air pollution and its negative effects on human health and the environment.

As discussed by Tanrikulu et al. 2000, Deng et al. 2004, and Otte et al. 2008a and b, accurate meteorological information is critically important in air quality modeling. Meteorological models commonly used in air quality studies include the Penn State/NCAR fifth generation of mesoscale model (MM5, Grell et al. 1994), and the recently-developed Weather Research and Forecasting model (WRF, Skamarock 2008). These models generate gridded meteorological fields that can be used to drive air quality models. Using source information and the generated meteorological fields, photochemical air quality models, such as the Community Multiscale Air Quality (CMAQ) Model (Byun and Schere, 2006) and the ENVIRON International Corporation's Comprehensive Air Quality Model with Extensions (CAMx) (Kumar and Lurmann, 1997), are able to simulate particulate matter (PM), air toxics, and ozone concentrations.

Soong et al. (2006) evaluated the performance of MM5 and WRF for an ozone episode in central California from July 31 to August 2, 2000. Both models simulated the wind and daytime temperatures quite well. In the SFBA region, both models overestimated the temperature along the coast by about 5 ℃ and underestimated it in the Bay Area inland valleys by 3-5 ℃. One shortcoming of the WRF model was the over prediction of the nighttime temperatures, which were about 5 ℃ too warm in most areas, which is likely due to the fact that WRF experiment did not use four dimensional data assimilation (FDDA) because it was not available at the time of the study. FDDA, as adapted from Stauffer and Seaman (1994), is a method to incorporate observed weather data into a running model without causing negative disruptions. FDDA capabilities were recently implemented by Penn State (Deng et al. 2009).

WRF-ARW was recently used to study the air quality issues in California. A recent study undertaken by Bao et al. (2008) investigated the ability of the WRF model to accurately simulate the near surface winds in the Central Valley region of California for a summer ozone study. The results showed that WRF was capable of simulating many of the low level flow features found in the Central Valley, even though FDDA was not used because it was not available at the time. The flow features that WRF was able to simulate are: 1) westerly marine airflow through the Carquinez Strait and into the Sacramento River delta located east of the San Francisco Bay Area, 2) upslope and downslope flows in the Central Valley, 3) up-valley and down-valley flows along the Sacramento Valley, 4) the nocturnal low level jet in the San Joaquin Valley, and 5) the Fresno and Schultz eddies. Figure 1-1 shows the conceptual model of the flows in the Central Valley for both day and night.
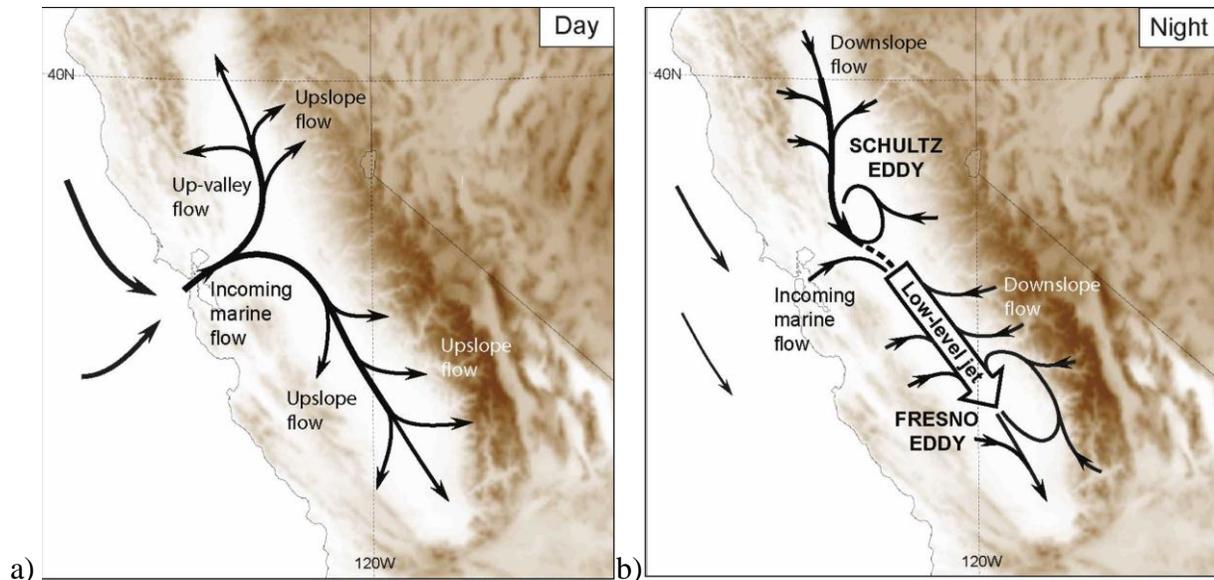
Figure 1-1: Conceptual model of the summertime low-level winds in the Central Valley region of California. a) The daytime. b) The nighttime. Figure from Bao et al. (2008).

As development of MM5 has been discontinued, the BAAQMD is interested in transitioning to the WRF model. As the new state-of-the-science mesoscale numerical weather prediction model (NWP), WRF is under continuous development by the user community. The plan is to transition to WRF when it can perform as well as MM5.

The purpose of this study is to determine the optimal WRF model configuration to air quality modeling in the SFBA and Central Valley (CV) regions. Both summer and winter cases are modeled. Evaluation of WRF performance using different model physics and FDDA strategies are conducted. The ability of WRF to accurately simulate meteorological variables important for air quality will be examined using different land surface models, atmospheric radiation schemes, and FDDA strategies. The use of different data analyses for the initial conditions (IC) and lateral boundary conditions (BC) and different methods of objective analysis and quality control of observations used for data assimilation will also be examined. Two high pollution cases are chosen to evaluate the ability of WRF to simulate the meteorology in the SFBA and CV regions during both the summer and the winter.

2 MODEL DESCRIPTIONS

The meteorological model used in this study is the advanced research dynamics version of the WRF model (WRF-ARW, Skamarock et al. 2008). A complete description of the WRF-ARW can be found in Skamarock et al. (2008) although some of the fundamental aspects of the WRF model are described below.

Similar to MM5, WRF-ARW is a nonhydrostatic, fully compressible three dimensional primitive equation model with terrain-following, hydrostatic pressure vertical coordinates. The WRF-ARW core contains predictive equations for three wind components (u, v, and w), potential temperature, geopotential height and mixing ratio of moisture species including water vapor and cloud hydrometeor species.

WRF-ARW has a variety of physics options for microphysics, cumulus parameterization, atmospheric radiation, and planetary boundary layer (PBL)/turbulence processes that can interact with the model's dynamics and thermodynamics. WRF-ARW also has several land-surface models (LSMs) that use information from the other WRF physics schemes in combination with information on the land's state variables and land-surface properties, to predict heat and moisture fluxes to the atmosphere. WRF-ARW also has FDDA capabilities similar to that in MM5.

2.1 Relevant Model Physics

The WRF Single-Moment 3-class (WSM3) simple ice scheme (Hong et al. 2004) was chosen as the microphysics option for the model simulations in this study. This scheme calculates ice number concentration from ice mass instead of temperature and predicts three categories of hydrometeors: water vapor, cloud water or cloud ice, and rain or snow. The scheme assumes cloud water and rain exist above freezing temperatures, and cloud ice and snow exist below freezing temperatures. The scheme also does not allow mixed phases of hydrometeors, such as the existence of ice and rain together in the same grid cell. The WSM3 scheme, rather than a more complicated microphysics scheme, was chosen because both cases in this study had very little precipitation, as a ridge of high pressure was over the model domain.

To determine a suitable radiation scheme for the simulations, two methods were considered: 1) Rapid Radiative Transfer Model (RRTM, Mlawer et al. 1997), and 2) the Rapid Radiative Transfer Method for general circulation models (GCMs) (RRTMG). RRTM calculates longwave radiation fluxes using a correlated-k method. Radiative fluxes and cooling rates for inhomogeneous atmospheres are approximated (Mlawer et al. 1997, Iacono et al. 2008). RRTM uses pre-set tables to represent longwave processes due to water vapor, ozone, carbon dioxide, and other trace gases (if present), as well as accounting for cloud optical depth. As in MM5, the RRTM scheme is used for longwave radiation in combination with the Dudhia shortwave scheme (Dudhia 1989). This shortwave scheme calculates solar flux and accounts for clear air radiative scattering, water vapor absorption and cloud albedo and absorption. The current RRTM scheme in WRF-ARW also accounts for terrain slope and shadowing effects on the surface solar flux. The RRTMG scheme (Iacono et al. 2008) that was originally designed for GCMs, recently implemented in WRF-ARW, was also used for both longwave and shortwave radiation in our study.

To determine the best LSM for the simulations, the 5-layer thermal diffusion scheme was evaluated, along with three other LSMs: 1) Noah LSM (Chen and Dudhia 2001), 2) Rapid Update Cycle (RUC, Smirnova et al., 1997, 2000) LSM, and 3) Pleim-Xiu (PX) LSM (Pleim and

Xiu, 1995; Xiu and Pleim, 2001). The 5-layer thermal diffusion scheme is the MM5 5-layer soil temperature scheme with the layers 1, 2, 4, 8 and 16 cm thick. The energy budget includes radiation, sensible, and latent heat flux. Soil moisture is fixed with a constant value. The Noah LSM is 4-layer soil temperature and soil moisture model with layer thicknesses of 10, 30, 60, and 100 cm. It includes root zone, evapotranspiration, soil drainage, and runoff. The Noah LSM also takes into account vegetation categories, monthly vegetation fraction, and soil texture. The scheme provides sensible and latent heat fluxes to the PBL calculations and can predict soil ice and snow cover. The RUC LSM is a multi-level soil model that contains 6 levels as its default, but can have 9 or more layers. Layer thicknesses for the default 6 levels are 0, 5, 20, 40, 160, and 300 cm. The model solves heat diffusion and Richards' moisture transfer equations, while accounting for phase changes of soil moisture in the winter. The RUC LSM has a complex multi-layer snow model that accounts for changing snow density, snow depth, refreezing of liquid water, and fractional snow cover. The PX LSM is a two-layer soil temperature and moisture model, with layers that are 1 cm and 99 cm thick. There are three pathways for moisture fluxes: evapotranspiration, soil evaporation, and evaporation from canopies. Two indirect nudging schemes are involved in the PX LSM. In the first scheme, soil moisture is nudged according to biases between model and observation based analyses of temperature and relative humidity (RH) at 2 meters (Pleim and Xiu, 2003). In the second scheme, deep soil temperature in the soil temperature force–restore (FR) model is also nudged according to the model bias of air temperature at 2 meters, but only during nighttime (Pleim and Gilliam, 2009).

For this study the model configuration was comprised of three domains, with 36-km, 12-km, and 4-km grid spacing. The Kain-Fritsch (KF) scheme (Kain and Fritsch 1990, Kain 2004) was used for the cumulus parameterization on the 36- and 12-km grids. The KF scheme is a simple cloud model that takes in account moist updrafts and downdrafts, having the ability to detrain and entrain clouds. The scheme imposes a minimum entrainment rate to suppress convection in unstable, dry environments, where the entrainment rate varies as a function of low level convergence. The scheme also allows for shallow convection without precipitation. Note that since both cases chosen are weakly-forced with very little precipitation, we simply choose a cumulus parameterization scheme that has been shown to perform well.

Two different PBL physics schemes are used in this study: the Turbulent Kinetic Energy (TKE) predicting Mellor-Yamada Level 2.5 turbulent closure scheme (MYJ PBL) (Janjic 1996, 2002) and the asymmetrical convective model version 2 (ACM2), which is designed for the PX physics suite (Pleim, 2007). The MYJ PBL scheme defines a PBL top that is dependent on the TKE, buoyancy, and shear of the driving flow. Also, in this scheme unstable mixing in the PBL is done by eddy diffusion calculated from the TKE and entrainment of dry air at the PBL top. The ACM2 PBL scheme is a combination of the original ACM and an eddy diffusion model. In convective conditions the ACM2 can simulate rapid, nonlocal upward transport in buoyant plumes and local shear induced turbulent diffusion. The scheme can transition from local eddy diffusion under stable conditions to combined local and non-local transport in unstable conditions. The ACM2 can consistently transport any atmospheric quantity (meteorological and chemical trace species) within the PBL.

2.2   Four-Dimensional Data Assimilation

FDDA used in this research was originally developed at Penn State (Stauffer and Seaman 1990, 1994) and was recently enhanced and implemented into WRF-ARW (Deng et al. 2009). In

FDDA nudging, the model state is relaxed continuously toward the observed state at each time step by adding an artificial tendency term to the prognostic equations, which is based on the difference between the two states. Data assimilation can be accomplished by either nudging the model solutions toward gridded analyses based on observations (analysis nudging), or by nudging directly toward the individual observations (obs nudging). Within a multiscale grid-nesting assimilation framework, a combination of analysis and observational (obs) nudging is often employed.

Further development of obs nudging in WRF-ARW has brought more flexibility in how surface observations are extended in the vertical. As illustrated in Figure 2-1, WRF-ARW users have freedom to choose different vertical weighting functions for the surface observations. In contrast, the MM5 obs nudging defaults to surface winds spread through the lowest three model layers with linearly decreasing weights for all PBL regimes (column 5). Under an unstable PBL regime (regime 4, column 3), the WRF-ARW default allows the surface observations to be spread through the entire depth of the PBL at full strength, decreasing linearly to zero 50 meters above the PBL top. . For the stable PBL regimes (regimes 1 and 2), the WRF-ARW default allows the surface observations to be spread upward to 50 m at full strength, then linearly decreases to zero for the next 50m. The default surface data weighting functions are used for this study.

Column #	1	2	3	4	5	6	7	8	Column #
Ramp Layer	0	-5000	50	0	50	50	-5000	-5000	Ramp Layer
Full Layer	-5000	0	-5000	50	0	50	50	-4950	Full Layer
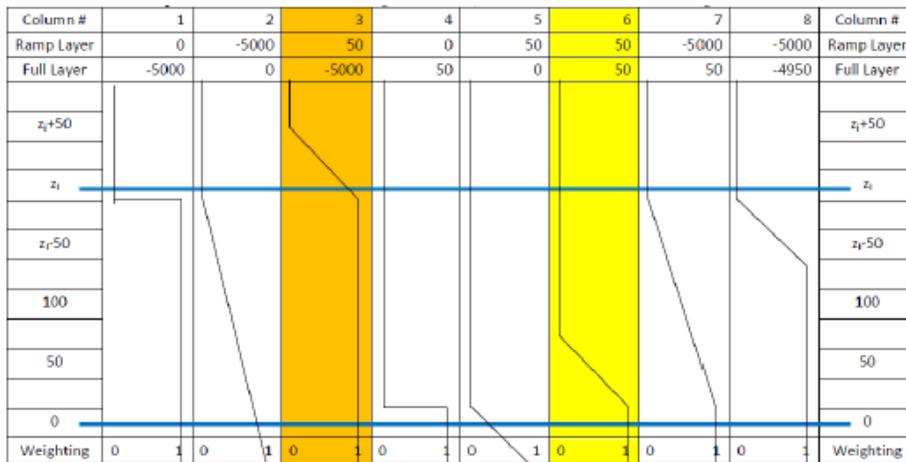
Figure 2-1: Illustration of possible vertical weighting functions for surface observations. For each of the eight examples, the horizontal axis is the weight (from zero to one) and the vertical axis is height from 0 (the ground) to z_i+50 (50 m above the top of the PBL). The settings used to produce the vertical weighting function are indicated in the second two rows. The blue horizontal lines indicate the surface and the PBL top. For WRF-ARW, column 6 is the default for stable PBL regimes (regime 1 and 2), and column 3 is the default for the unstable PBL regime (regime 4). Figure from Deng et al. (2008).

3   CASE DESCRIPTIONS

Two real-world air pollution episodes were chosen for this modeling study: a winter PM case during the middle of December, 2000 and a summer ozone case during the end of July and beginning of August, 2000. The following is an overview of the two cases.

3.1   Winter PM Case

The winter case began at 1200 UTC on December 16, 2000 with a surface high over northern Nevada and an inverted trough just off the coast of California (Figure 3-1a). The temperatures were cooler over the land than over the ocean (Figure 3-1b) which resulted in a decreasing pressure gradient toward the ocean. Due to this pressure difference, winds along the California coast became easterly, producing off shore (not shown). Winds over the interior SFBA and within the Central Valley were very light (5 kts) and northerly to northwesterly (Figure 3-1c). At the 850 millibar (mb) level, a high was centered over western Nevada (Figure 3-2a), leading to stable synoptic conditions over California with light surface winds over the SFBA. An upper level ridge at the 500 mb level was over the entire western coastline of the U.S., helping to reinforce the stable conditions at the surface (Figure 3-2b). The stable surface conditions allowed for the development of a high PM episode.

Figure 3-1: Surface observations at 1200 UTC December 16, 2000. a) Sea level pressure. b) Temperature. C) Winds. Maps made at Plymouth State University Make Your Own Map website.
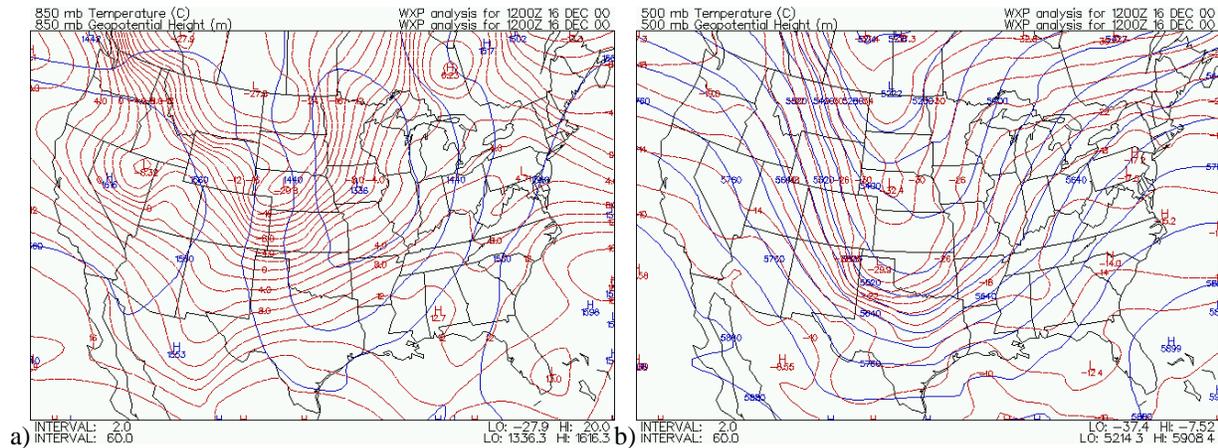
Figure 3-2: Upper air observations at 1200 UTC December 16, 2000. a) 850 mb geopotential height and temperature. b) 500 mb geopotential height and temperature. Maps made at Plymouth State University Make Your Own Map website.

At 1200 UTC December 17, the surface high moved southeast and was centered over the Four Corners region in the southwestern U.S (Figure 3-3a), with the inverted trough, though smaller, remaining off the coast. A surface high was also centered over the Pacific Ocean west of the Oregon-California border. Between the two highs a trough extended through Nevada and into California. South of this trough, light southerly winds occurred over the SFBA, while the Central Valley experienced southeasterly flow (not shown). North of the trough the winds were northerly, suggesting the existence of a weak cold front. The temperature pattern (Figure 3-3b) was similar to the previous day. At 850 mb the high also moved southeast and was centered over the Four Corners region (Figure 3-4a), while another ridge moved over California from the ocean. This kept weak synoptic conditions in place over the SFBA with the winds generally northwesterly at around 15 kts at this level (Figure 3-4b). At 500 mb the ridge had progressed eastward and an upper level trough moved over the northwestern U.S. (Figure 3-4c). The trough did not push down into California; therefore the winds were relatively light for this level at 50 knots (Figure 3-4d).

c)

Figure 3-3: Surface observations at 1200 UTC December 17, 2000. a) Sea level pressure. b) Temperature. C) Winds. Maps made at Plymouth State University Make Your Own Map website.

a)

b)

c)

d)
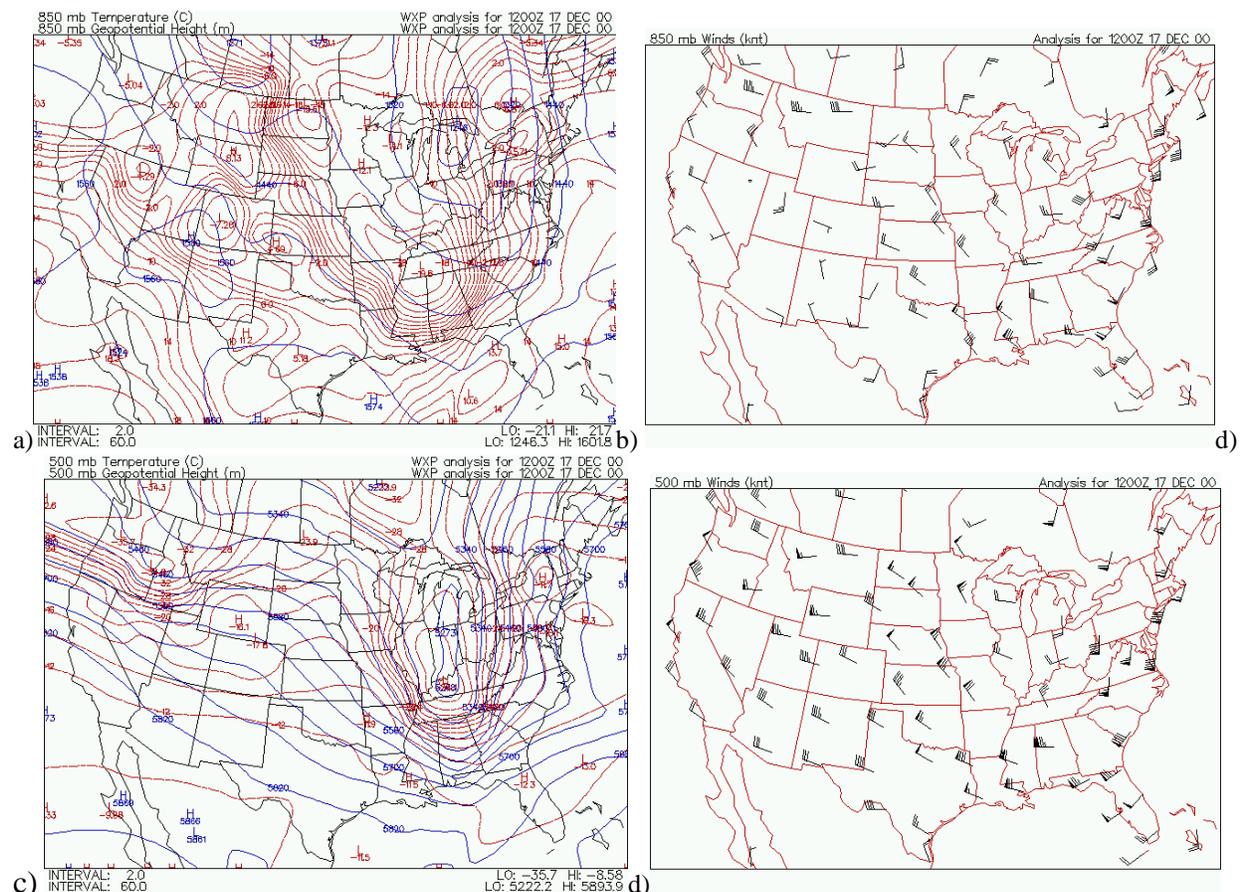
Figure 3-4: Upper air observations at 1200 UTC December 17, 2000. a) 850 mb geopotential height and temperature. b) 850 mb winds. c) 500 mb geopotential height and temperature. d) 500 mb winds. (Maps source: Plymouth State University Make Your Own Map website.)

The surface high over the ocean moved onshore over Idaho by 1200 UTC December 18 (Figure 3-5a) bringing calm and stable conditions to many of the western states. The temperature gradient over California was stronger than it was on the first day (Figure 3-5b) resulting in

stronger and more prevalent offshore, easterly flow (Figure 3-5c). At 850 mb (Figure 3-6a) a high pressure moved over northern Nevada, maintaining the stable conditions with easterly winds over the SFBA. At 500 mb the trough had deepened over the Midwest and a ridge had moved onshore over the western coast (Figure 3-6b) reinforcing the weak synoptic conditions over California.

a)

b)

c)

Figure 3-5: Surface observations at 1200 UTC December 18, 2000. a) Sea level pressure. b) Temperature. c) Winds. Maps made at Plymouth State University Make Your Own Map website.

Figure 3-6: Upper air observations at 1200 UTC December 18, 2000. a) 850 mb geopotential height and temperature. b) 500 mb geopotential height and temperature. Maps made at Plymouth State University Make Your Own Map website.

The surface high that was over Idaho had moved southeast over the Utah-Colorado border by 1200 UTC December 19 (Figure 3-7a), but still influenced calm and stable conditions to persist over California and the southwestern U.S. The temperature pattern remained fairly static from the previous day (Figure 3-7b), as offshore flow persisted along the coast, light southeasterly winds dominated over the interior SFBA, and northwesterly winds developed in the Central Valley. At 850 mb the high pressure center was directly positioned above the surface high over eastern Utah, allowing the weak synoptic conditions at this level to continue (Figure 3-8a). At 500 mb the ridge remained along the western states (Figure 3-8b).
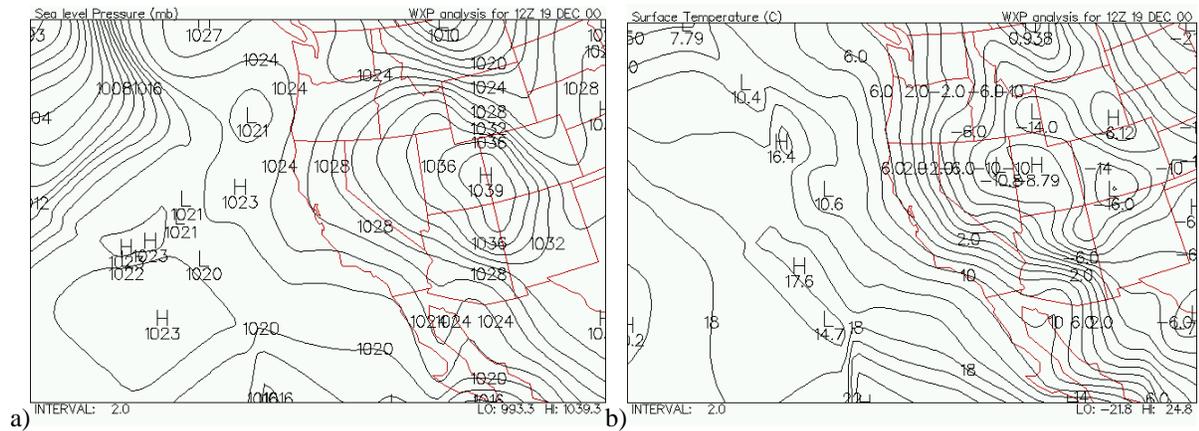
Figure 3-7: Surface observations at 1200 UTC December 19, 2000. a) Sea level pressure. b) Temperature. Maps made at Plymouth State University Make Your Own Map website.
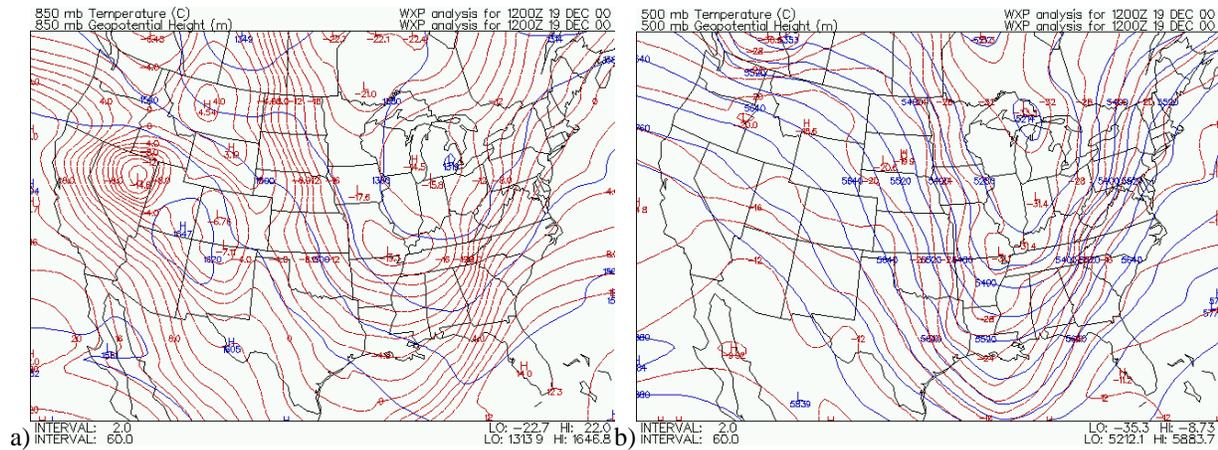
Figure 3-8: Upper air observations at 1200 UTC December 19, 2000. a) 850 mb geopotential height and temperature. b) 500 mb geopotential height and temperature. Maps made at Plymouth State University Make Your Own Map website.

By 1200 UTC December 20, the surface high that was over Utah the previous day had moved further east, while a ridge extended over California from a surface high located over southern Canada (Figure 3-9a). As with previous days, the temperature pattern remained fairly static (Figure 3-9b), easterly offshore flow persisted along the coast, but winds became calm in the Central Valley (Figure 3-9c). At 850 mb the closed high that was over Utah the previous day broke down, while a ridge formed over California, maintaining stable conditions (Figure 3-10a). At 500 mb the ridge persisted over the western coast of the U.S., reinforcing the stable conditions at the surface (Figure 3-10b).
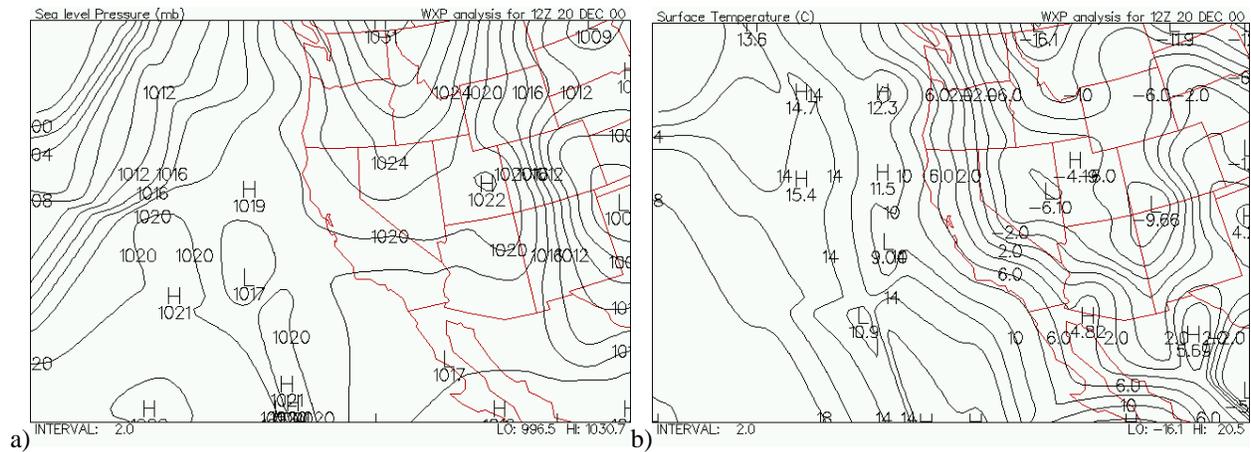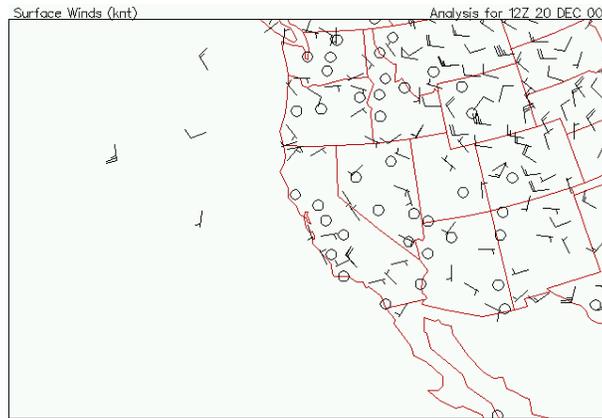
c)

Figure 3-9: Surface observations at 1200 UTC December 20, 2000. a) Sea level pressure. b) Temperature. c) Winds. Maps made at Plymouth State University Make Your Own Map website.

a) b)
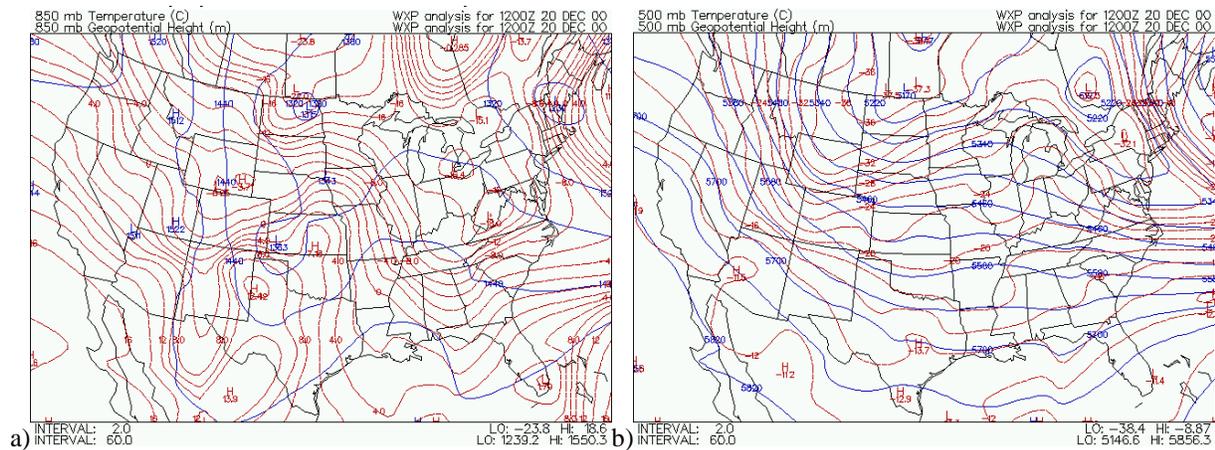
Figure 3-10: Upper air observations at 1200 UTC December 20, 2000. a) 850 mb geopotential height and temperature. b) 500 mb geopotential height and temperature. Maps made at Plymouth State University Make Your Own Map website.

By 1200 UTC December 21 the surface high moved over northern Colorado, with its associated ridge extending from the high to over California (Figure 3-11a). This was the last day of the 5-day study period as the synoptic pattern was about to change due to a low pressure system over the eastern Pacific Ocean that was moving toward the coast. The temperature gradient over California also weakened (Figure 3-11b) and the coastal winds became westerly (onshore) over the SFBA (Figure 3-11c). At 850 mb the ridge moved further east and a trough extended onshore from over the ocean (Figure 3-12a). At 500 mb the ridge continued over the western U.S., maintaining weak synoptic conditions at the surface (Figure 3-12b).

12

a)

b)

c)
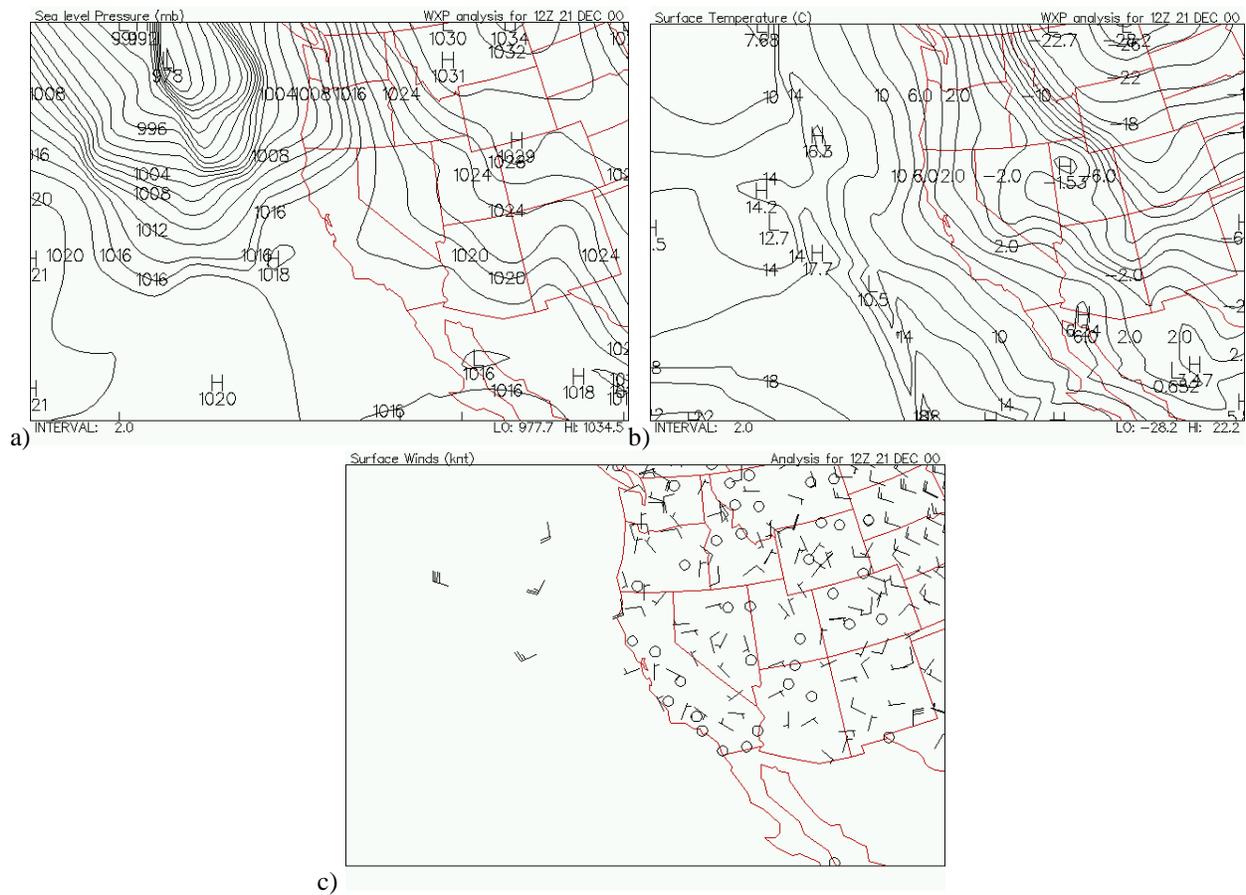
Figure 3-11: Surface observations at 1200 UTC December 21, 2000. a) Sea level pressure. b) Temperature. c) Winds. Maps made at Plymouth State University Make Your Own Map website.

a)

b)
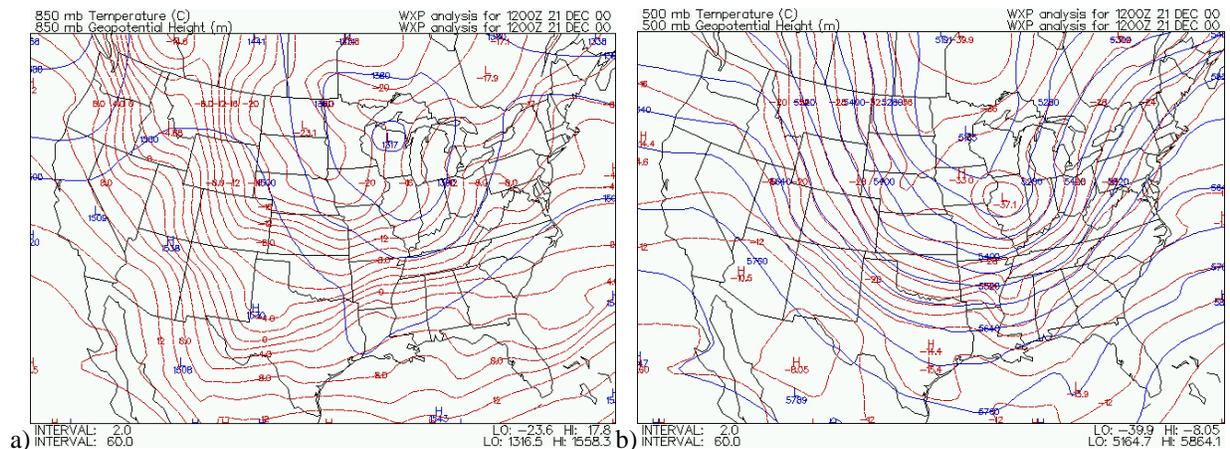
Figure 3-12: Upper air observations at 1200 UTC December 21, 2000. a) 850 mb geopotential height and temperature. b) 500 mb geopotential height and temperature Maps made at Plymouth State University Make Your Own Map website.

3.2   Summer Ozone Case

Many previous modeling studies focused on the summertime air quality applications (Seaman et al. 1995, Bao et al. 2008) due to general ozone concerns. Summertime weather

conditions usually consist of a high pressure ridge off the California coast and low pressures inland due to strong surface heating (Seaman et al. 1995). The ridge and the inland lows develop a pressure gradient that produces winds that flow onshore. In the afternoon a sea breeze often develops due to this pressure difference, strengthening onshore flow. The onshore winds bring cooler marine air over coastal areas such as the SFBA and Monterey, keeping temperatures over San Francisco and other coastal areas relatively cool. The marine air speeds up as it passes through the Carquinez Strait and other narrow gaps in the Coastal Ranges into the Central Valley. The air is modified by surface heating during day and the mixed layer increases in height from 200-600 m to 400-1200 m in the Central Valley (Seaman et al. 1995).

Summertime temperatures in the Central Valley can reach 40°C or higher, creating low pressure areas near Redding in northern Sacramento Valley and Bakersfield in southern San Joaquin Valley (Seaman et al. 1995). These low pressures create pressure gradients that cause the incoming marine flow to split near the Sacramento delta with southerly flow into the Sacramento Valley and northerly flow into the San Joaquin Valley. During the day, upslope winds along the Coastal Ranges and the Sierra Nevada cause divergence of the low level flows. At night, downslope winds from the mountains converge with the low level flows and, during the summer months, can produce a nocturnal low level jet at around 400 m above the surface that flows parallel to the San Joaquin Valley (Seaman et al. 1995). This jet develops because of a decoupling from the surface friction, allowing the air to accelerate.

The summer ozone case chosen for this study is the same case used by Bao et al. (2008). This period began at 12 UTC on July 29, 2000 with a surface thermally induced low south of the San Joaquin Valley and an inverted trough extending from the low northward over the western U,S. (Figure 3-13a). The temperatures were much warmer over the land than over the ocean, with the warmest temperatures under the thermal low (Figure 3-13b). This synoptic pattern lead to surface onshore flow along the coast with westerly winds averaging 5 kts over the SFBA and northwesterly winds within the Central Valley (Figure 3-13c). This westerly flow has the potential to transport ozone and it's precursors from the SFBA toward the Central Valley. At 850 mb high pressure was over northern California, leading to stable conditions over the state (Figure 3-14a) with very light offshore winds (Figure 3-14b). At 500 mb a broad high pressure area was centered over northern Arizona (Figure 3-14c), which sustained the weak synoptic conditions with very light winds over the region (Figure 3-14d).

At 1200 UTC July 30, conditions were similar to those on the 29th, with a thermal low over southern California (Figure 3-15a) and a positive temperature gradient from the cooler ocean to warmer land (Figure 3-15b). Conditions were much the same at the 850 and 500 mb levels as were observed on the 29th (Figure 3-16a, b). The same weakly-forced synoptic conditions persisted over the next several days (not shown), with the surface thermal low located over California, large high pressure system at the 850 and 500 mb levels over west coast. These weak synoptic conditions allowed for high concentrations of ozone to persist over the SFBA and Central Valley for multiple days.
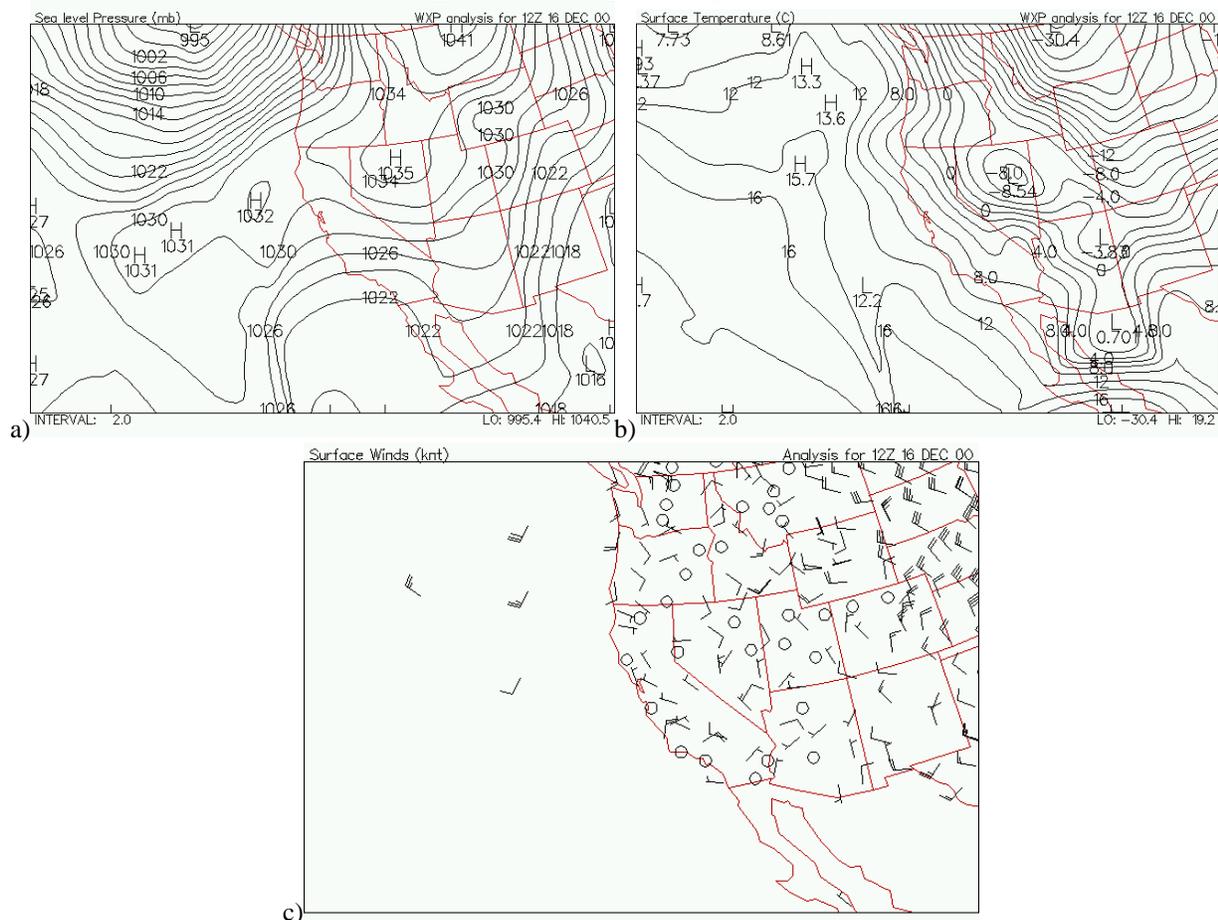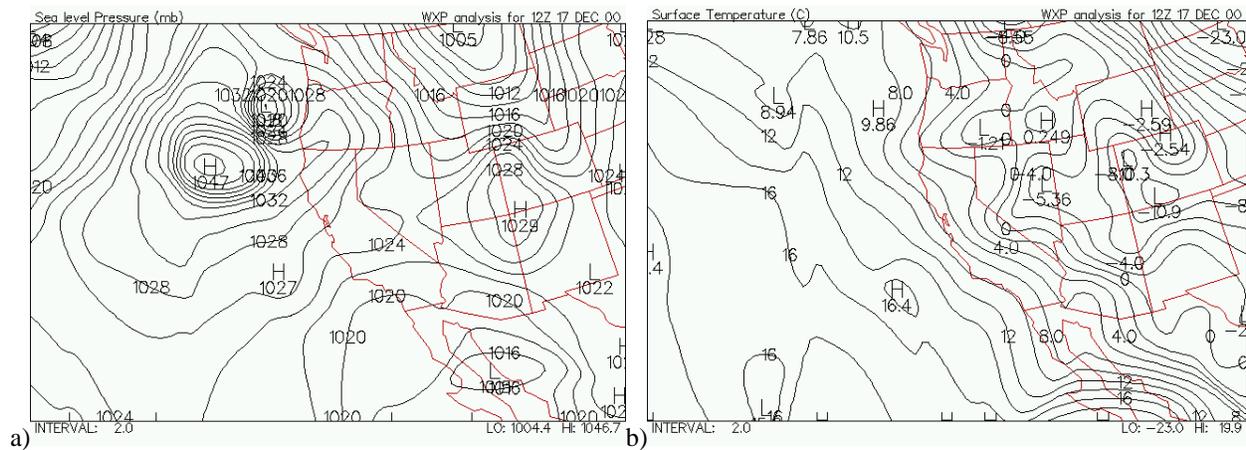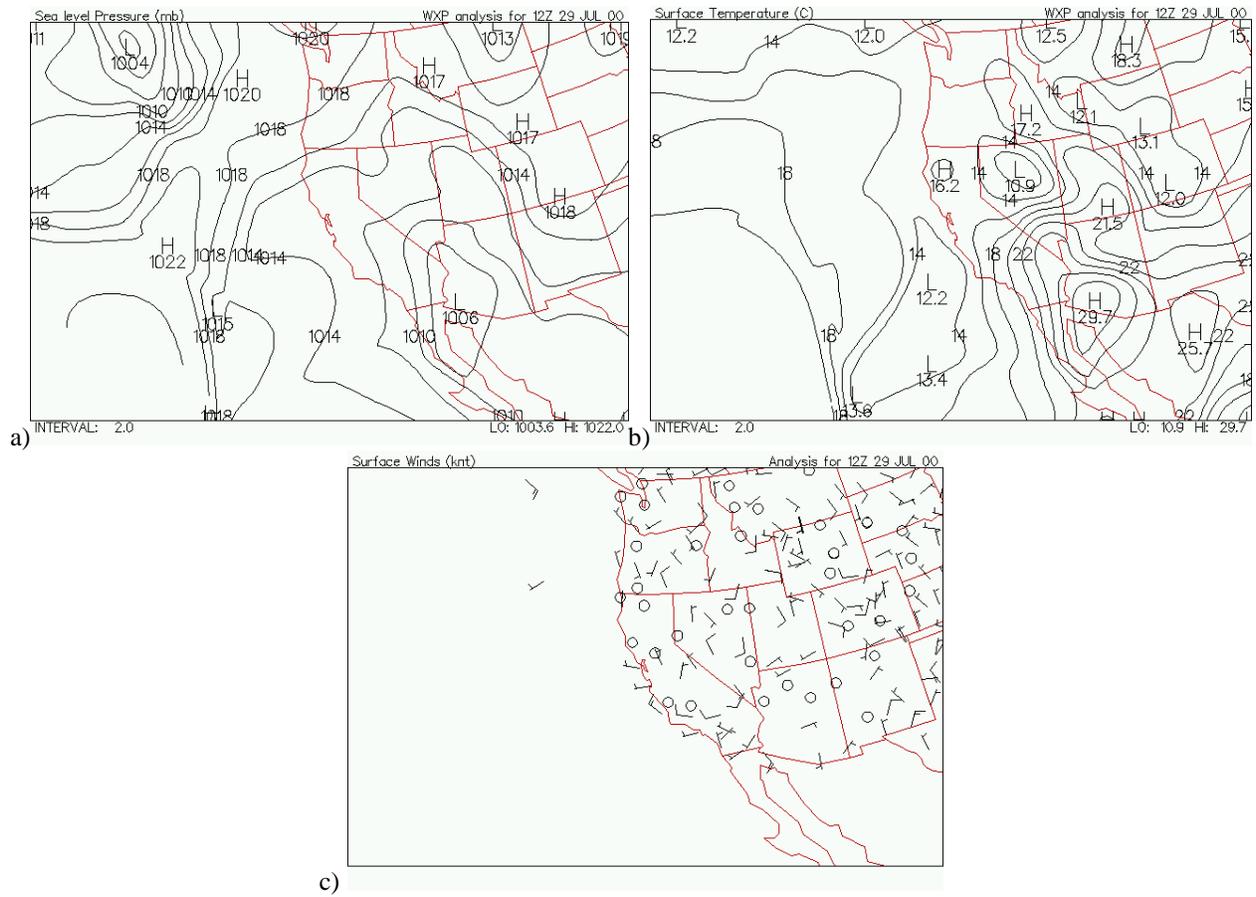
Figure 3-13: Surface observations at 1200 UTC July 29, 2000. a) Sea level pressure. b) Temperature. c) Winds. Maps made at Plymouth State University Make Your Own Map website.

c)

d)
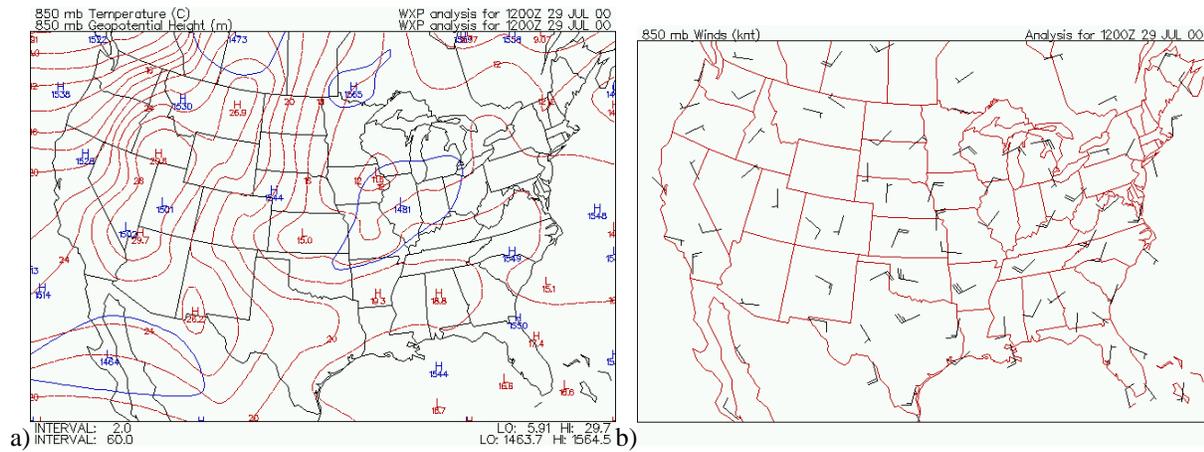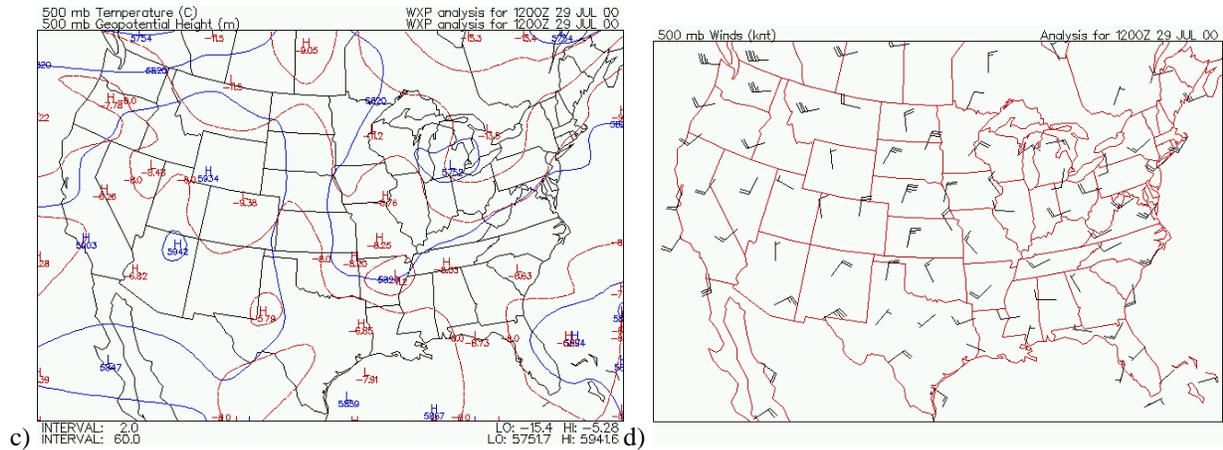
Figure 3-14: Upper air observations at 1200 UTC July 29, 2000. a) 850 mb geopotential height and temperature. b) 850 mb winds. c) 500 mb geopotential height and temperature. d) 500 mb winds. Maps made at Plymouth State University Make Your Own Map website.
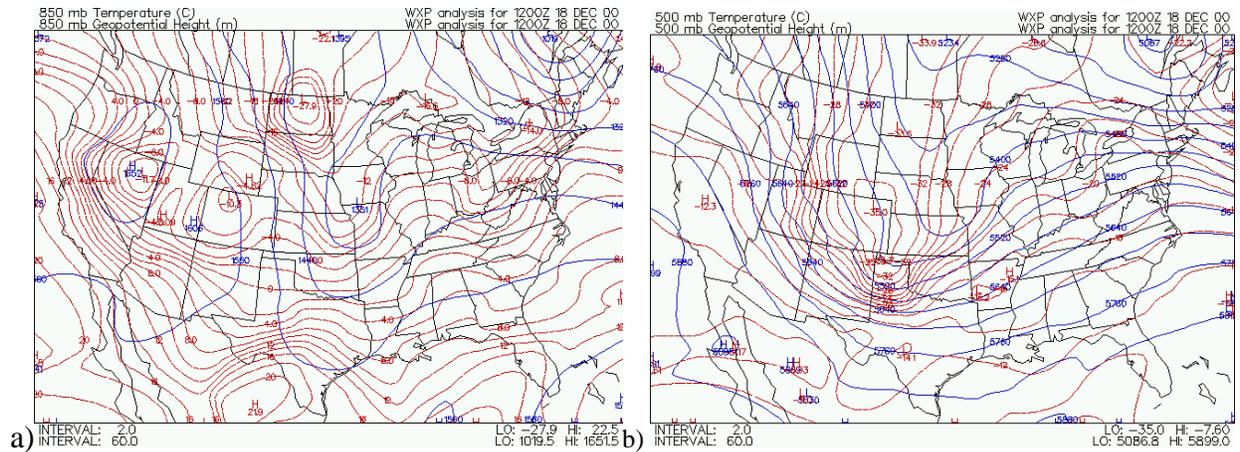

a)

b)

c)
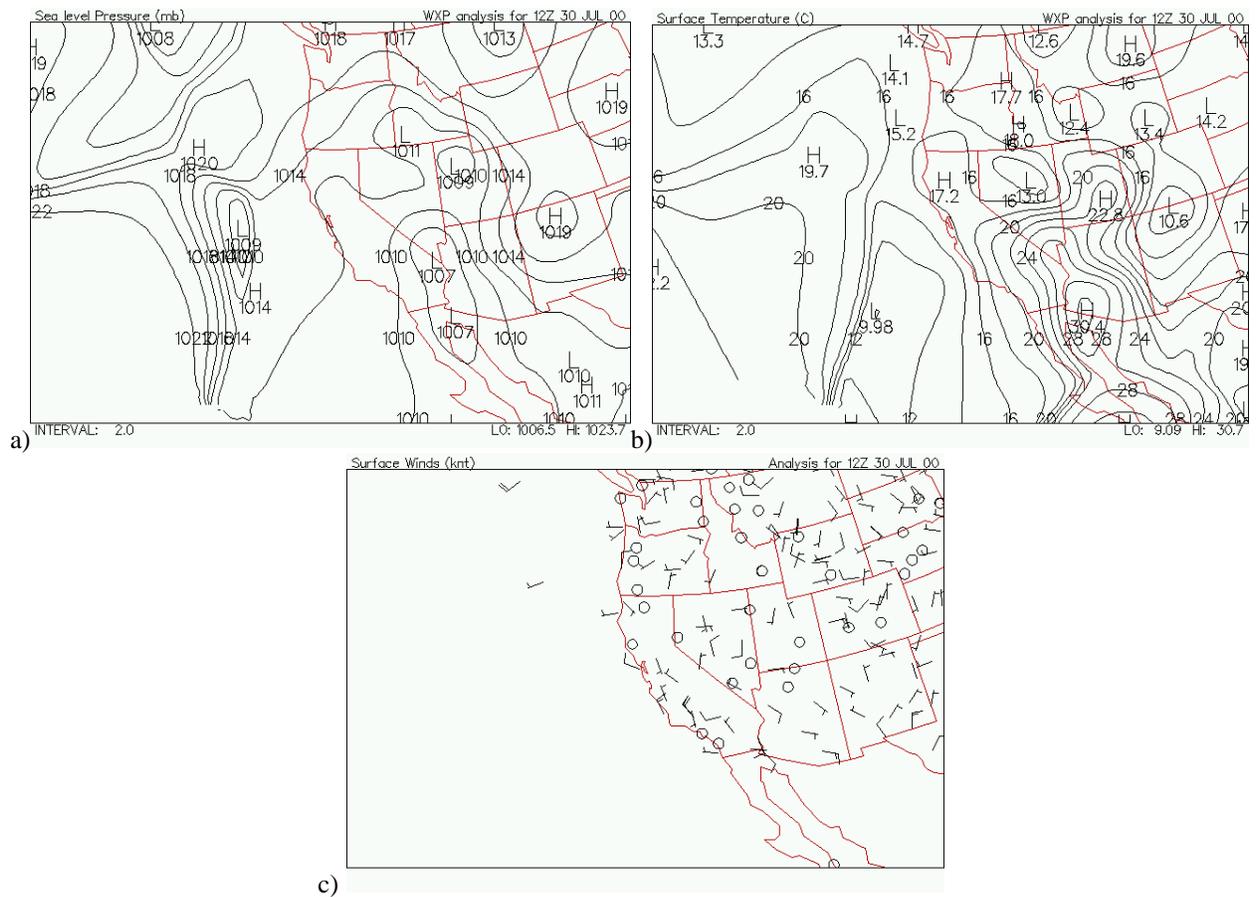
Figure 3-15: Surface observations at 1200 UTC July 30, 2000. a) Sea level pressure. b) Temperature. c) Winds. Maps made at Plymouth State University Make Your Own Map website

16

Figure 3-16: Upper air observations at 1200 UTC July 30, 2000. a) 850 mb geopotential height and temperature. b) 500 mb geopotential height and temperature. Maps made at Plymouth State University Make Your Own Map website.

4   EXPERIMENTAL DESIGN

### 4.1   *WRF-ARW Modeling Configuration*

For this study the model configuration was comprised of three domains: 36-km, 12-km, and 4-km grid spacing (Figure 4-1). The 36-km domain, with a mesh of 91x95 grid points, contained the entire western United States, parts of Mexico and Canada, and a large area of the eastern Pacific Ocean. The 12-km domain, with a mesh of 157x151 grid points, contained the entire state of California, the states of Oregon and Nevada, parts of Idaho, Utah, Wyoming, Arizona, and Montana, parts of Mexico and the Pacific Ocean. The 4-km domain, with a mesh of 190x190 grid points, contained the entire central California air quality modeling domain. It consisted of the SFBA and the Central Valley region that contained both the Sacramento Valley and the San Joaquin Valley. Fifty (50) vertical η layers were used in all numerical experiments for all grids (Table 4-1). The lowest half layer was located at approximately 12 m above ground level (AGL). The thickness of the layers increased gradually with height, with 27 layers below 850 mb (~1550 m AGL). The top of the model was set at 100 mb. One-way nesting was used for all experiments so that information from the coarse domains translated to the fine domains, but no information from the fine domains translated out to the coarse domains.
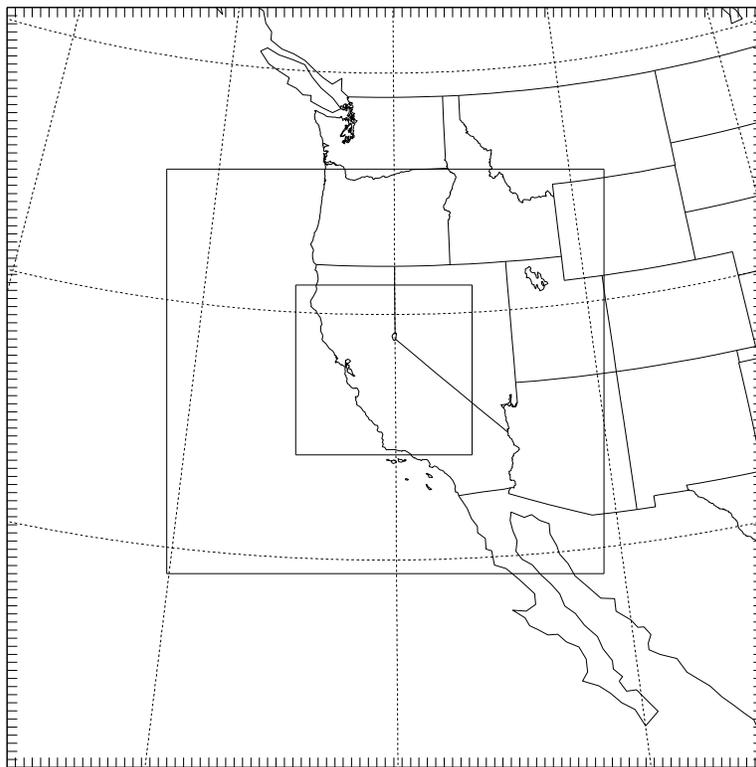
Figure 4-1: Nested domains for the model simulations showing the 4-km (innermost), 12-km (middle) and 36-km (outermost) domains.

18

Table 4-1: Distribution of the η layers

K	Eta	P(mb)	H_AGL(m)	dz(m)
50	0.0181	116.76	15400.39	1718.14
49	0.0583	153.98	13737.73	1607.19
48	0.1072	199.22	12170.61	1527.06
47	0.1661	253.76	10644.04	1526.07
46	0.2270	310.20	9294.66	1172.68
45	0.2839	362.85	8200.78	1015.09
44	0.3388	413.71	7253.09	880.28
43	0.3909	461.99	6431.96	761.98
42	0.4395	507.13	5722.33	657.28
41	0.4845	548.79	5109.89	567.60
40	0.5256	586.90	4581.26	489.67
39	0.5629	621.50	4125.60	421.64
38	0.5964	652.64	3733.58	362.41
37	0.6265	680.49	3396.33	312.07
36	0.6537	705.79	3100.36	279.88
35	0.6790	729.24	2834.35	252.14
34	0.7023	750.93	2595.06	226.44
33	0.7238	770.91	2380.08	203.52
32	0.7436	789.30	2186.39	183.88
31	0.7617	806.17	2012.35	164.20
30	0.7783	821.59	1856.21	148.09
29	0.7936	835.71	1715.78	132.77
28	0.8075	848.63	1589.44	119.92
27	0.8201	860.38	1476.12	106.72
26	0.8316	871.06	1374.45	96.61
25	0.8421	880.82	1282.67	86.95
24	0.8519	889.93	1197.82	82.76
23	0.8614	898.75	1116.25	80.37
22	0.8706	907.35	1037.47	77.19
21	0.8796	915.67	961.84	74.09
20	0.8882	923.71	889.27	71.04
19	0.8966	931.52	819.28	68.94
18	0.9048	939.10	751.79	66.05
17	0.9126	946.40	687.15	63.21
16	0.9202	953.47	625.02	61.06
15	0.9276	960.30	565.49	57.99
14	0.9347	966.91	508.60	55.79
13	0.9416	973.33	453.88	53.66
12	0.9482	979.42	402.40	49.31
11	0.9543	985.14	354.31	46.87
10	0.9601	990.54	309.01	43.72
9	0.9656	995.61	266.48	41.35
8	0.9707	1000.36	226.70	38.20
7	0.9754	1004.78	189.72	35.76
6	0.9799	1008.97	154.80	34.10
5	0.9841	1012.88	122.29	30.92
4	0.9880	1016.51	92.18	29.30
3	0.9917	1019.95	63.68	27.69
2	0.9952	1023.21	36.79	26.09
1	0.9985	1026.24	11.87	23.74

4.2 Data Sources and Description

In order to provide an effective FDDA on the 4-km domain, observational data with meso-beta (20-200 km) scale resolution is needed. The World Meteorological Organization (WMO) observations from the National Weather Service (NWS) have sufficient resolution for synoptic scale and meso-alpha (200-2000 km) scale phenomena, but not for the meso-beta scale phenomena that can occur within the Central Valley and over the SFBA. Therefore special observations that are collected by the BAAQMD are needed in addition to the WMO data.

The BAAQMD collected weather observations from networks of individual monitoring stations within each of the subregions: the SFBA, the Sacramento Valley, and the San Joaquin Valley (Beaver, 2008). No upper-air observations from these networks were used. In the SFBA region, 12 meteorological monitors provide hourly quality assured surface wind speed, wind direction, and temperature measurements. These monitors were operated by the BAAQMD and no additional quality control was necessary for these data.

In the Sacramento Valley region, observations were queried from the Aerometric Information Retrieval System (AIRS)/Air Quality System (AQS) database. All of the queried stations were operated by the Air Resources Board (ARB). There were 45 monitors that measured wind speed, wind direction, and temperature at the surface. For the San Joaquin Valley region, combinations of databases were queried to procure weather data from AQS, the California Irrigation Management Information System (CIMIS), the Remote Automated Weather Stations (RAWS) archives, and the National Climatic Data Center (NCDC). CIMIS and RAWS operate their own monitoring sites. NCDC is a repository of weather data from a variety of monitoring networks that includes NWS and Federal Aviation Administration (FAA) meteorological sites. Data from 18 monitors were queried from AQS, 19 from CIMIS, 11 from RAWS, and 8 from NCDC. The model simulations use the National Centers for Environmental Prediction -NCAR (NCEP-NCAR) 40-km Eta analyses for the IC/LBCs. The initial condition fields were further enhanced by radiosonde (balloon lifted weather package) and surface data through the WRF-ARW objective analysis process, OBSGRID (Deng et al. 2009), using a modified Cressman analysis (Benjamin and Seaman 1985).

4.3 Enhanced IC/LBCs with Observations

The Cressman analysis assigns a particular radius of influence to all the observations used for objective analysis. The first guess field at each grid point is adjusted for all the observations that influence the grid point. After the differences are calculated between the first guess fields and the observations, a distance-weighted average of these differences are added to the first guess fields at each grid point. After all the grid points are adjusted the new field is used as the first guess field for another analysis cycle. Subsequent cycles use a smaller radius of influence for the observations. The default Cressman scheme has explicitly assigned scale factors that determine the radii of influence (RIN) for the observations. The default scale factors are 5, 4, 3, and 2 grid increments. For each pass the scale factor is multiplied with the grid spacing on each domain to determine the radius of influence that each observation has.

For this study, the default scheme tended to give unrealistic patterns for the prognostic meteorological variables such as temperature (Figure 4-2a). The default scheme gave a bull's eye pattern for the 700 mb temperature innovation, or the difference between the observations and the first guess field. Not only did this pattern have temperature anomalies too small in area

coverage, but the pattern also did not reflect the anomalies across the entire domain (i.e., there were no temperature measurements along the California coast). Instead of using the default RINs, a new approach was used. This new approach was similar to that used in MM5 RAWINS and was recently implemented into OBSGRID (Deng et al. 2008). It automatically assigned the scale factors for each analysis cycle by explicitly setting the radius of influence to zero. For this study the scale factors that were automatically assigned by the OBSGRID program were 15 (or 540 km), 11, 8, and 6. The modified scheme gave more realistic patterns (Figure 4-2b). The temperature anomalies were more spread out and covered the entire domain.
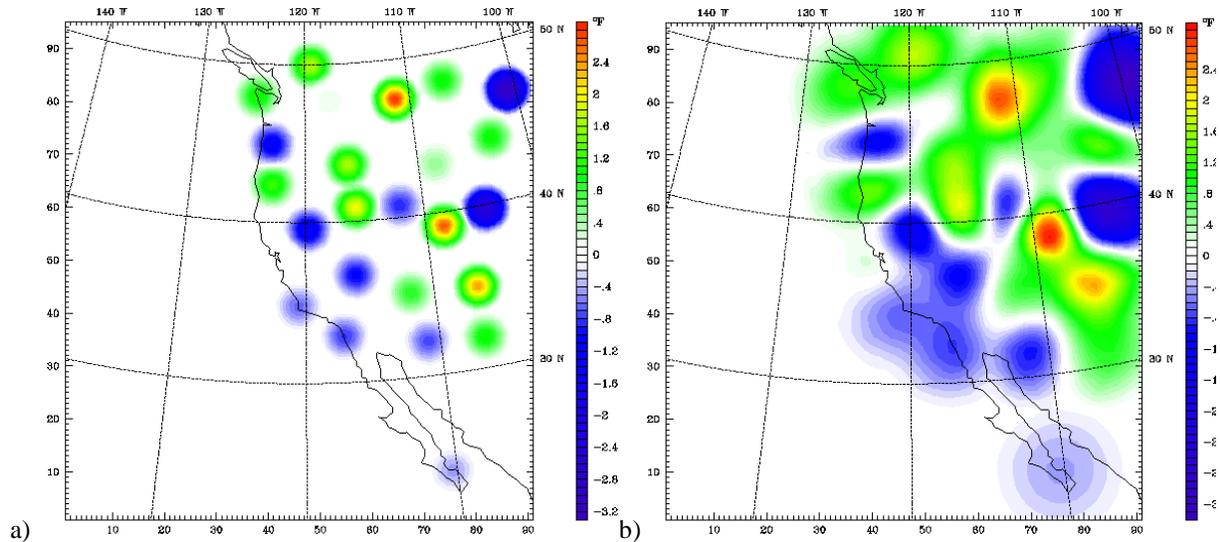
Figure 4-2: The 700 mb temperature innovation using the Cressman scheme in OBSGRID. a) Default scheme; (The scale factors for each analysis cycle: 5, 4, 3, 2). b) Modified Cressman scheme developed by Penn State; (The scale factors for each analysis cycle: 15, 11, 8, 6).

The lateral boundary conditions and three-dimensional (3D) analyses used for analysis FDDA w ere also enhanced by the objective analysis process and were defined at six hour intervals. Surface analysis fields used for surface analysis FDDA were generated by OBSGRID at three hour intervals. The 12 hourly upper air observations provided by NWS radiosondes, the surface observations from the NWS, and the special surface observations from about 90 stations located in the valleys provided by BAAQMD were quality-checked (QC) for erroneous data and observations that were not useful. In addition, all of these data were QC-ed by OBSGRID (using a high-resolution version of the WPS/UNGRIB software that included more pressure levels on which data were interpolated).These QC-ed observations (Figure 4-3) were needed for both obs nudging and model verification.
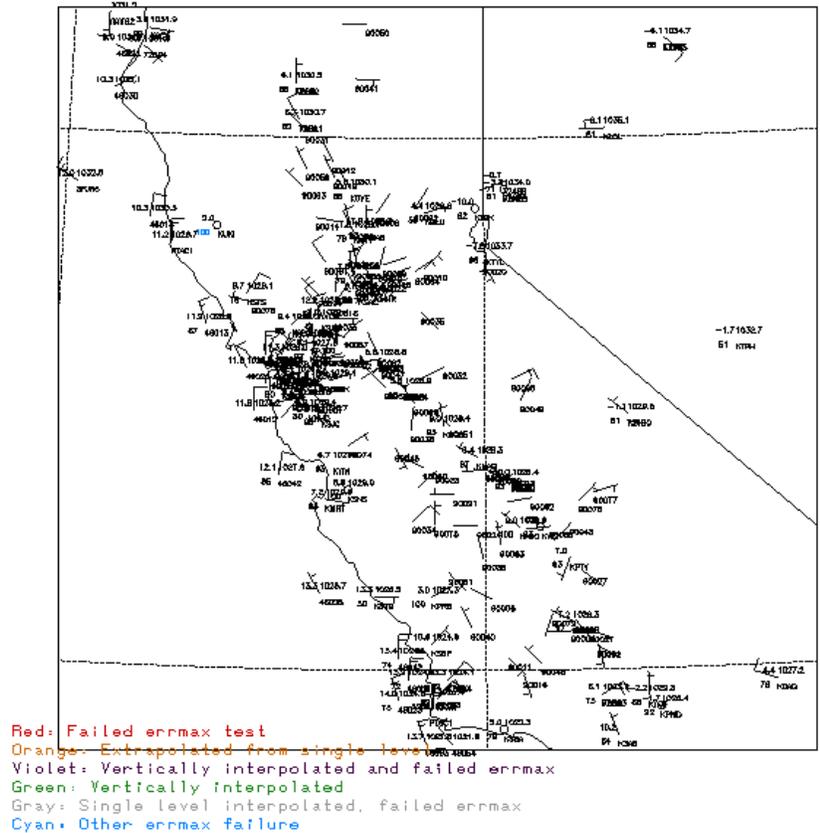
Figure 4-3: Observational data used for data assimilation for the winter PM case. WMO data and BAAQMD surface data for the surface.

4.4 Model Experiments

The purpose of this study was to determine the optimal WRF-ARW model configuration to provide the most accurate meteorological information for the Bay Area and the Central Valley regions. To achieve this goal various WRF-ARW experiments, with varying model physics and FDDA options, were conducted. The investigation started with comparing the use of two commonly used atmospheric radiation schemes, RRTM/Dudhia and RRTMG. The reason that the RRTM radiation was selected was that it was a mature scheme and was used by Penn State for many previous projects. As indicated later in the results section, it was decided that the RRTM scheme would be used for all the rest of the numerical experiments. The following step was done to determine an optimal LSM for the region since the CMAQ model used at the BAAQMD was customized to use the land surface fields as input to the air chemistry model. It was found for the winter PM case period that the PX physics had a clear advantage (see details later in the results section). For the summer the ozone case period, the results based on different LSM were quite mixed. Based on the BAAQMD's previous experience with PX physics, a decision was made to use the PX for all the rest of the experiments involving FDDA. As indicated in Section 2, for both the winter PM and the summer ozone cases, all FDDA

experiments used WSM3 simple ice microphysics, K-F cumulus parameterization on the 36- and 12-km grids, and ACM2 PBL scheme as part of the PX physics suite.

4.5   FDDA Experimental Design for the Winter PM Case

Using the best model physics found as the result of sensitivity study for atmospheric radiation and land surface processes as the baseline model configuration, a set of six model simulations were performed for the winter PM case period (Table 4-2): 1) NOFDDA, no data assimilation of any form was used; 2) GFDDA, 3D (excluding surface) analysis nudging was used on the 36-km and 12-km domains; 3) OFDDA, only obs nudging was used on all three domains, assimilating WMO and special wind profiler data; 4) MFDDA, an experiment combining 3-D analysis nudging (on the 36- and 12-km domains) and obs nudging (on all domains) in a multiscale FDDA framework; 5) MFDDA2, same as MFDDA experiment except surface analysis nudging was used, including the soil temperature nudging (Pleim and Gilliam 2009) that was automatically activated when the surface analysis nudging was used with PX physics; and 6), MFDDA2SP, same as MFDDA2 except soil temperature nudging were turned off. The purpose of the MFDDA2SP experiment was to evaluate the effects of assimilating the special surface observations of the BAAQMD meteorological network in addition to the WMO observations. The experimental design was motivated by the Penn State previous experience with FDDA in the MM5 modeling system that was applied over the SFBA region. It was necessary to evaluate each of the FDDA capabilities in the WRF-ARW model to come up with the best FDDA configuration for the region.

Table 4-2: FDDA configuration for the six FDDA model simulations for the winter case

Exp. Name	36 km		12 km		4 km	
	Analysis Nudging	OBS Nudging	Analysis Nudging	OBS Nudging	Analysis Nudging	OBS Nudging
NOFDDA	NO	NO	NO	NO	NO	NO
GFDDA	YES	NO	Yes	NO	NO	NO
OFDDA	NO	YES	NO	YES	NO	YES
MFDDA	YES (3D)	YES	YES (3D)	YES	NO	YES
MFDDA2	YES (3D+Sfc)	YES	YES (3D+Sfc)	YES	NO	YES
MFDDA2SP	YES (3D+Sfc)	YES	YES (3D+Sfc)	YES	NO	YES

The parameters used in the FDDA experiments are shown in Table 4-3. Nudging of the wind field was applied through all model layers, but nudging for the mass field was only allowed above the model-simulated PBL so that the PBL structure produced by the model was dominated by the model physics. A time window of two hours was used in obs nudging for upper air observations, with a reduced window of one hour at the surface. The radius of influence for surface data was reduced as suggested by Seaman et al. (1995) (i.e. by multiplying a factor of 0.5 to the specified value in Table 4-3) during the obs nudging process. Note that as indicated in Table 4.4, in the multiscale FDDA framework, the analysis nudging was applied on the 12-km grid with reduced strength (from 0.0003 to 0.0001). This was done to allow the FDDA to be effective in reducing the model errors on the 4-km domain but at the same time not cause the meso-beta scale features to be smoothed out by the FDDA.

23

Table 4-3: FDDA parameters used for the winter case. Note that these parameters were also used for the summer case presented later except that there was no observational nudging on the 36-km domain

	Analysis Nudging			OBS Nudging		
	36km	12km	4km	36km	12km	4km
G (1/sec)	0.0003	0.0001	N/A	0.0004	0.0004	0.0004
3-D wind field	Nudging all layers	Nudging all layers	N/A	Nudging all layers	Nudging all layers	Nudging all layers
3-D mass field	Nudging above PBL	Nudging above PBL	N/A	Nudging above PBL	Nudging above PBL	Nudging above PBL
Sfc wind field	Used within PBL	Used within PBL	N/A	Used within PBL	Used within PBL	Used within PBL
Sfc mass field	Not used	Not used	N/A	Not used	Not used	Not used
RINXY (km)	N/A	N/A	N/A	150	100	100
TWINDO (hr)	N/A	N/A	N/A	2	2	2
dt (sec)	N/A	N/A	N/A	180	60	20

4.6 FDDA Experimental Design for the Summer Ozone Case

Based on the lessons learned from the winter case, there was no need to perform the analysis nudging only and obs nudging only experiments. A set of three model simulations were conducted for the summer ozone case (Table 4-4): 1) NOFDDA, no data assimilation of any form was used; 2) MFDDA4, similar to the MFDDA2 experiment from the winter case, and only the WMO observations were assimilated; and 3) MFDDA5, same as MFDDA4, except the simulation also assimilated the BAAQMD special surface observations in addition to the WMO observations. The MFDDA5 experiment was designed to show the added value of assimilating the special surface observations of the BAAQMD meteorological network.

Table 4-4: FDDA configuration for three FDDA model simulations for the summer case.

Exp. Name	36 km		12 km		4 km	
	Analysis Nudging	OBS Nudging	Analysis Nudging	OBS Nudging	Analysis Nudging	OBS Nudging
NOFDDA	NO	NO	NO	NO	NO	NO
MFDDA4	YES (3D+Sfc)	NO	YES (3D+Sfc)	YES	NO	YES
MFDDA5	YES (3D+Sfc)	NO	YES (3D+Sfc)	YES	NO	YES

The FDDA parameters were the same for the winter case shown in Table 4-3 except there was no observational nudging on the 36-km domain. Such experimental design was based on the request from BAAQMD because obs nudging was not usually used on their 36-km coarse domain. The observational nudging was turned off on this domain to gauge how the 4-km domain was affected when there was no observational nudging information being passed from the 36-km domain to the 12-km domain, which in turn passed information to the 4-km domain. As expected, the difference in the 4-km WRF-ARW solutions between the experiment with and without 36-km obs nudging was minimal and could be neglected.

24

5   METEOROLOGICAL MODEL RESULTS

The results of the numerical simulations will be examined in this section.   The verification strategies used for objective analysis and subjective analysis will be presented first. This will be followed by the objective verification of the model parameters such as RH, temperature, wind speed, and wind direction. The subjective, visual comparisons of the simulations with the observations will be presented last.   The model configuration that has the smallest statistic error and overall best subjective verification will be considered to be the best configuration to keep.

5.1   Verification Strategy

5.1.1 Objective Verification

The SCATTERPLOTS statistical software developed at NCAR was used for the objective analysis. This program took the model output and the observations used for objective analysis and data assimilation, and created pairs of the observed values and model values that were interpolated to the locations of the observations. Mean absolute error (MAE) and mean error (ME) statistics were made from the pairs.

The MAE and ME equations can be found in APEXDIX of this report. The MAE was used to measure how close the model values were compared to the observed values. The ME measured the bias of the model values compared to the observed values. Theses biases were calculated for the RH, temperature, wind speed, and wind direction.

For this study the original SCATTERPLOTS software was expanded to include wind direction statistics. Details of this implementation can be found in APPENDIX. A calm wind threshold was used for this study to remove calm winds or very light winds (less than or equal to 1 m s^{-1}) for wind direction statistics calculation because the wind direction for near calm wind was uncertain and could produce large wind direction errors.

5.1.2 Subjective Verification

Horizontal charts of the simulated winds, temperature, pressure and geopotential heights were made at the surface, 850 mb, and 500 mb in order to subjectively analyze how well the model was able to simulate these meteorological variables. The analysis began with comparing the control simulation, which did not use any data assimilation and only included the observations for the IC/LBCs with the observations first in order to establish a baseline model configuration, followed by comparisons of the FDDA simulations with the observations and the baseline.

The simulations were also analyzed to determine if the model could capture and reasonably simulate the mesoscale features that were specific to the study region, such as the upslope and downslope flows along the valley walls and air flows into and out of the Central Valley. The winter PM case and the summer ozone case each had different meteorological features that were important either high PM or ozone. The features that were analyzed for the winter case were easterly flow from the Central Valley into the Bay Area, downslope flow over the Central Valley walls to the valley floor, and the diurnal cycle of the surface temperatures throughout the study domain. The features important for the summer case were westerly flow

from the Bay Area into the Central Valley, westerly flow between the Pacific Ocean and the Central Valley, the mean flows along the major axes of the Central Valley, and the Fresno Eddy and Schultz Eddy (Lin and Jao 1995, Bao et al. 2008). Close-up charts for specific locations within the study domain such as the Bay Area, the Sacramento Valley, and the San Joaquin Valley were made to compare the simulated winds and temperature to the specific observation stations. These close-up plots also helped with the analysis of the mesoscale features in these specific regions. Time series plots of simulated and observed winds and temperature were made to help measure how well the modeled and observed values compared at individual locations.

5.2   Winter PM Case Results

5.2.1 Objective Analysis Results

Although subjective analysis qualitatively compares the model simulations with the observations, the simulations also need to be evaluated quantitatively to measure the model performance. This section provides the statistical (MAE and ME) comparisons between the model experiments and the observed data (including both assimilated and the independent obs). These statistics comparisons are used to determine the optimal model configuration in our study.

The investigation begins with comparing the WRF-ARW solutions between using the RRTM and RRTMG radiations schemes because radiation plays a very important role in air pollution concentrations. The heating of the atmosphere and the surface is determined by the amount of incoming solar (shortwave) radiation that is received and the amount of longwave radiation that is absorbed and emitted by gases and the surface. The solar radiation and the temperature determine the amount of volatile organic compounds (VOCs) released by plants to interact with ozone or other particulate matter (PM) to create secondary aerosols and other compounds. Therefore it is important to accurately calculate the flux of shortwave and longwave radiation.

5.2.1.1 Entire Domain Verification

Table 5-1 shows the MAE of the WRF-simulated surface-layer RH, temperature, wind direction and wind speed between the simulation using RRTM radiation scheme and simulation using the RRTMG radiation scheme. The statistics were averaged over the entire domains and simulation period. The model results were compared against the NWS surface and radiosonde observations. Both of the experiments used the identical physics (e.g. MYJ PBL scheme, WMS3 simple ice microphysics, and K-F CPS on the two coarser domains) except in the radiation schemes. Neither of the experiments used FDDA and both used objective analysis for the IC/BCs. It was found that both radiative schemes produced similar results for all verification fields (i.e. wind speed, wind direction, temperature and water vapor mixing ratio), with a slight degradation shown in some fields in the RRTMG experiment. For example, both experiments produced RH errors between 16% and 18% with the RRTMG experiment performing slightly better than the RRTM experiment on all three domains (Table 5-1). Both experiments had about the same temperature errors on the 36-km and 4-km domains with the RRTM experiment performing slightly better than the RRTMG experiment on the 36- and 12-km domain. The results for the wind direction and the wind speed were also similar between the two simulations, with the RRTMG experiment performing slightly better than the RRTM experiment on all three

domains. Both experiments produced wind direction errors around 50 degrees and wind speed errors between 1.6 m s^{-1} and 2.5 m s^{-1}.

Table 5-1: Surface MAE Statistics comparing two atmospheric radiation schemes

	36 km	*12 km*	*4 km*
Relative Humidity			
RRTM	16.7	18.4	16.9
RRTMG	16.5	17.9	16.1
Temperature			
RRTM	2.8	2.6	2.5
RRTMG	2.8	2.7	2.5
Wind Direction			
RRTM	49.4	52.8	48.4
RRTMG	49.1	52.6	48.2
Wind Speed			
RRTM	2.5	1.9	1.6
RRTMG	2.5	1.9	1.6

Table 5-2: Upper Air MAE Statistics comparing two atmospheric radiation schemes

	36 km	**12 km**	**4 km**
Relative Humidity			
RRTM	15.1	14.6	12.4
RRTMG	14.0	13.9	12.6
Temperature			
RRTM	1.5	1.5	1.4
RRTMG	1.7	1.6	1.5
Wind Direction			
RRTM	18.6	20.0	21.3
RRTMG	15.9	17.0	17.3
Wind Speed			
RRTM	3.4	3.0	2.7
RRTMG	3.7	3.2	2.8

Table 5-2, similar to Table 5-1, shows the WRF-ARW-simulated MAE statistics of the radiation schemes for the upper air. The upper air statistics were averaged over the entire domain every twelve hours because radiosonde observations were only taken twice a day. The differences in the errors between the two experiments were also similar for the upper air as at the surface. The RRTMG experiment performed better than the RRTM experiment on the 36-km domain and the 12-km domain for the RH, but slightly worse on the 4-km domain. Both experiments produced RH errors between 13% and 15%. For temperature, the RRTM experiment performed better than the RRTMG experiment on all three domains for the temperature, with both experiments producing temperature errors around 1.5 K. For wind direction, the RRTMG experiment performed better than RRTM experiment for the wind direction on all the domains, with both experiments producing wind direction errors between 15 degrees and 20 degrees. The wind speed results were similar to the temperature results with the RRTM experiment performing better than the RRTMG experiment on all three domains. The

range of the wind speed errors for both experiments was between 2.7 m s^{-1} and 3.7 m s^{-1}.

Each radiation scheme performed better than the other for a given meteorological field; therefore the choice of scheme was dependent on which fields were more important for this study. Because most high pollution episodes occurred during weakly-forced, fair weather conditions, WRF-ARW with varying atmospheric radiation schemes was less likely to produce dramatically different solutions. Therefore the RRTM radiation scheme, in combination with the Dudhia shortwave scheme (Dudhia 1989), was chosen as the atmospheric radiation scheme for all the numerical experiments made during this study.

The evaluation of land surface models in this study was critical, as calculated surface heat and moisture fluxes, determined the amount of vertical transport of pollutants, heat, and moisture within and potentially above the PBL, possibly leading to long range transport. To determine an optimal LSM to use, the MAEs of the WRF-ARW-simulated fields were compared among the 5-layer thermal diffusion scheme, the Noah LSM, the RUC LSM, and the Pleim-Xiu (PX) LSM.

Table 5-3 shows the MAE statistics of the WRF-ARW-simulated surface-layer fields when land surface models (LSMs) were used. All the experiments used observations included in the initial condition though objective analysis, the RRTM scheme, and FDDA were not used. For surface RH, the 5-layer thermal diffusion scheme produced the largest errors on the 36-km and 12-km grids. The Pleim-Xiu (PX) LSM produced the smallest errors on all domains, except on the 4-km domain, where RUC LSM outperformed the other schemes. The Noah LSM performed better than the thermal diffusion LSM on the 36-km and 12-km domains and was superior to the RUC on the 36-km domain; however, it produced the largest error on the 4-km domain. Overall, RH errors were between 14% and 20%, generally increasing as the grid resolution increased.

Table 5-3: Surface MAE Statistics comparing four land surface models

	36 km	12 km	4 km
Relative Humidity			
Thermal Diff.	16.7	18.4	16.9
Noah	14.4	16.3	19.1
RUC	14.9	15.8	16.6
Pleim-Xiu	13.9	15.3	17.8
Temperature			
Thermal Diff.	2.8	2.6	2.5
Noah	2.9	2.7	2.5
RUC	2.8	2.6	2.5
Pleim-Xiu	2.7	2.4	2.0
Wind Direction			
Thermal Diff.	49.4	52.8	48.4
Noah	49.1	52.5	47.3
RUC	48.9	52.4	48.7
Pleim-Xiu	49.0	52.1	46.8
Wind Speed			
Thermal Diff.	2.5	1.9	1.6
Noah	2.6	2.0	1.6
RUC	2.4	1.8	1.6
Pleim-Xiu	2.2	1.7	1.5

For surface temperature, the PX LSM produced the smallest errors on all three domains. The other three LSMs produced about the same magnitude of error on all the domains with the Noah LSM having the largest error on the 36-km and 12-km domains. The temperature errors were between 2 K and 3 K.

As shown in Table 5-3, the PX LSM again produced the smallest error for wind direction on the 12-km and 4-km domains, while the RUC LSM produced the smallest error on the 36-km domain, but the largest error on the 4-km domain. Wind direction errors were between 46° and 53°. Wind direction errors were largest on the 12-km domain. For wind speed, the PX LSM once again produced the smallest error on all the domains, but only by a slight margin. Wind speed errors were between 1.5 m s^{-1} and 2.5 m s^{-1}.

Table 5-4 shows the MAE statistics of the WRF-ARW-simulated upper-air fields for the four LSMs. The PX LSM produced similar error statistics for RH as compared to the other LSMs, with a range of 12% to 15%. The temperature MAEs for all the LSMs were similar, falling between 1.5 K and 1.7 K, with the Noah and RUC LSM having the larger errors. For wind direction, the PX LSM again produced the smallest errors on all three domains. The thermal diffusion LSM again produced the largest errors on all the domains. Wind direction errors were between 15 degrees and 22 degrees. For wind speed, the thermal diffusion LSM produced the smallest errors on all the domains, similar to its performance for temperature. The other three LSMs performed similarly, all performing only slightly worse than the thermal diffusion LSM. The wind speed errors were between 2.7 m s^{-1} and 3.7 m s^{-1}.

Table 5-4: Upper Air MAE Statistics comparing four land surface models

	36 km	12 km	4 km
Relative Humidity			
Thermal Diff.	15.1	14.6	12.4
Noah	13.7	13.3	12.2
RUC	13.8	13.3	12.1
Pleim-Xiu	13.6	13.3	12.3
Temperature			
Thermal Diff.	1.5	1.5	1.4
Noah	1.7	1.6	1.5
RUC	1.7	1.6	1.5
Pleim-Xiu	1.6	1.6	1.4
Wind Direction			
Thermal Diff.	18.6	20.0	21.3
Noah	16.0	17.0	17.9
RUC	15.8	16.7	17.9
Pleim-Xiu	15.8	16.6	17.6
Wind Speed			
Thermal Diff.	3.4	3.0	2.7
Noah	3.7	3.3	2.8
RUC	3.7	3.2	2.8
Pleim-Xiu	3.6	3.2	2.8

Comparing the PX results with those from the thermal diffusion, Noah, and the RUC schemes revealed that using the PX physics generally produced better surface statistics. For

upper air, the PX, Noah, and RUC schemes all performed better than the 5-layer thermal diffusion scheme for RH and wind direction. Overall, the PX LSM method outperformed the other LSM options. Therefore, for this study, the model simulations using FDDA used the PX LSM as part of the baseline configuration, along with the RRTM radiation scheme.

FDDA can improve the model simulations by using analyses and observations to suppress the model error growth. A combination of nudging methods can be utilized to improve the model results: (1) analysis nudging can help accurately capture the synoptic flows by using gridded analyses to keep the model close to the observed state, (2) obs nudging can help to accurately capture the mesoscale flows by using individual surface and upper air observations to keep the model close to the observed state, and (3) multiscale nudging, a combination of analysis nudging, applied on the coarser grids to ensure the model solution is in phase with the larger-scale analysis and obs nudging, used on finer grids, to capture/retain mesoscale features.

Figure 5-1 shows the WRF-simulated surface MAE statistics and the surface ME statistics for RH for six numerical experiments that used the combination of nudging FDDA methods described above, averaged over each domain. The figure is an improvement trend graph where the worst model simulation is on the left and the best is on the right. All of the experiments that used FDDA assimilated individual WMO observations in obs nudging and/or assimilated analyses generated by OBSGRID using the WMO observations and 40-km Eta analysis. All six experiments were compared against both the WMO observations and the data from the BAAQMD's network. Note that only MFDDA2SP assimilated both WMO and the BAAQMD special surface data (Table 4-2). The MAE shows the average error each experiment produced compared to the observations, while the ME indicates the bias that each experiment had compared to the observations.

Figure 5-1a shows that for the MAE statistics there was an improvement in results from the use of FDDA, even though nudging of the mass fields were excluded from the PBL. Figure 5-1b shows that there was a negative (drier) bias for RH on all three domains for RH, except for the GFDDA and the MFDDA experiments on the 12-km domain. Note that the biases were only a few percent on the 36- and 12-km domains and about 10 percent on the 4-km domain. This meant that the experiments tended to be drier than the observations, which lead to the development of fewer clouds, but higher ozone and PM concentrations.
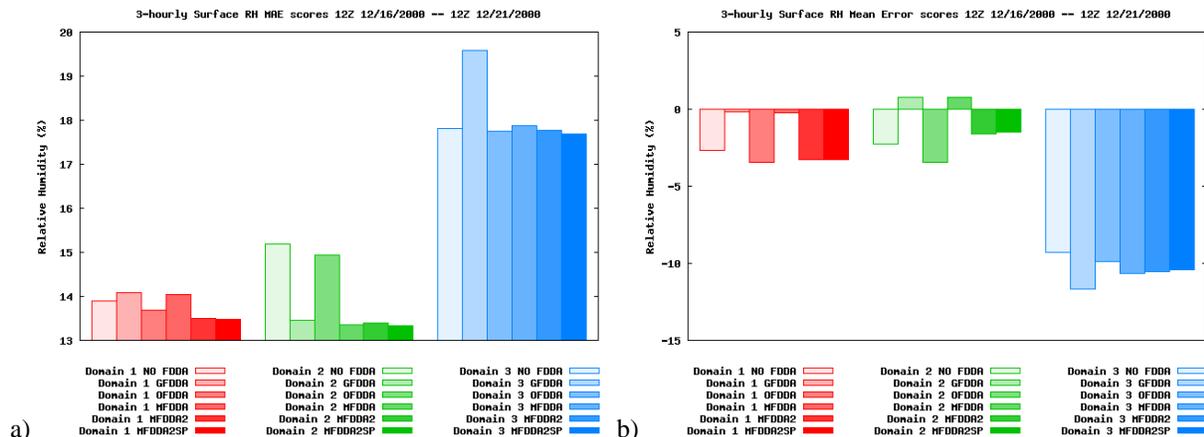
Figure 5-1: Surface statistics for the FDDA numerical experiments for RH. a) MAE. b) ME.

Figure 5-2 shows the surface MAE and ME statistics for temperature. Although there was no assimilation of surface temperature, the use of FDDA reduced the model errors on all the domains except for the GFDDA and the MFDDA experiments on the 36-km domain (Figure 5-2a). Using multiscale FDDA, analysis nudging (3D & surface) and obs nudging produced the best results, as shown by the MFDDA2 and MFDDA2SP experiments. The largest error was only 0.4 K, which was due to temperature is not being assimilated at the surface. The ME for temperature (Figure 5-2b) shows that the MFDDA2 and MFDDA2SP tended to be warmer on average than the observations on all the domains. However, the warmer temperature biases were quite small on the 12- and 4-km domains. The GFDDA and MFDDA experiments had larger cold biases than the warm biases on the 36-km and 12-km domains, although no temperature field was nudged at the surface in both experiments. Note that these statistics from domain to domain were not based on geographical location.
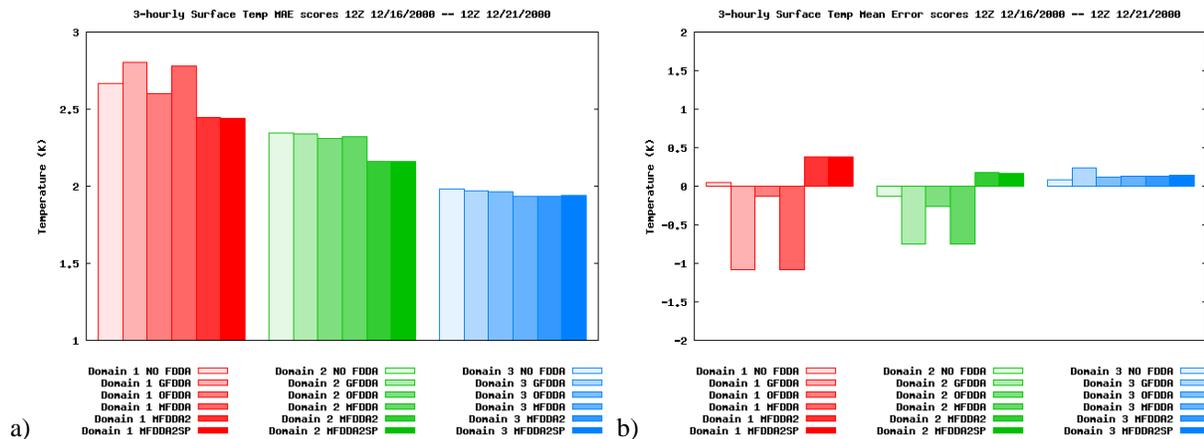
Figure 5-2: Surface statistics for the FDDA numerical experiments for temperature. a) MAE. b) ME.

Figure 5-3 shows the surface MAE and ME statistics for wind direction. The use of FDDA reduced the wind direction MAE errors (Figure 5-3a) on all three domains except for the GFDDA experiment on the 4-km domain where analysis nudging was not used. Using multiscale FDDA significantly reduced the wind direction MAE errors compared to the NOFDDA experiment; the MFDDA2SP experiment reduced errors by about 15 degrees. The errors were significantly reduced because the winds were assimilated at the surface, unlike moisture and temperature. The ME shows that all six experiments tended to have wind directions greater than what was observed (Figure 5-3b). A wind direction error of a few degrees could potentially have a significant impact on the transport of ozone and PM and misplace these pollutants by hundreds of kilometers.

Figure 5-3: Surface statistics for the FDDA numerical experiments for wind direction. a) MAE. b) ME.

Figure 5-4 shows the surface MAE and ME statistics for the wind speed. Since this was a synoptic weakly-forced case with light winds, the wind speed MAE errors were generally small, as expected, even without FDDA (Figure 5-4a). FDDA further reduced MAE errors, except in the OFDDA simulation on the 36- and 12-km domains. The degradation was likely caused by quality issues with the observed data, discussed later in the report. Further investigation is needed. The multiscale experiment, MFDDA2SP, produced the smallest wind speed MAE errors. The ME scores (Figure 5-4b) show that the 12-km and 4-km experiments in general tended to be faster than the observations. This could lead to ozone and PM being transported farther downstream than where they actually would occur. However, the MFDDA2SP experiment tended to have wind speeds slower than the observations on the 36-km and the 12-km domains and slightly faster on the 4-km domain, and had the smallest wind speeds bias on the 4-km domain.
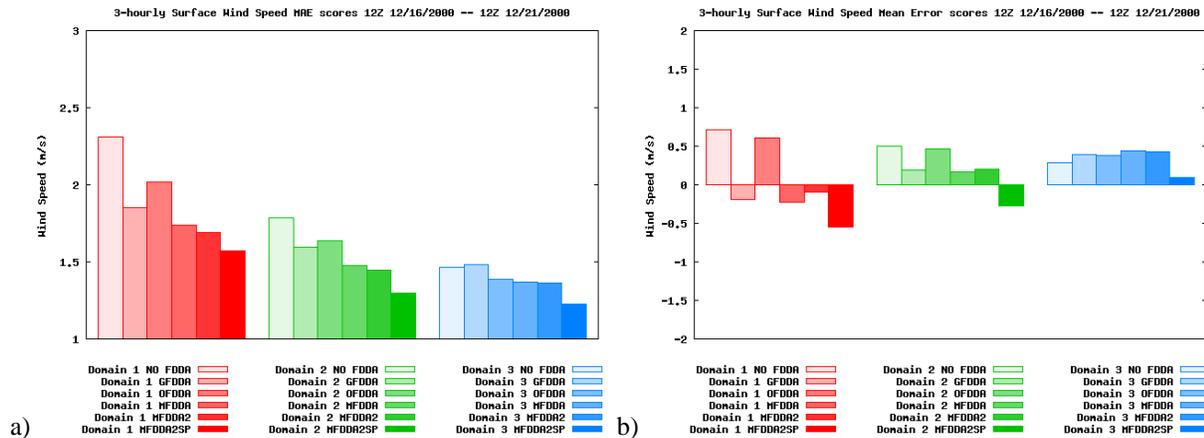
Figure 5-4: Surface statistics for the FDDA numerical experiments for wind speed. a) MAE. b) ME.

MFDDA2SP proved to be the best configuration to use as it included both multiscale FDDA and surface data assimilation. Overall, at the surface, the model did a reasonable job of simulating moisture, temperature, and winds. Using FDDA further reduced the errors created within the model and improved the simulations as a whole. Analysis nudging and obs nudging both improved the model simulations, depending on which domain either was used on. In addition, assimilating data at the surface further improved the results.

32

The six experiments that were run for the surface and PBL were also applied to upper air or mid-tropospheric levels to determine the effectiveness of FDDA in modeling turbulent mixing of pollutants in the PBL up to the free atmosphere and transport by large scale horizontal flow. How FDDA affected RH and temperature at the upper levels were also evaluated because of their importance to cloud formation and the production of pollutants. Figure 5-5 shows upper air MAE and ME statistics for RH for the six experiments. Using FDDA improved the model results, with the multiscale experiments producing the best results (Figure 5-5a). In the upper air the moisture was assimilated, unlike at the surface where it was not; therefore, all the experiments using FDDA significantly reduced RH errors compared to the NOFDDA experiment. A similar conclusion was seen in the ME statistics where, the FDDA experiments showed very small positive moisture biases on all three domains (Figure 5-5b) except for the GFDDA experiment on the 4-km domain where analysis nudging was not used. Having positive moisture biases suggests that the models tended to produce higher RH which will lead to clouds affecting the radiative flux and the production of pollution.

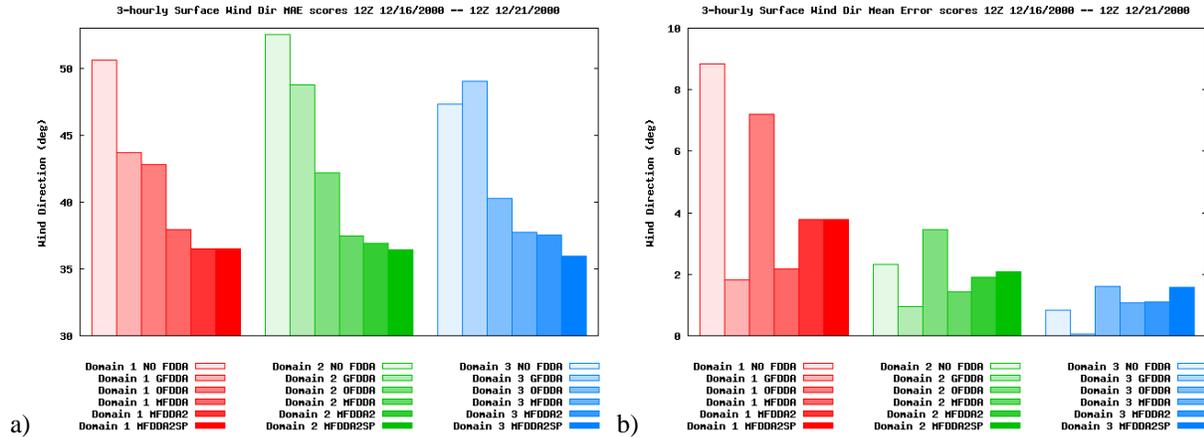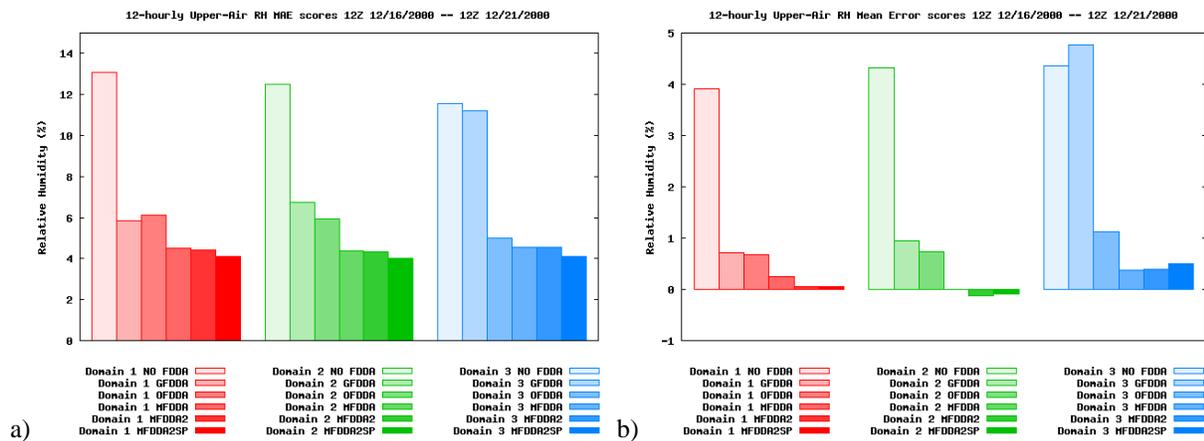**Figure 5-5: Upper air statistics for the FDDA numerical experiments for RH. a) MAE. b) ME.**

Figure 5-6 shows the MAE and ME statistics for WRF-ARW-simulated upper-air temperature. All of the FDDA experiments showed improvement compared to the NOFDDA experiment, where the multiscale FDDA experiments again produced the best results and (Figure 5-6a). Temperature was assimilated in the upper air similar to the moisture. The ME statistics (Figure 5-6b) show that the temperature biases for all experiments were quite small. The upper air statistics for the wind direction (Figure 5-7) and the wind speed (Figure 5-8) also show the positive effects of using FDDA. Because winds were assimilated in the upper air, similar to the surface, using FDDA improved the results significantly compared to the NOFDDA experiment, with the multiscale FDDA experiments again having produced the best results. All six of the experiments show positive wind direction biases on the 36-km and 12-km, but they all had negative biases on the 4-km domain, unlike at the surface. Since the wind speeds were generally small, the wind speed biases were also quite small, with the biases further reduced in the multiscale FDDA runs.
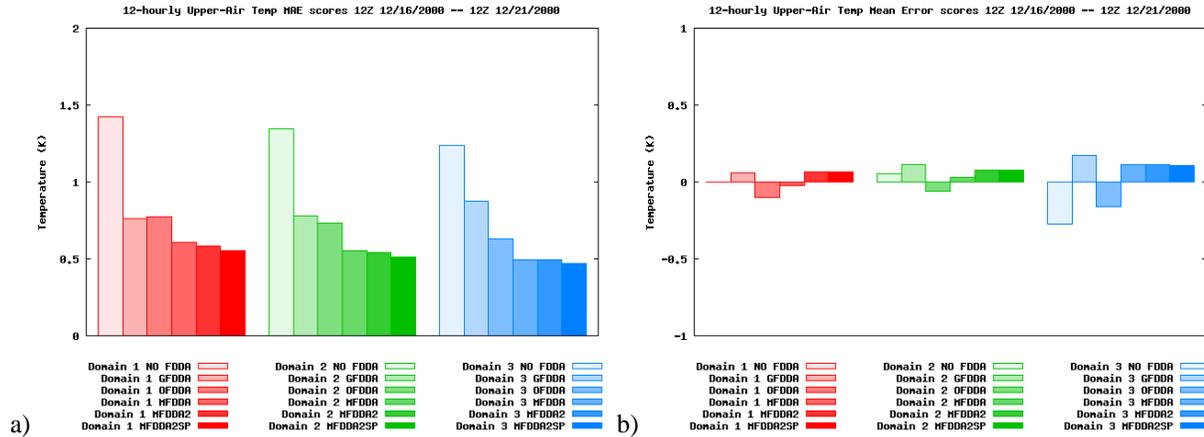
Figure 5-6: Upper air statistics for the FDDA numerical experiments for temperature. a) MAE. b) ME.
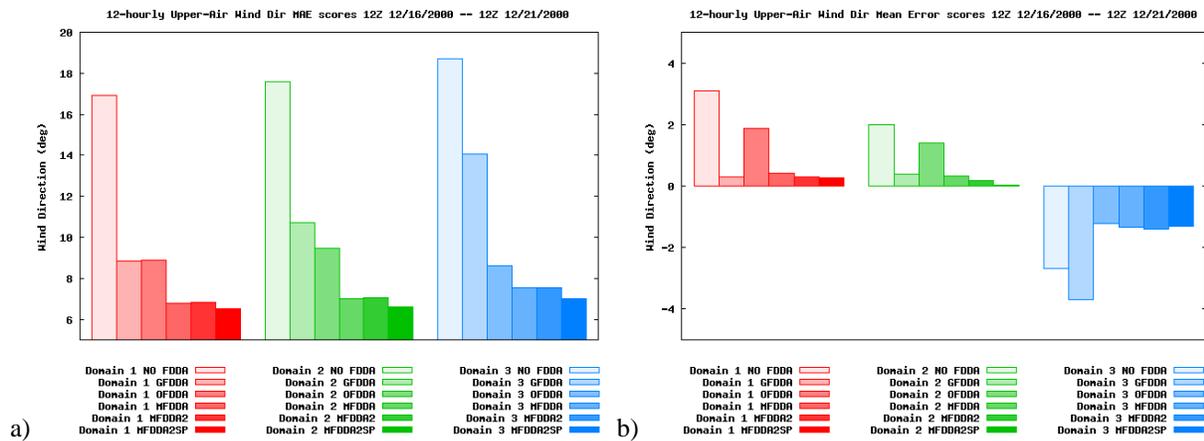
Figure 5-7: Upper air statistics for the FDDA numerical experiments for wind direction. a) MAE. b) ME.
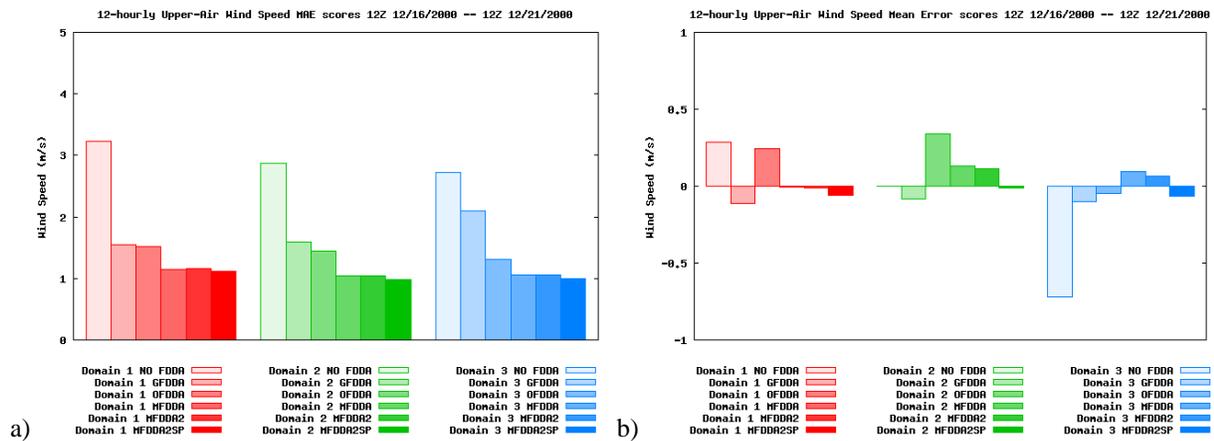
Figure 5-8: Upper air statistics for the FDDA numerical experiments for wind speed. a) MAE. b) ME.

Overall, the upper air results were qualitatively similar to those for the surface. FDDA improved the results compared to the NOFDDA experiment where data were not assimilated. Using a multiscale FDDA method that combined the analysis nudging on the 36-km and 12-km domains with the obs nudging on all three domains produced the best results. The inclusion of special surface wind observations from an independent network for data assimilation in addition

34

to the WMO observations further improved the results. Therefore, MFDDA2SP was the best configuration to use for the upper air as well.

5.2.1.2 Subregion Verification

The conclusion that the MFDDA2SP model was the best configuration for the winter case was based on the surface and upper air statistics for the entire 4, 12, or 36-km domain. However, the model results may not be similar for specific areas within each domain. For this study, there were three key regions that were of interest (Figure 5-9): 1) the SFBA; 2) the Sacramento Valley; and 3) the San Joaquin Valley. The sample data distribution shown in Figure 5-9 indicated that the data density was quite good for each region, especially for the SFBA region.
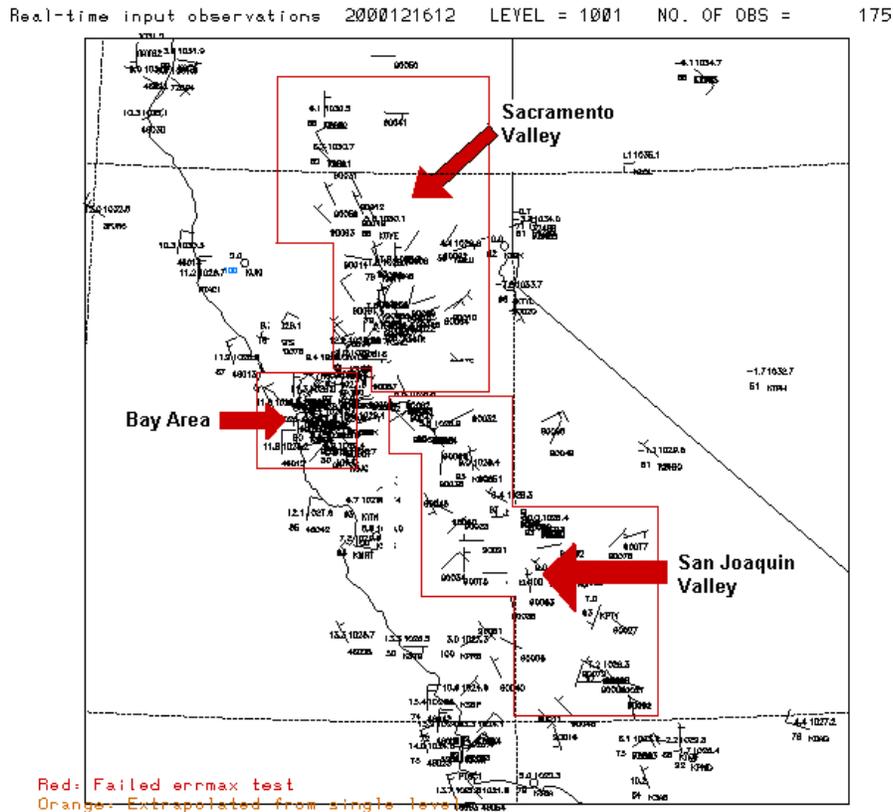
Figure 5-9: Map of the 4-km domain showing the three subregions, the WMO observations, and the BAAQMD special surface wind observations used for data assimilation.

Table 5-5 shows the surface MAE statistics for the SFBA. The statistics were made and compared in similar fashion to the statistics for the FDDA simulations previously discussed. However, these statistics only compared the best experiment (MFDDA2SP) with the baseline experiment (NOFDDA). The MFDDA2SP experiment showed improvement over NOFDDA on all four meteorological fields in the Bay Area, similar to the surface results on the entire domains; however, the MFDDA2SP experiment did not improve the wind directions for this subregion as much as for the entire domain (i.e., a 7 degree improvement for the Bay Area but a 11 degree improvement for the entire 4-km domain). This confirmed that although the model could show great improvement over an entire model domain, the model may not show the same improvement for a specific subregion or a specific location within the model domain. The upper

air MAE statistics for the Bay Area (not shown) also showed the MFDDA2SP experiment improved the errors for the subregion compared to the NOFDDA.

Table 5-5: Surface MAE Statistics for the Bay Area region comparing NOFDDA and MFDDA2SP

	36 km	12 km	4 km
Relative Humidity			
NOFDDA	14.8	20.0	20.8
MFDDA2SP	9.4	12.0	16.6
Temperature			
NOFDDA	2.3	1.8	1.5
MFDDA2SP	1.8	1.7	1.4
Wind Direction			
NOFDDA	56.7	52.7	49.1
MFDDA2SP	46.2	44.4	41.7
Wind Speed			
NOFDDA	2.1	1.7	1.5
MFDDA2SP	1.6	1.4	1.3

Table 5-6 shows the surface MAE statistics for the Sacramento Valley region. The MFDDA2SP experiment showed the same improvement in wind direction and wind speed errors as was the case for the full domains. However, the experiment actually showed slight degradation in the RH errors on the 4-km domain and degradation in the temperature errors on both the 12-km and 4-km domains compared to the NOFDDA experiment. Therefore, although the MFDDA2SP experiment was the best model configuration for this case, it did not actually help improve the RH and temperature patterns in the Sacramento Valley. However, the degradation in the errors by the MFDDA2SP was very small. Upper air statistics (not shown) showed similar improvements in wind direction and wind speed errors. No statistics were calculated for the RH and the temperature because only two upper air observation sites were located along the California coast within the 4-km domain (one in Oakland in the SFBA, and one at the Vandenberg Air Force Base located near the southern edge of the domain).

Table 5-7 shows the surface MAE statistics for the San Joaquin Valley region. The MFDDA2SP experiment showed improvement in the wind direction and wind speed errors. However, the experiment had a very slight degradation in the RH and the temperature errors on the 4-km domain, similar to the Sacramento Valley. The upper air statistics (not shown) were similar to the statistics for the Sacramento Valley. The MFDDA2SP experiment improved the wind direction and wind speed errors, and there were no statistics for the RH and the temperature because there were no upper air observation sites within the subregion.

Table 5-6: Surface MAE Statistics for the Sacramento Valley region comparing NOFDDA and MFDDA2SP

	36 km	*12 km*	*4 km*
Relative Humidity			
NOFDDA	26.7	24.6	19.0
MFDDA2SP	17.7	17.3	24.4
Temperature			
NOFDDA	2.2	1.8	1.5
MFDDA2SP	1.9	1.9	1.7

Wind Direction			
NOFDDA	50.8	45.8	40.7
MFDDA2SP	37.7	36.0	34.5
Wind Speed			
NOFDDA	2.0	1.6	1.4
MFDDA2SP	1.4	1.2	1.2

Table 5-7: Surface MAE Statistics for the San Joaquin Valley region comparing NOFDDA and MFDDA2SP

	36 km	**12 km**	**4 km**
Relative Humidity			
NOFDDA	41.3	26.5	16.3
MFDDA2SP	30.0	20.9	16.8
Temperature			
NOFDDA	3.3	2.4	2.1
MFDDA2SP	2.8	2.2	2.2
Wind Direction			
NOFDDA	53.5	62.3	53.3
MFDDA2SP	41.1	41.1	41.6
Wind Speed			
NOFDDA	2.1	1.5	1.1
MFDDA2SP	1.4	1.1	0.9

5.2.1.3 Independent Verification

The statistics previously discussed proved that using surface and upper air observations along with special network of data for data assimilation significantly improved the model simulations for these two cases. The model results previously discussed were verified against the entire set of WMO observations and the special surface wind observation data network that were assimilated. However, to verify how well the model compared with the observations, the simulations should be compared with a set of observations that were not used for data assimilation.

The data used for independent verification were the surface observations from the meteorological monitors from the AQS dataset that the BAAQMD provided. These observations were not used for data assimilation (sample data shown in Figure 5-10).

Table 5-8 shows the surface MAE statistics for the NOFDDA and the MFDDA2SP experiments. The MFDDA2SP experiment showed improvements in the temperature errors, wind direction errors, and wind speed errors on all the domains, with the exception of temperature on the 4-km domain which showed a very small degradation over NOFDDA. There were no statistics for RH because RH observations were not available.
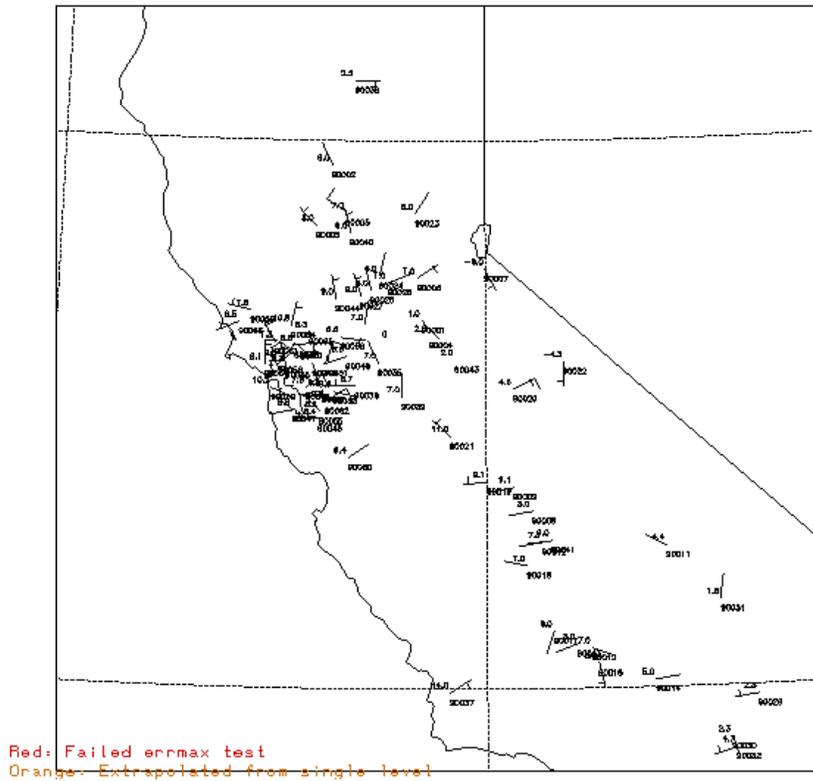
Figure 5-10: Map of the meteorological monitors from the AQS dataset used for independent verification.

Table 5-8: Surface MAE Statistics comparing NOFDDA and MFDDA2SP against an independent dataset.

	36 km	12 km	4 km
Temperature			
NOFDDA	2.8	2.4	2.2
MFDDA2SP	2.6	2.3	2.2
Wind Direction			
NOFDDA	66.6	54.8	49.8
MFDDA2SP	56.2	47.6	44.2
Wind Speed			
NOFDDA	1.7	1.5	1.3
MFDDA2SP	1.3	1.3	1.2

5.2.2 Subjective Analysis Results

The numerical simulations were examined to subjectively verify the model with the observations. Only the baseline experiment (NOFDDA) and the best experiment (MFDDA2SP) were examined. Analysis of the results on the 36-km domain will be presented first followed by analysis of the results on the 4-km domain.

5.2.2.1 36-km Domain

Figure 5-11 shows the model initial conditions on the 36-km domain, at 1200 UTC 16

Dec 2000. The model surface initial conditions (Figure 5-11a) showed a low over southern California. As observed, the strong pressure gradient along the coast of northwest United States and southern Canada was well represented (see Section 3). In agreement with the observations, easterly flow offshore of northern California and northerly flow over the Bay Area and along the coast were present. The northerly flow in the Central Valley was also well represented in the model initial conditions.
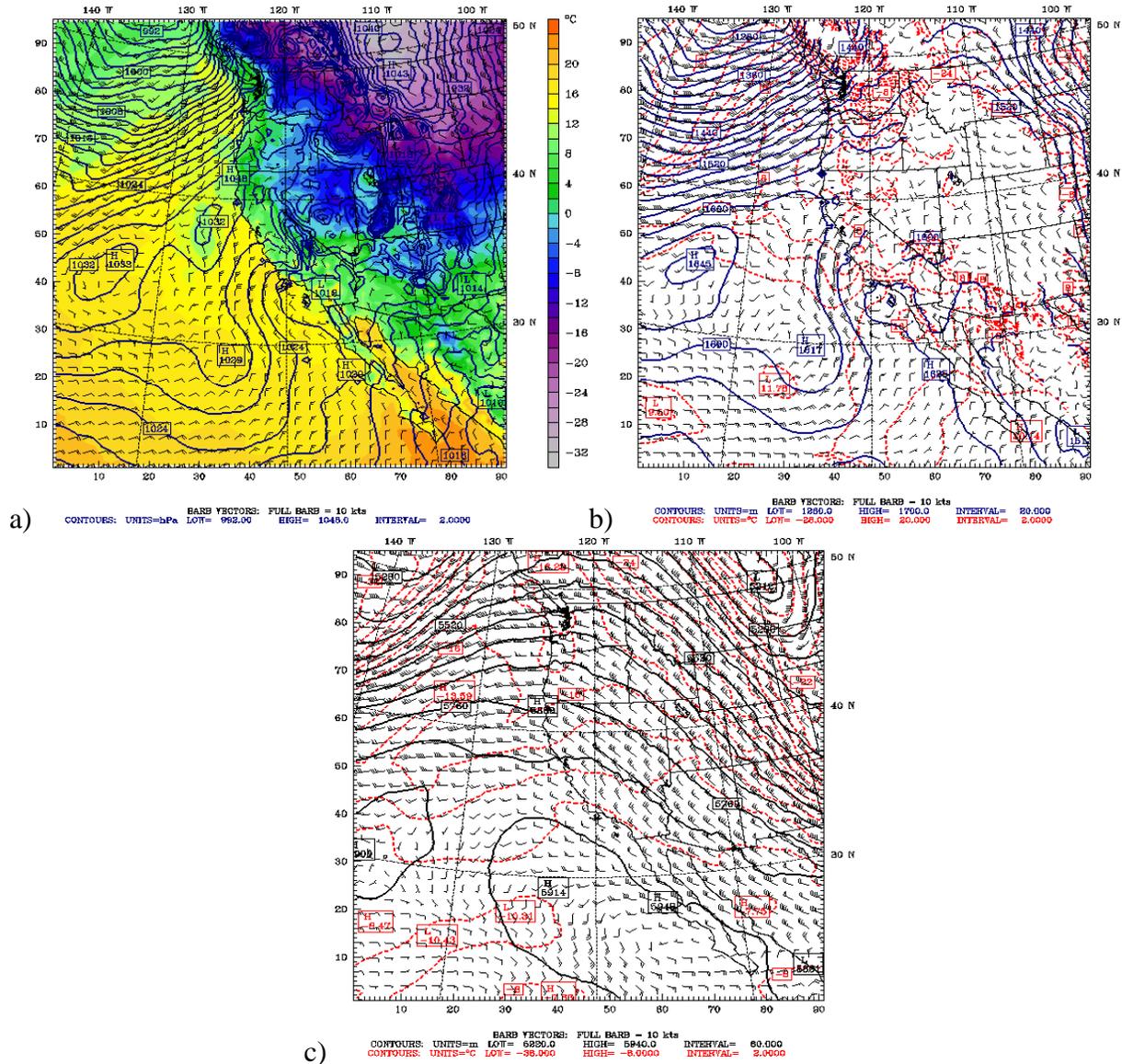
Figure 5-11: Plots of the NOFDDA simulation at 1200UTC 16 Dec 2000. a) Surface. b) 850 mb. c) 500 mb.

At the 850 mb level the model initial conditions at 1200 UTC 16 Dec 2000 showed the strong height gradient that was present over the ocean, Washington, Oregon, and southern British Columbia (Figure 5-11b). Consistent with the observed surface inverted trough, offshore flow over the California coast was reproduced, along with the strong onshore flow over Oregon and Washington. At the 500 mb level, model initial conditions showed a large ridge over the western coast of the U.S. and Canada (Figure 5-11c).

39

Over the next three days conditions did not change much as the high remained over the western United States and calm conditions persisted. Although the NOFDDA and MFDDA2SP simulations were quite comparable, there were some noticeable differences, and overall MFDDA2SP performed better. For example, at 1200 UTC 17 Dec 2000, at 24 hours into the simulation, at the 850 mb level, both simulations developed a ridge offshore of the Pacific Northwest (Figure 5-12), which agreed well with observations (Figure 3-4a). The location of the high over the ocean west of California in both experiments also agreed well with observations. The MFDDA2SP (Figure 5-12b) resolved the short wave that was observed over northern California and the strong temperature gradient that was present just off the coast, in agreement with the observed surface cold (not shown). The temperature gradient was not as strong in the NOFDDA simulation. The simulated winds in both experiments agreed quite well with observations. Both developed the onshore flow over California, Oregon, and Washington, but the MFDDA developed the northwesterly flow observed at the Vandenberg Air Force Base (not shown) while the NOFDDA had the winds more northerly in direction. Overall, both experiments simulated the heights, heights and temperature well, but the MFDDA experiment did a better job with simulating strong temperature gradients that signified a front.
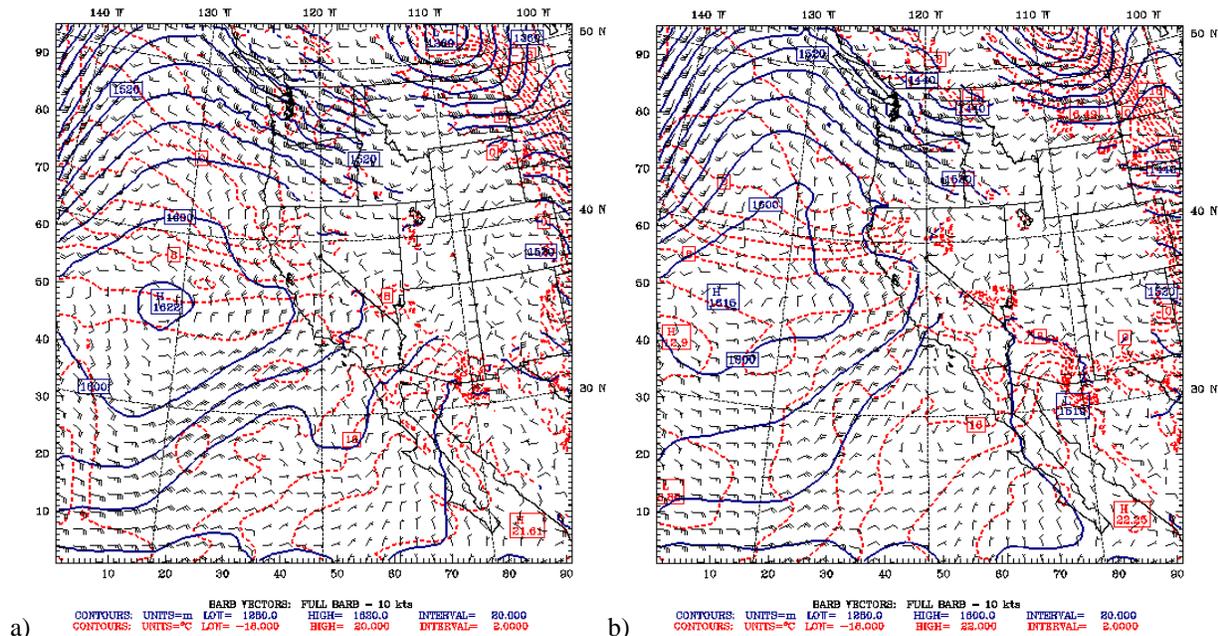
Figure 5-12: 850 mb plots at 1200 UTC 17 Dec 2000. a) NOFDDA. b) MFDDA2SP.

72 hours into the simulation, at 1200 UTC 19 Dec 2000, at the 500 mb level, both experiments developed a trough over southern British Columbia and northern Washington (Figure 5-13a and b) which agreed well with the observations (Figure 5-13c). However, the MFDDA (Figure 5-13bFigure 5-15) developed the short wave feature observed on the 5640 m isohypse over the ocean west of Washington while the NOFDDA (Figure 5-13a) did not fully develop this feature. The temperature patterns in both experiments agreed well with the observations. The simulated onshore flow over the western U.S in both experiments also agreed well with observations. Both experiments did a good job simulating the heights, winds, and temperature at this level, but the MFDDA did a better job of fully developing short wave patterns.
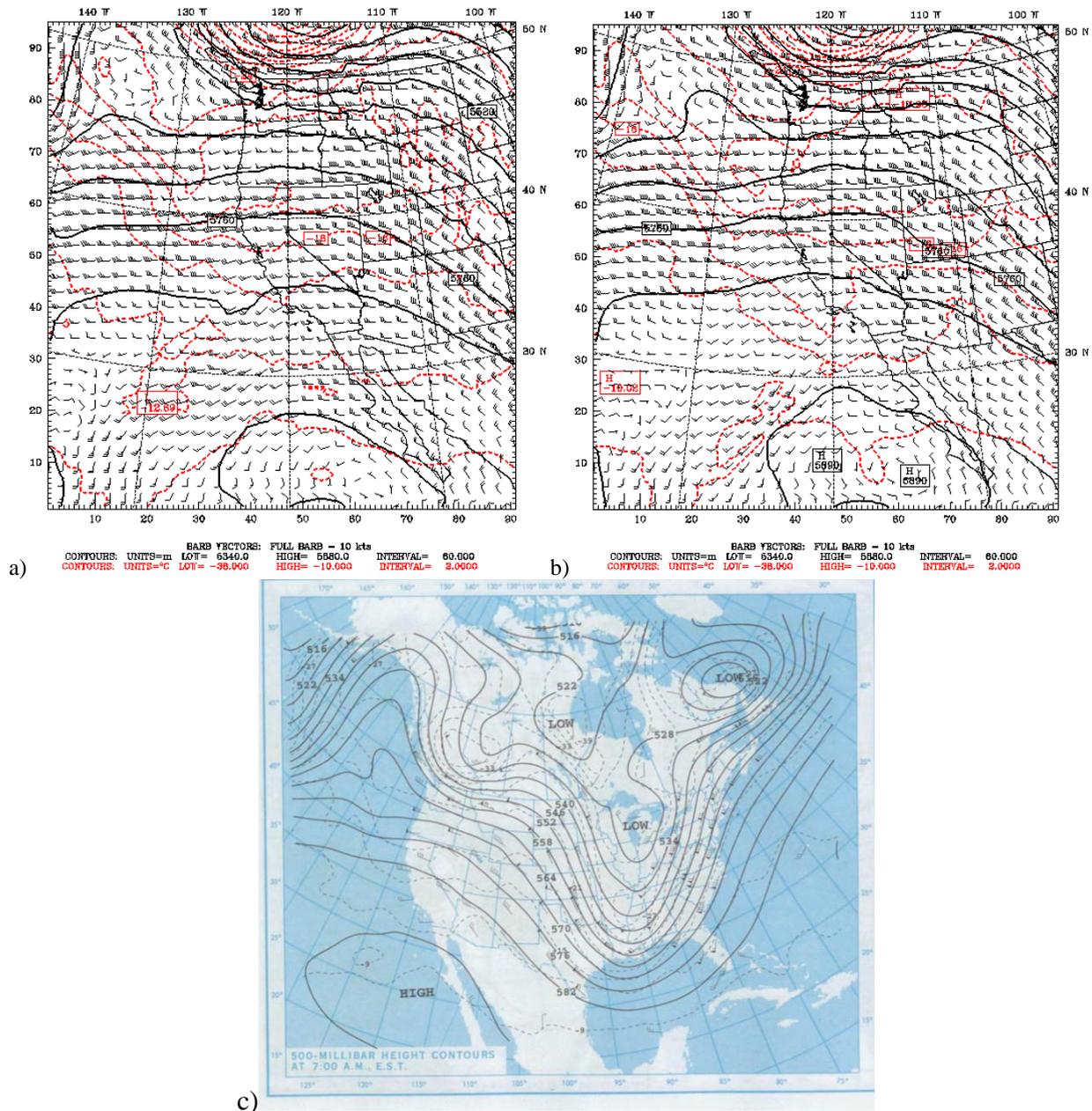
Figure 5-13: 500 mb plots at 1200 UTC 19 Dec 2000. a) NOFDDA. b) MFDDA2SP. C) 500 mb analysis (Taken from *Daily Weather Maps* from the NOAA Central Library Data Imaging Project)

96 hours into the simulation, at 1200 UTC 20 Dec 2000, the surface observations showed a high in Canada with a ridge extending south through the western United States through California, Nevada, and Arizona (Figure 3-9a). Both the NOFDDA and the MFDDA2SP developed the pressure gradient along the coast of Canada and Washington. The NOFDDA experiment developed a high just off the coast of California, which was not found in the observations (Figure 5-14a). However, the MFDDA2SP experiment did not develop this high either and instead developed a high further to the southwest (Figure 5-14b). The NOFDDA was able to develop the offshore flow over southern California and the southerly flow over the Bay

41

Area in agreement with observations, but it did not develop the observed northeasterly flow offshore over northern California and simulated southerly flow instead. The MFDDA2SP experiment did develop the offshore flow over northern California, but the flow was easterly instead of northeasterly as observed. MFDDA2SP also developed the easterly offshore flow over southern California, but did not develop the southerly flow over the Bay Area, simulating easterly offshore flow. Overall, both experiments did a good job in simulating the surface winds and sea level pressure. However, the MFDDA2SP did better with the placement of the highs over the ocean.

Figure 5-14: Surface plots at 1200UTC 20 Dec 2000. a) NOFDDA simulation. b) MFDDA2SP simulation.
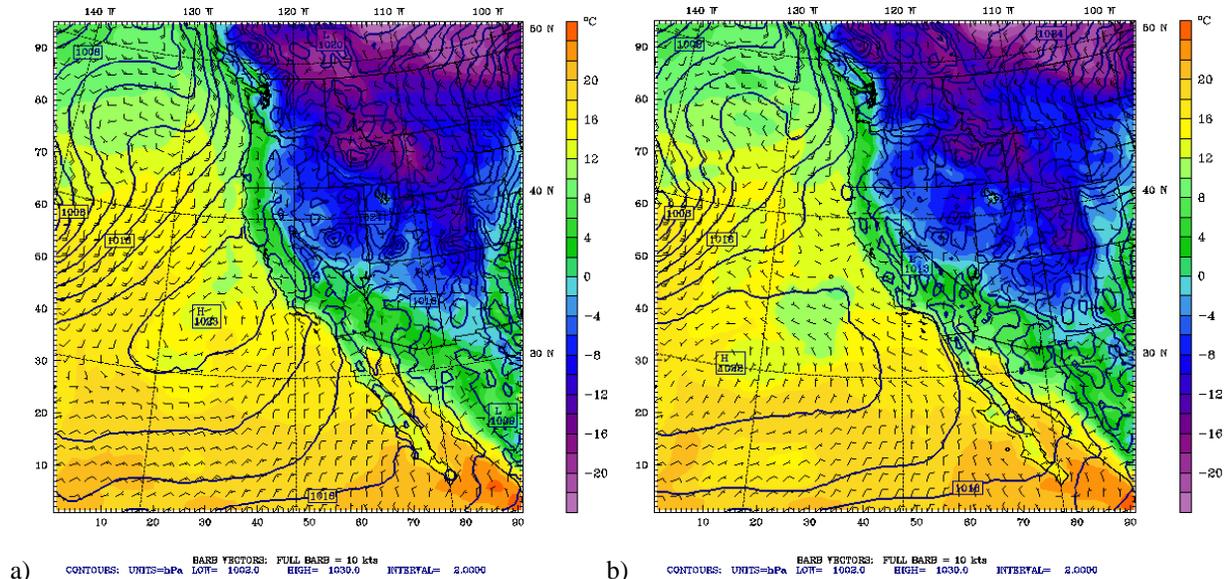
5.2.2.2 4-km Domain

On the 4-km domain the effects of FDDA were quite evident. Obs nudging allowed the model to assimilate individual observations from surface and radiosonde stations. This assimilation of data helped to reduce the model errors on this domain with high resolution. When obs nudging was used in combination with the analysis nudging on the coarser domains, which helped to improve the synoptic scale through the assimilation of gridded analyses, the model results were significantly improved when compared to a simulation that did not use data assimilation.

Figure 5-15 shows plots of the simulated wind field on the entire 4-km domain 12 hours into the simulation, at 0000 UTC 17 Dec 2000 (4:00 P.M. Pacific Standard Time [PST] on 16 Dec 2000). The domain consisted of the Bay Area and the Sacramento and San Joaquin Valleys that make up the Central Valley region of California. The plots were overlaid with the WMO observations taken at this time for comparison with the model results. The plots were made 12 hours into the simulation to allow the FDDA to influence the model. The simulated winds for the NOFDDA experiment showed that even without FDDA the model did a reasonable job of resolving the mesoscale features that comprised this specific region (Figure 5-15a). The simulation showed the divergent flow as it traveled both up the Sacramento Valley and down the San Joaquin Valley. The winds were more divergent in the San Joaquin Valley than in the Sacramento Valley. The divergent flow became upslope flow along the Coastal Ranges and the

42

Sierra Nevada Mountains due to the lower pressure along the mountaintops which allowed the air in the valley to flow up the mountain slopes. The simulated northerly flow along the coast and over the southern Bay Area near San Jose (KSJC) agreed with the observations. However, the NOFDDA simulation showed big wind direction errors over the Central Valley. At some locations the directions of the winds varied by as much as 90 degrees, such as winds near Bakersfield (KBFL) in the southern San Joaquin Valley. Winds over Sacramento varied about 45 degrees.

In the MFDDA2SP simulation (Figure 5-15b), the overall modeled wind patterns had better agreement with the observations than the NOFDDA. The MFDDA2SP winds over the Sacramento Valley varied by about 30 degrees, while the winds near Bakersfield in the San Joaquin Valley varied by about 45 degrees. While these wind direction variations seemed large, they were as much as 50% less than the direction errors from the NOFDDA, especially at Bakersfield. It must be noted that sometimes the FDDA caused degradation in the wind direction at some locations. For example, winds along the axis of the San Joaquin Valley became more northerly due to the FDDA while they were easterly in the NOFDDA simulation, which agreed better with observations.
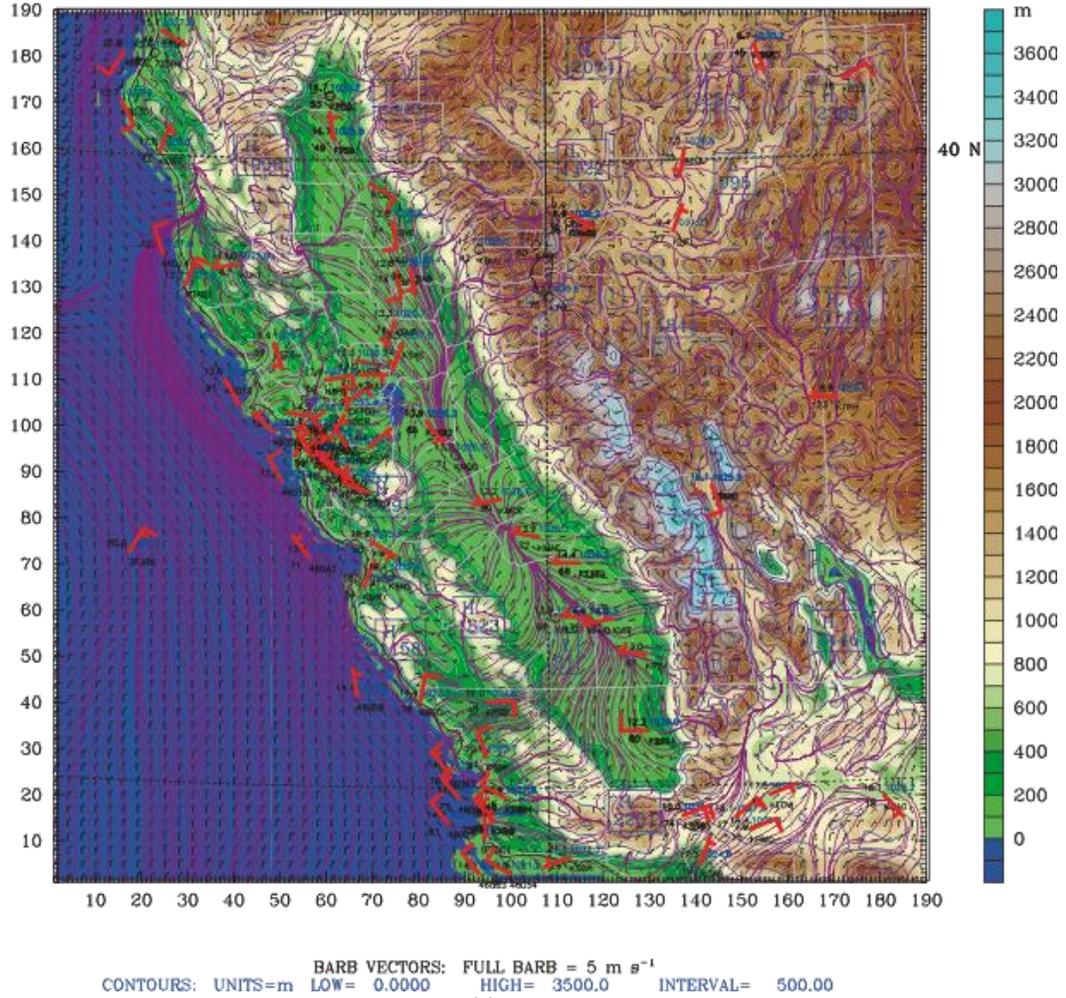
Figure 5-15: Simulated winds overlaid with WMO observations at 0000 UTC 17 Dec 2000. a) NOFDDA. b) MFDDA2SP

Dataset: d03   RIP: 2000-12 16-21 OFDDA2 spec re   Init: 1200 UTC Sat 16 Dec 00
Fcst: 12.00 h   Valid: 0000 UTC Sun 17 Dec 00 (1600 PST Sat 16 Dec 00)
Terrain height AMSL
Terrain height AMSL
Horizontal wind vectors   at k-index = 50
Horizontal wind streamlines

(b)
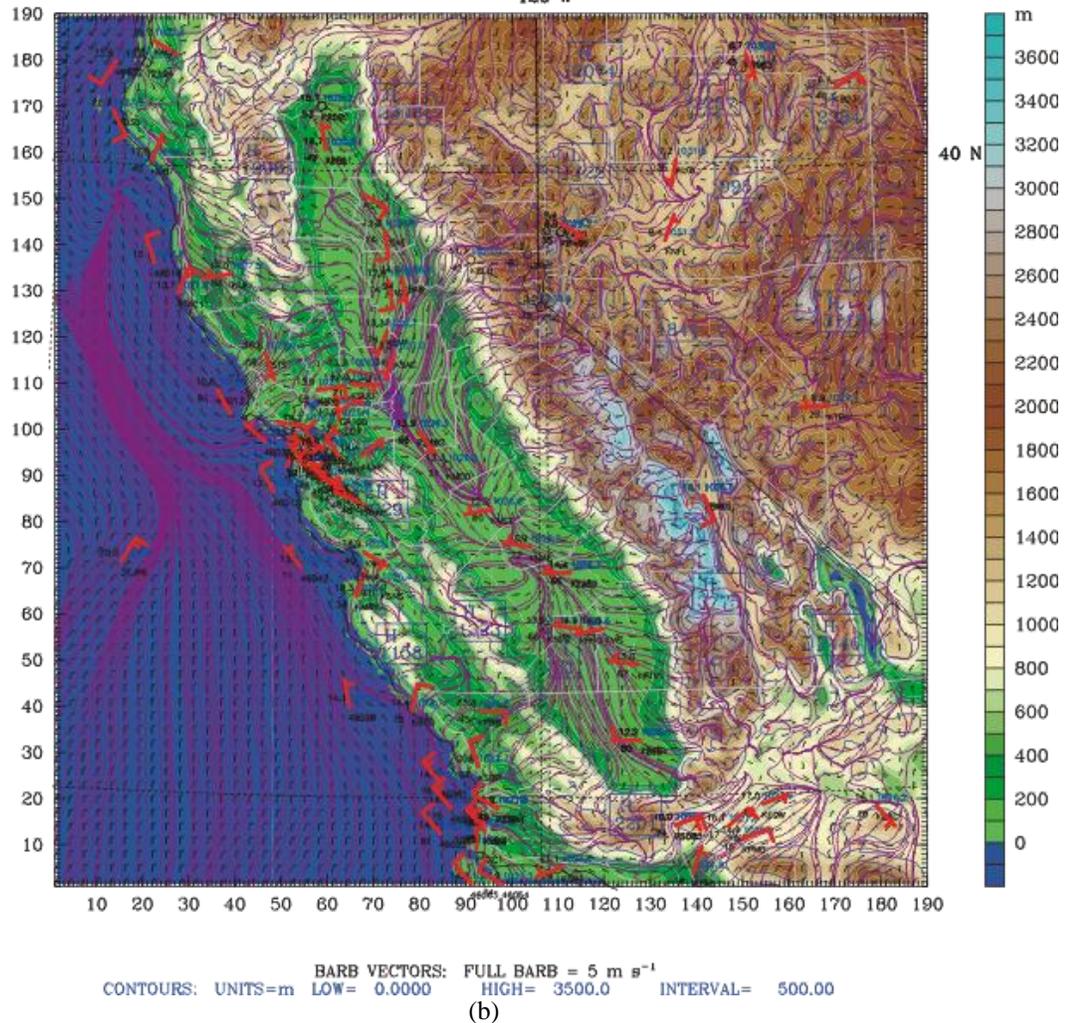
Fig 5.15) cont'd.

Figure 5-16 displays the same fields as in Figure 5-15, except that the plots are for 24 hours into the simulation, at 1200 UTC 17 Dec 2000 (4:00 A.M. PST).   The NOFDDA experiment again did a reasonable job of resolving the mesoscale features (Figure 5-16a). The simulation developed the convergent flow within the valleys caused mostly by downslope, or drainage, flow off the higher terrain surrounding the valleys. Again the simulated flow along the coast was well represented. However, large wind direction errors also existed in this simulation, with direction errors south of Sacramento as large as 90 degrees. One area of very large wind direction errors was in the southern SFBA, where the wind direction error in this area was about 180 degrees. The simulated winds were northerly while the observations show southerly flow. The MFDDA2SP experiment again agreed better with the observations (Figure 5-16b).   The wind direction error near Sacramento was absent, but the large wind direction errors over the southern SFBA still existed (>150 degrees). It should be noted that there were fewer WMO

45

observations available at night compared to the daytime.

(a)
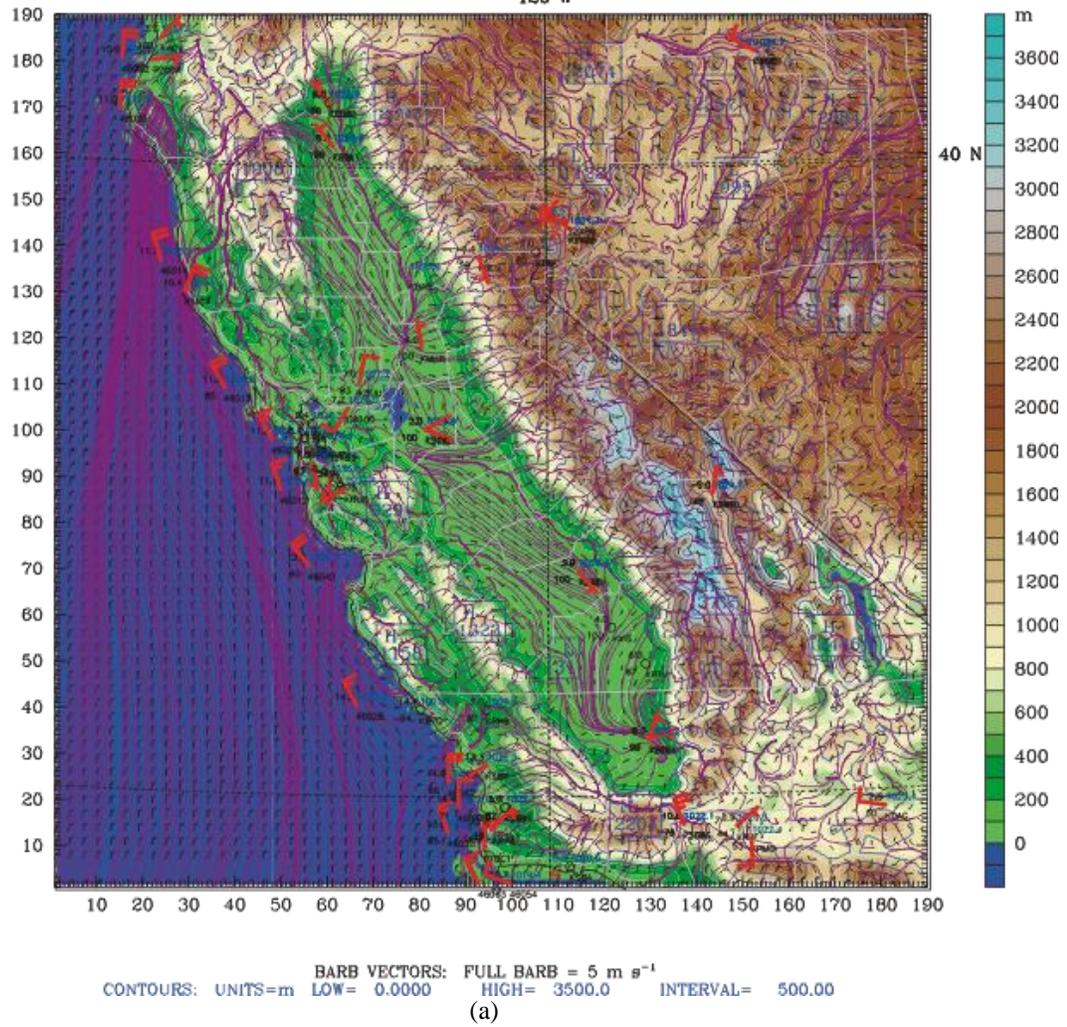
Figure 5-16: Simulated winds overlaid with WMO observations at 1200 UTC 17 Dec 2000. a) NOFDDA. b) MFDDA2SP

Fig 5.16) cont'd.

5.2.2.3 Subregion Mesoscale Analysis

As mentioned in the previous section, FDDA can have less impact on particular locations within a model domain when working to reduce the model errors. It was important to find how FDDA effects varied with space within the SFBA, Sacramento Valley, and San Joaquin Valley.

Figure 5-17 shows plots of the simulated wind field over the Bay Area region 84 hours into the simulation, at 0000 UTC 20 Dec 2000 (4:00 P.M. PST). The plots were overlaid with the observations from the AQS dataset. This time was chosen because it showed key mesoscale features that were of interest for this region. Once again the observations were from the AQS dataset that were used for the independent verification because they were fewer in number and provided a clearer picture than the many WMO observations within this region. Both simulations were very similar to each other and represented the air flow over the region very well

47

when compared to the observations. Both simulations developed the easterly flow from the Central Valley into the Bay Area, where the air flow continued through the Golden Gate Bridge and out over the Pacific Ocean. The air flows into the Livermore Valley from the Central Valley and through the I-680 corridor were well represented in both plots. The air flowed from the Central Valley through the Carquinez Strait, northeast of Concord (CCR), before turning toward the south to pass through the I-680 corridor. The air then converged with the easterly flow that had flowed over the Altamont Pass and Livermore from the Central Valley. The air flow then continued westward and converged with northerly flow near San Jose (SJC).

However, comparison between the two simulations demonstrated a clear overall degradation in wind direction in the FDDA simulation (although time series analysis shown later at KSJC did show benefit of FDDA in improving the temperature). The only area where FDDA showed the improved winds was near Vallejo (VAL). In the NOFDDA simulation (Figure 5-17a), the area around VAL had northeasterly flow heading offshore, which did not agree with the observations, while the MFDDA2SP (Figure 5-17b) had southeasterly flow that better agreed with the observation at VAL.

a)

48

b)
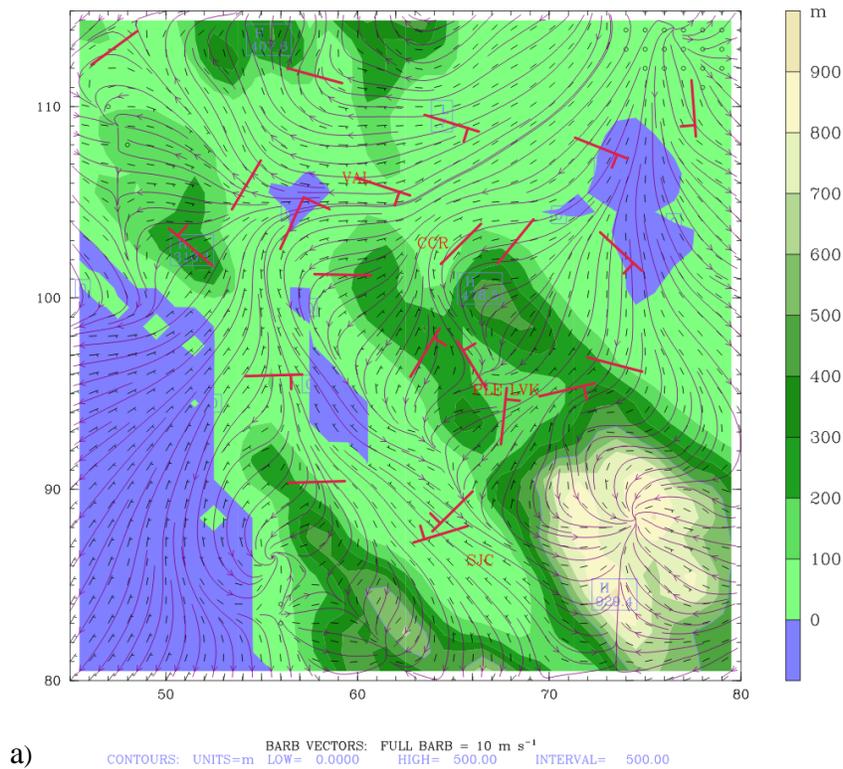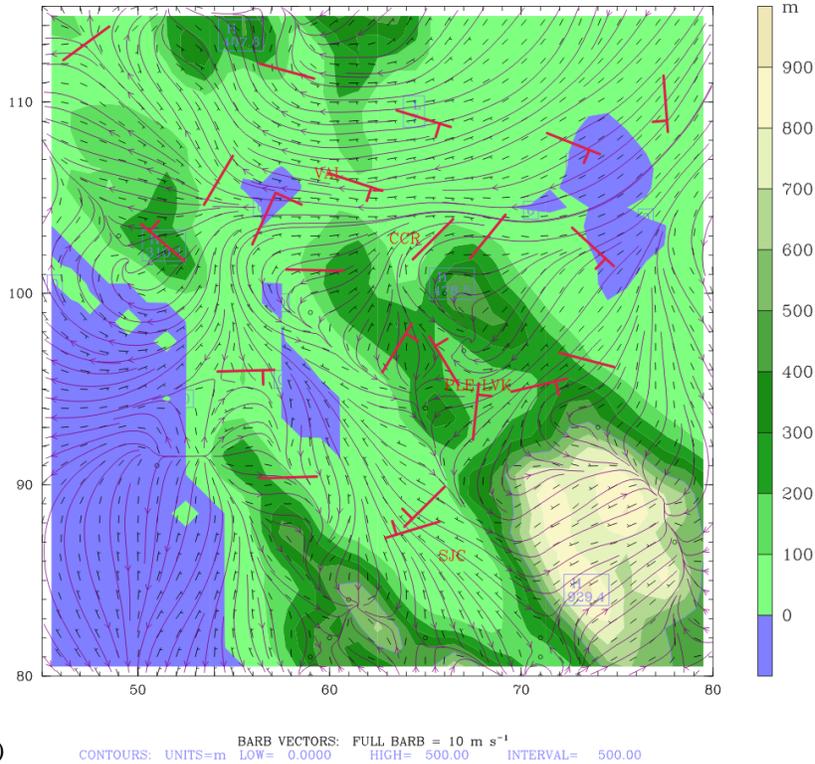
Figure 5-17: Surface winds in the Bay Area region at 0000 UTC 20 Dec 2000. a) NOFDDA. b) MFDDA2SP.

For the Sacramento Valley subregion, both simulations modeled the observed winds fairly well. For example, at 0600 UTC 19 Dec 2000 (10:00 P.M. PST), which was 66 hours into the simulation, the MFDDA2SP simulation showed downslope flow off the Sierra Nevada Mountains to the east and other mountains to the west (Figure 5-18). This agreed well with the observed drainage flow shown toward the south end of the valley. The drainage flows converged along the center of the valley where they flowed south passing west of Sacramento (SAC) and out of the valley, before turning toward the west and flowing through the Carquinez Strait. The NOFDDA simulation was very similar to the MFDDA2SP (not shown).

The modeled winds in the San Joaquin Valley subregion were again fairly well reproduced by both simulations. For example, at 0600 UTC 19 Dec 2000 the MFDDA2SP simulation showed drainage flows all along the Sierra Nevada Mountains to the east, the Tehachapi Mountains to south, and the Coastal Ranges to the west (Figure 5-19). The drainage flows converged on the west side of the valley along the Coastal Ranges where it flowed to the north, passed west of Fresno (FAT) and Modesto (MOD), and headed out of the valley by turning west through the Carquinez Strait. The observations are all along the axis of the valley and showed easterly flow, agreeing with the simulation. The MFDDA2SP simulation showed the mean flow passing closer to Fresno than in the NOFDDA simulation (not shown).

49

Dataset: d03   RIP: 2000-12 16-21 MFDDA2SP d03 w   Init: 1200 UTC Sat 16 Dec 00
Fcst: 66.00 h   Valid: 0600 UTC Tue 19 Dec 00 (2200 PST Mon 18 Dec 00)
Terrain height AMSL
Terrain height AMSL
Horizontal wind vectors   at k-index = 50
Horizontal wind streamlines   at k-index = 50
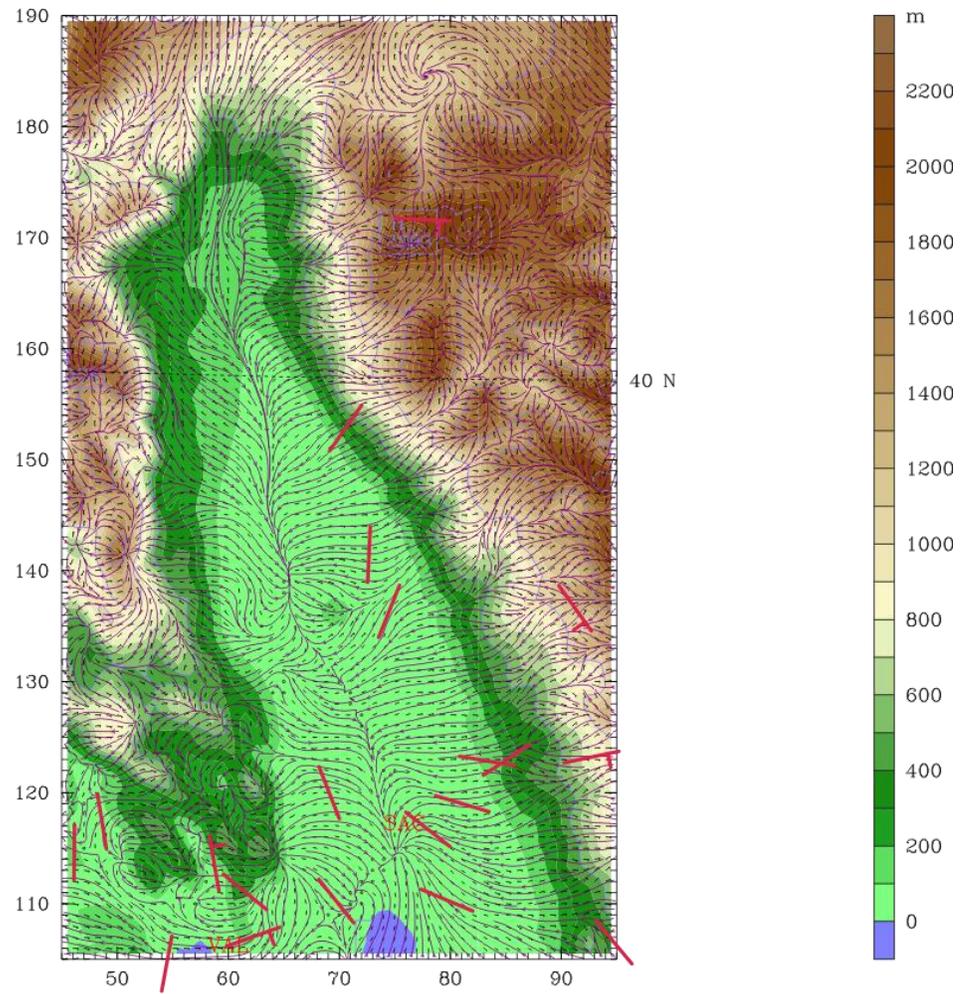
BARB VECTORS: FULL BARB = 10 m s⁻¹
CONTOURS: UNITS=m LOW= 0.0000   HIGH= 2000.0   INTERVAL=   500.00

Figure 5-18: Winds in the Sacramento Valley for the MFDDA2SP simulation at 0600 UTC 19 Dec 2000.

Dataset: d03 RIP: 2000-12 16-21 MFDDA2SP d03 w   Init: 1200 UTC Sat 16 Dec 00
Fcst: 66.00 h   Valid: 0600 UTC Tue 19 Dec 00 (2200 PST Mon 18 Dec 00)
Terrain height AMSL
Terrain height AMSL
Horizontal wind vectors   at k-index = 50
Horizontal wind streamlines   at k-index = 50

BARB VECTORS: FULL BARB = 10 m s⁻¹
CONTOURS: UNITS=m LOW= 0.0000   HIGH= 3500.0   INTERVAL= 500.00
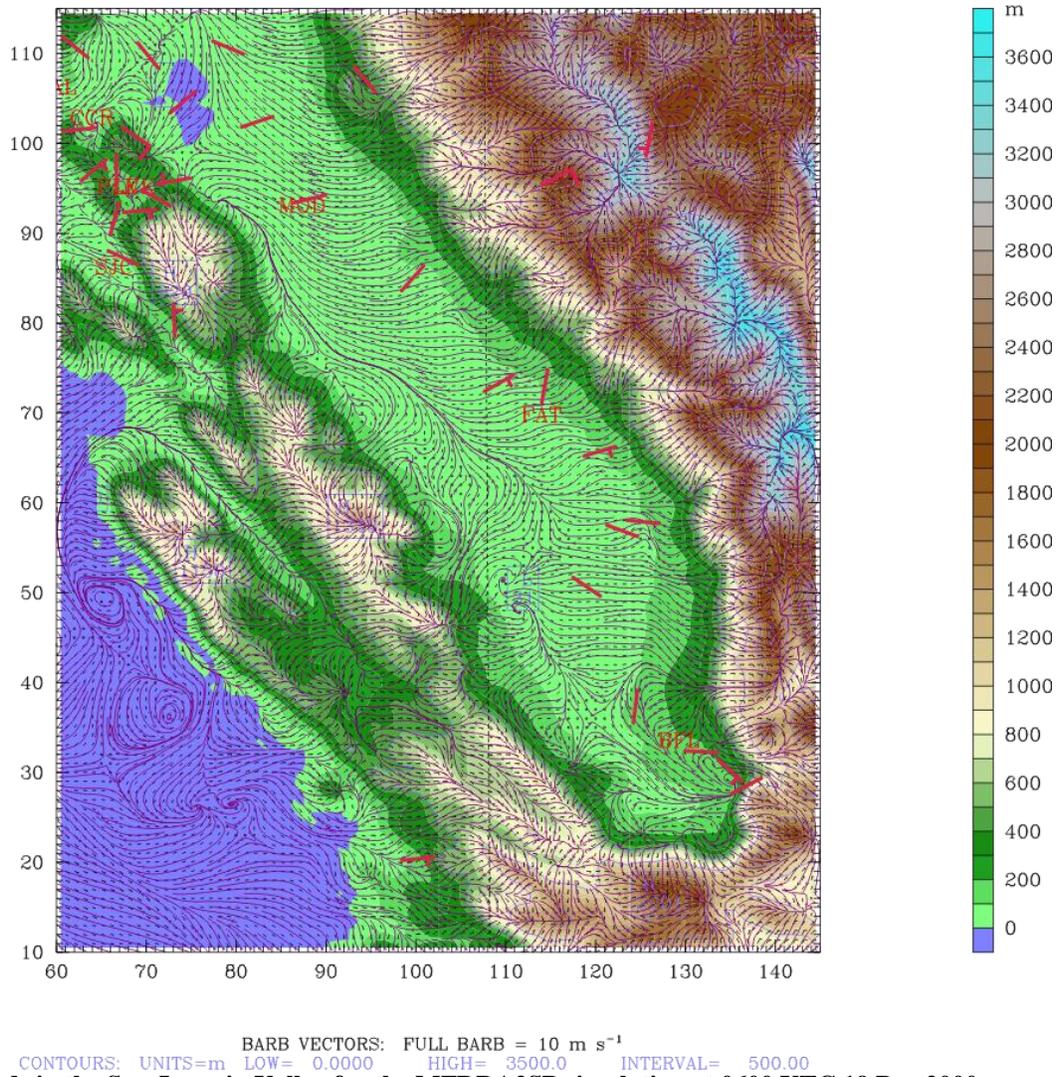
Figure 5-19: Winds in the San Joaquin Valley for the MFDDA2SP simulation at 0600 UTC 19 Dec 2000.

5.2.2.4 Time Series Analysis

As shown above, the effects of FDDA had an overall positive impact on the entire 4-km domain. Certain key observation sites were important to understanding the meteorology of the region those sites were located. Thus, it was important that the simulations performed well when comparing the results to the observations.

Figure 5-20 shows the WRF-simulated time series of the WRF-predicted first model layer (~10 m AGL) temperature and wind fields, comparing results from the NOFDDA and MFDDA2SP experiments to the observed values at San Jose (KSJC) in the Bay Area subregion. For temperature, both simulations did a good job of simulating the diurnal temperature cycle at San Jose (KSJC) (Figure 5-20a), with the simulated temperatures agreeing well with the

51

observations for the first 40 hours. Further into the simulation period, while the observed night-time temperatures became cooler, the simulations produced temperatures that were consistently warmer. Thus the model was biased to produce warmer temperatures, especially at the night. The biggest difference (about 5 °C) between the model and the observations occurred on the fourth night (between 70 and 76 hours). Note that the MFDDA simulation did a better job trying to correct the model temperature bias as shown on the third night (between 72 and 76 hours) and also forth night (between 90 and 105 hours). The gap in the observations between 21 and 25 hours were missing observations. The fastest observed wind speeds (Figure 5-20b) occurred during the first 48 hours, after which winds were between 0 and 2 m s^{-1} for most of the time as high pressure settled over the region. Both simulations were able to track the observed speeds for most of the time. The observed wind speed oscillations (between 0 and 1 m s^{-1}) were due to the speed sensor's limitation of not being able to measure velocities below 1 m/s at KSJC. Many of the wind direction observations were missing and the calm winds were removed from the time series plots to provide a clearer picture (Figure 5-20c). Both simulations showed westerly flow for the first 60 hours before the winds started to become easterly, which corresponded to the northwesterly flow in the Santa Clara valley shown in Figure 5-17. After 60 hours there was much variability in the wind direction, which corresponded with the observed light wind speeds that were associated with the high pressure. The effect of FDDA on winds was not evident at San Jose.

At Livermore (KLVK), the split between the temperature observations and simulated values occurred by the 18th hour of the period. Once again the nighttime warm bias in the model was apparent (Figure 5-21a). As with the KSJC simulations, the MFDDA simulation did a better job to correct the bias. Both simulations did a good job of representing the diurnal temperature cycle. The winds speeds observed were similar to those at San Jose, but the highest wind speeds occurred on the third day, the day before the winds became near calm for most of the remainder of the period that was dominated by the high pressure (Figure 5-21b). There were not many differences in wind direction between the simulations over the entire period. The winds were northeasterly for most of the period in the simulations (Figure 5-21c), which corresponded to the air flow from the Central Valley through the Altamont Pass, the Carquinez Strait, and the I-680 corridor (shown in Figure 5-17).

At Vallejo, both temperature simulations agreed well with the observations throughout the entire period, except for the first night when the model was too warm by 5-6 degrees, and the last day when the model was too cold by 3 degrees (Figure 5-22a). Both simulations had light winds for the first 40 hours before increased the wind speeds developed and persisted for the rest of the period. The model and observations agreed will for about the first 60 hours, after which differences in wind speed became large. The model simulated wind speeds were too fast compared with the observations, with the largest speed bias of about 6 m/s occurring near hour 70 (Figure 5-22b).   Modeled wind directions oscillated between westerly and easterly during the first 48 hours, before becoming easterly for the remainder of the period (Figure 5-22c). This was in good agreement with the observed directions and consistent with the easterly flow shown in Figure 5-17 near VAL. This easterly flow (both modeled and observed) was from the Central Valley, which passed through the Carquinez Strait, and continued toward the ocean.

In the Sacramento Valley, at Sacramento (KSAC, Figure 5-23), the simulations developed diurnal temperature cycles and matched up well with observations (Figure 5-23a). Observed wind speeds in the valley were nearly calm for most of the period, except on the second day when speeds were as high as 12 m/s due to a front that passed through the state

(Figure 5-23b). The near calm winds after the second day occurred during the time when the high PM episode occurred in the Central Valley. Both simulations agreed with the observations, with the exception of slower than observed wind speeds during the frontal passage. Note that the FDDA effect did simulate faster wind speeds than NOFDDA. The wind direction time series show both simulations had the winds change between northwesterly and northeasterly constantly, except for the second day when the winds were northerly during the frontal passed (Figure 5-23c). The changes in wind direction f appeared to show large swings in direction, but in reality, the change was smaller as the wind likely swung back and forth between northwesterly to northeasterly across 360 degrees. The directions only span between 0 degrees and 360 degrees and therefore it appears the wind may have changed significantly when it more likely only changed by about 20 degrees. Unfortunately, too many observations were missing or calm to provide a clear comparison. However, Even though there were a lot of missing observations, there was some indication of the wind direction adjustment when FDDA was used.

In the San Joaquin Valley, at Modesto (KMOD, Figure 5-24), both simulations developed the temperature diurnal cycles seen in the observations (Figure 5-24a). The daytime temperature from both simulations for the first two days and the last day agreed well with the observation, while simulated values were warmer than observations on days 3 and 4. The modeled nighttime temperatures were warmer than the observed, which confirmed a warm bias in the model for the Modesto area. Wind speeds were nearly calm throughout the period, except for the second day when the front passed through, which both simulations reproduced fairly well (Figure 5-24b). There was a clear advantage of using FDDA at this site as the FDDA curve was closer to the observed. The winds were southerly for half of the time period, switching to northerly the front passed (Figure 5-24c). During the fourth day the MFDDA simulation shows that northwesterly and northeasterly flow persisted over the area, in relative agreement with observations. As with wind speed, there was an advantage to using FDDA.

At Fresno (KFAT), both simulations agreed well with the temperature observations (Figure 5-25a). During the second day the front kept temperatures in the area cool. Both simulations produced cooler temperatures than the observations for about the first 70 hours. Although no temperature was nudged in PBL, FDDA seemed to help during daytime conditions on day1, 4, and 5. The observed wind speeds oscillated between calm and non-calm nearly every hour throughout the period (Figure 5-25b), with the fastest observed winds occurring during the frontal passage on day 2. The oscillating patterns seemed to indicate possible problems with the wind speed instrument, even at higher speeds. Both simulations showed a diurnal signal with wind directions (Figure 5-25c). Winds were easterly at night, which represented the downslope flow off the Sierra Nevada Mountains, and westerly during the day, representing the upslope flow.

At Bakersfield (KBFL, Figure 5-26), at the southern end of the San Joaquin Valley, the simulated temperatures agreed well with observations for the first 60 hours before the model produced consistently warmer results than the observations. The largest warm bias neared 10°C on the third night (Figure 5-26a). The observed calms that occurred throughout the period were not reproduced by the simulations (Figure 5-26b). The frontal passage of day 2 was not handled well by either simulation, but the MFDDA2SP run did produce higher speeds than NOFDDA. The simulations reproduced the upslope and downslope winds as at Fresno (Figure 5-26c). Once again, the MFDDA2SP results were in better agreement with the observations.
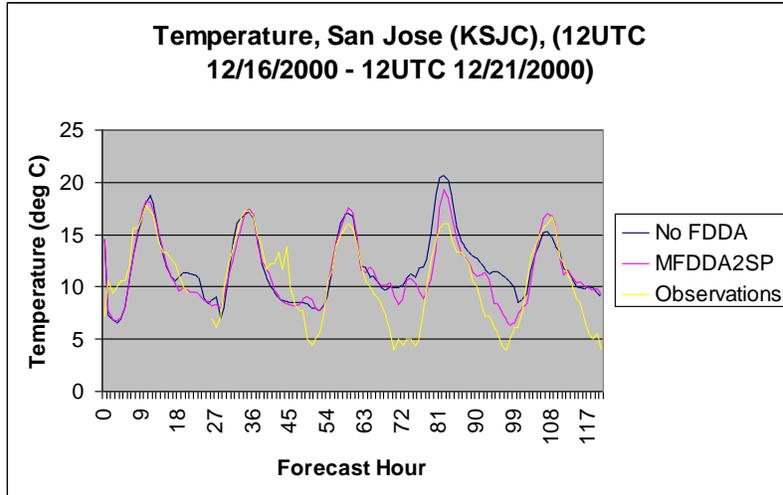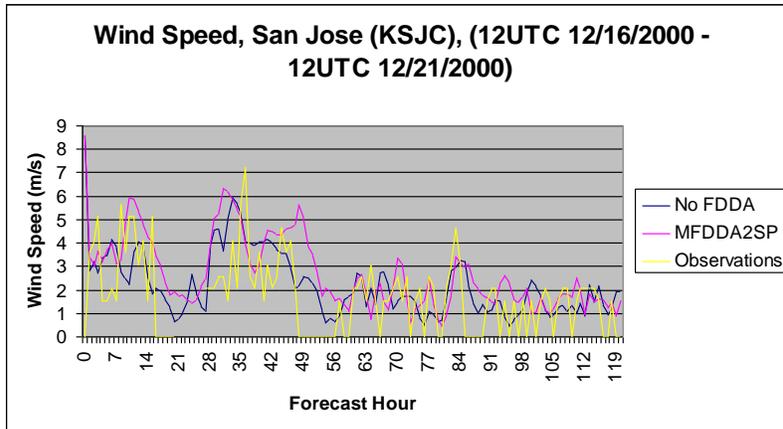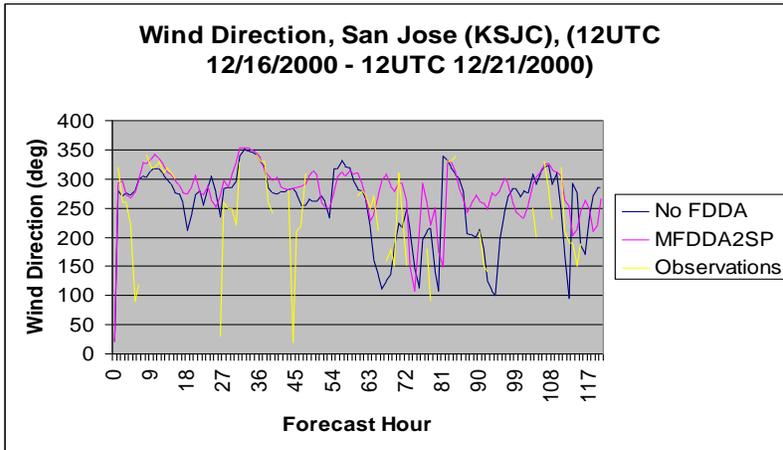
Figure 5-20: Time series plots for San Jose. a) Temperature. b) Wind Speed. c) Wind Direction.

Figure 5-21: Time series plots for Livermore. a) Temperature. b) Wind Speed. c) Wind Direction.

Figure 5-22: Time series plots for Vallejo. a) Temperature. b) Wind Speed. c) Wind Direction.

56

Figure 5-23: Time series plots for Sacramento. a) Temperature. b) Wind Speed. c) Wind Direction.

57

a)

b)
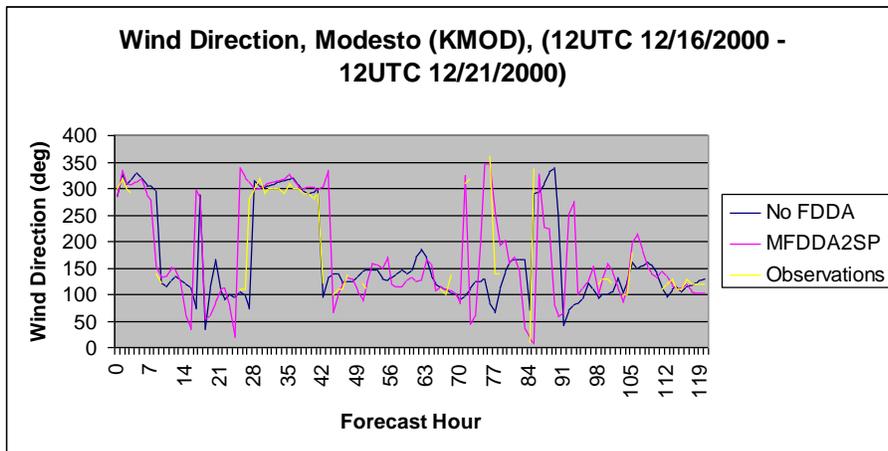
c)

Figure 5-24: Time series plots for Modesto. a) Temperature. b) Wind Speed. c) Wind Direction.
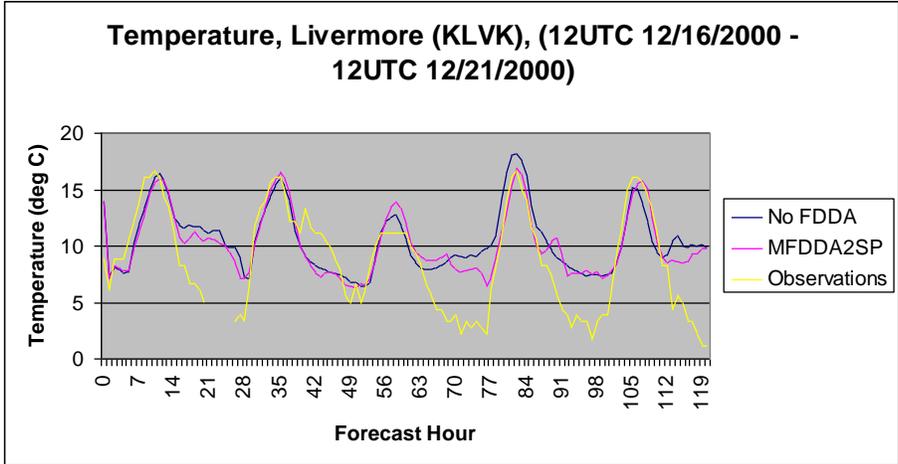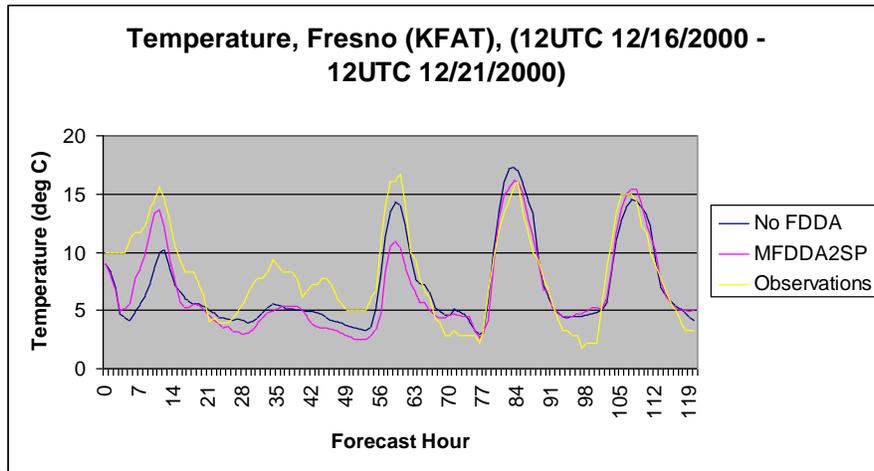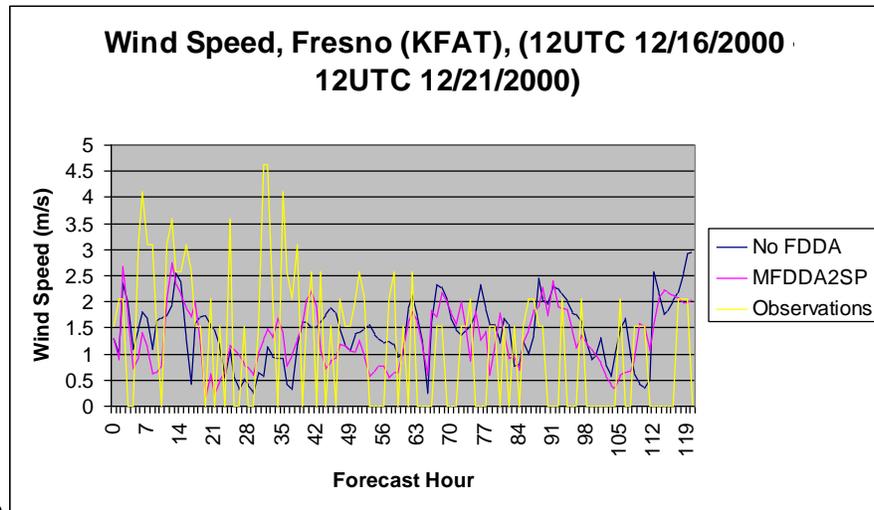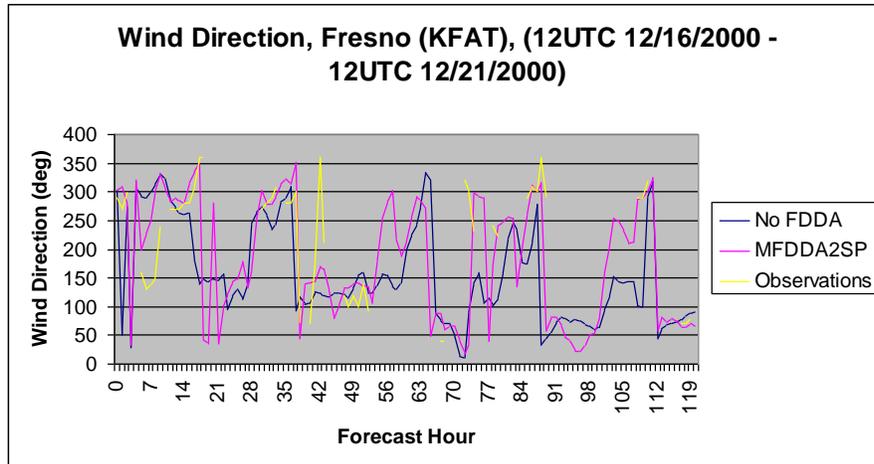
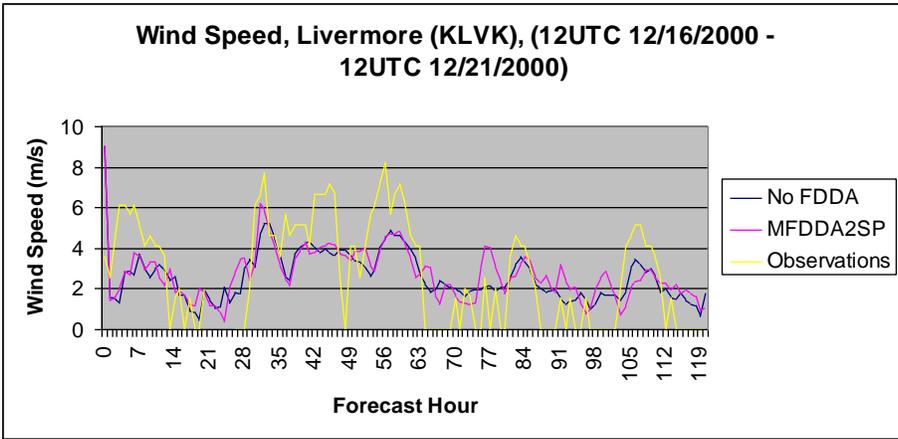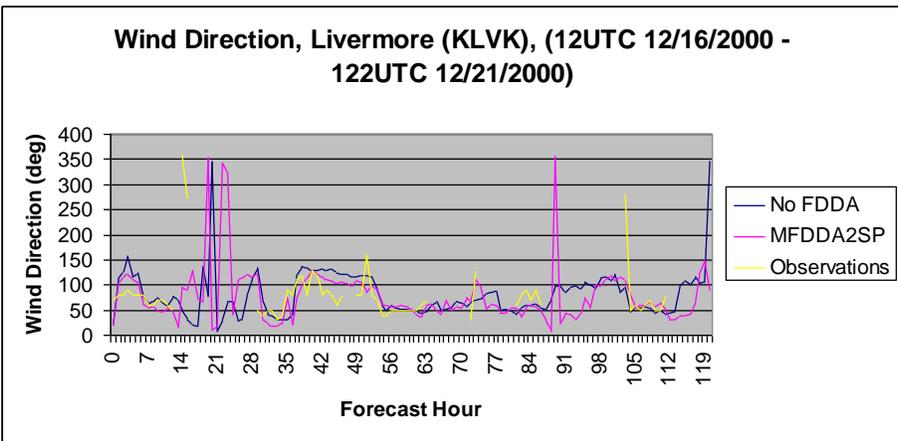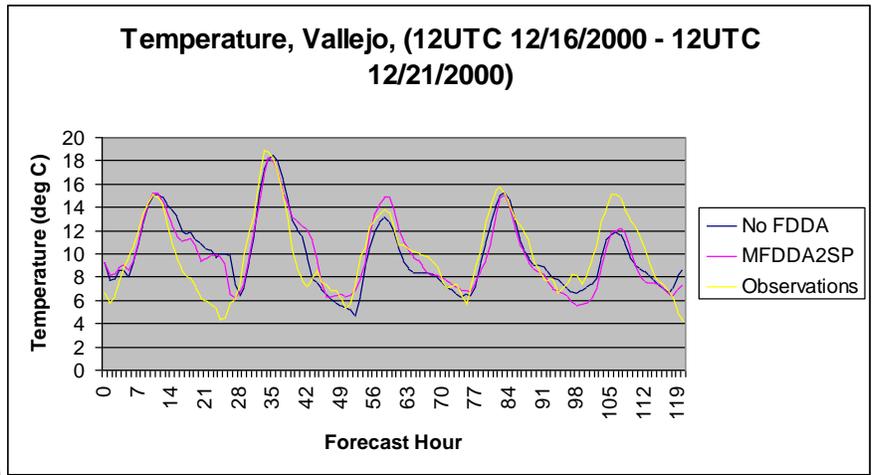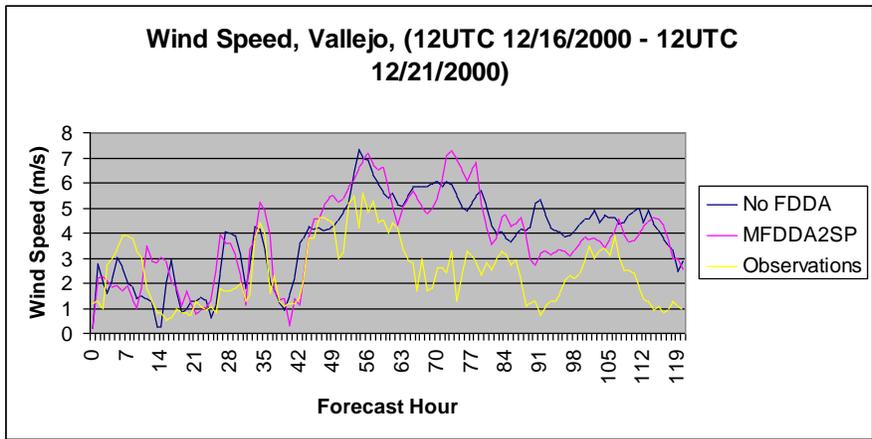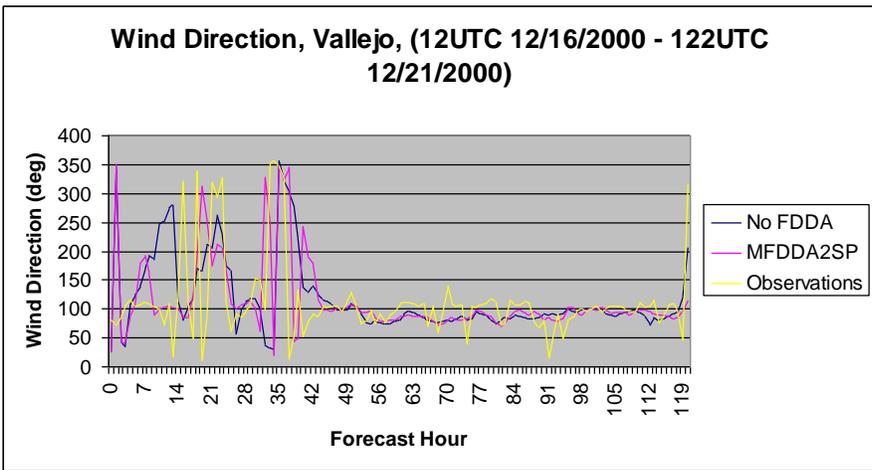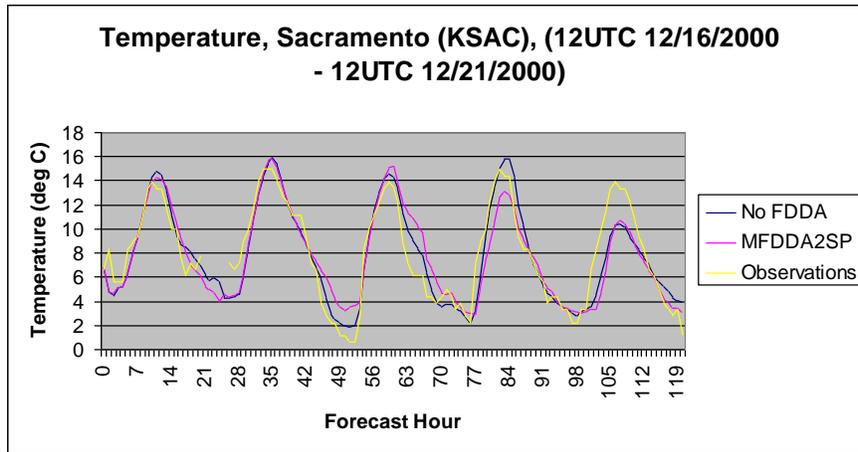**Figure 5-25: Time series plots for Fresno. a) Temperature. b) Wind Speed. c) Wind Direction.**

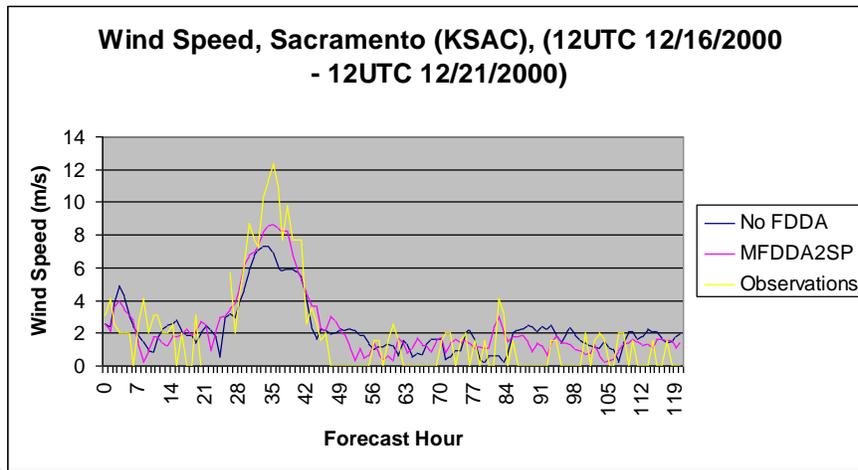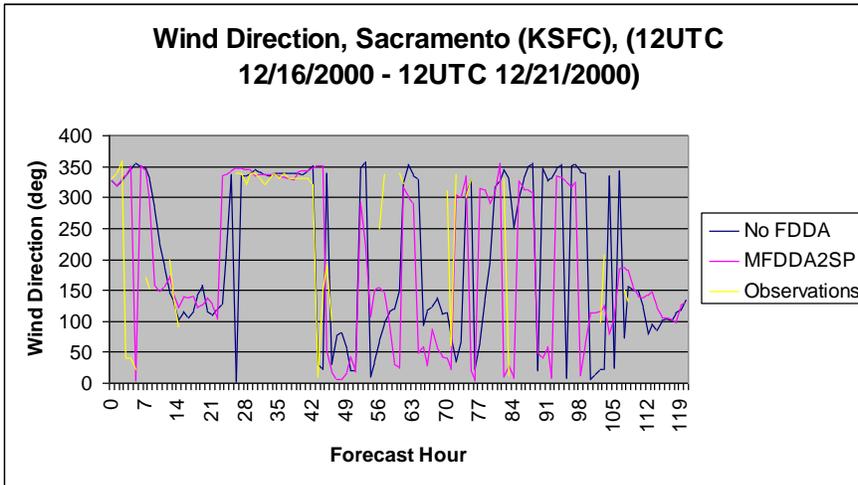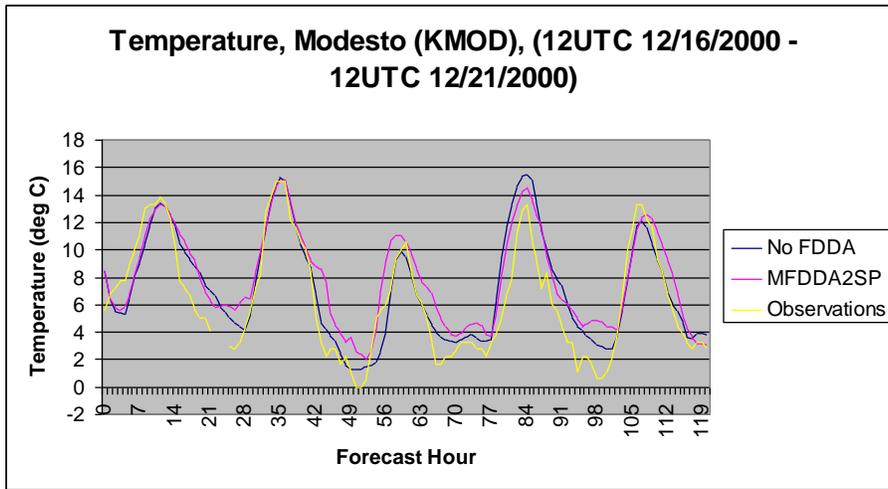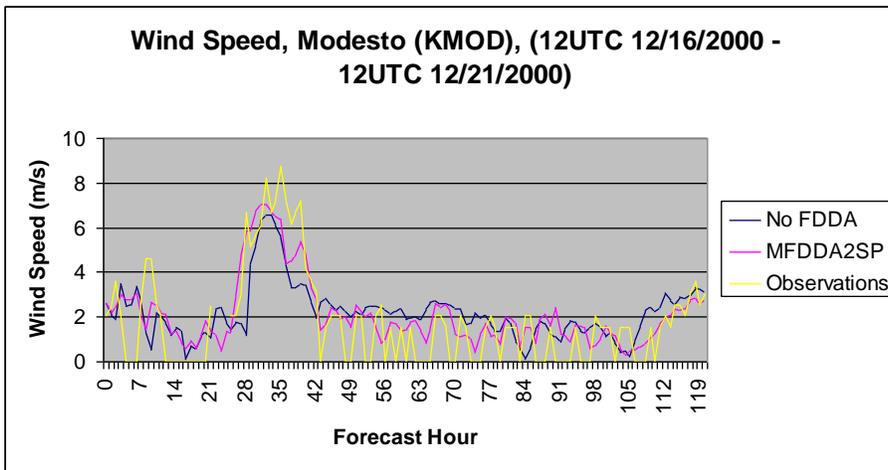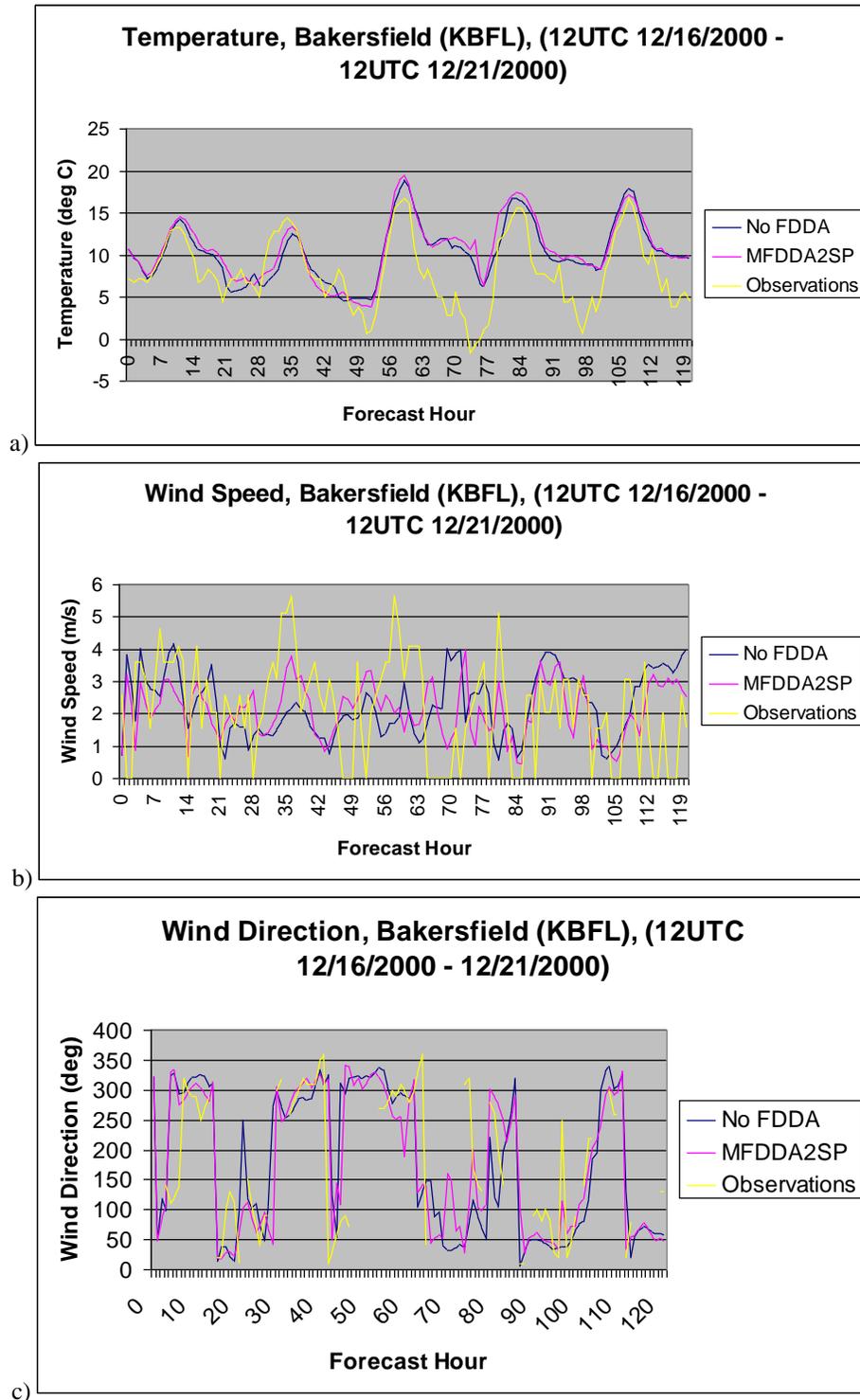**Figure 5-26: Time series plots for Bakersfield. a) Temperature. b) Wind Speed. c) Wind Direction.**

5.3   Summer Ozone Case Results

5.3.1 Objective Analysis Results

All statistical comparisons made for the summer ozone case were similar to the statistics for the winter case (averaged over each domain, averaged every three hours for the surface, averaged every twelve hours for upper air, etc.). This case had similar synoptic conditions with the winter case where there was very little moisture and precipitation throughout the time period. Therefore, the same radiation scheme used for the winter case (RRTM) was also used for the summer. As with the winter case, three LSMs were evaluated.

5.3.1.1 Entire 4-km Domain Verification

For the summer case, the 5 layer thermal diffusion scheme was not tested or compared with the other three LSMs because it produced the largest errors on the three domains for the four meteorological fields (not shown). Table 5-9 shows the surface MAE statistics comparing the NOAH, RUC, and the Pleim-Xiu (PX) land surface models (LSMs). All three LSMs performed similarly for all the fields with only 2% difference in RH, less than 0.5 ºC difference in temperature, about 1-2 degrees difference in wind direction, and all three produced nearly the same errors for the wind speed. Each LSM performed slightly better than the others for a particular field such as the NOAH LSM was the best for RH, while the RUC LSM was the best for temperature.

Table 5-9: Surface MAE Statistics comparing three land surface models

	36 km	12 km	4 km
Relative Humidity			
Noah	12.1	12.5	11.8
RUC	13.0	12.8	14.1
Pleim-Xiu	12.2	13.2	12.4
Temperature			
Noah	2.7	3.0	2.9
RUC	2.5	2.7	2.7
Pleim-Xiu	2.7	3.2	3.0
Wind Direction			
Noah	56.2	54.3	48.5
RUC	56.2	54.7	48.9
Pleim-Xiu	56.7	54.6	47.5
Wind Speed			
Noah	1.6	1.6	1.5
RUC	1.7	1.6	1.6
Pleim-Xiu	1.6	1.6	1.7

For upper air (Table 5-10), similar to the surface, all three LSMs performed about the same for all the fields. Overall, no particular LSM was better than the others for the summer case. Since the PX scheme was developed by the U.S. EPA and has been widely used for air quality applications, it was chosen, along with the RRTM radiation scheme, as the land surface model for use in the FDDA simulations for this case.

Table 5-10: Upper Air MAE Statistics comparing three land surface models

	36 km	12 km	4 km
Relative Humidity			
Noah	13.1	12.8	13.9
RUC	13.0	13.0	14.2
Pleim-Xiu	12.7	12.7	13.5
Temperature			
Noah	1.2	1.2	1.2
RUC	1.2	1.2	1.2
Pleim-Xiu	1.2	1.3	1.1
Wind Direction			
Noah	26.9	28.5	26.2
RUC	26.7	28.3	26.9
Pleim-Xiu	27.5	30.3	26.7
Wind Speed			
Noah	2.6	2.5	2.2
RUC	2.6	2.5	2.2
Pleim-Xiu	2.6	2.5	2.2

Figure 5-27 shows the surface MAE statistics comparing the experiment that did not use FDDA to two simulations that both used the multiscale FDDA method. The MFDDA4 experiment assimilated only WMO observations, while the MFDDA5 experiment assimilated both WMO observations and the same kind of special surface wind observations as used in the winter case. The special surface wind dataset used in this study consisted of data from the BAAQMD, AQS, and NCDC. Figure 5-27 a and b shows that using FDDA reduced the temperature and RH errors, even though no FDDA was applied for mass fields within the PBL. The largest improvement attributed to the use of FDDA was in the wind direction field (Figure 5-27c). Only a slight improvement was seen in the wind speed errors (Figure 5-27d) by the two MFDDA experiments which were expected since the wind was generally light in this ozone case. The MFDDA5 had the best fit statistically out of the three as was expected for the surface.

Figure 5-28 shows the surface ME comparison statistics of the three experiments. The MFDDA experiments had a slightly higher moisture bias than the NOFDDA experiment (Figure 5-28a). All three experiments had a cold bias of approximately 0.5-1 K (Figure 5-28b). The wind direction biases were quite small (Figure 5-28c). Wind speeds also slowed slightly (Figure 5-28d).

Figure 5-29 is similar to Figure 5-27, except now showing upper air MAE statistics for the three FDDA experiments. The assimilation of moisture and temperature data in the upper air allowed the MFDDA experiments to significantly reduce the RH errors (Figure 5-29a), while the temperature errors dropped by approximately 0.7 K compared to the NOFDDA experiment (Figure 5-29b). Both MFDDA experiments also significantly reduced the wind direction errors by approximately 15-22 degrees (Figure 5-29c) and wind speed errors by about 1.5 m/s (Figure 5-29d) compared to the NOFDDA simulation, as was the case for the wind direction and wind speed statistics for the surface. Once again the MFDDA5 experiment was the best experiment of the three for the upper air.
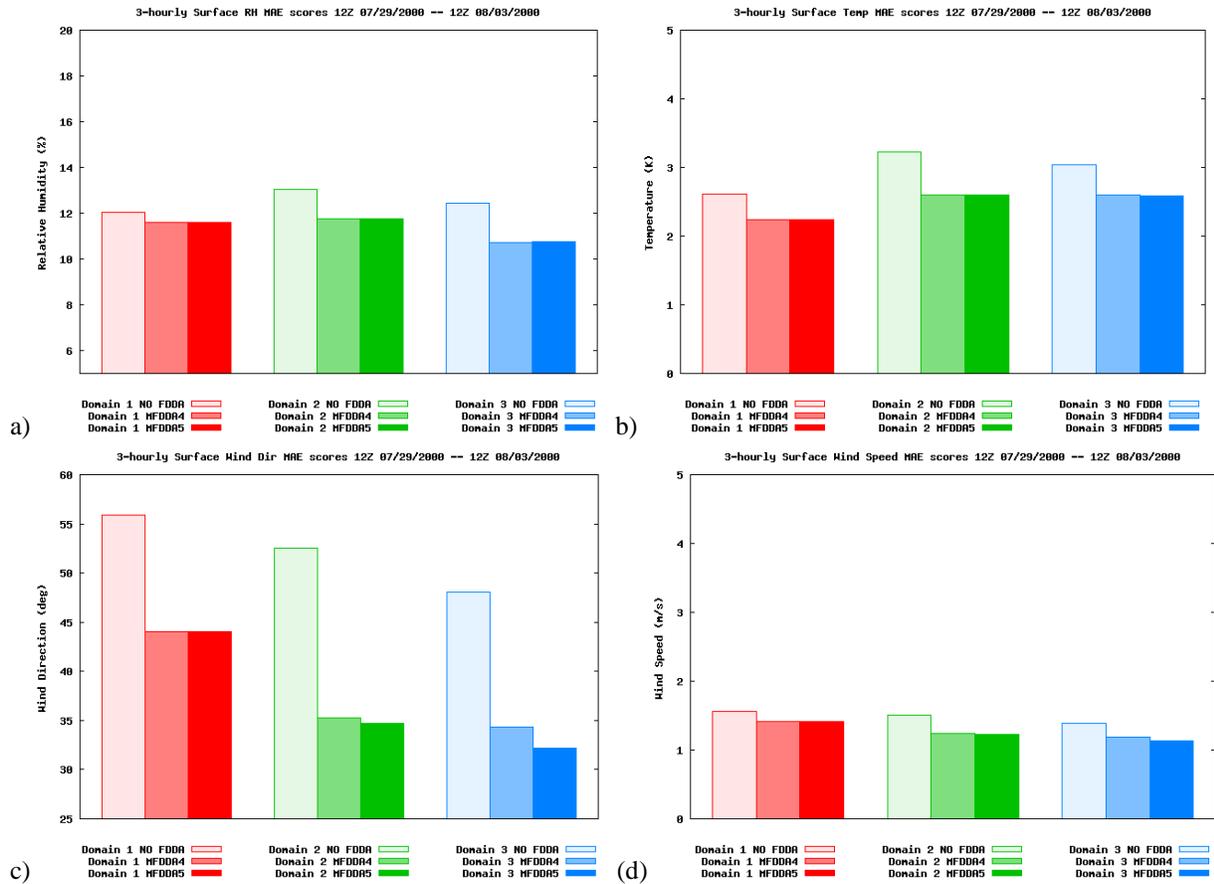
Figure 5-27: Surface MAE statistics comparing three FDDA experiments. a) RH. b) Temperature. c) Wind Direction. d) Wind Speed.
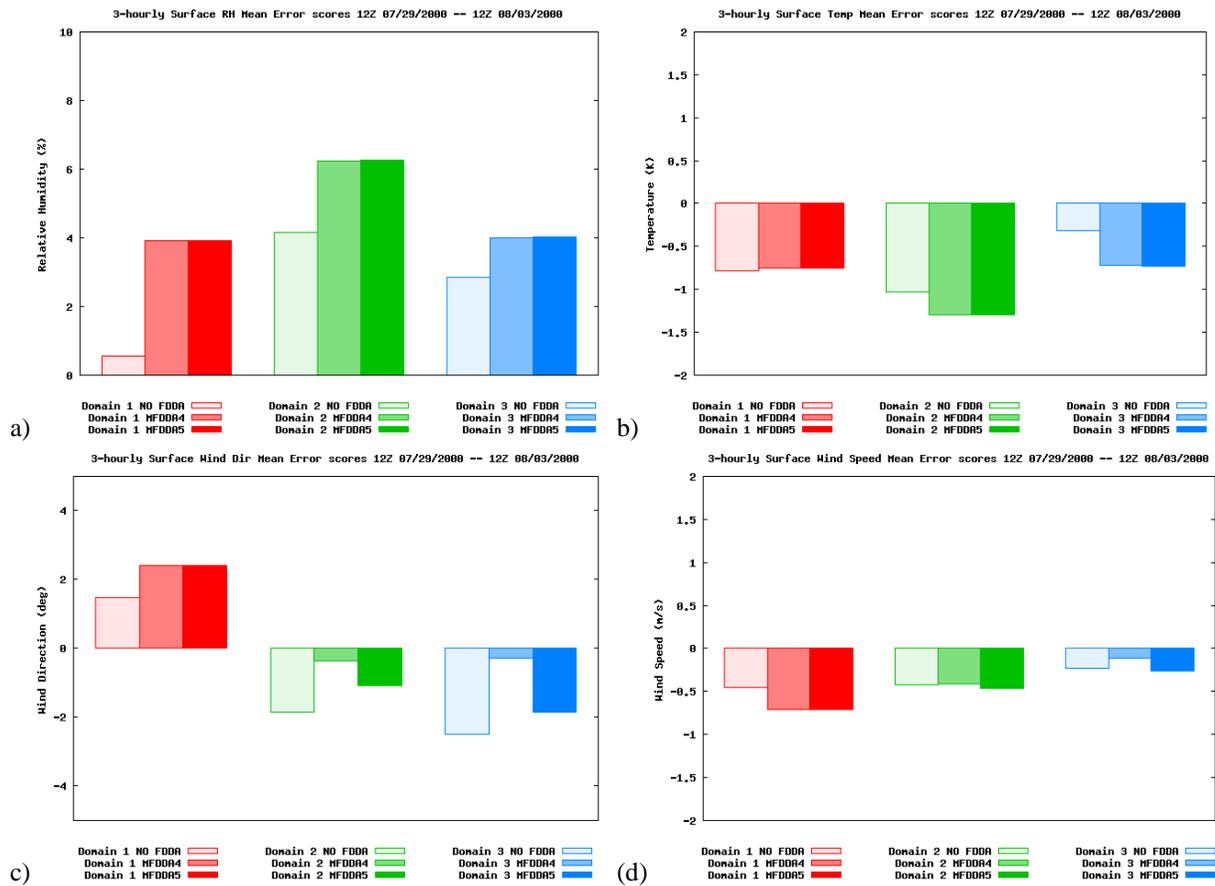
Figure 5-28: Surface ME statistics comparing three FDDA simulations. a) RH. b) Temperature. c) Wind Direction. d) Wind Speed.
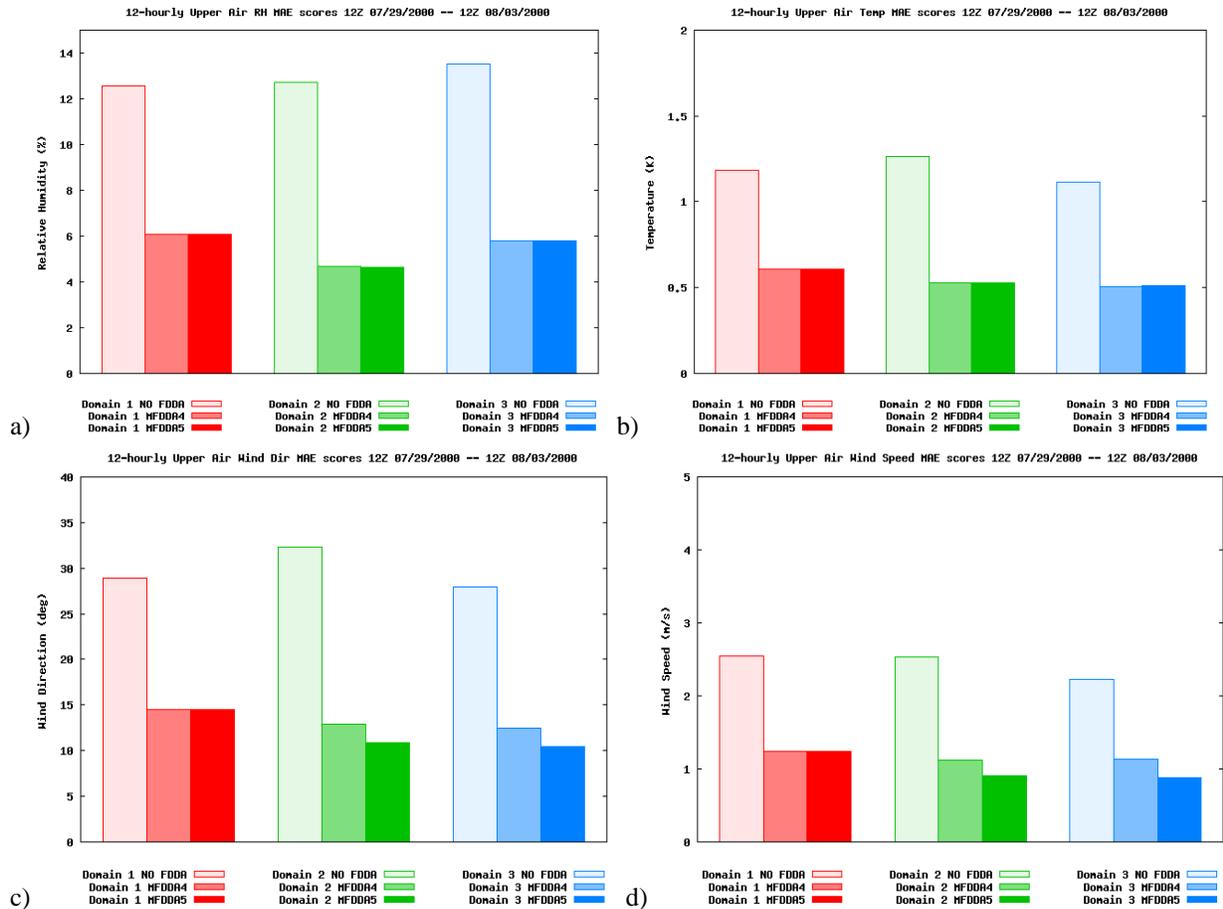
Figure 5-29: Upper air MAE statistics comparing three FDDA experiments. a) RH. b) Temperature. c) Wind Direction. d) Wind Speed.

Figure 5-30 shows the upper air ME statistics for the three experiments. The upper air biases for all four meteorological fields were much smaller than the surface biases (Figure 5-28). The MFDDA experiments essentially did not have any temperature (Figure 5-30b), wind direction (Figure 5-30b), or wind speed (Figure 5-30b) biases except for the small wind direction bias on 12-km domain. Note that there were only two upper air stations within the 4-km domain, both used in FDDA. The results by no means implied that the model was nearly perfect. Both MFDDA experiments overall performed better than the NOFDDA experiment, with the MFDDA5 experiment having performed the best of all three. Therefore, the MFDDA5 experiment would be compared to the NOFDDA experiment for the subregion statistical comparisons and for the independent verification.
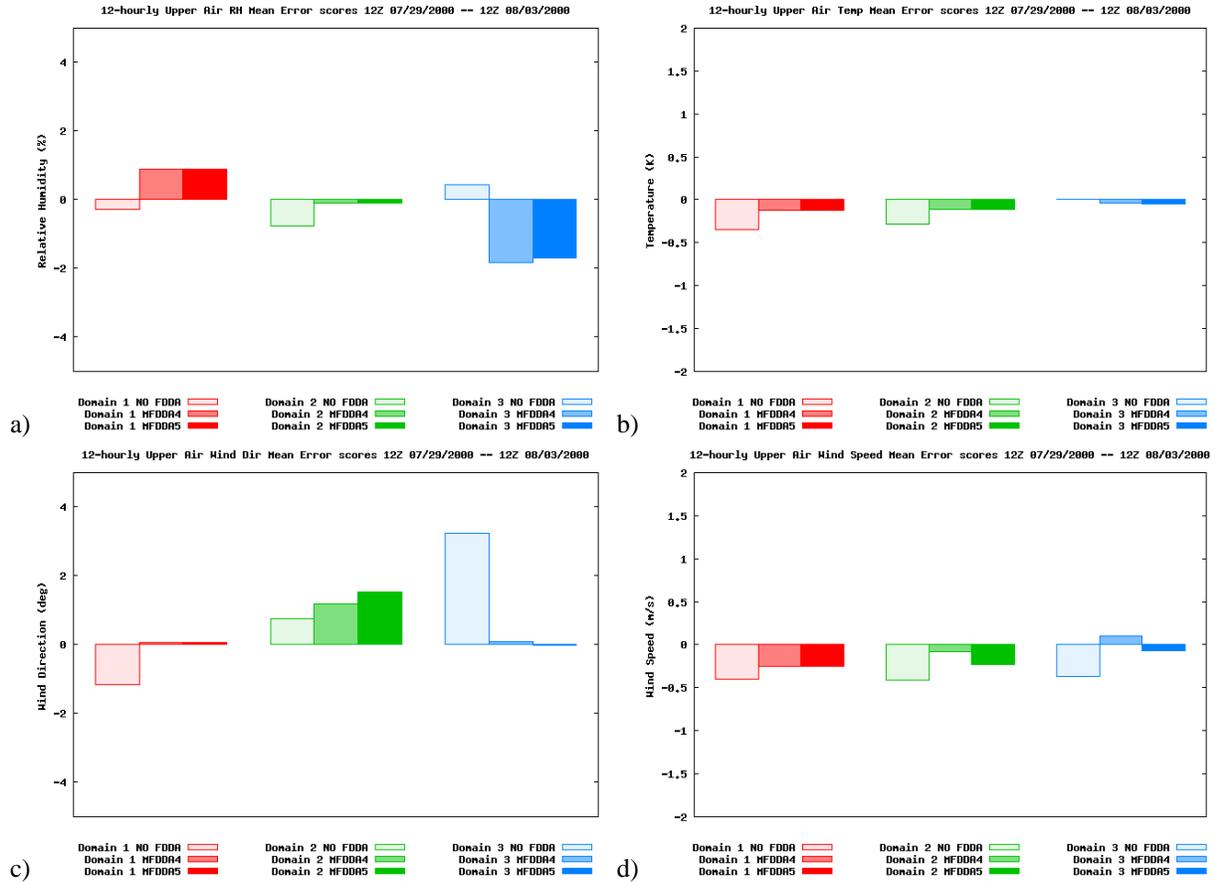
Figure 5-30: Upper air ME statistics comparing three FDDA experiments. a) RH. b) Temperature. c) Wind Direction. d) Wind Speed.

5.3.1.2 Subregion Verification

Similar to the winter case, statistical evaluations were performed for each of the three subregions: Bay Area, Sacramento Valley, and San Joaquin Valley. The surface MAE statistics for the Bay Area region are shown in Table 5-11. Both experiments produced nearly the same errors for the RH and temperature, with some slight degradation for MFDDA5. Note that there was no mass field assimilation within the PBL. The MFDDA5 experiment had smaller errors than the NOFDDA experiment for wind direction and wind speed due to assimilated of surface wind observations. The upper air MAE statistics for the Bay Area region are shown in Table 5-12. The MFDDA5 experiment did much better than the NOFDDA experiment and reduced errors by at least 50% on all four fields.

Table 5-11: Surface MAE Statistics for the Bay Area region comparing NOFDDA to MFDDA5.

	36 km	**12 km**	**4 km**
Relative Humidity			
NOFDDA	9.8	11.3	11.2
MFDDA5	10.4	10.1	10.3
Temperature			
NOFDDA	2.3	2.1	2.0

MFDDA5	2.4	2.2	1.9
Wind Direction			
NOFDDA	48.1	40.5	35.9
MFDDA5	40.7	34.8	28.8
Wind Speed			
NOFDDA	1.7	1.6	1.6
MFDDA5	1.5	1.4	1.2

Table 5-12: Upper Air MAE Statistics for the Bay Area region comparing NOFDDA to MFDDA5.

	36 km	12 km	4 km
Relative Humidity			
NOFDDA	12.1	12.4	12.6
MFDDA5	4.7	3.5	4.2
Temperature			
NOFDDA	1.1	1.1	1.1
MFDDA5	0.5	0.4	0.5
Wind Direction			
NOFDDA	27.1	27.4	28.0
MFDDA5	11.4	8.0	9.7
Wind Speed			
NOFDDA	2.2	2.2	2.2
MFDDA5	0.9	0.6	0.8

In the Sacramento Valley (Table 5-13) the MFDDA5 experiment performed better than the NOFDDA on all the domains at the surface for RH and temperature, unlike in the Bay Area region. Therefore, although there was no mass field assimilation, the effect of FDDA was evident in the Sacramento Valley. MFDDA5 significantly improved the wind speed and direction errors. There were no upper air statistics available for RH and temperature because the only upper air stations within the 4-km domain were outside the region. The upper air wind direction and wind speed statistics were similar to statistics for the surface and, therefore, are not shown.

In the San Joaquin Valley (Table 5-14) the MFDDA5 experiment performed better than the NOFDDA experiment at the surface for all the fields, except on the 36-km domain for the RH, where results were only slightly worse. As in the Sacramento Valley, the MFDDA5 only slightly improved RH, temperature, and wind speed errors, while wind direction errors were significantly improved. The upper air statistics were not available for RH and temperature because no upper air stations were located within the region. The upper air statistics for wind direction and wind speed were similar to statistics for the surface and, therefore, are not shown.

Table 5-13: Surface MAE Statistics for the Sacramento Valley region comparing NOFDDA to MFDDA5.

	36 km	12 km	4 km
Relative Humidity			
NOFDDA	11.2	14.5	11.8
MFDDA5	10.1	13.3	10.1
Temperature			
NOFDDA	2.7	3.2	2.6
MFDDA5	2.3	2.8	2.3
Wind Direction			
NOFDDA	42.5	40.3	48.1
MFDDA5	35.4	29.9	28.6
Wind Speed			
NOFDDA	1.3	1.3	1.3
MFDDA5	1.2	1.1	1.1

Table 5-14: Surface MAE Statistics for the San Joaquin Valley region comparing NOFDDA to MFDDA5.

	36 km	12 km	4 km
Relative Humidity			
NOFDDA	9.1	12.5	17.8
MFDDA5	10.4	12.3	14.8
Temperature			
NOFDDA	2.4	3.4	3.7
MFDDA5	2.3	2.8	3.4
Wind Direction			
NOFDDA	51.6	45.8	50.9
MFDDA5	45.4	32.0	33.6
Wind Speed			
NOFDDA	1.3	1.2	1.1
MFDDA5	1.1	1.0	1.0

5.3.1.3 Independent Verification

Similar to the winter case, the NOFDDA experiment and the MFDDA5 experiment were compared with an independent set of observations to judge the performance of the model. The independent data consisted of observations taken from the EPA's AQS database. Table 5-15 shows the surface MAE statistics averaged over the entire 4-km domain. The MFDDA5 experiment performed better than the NOFDDA for all the fields, with a slight improvement in temperature and wind speed errors, and somewhat larger improvements in the wind direction errors. This independent verification demonstrated again that multiscale FDDA was an effective technique to reduce model errors and produce high quality gridded mesoscale analysis for use in air quality applications.

Table 5-15: Surface MAE Statistics comparing NOFDDA to MFDDA5 against an independent dataset.

	36 km	12 km	4 km
Temperature			
NOFDDA	3.0	3.1	2.9
MFDDA5	2.8	2.8	2.7
Wind Direction			
NOFDDA	54.9	48.1	50.6
MFDDA5	50.6	38.4	39.3
Wind Speed			
NOFDDA	1.4	1.4	1.3
MFDDA5	1.3	1.2	1.2

5.3.2 Subjective Analysis Results

5.3.2.1 36-km Domain Verification

Subjective analyses of the numerical simulations for the summer case were conducted for the 36-km and 4-km domains, subregions, and individual observation sites. Only the baseline (NOFDDA) and the best FDDA experiments (MFDDA5) were examined.

On the 36-km domain, at 1200 UTC 29 July 2000, the model initial condition at the surface (Figure 5-31a) showed the trough that was observed over California (Figure 3-13a). A low pressure center was located over western Arizona while a pressure high was located over central Colorado. There was north-northwesterly flow along the California coast. The 850 mb model initial condition (Figure 5-31b) showed that over the ocean, there was a low pressure system to the north and high pressure system to the south, which produced the on-shore flow along the coast of northern California, Oregon, and Washington. Over the southern tip of Nevada, a cyclonic eddy caused offshore flow south of the Bay Area. The 500 mb initial conditions (Figure 5-31c) showed a high pressure system over central Arizona with the western edge of the high over California.
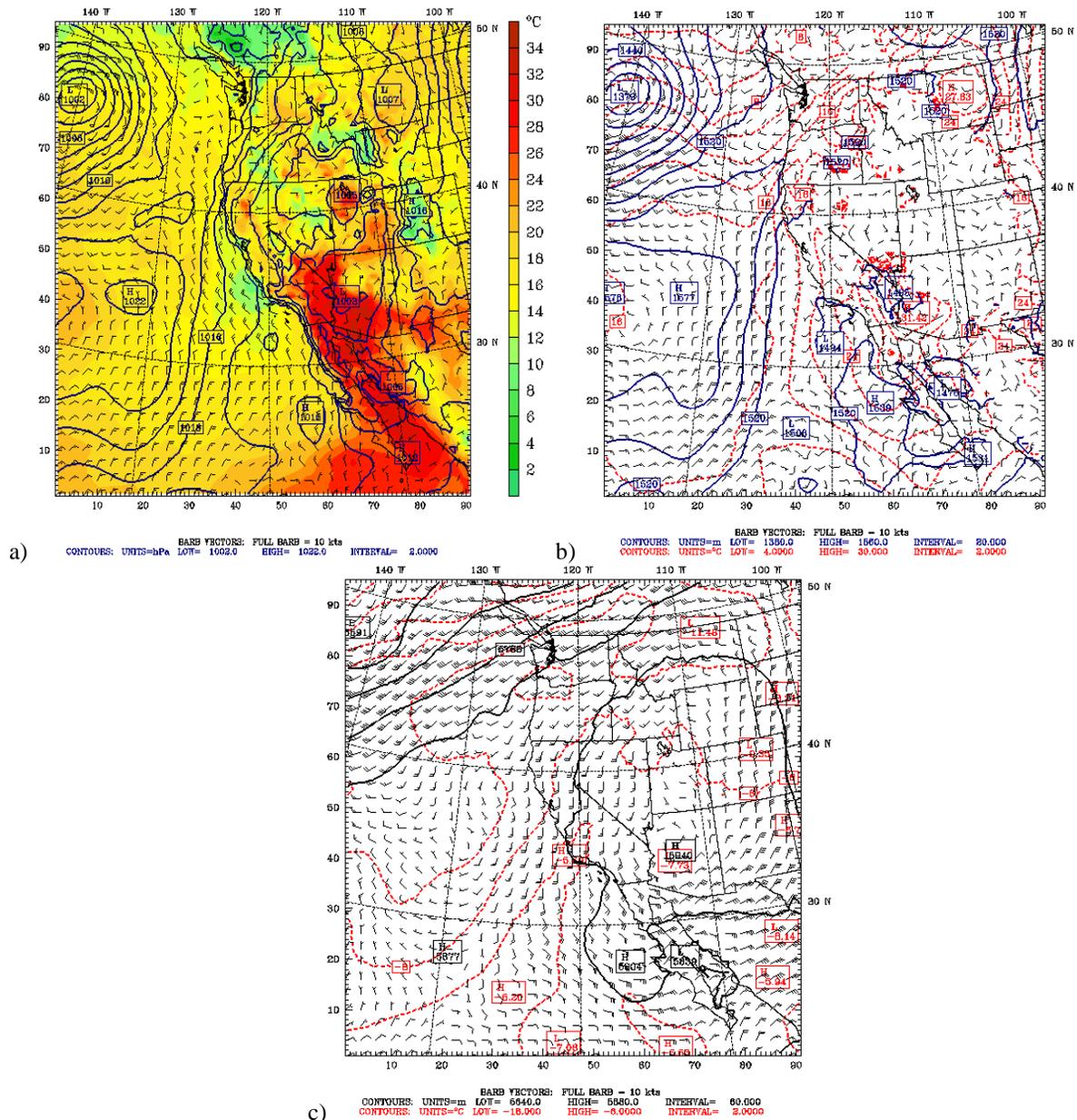
Figure 5-31: Plots of NOFDDA simulation at July 29, 2000. a) Surface. b) 850 mb. c) 500 mb.

Over the next 72 hours the observed synoptic conditions did not change much. The observed surface heat lows and their associated troughs remained over California and the southwest U.S., while a high remained centered over Colorado (Figure 3-15). Model simulations with and without FDDA did a good job in simulating the synoptic conditions. For example, at 1200 UTC 1 Aug 2000, 72 hours into the simulation, both the NOFDDA and the MFDDA5 simulations placed a surface high over the Pacific Ocean west of Washington and Oregon (Figure 5-32a and b) which agreed with observations (not shown). The pressure gradient that was simulated along the coast and the high centered over southern Colorado also agreed with observations. Both simulations reproduced the observed winds. Both had the observed onshore winds over the Bay Area and westerly flow through the Carquinez Strait into the Central Valley. The observations showed southerly winds over Fresno, which both simulations did pick up,

along with the observed northerly flow over the Sacramento Valley.

On the 850 mb level, the MFDDA5 simulation did a superior job over the NOFDDA run in producing the strong pressure gradient around the high that was observed off the coast of California (Figure 5-33). MFDDA5 also represented the ridge over Oregon better as well. Both simulations developed the temperature pattern observed over the ocean west of Washington and Oregon (not shown). MFDDA5 simulated winds were much better because it did a superior job in resolving the observed easterly and northeasterly offshore flow over the Bay Area. In contrast, the NOFDDA results had westerly onshore flow over the Bay Area. This comparison indicated that MFDDA5 did a better job on the 850 mb level than NOFDDA.

On the 500 mb level, at 1200 UTC 1 Aug 2000, the observed high progressed to the north and was centered over northern Nevada (not shown). The NOFDDA extended the high over the ocean, which agreed with the observations, although the observed high extended a little farther out to sea. In contrast, MFDDA5 only extended the edge of high to the coast (Figure 5-34). Both simulations did a good job reproducing the winds, especially the southerly and southeasterly flows observed over the Bay Area and Central Valley. Overall, both the NOFDDA and the MFDDA5 simulations did a reasonable job of representing the 500-mb observations on the 36-km domain.
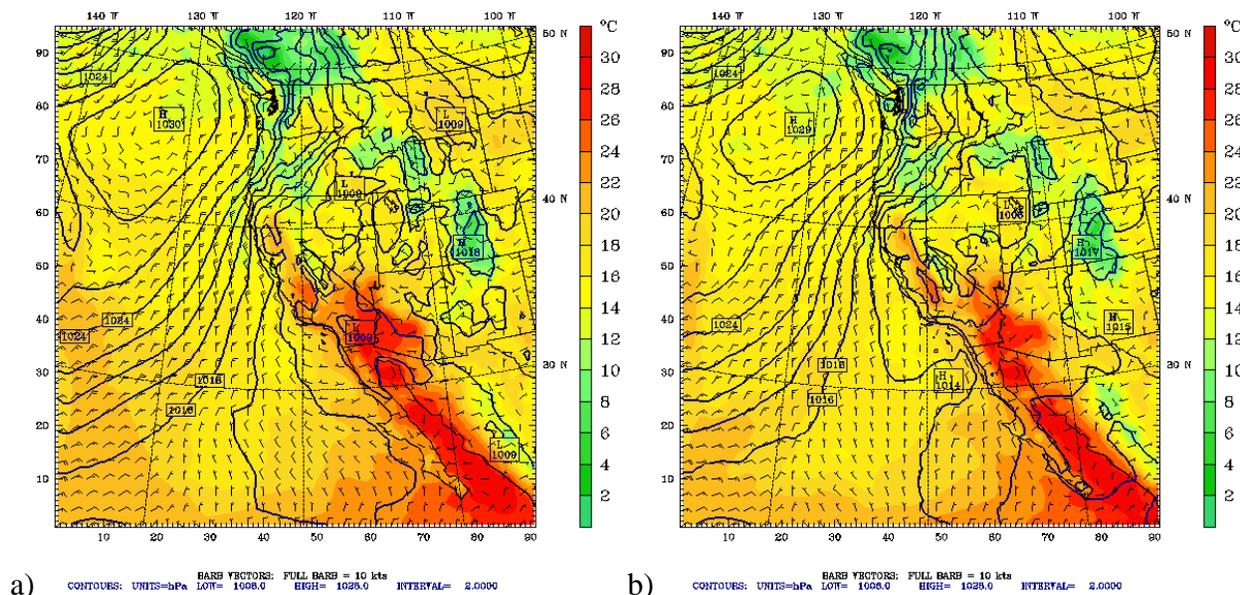
Figure 5-32: Surface plots at 1200 UTC August 1, 2000: a) NOFDDA. b) MFDDA5.
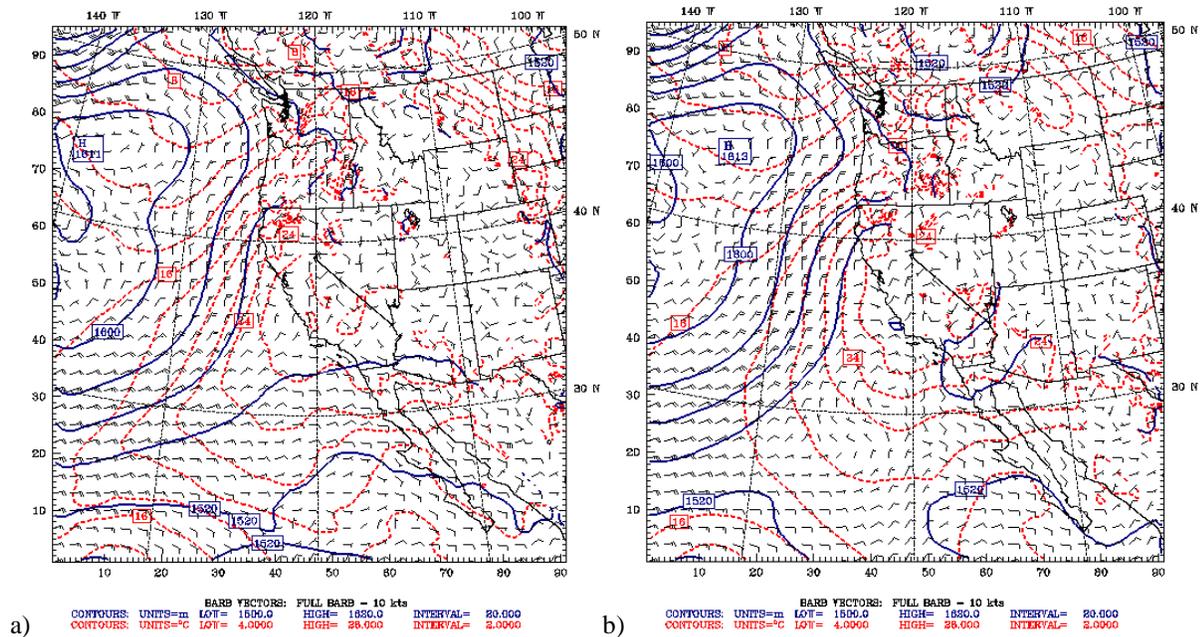
Figure 5-33: 850 mb plots at 1200 UTC August 1, 2000: a) NOFDDA. b) MFDDA5.
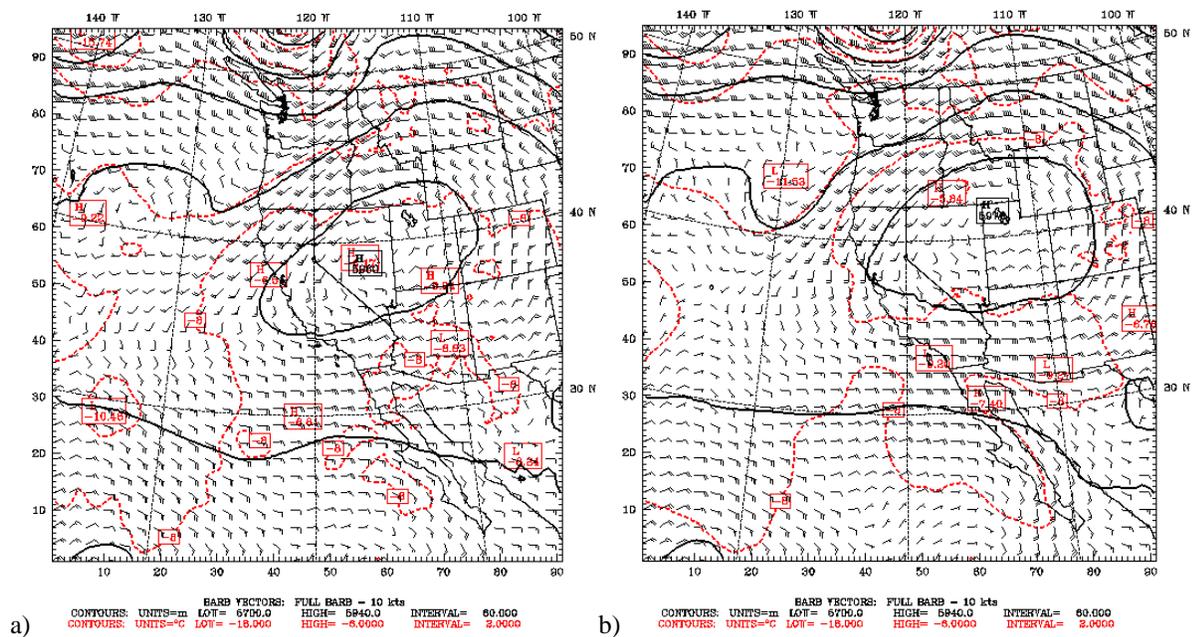
Figure 5-34: 500 mb plots at 1200 UTC August 1, 2000: a) NOFDDA. b) MFDDA5.

5.3.2.2 4-km Verification

As with the winter case, the application of FDDA improved the WRF-ARW simulation of the summer case. Figure 5-35 shows the WRF-simulated surface-layer winds on the 4-km domain at 0000 UTC 30 July 2000, 12 hours into the simulation. The simulated winds from both simulations (only MFDDA simulation is shown in Figure 5-35) did a fair job in confirming the conceptual model from Bao et al. (2008) (Figure 1-1a) for the daytime. The westerly onshore winds over the Bay Area represented the incoming marine flow. Ozone and PM concentrations

72

could possibly be carried away from the Bay Area by this flow. The marine flow in the model passed through the Carquinez Strait and into the Central Valley, where it split into two mean branches (represented by the heavy purple shading of the streamlines): 1) southerly flow up the Sacramento Valley and 2) northerly flow down the San Joaquin Valley (Figure 5-35). As a result of this splitting, ozone and PM being transported to cities such as Sacramento, Fresno, and Bakersfield existed. The heating of the mountaintops all along the Central Valley created upslope flows that caused smaller divergences from the two mean flows. At the southern end of the San Joaquin Valley, the northerly mean flow left the valley through the Tehachapi Mountains. The observations showed that the simulated winds were in agreement.
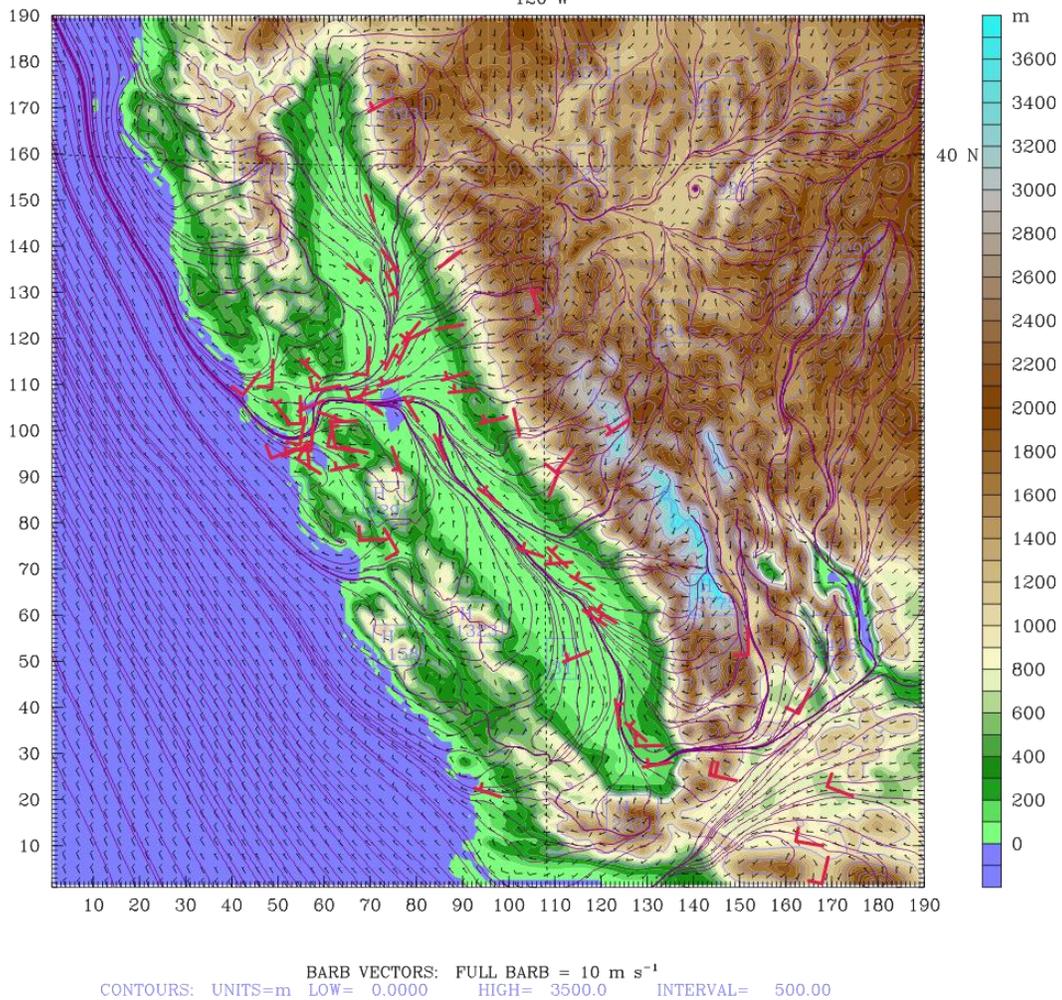
Figure 5-35: Plot of surface winds for the MFDDA5 simulation.

Twelve hours into the simulation at 1200 UTC 30 July 2000, differences between the two simulations developed for the nighttime flow pattern (Figure 5-36). The winds shown in the
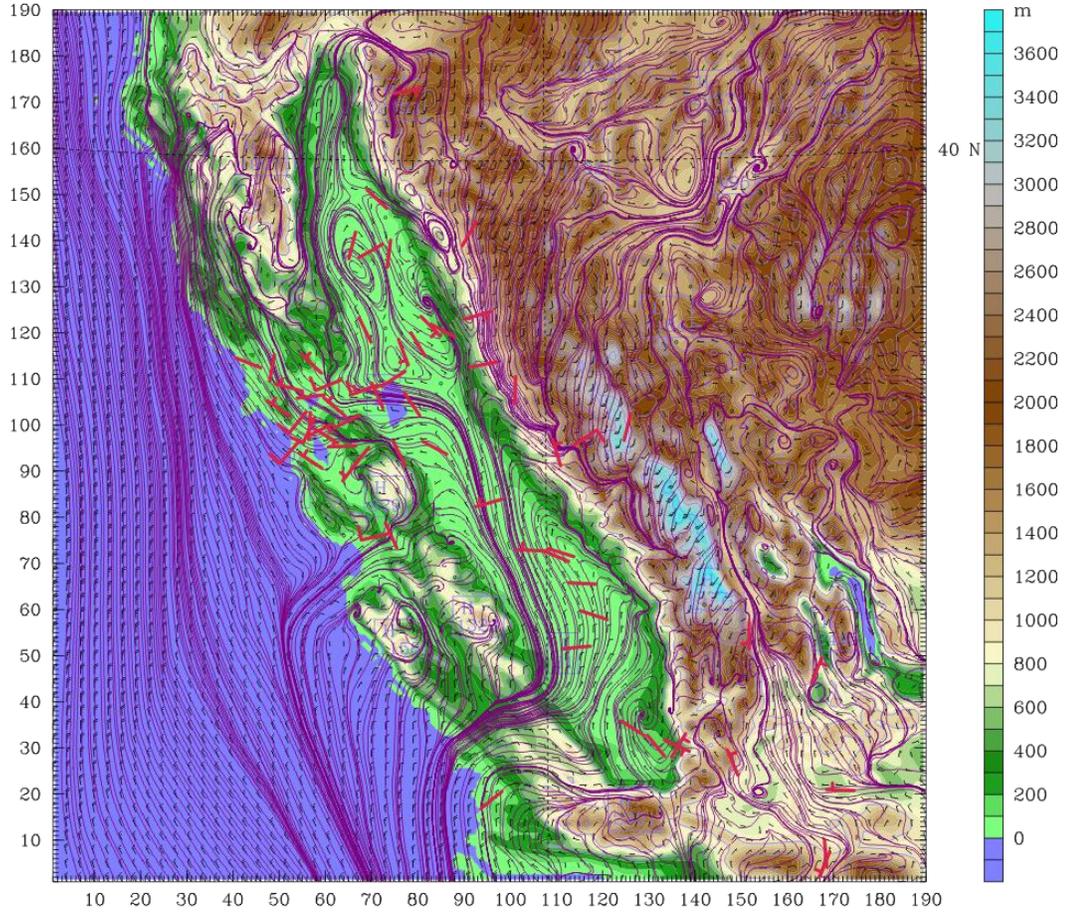
figure were about 300 meters above the ground (the 9th η level) corresponding to the mesoscale eddies indicated in previous papers including Seaman et al 1995, Liu and Jao 1995 and Bao et al. 2008. In the NOFDDA simulation (Figure 5-36a) the incoming marine flow over the Bay Area was not as distinguishable as in the MFDDA5 (Figure 5-36b). Both simulations developed the downslope flow from the mountains along the Central Valley. This downslope flow then converged with the main bulk northerly flow down both the Sacramento Valley and San Joaquin Valley. The downslope flow from the Sierra Nevada northeast of Sacramento (SAC) was southeasterly before crossing the valley and turning northerly to converge with the mean flow. This southerly flow and the northerly bulk flow developed the Schultz Eddy. The location of the Schultz Eddy was farther north in the NOFDDA run versus the MFDDA5 simulation, which had the location of the eddy just northwest of Sacramento. The location of the Schultz Eddy in the MFDDA5 agreed better with the conceptual model from Bao et al (2008) (Figure 1-1b). The downslope flow from the Sierra Nevada Mountains northeast of Bakersfield (BFL) was southerly as well. This flow crossed the valley before turning northerly, converging with mean flow. This southerly flow and the northerly mean flow created the Fresno Eddy. However, neither simulation developed the Fresno Eddy that was in the conceptual model. Instead both developed an eddy further south over Bakersfield (BFL). The northerly mean flow in the San Joaquin Valley left the valley over the Coastal Range's southern end, agreeing with the nighttime observed flow pattern. The observations showed the northerly bulk flow down the entire Central Valley and the downslope flow from the mountains around the valley as were simulated.

In order to demonstrate the existence of Fresno Eddy, a different nocturnal time was chosen. Figure 5-37 shows the simulated low-level winds (at the same 9th η level from the surface) from both simulations at 1200 UTC 1 Aug 2000, 72 hours into the simulation. Both simulations developed the Fresno Eddy at this time. However, the Schultz Eddy was not formed by this time by either simulation. The NOFDDA simulation also developed a secondary eddy over Bakersfield, while the MFDDA5 developed a secondary eddy near Modesto (MOD). Overall, the NOFDDA and the MFDDA5 simulated the winds reasonably well, but MFDDA5 did a better job with the location of the Schultz Eddy and developed a more pronounced marine flow than the NOFDDA.

Dataset: d03   RIP: 2000−07 29−03 NO FDDA d03 wi   Init: 1200 UTC Sat 29 Jul 00
Fcst: 24.00 h   Valid: 1200 UTC Sun 30 Jul 00 (0400 PST Sun 30 Jul 00)
Terrain height AMSL
Terrain height AMSL
Horizontal wind vectors   at k−index = 42
Horizontal wind streamlines   at k−index = 42

BARB VECTORS: FULL BARB = 10 kts
CONTOURS: UNITS=m LOW= 0.0000   HIGH= 3500.0   INTERVAL=   500.00

a)

Dataset: d03   RIP: 2000-07 29-03 MFDDA5 d03 win   Init: 1200 UTC Sat 29 Jul 00
Fcst: 24.00 h   Valid: 1200 UTC Sun 30 Jul 00 (0400 PST Sun 30 Jul 00)
Terrain height AMSL
Terrain height AMSL
Horizontal wind vectors   at k-index = 42
Horizontal wind streamlines   at k-index = 42

BARB VECTORS: FULL BARB = 10 kts
CONTOURS: UNITS=m LOW= 0.0000   HIGH= 3500.0   INTERVAL=   500.00

b)

Figure 5-36: Simulated winds at 1200 UTC 30 July 2000. a) NOFDDA. b) MFDDA5.

Dataset: d03   RIP: 2000-07 29-03 NO FDDA d03 wi   Init: 1200 UTC Sat 29 Jul 00
Fcst: 72.00 h   Valid: 1200 UTC Tue 01 Aug 00 (0400 PST Tue 01 Aug 00)
Terrain height AMSL
Terrain height AMSL
Horizontal wind vectors   at k-index = 42
Horizontal wind streamlines   at k-index = 42
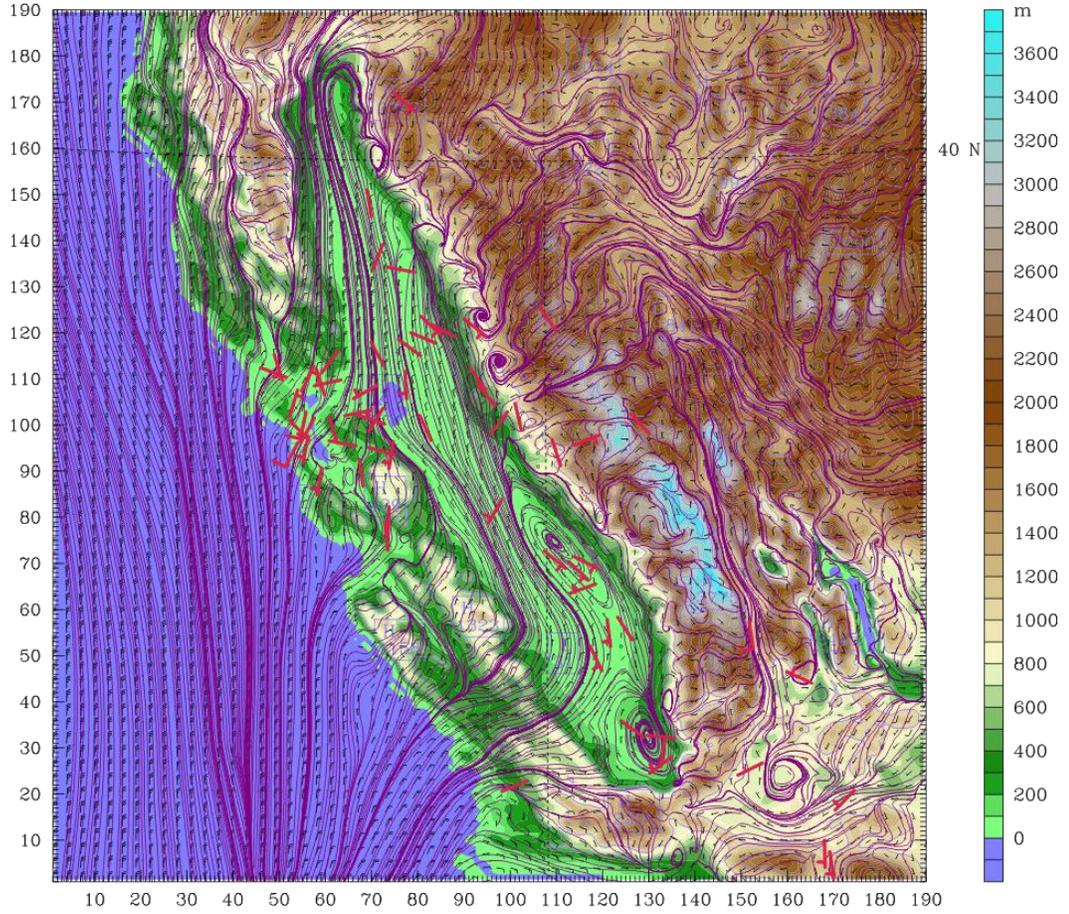
BARB VECTORS: FULL BARB = 10 kts
CONTOURS: UNITS=m LOW= 0.0000   HIGH= 3500.0   INTERVAL=   500.00

a)
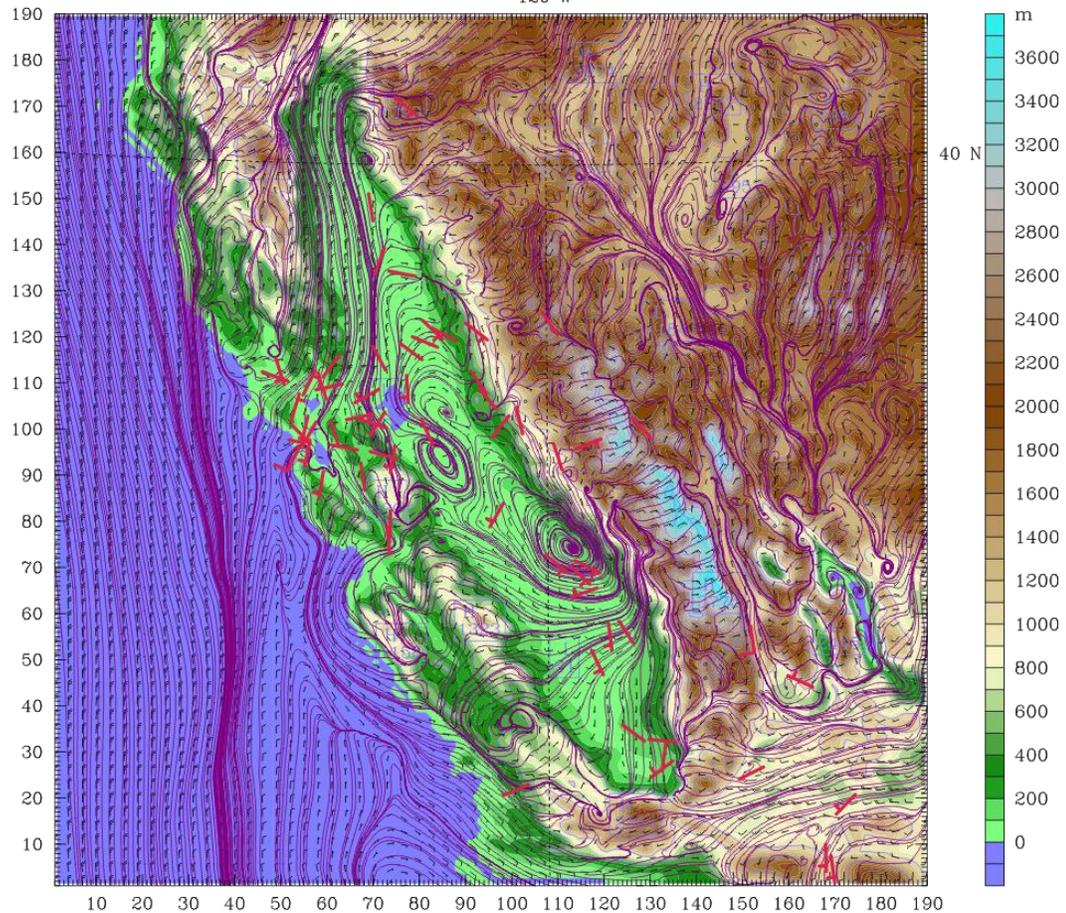
Figure 5-37: Simulated winds at 1200 UTC 1 Aug 2000. a) NOFDDA. b) MFDDA5.

5.3.2.3 Subregion Verification

In the Bay Area region, at 0600 UTC 30 July 2000 (10:00 PST July 29, 2000), 18 hours into the simulation, the incoming marine flow through the Golden Gate Bridge area was evident in both simulations (only the MFDDA5 simulation is shown in Figure 5-38). The incoming flow split over the Bay Area; one branch turned northward towards Vallejo (VAL), where it turned east through the Carquinez Strait. The second branch turned southerly toward Pleasanton (PLE), before turning east toward Livermore (LIV). The observations supported this splitting. The splitting of the flow occurred due to the cooling influence of the terrain which prevented the flow from going over the terrain. The split flows passed through the Carquinez Strait and the Livermore Valley before converging with the bulk flow in the Central Valley. The branch of the split flow that turned southerly (toward Pleasanton) also had a secondary split flow that continued south toward San Jose (SJC).

78

Thirty-Six hours into the simulation period, at 0000 UTC 31 July 2000 (local afternoon), the incoming marine flow did not split over the Bay Area and instead it flowed over the terrain as ground heating of the terrain allowed for upslope flow (Figure 5-39 a and b). The NOFDDA simulation produced some northerly flow through the I-680 corridor, where it converged over the Livermore Valley westerly flow. The flow over Livermore continued toward the south which did not agree with observations. The MFDDA5 showed the marine flow continuing eastward over the I-680 corridor toward the Central Valley. The northerly flow toward San Jose was produced in both simulations. Both simulations agreed well with observations for the 18 hour time, but the MFDDA5 simulation did better than NOFDDA at 36 hours.

Figure 5-38: Simulated winds in the Bay Area region for the MFDDA5 experiment at 0600 UTC 30 July 2000.

Dataset: d03 RIP: 2000-07 29-03 NO FDDA d03 wi   Init: 1200 UTC Sat 29 Jul 00
Fcst: 36.00 h   Valid: 0000 UTC Mon 31 Jul 00 (1600 PST Sun 30 Jul 00)
Terrain height AMSL
Terrain height AMSL
Horizontal wind vectors   at k-index = 50
Horizontal wind streamlines   at k-index = 50
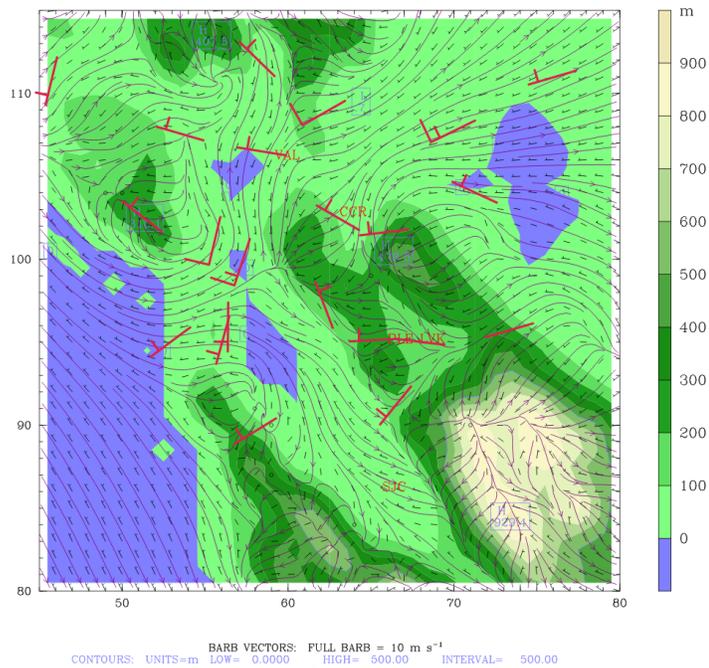
a)

Dataset: d03 RIP: 2000-07 29-03 MFDDA5 d03 win   Init: 1200 UTC Sat 29 Jul 00
Fcst: 36.00 h   Valid: 0000 UTC Mon 31 Jul 00 (1600 PST Sun 30 Jul 00)
Terrain height AMSL
Terrain height AMSL
Horizontal wind vectors   at k-index = 50
Horizontal wind streamlines   at k-index = 50

b)

Figure 5-39: Simulated winds in the Bay Area region at 0000 UTC 31 July 2000. a) NOFDDA. b) MFDDA5.

80

In the Sacramento Valley, the nighttime downslope flow off the mountains surrounding the valley began at 0600 UTC 30 Dec 2000 in both simulations (not shown). Six hours later at 1200 UTC 30 July 2000, 24 hours into the simulation, the nighttime pattern had developed, with northerly flow down the valley (Figure 5-40). The NOFDDA simulation developed numerous eddies within the valley (Figure 5-40a). The largest eddy in the simulation was located about 60 km north of Sacramento and had a diameter of about 80 km from north to south. This eddy was most likely the Schultz Eddy. The MFDDA5 had fewer eddies, with the largest eddy centered just northwest of Sacramento (Figure 5-40b), with a diameter of about 40 km. The location of this eddy agreed well with the conceptual model for the location of the Schultz Eddy and was consistent with observations.

a)

Dataset: d03 RIP: 2000-07 29-03 MFDDA5 d03 win   Init: 1200 UTC Sat 29 Jul 00
Fcst: 24.00 h   Valid: 1200 UTC Sun 30 Jul 00 (0400 PST Sun 30 Jul 00)
Terrain height AMSL
Terrain height AMSL
Horizontal wind vectors   at k-index = 42
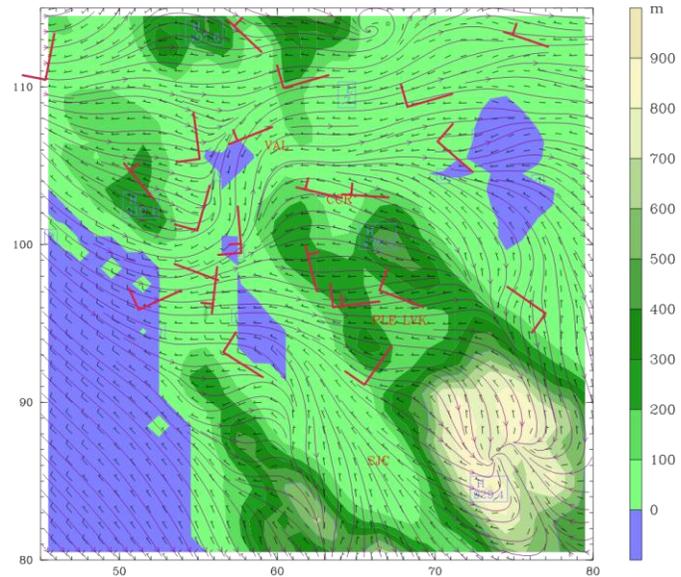Horizontal wind streamlines   at k-index = 42

BARB VECTORS: FULL BARB = 10 kts
CONTOURS: UNITS=m LOW= 0.0000   HIGH= 2000.0   INTERVAL= 500.00

b)

Figure 5-40: Simulated winds in the Sacramento Valley region at 1200 UTC 30 July 2000. a) NOFDDA. b) MFDDA5.

Six hours later, at 1800 UTC 30 July 2000, there was still northerly flow down the valley (not shown), but divergence of the flow dad developed as surface heating of the terrain created the upslope flow along the mountains. Six hours later, at 0000 UTC 31 July 2000, 36 hours into the simulation, the daytime pattern had developed. Southerly flow up the valley was present in both simulations (Figure 5-41), but the NOFDDA had created a divergence zone that spanned the valley (Figure 5-41a). This divergence was probably due to air moving away from the valley floor up into the mountains. North of this divergence zone there was southerly flow up the valley as expected. However, south of the divergence zone there was northerly flow down the valley toward Sacramento. The MFDDA5 simulation (Figure 5-41b) did not produce the divergence zone and showed southerly flow up the entire valley. Overall, for the entire 5-day period, both simulations developed the daytime pattern, with southerly flow up the valley and upslope along the mountains, and the nighttime pattern, with northerly flow down the valley and downslope flow along the mountains. However, the NOFDDA simulation created too many eddies and a divergence zone across the valley. The MFDDA5 simulation created fewer eddies and did not create the divergence zone, which agreed well with observations.

Dataset: d03   RIP: 2000-07 29-03 NO FDDA d03 wi   Init: 1200 UTC Sat 29 Jul 00
Fcst: 36.00 h   Valid: 0000 UTC Mon 31 Jul 00 (1600 PST Sun 30 Jul 00)
Terrain height AMSL
Terrain height AMSL
Horizontal wind vectors   at k-index = 50
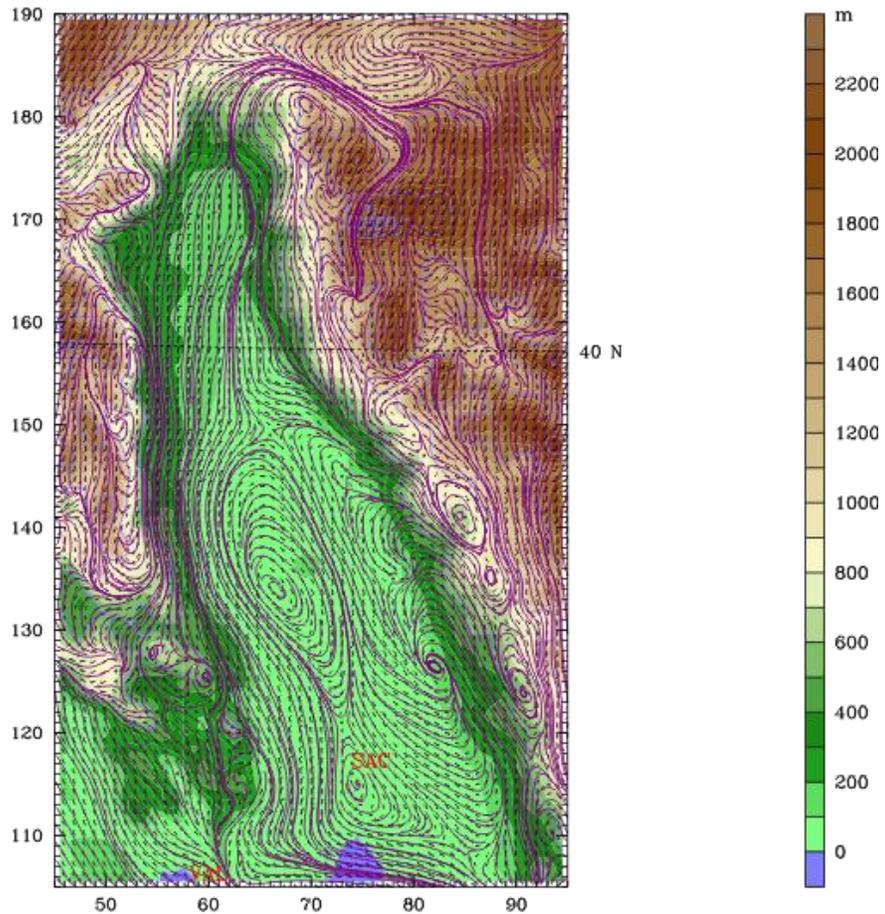Horizontal wind streamlines   at k-index = 50

BARB VECTORS: FULL BARB = 10 m s⁻¹

a)

CONTOURS: UNITS=m LOW= 0.0000   HIGH= 2000.0   INTERVAL=   500.00

83

Dataset: d03 RIP: 2000-07 29-03 MFDDA5 d03 win   Init: 1200 UTC Sat 29 Jul 00
Fcst: 36.00 h   Valid: 0000 UTC Mon 31 Jul 00 (1600 PST Sun 30 Jul 00)
Terrain height AMSL
Terrain height AMSL
Horizontal wind vectors   at k-index = 50
Horizontal wind streamlines   at k-index = 50

BARB VECTORS: FULL BARB = 10 m s⁻¹

b)   CONTOURS: UNITS=m LOW= 0.0000   HIGH= 2000.0   INTERVAL=   500.00

Figure 5-41: Simulated winds in the Sacramento Valley region at 0000 UTC 31 July 2000. a) NOFDDA. b) MFDDA5.

In the San Joaquin Valley, at 0600 UTC 30 July 2000, the downslope flow began to develop in both simulations (not shown). Six hours later, at 1200 UTC 30 July 2000, 24 hours into the simulation, the nighttime pattern (at the 300-m level) had developed, with northerly flow down the valley, along with downslope flow from the Sierra Nevada Mountains and the Coastal Ranges (Figure 5-42). Both simulations showed the nocturnal low level jet (LLJ) down the axis of the valley at a height of about 300 m (9th η level). The LLJ had a maximum wind speed of about 5-10 m/s at this level. The NOFDDA simulation showed southerly downslope flow just south of Fresno, but the flow did not create the Fresno Eddy (Figure 5-42a); the NOFDDA run developed an eddy over Bakersfield instead. The MFDDA5 simulation also did not develop the southerly downslope flow near Fresno (Figure 5-42b). MFDDA5 developed a small eddy just north of Bakersfield as well. Both simulations had the northerly mean flow exiting the valley

over the Coastal Ranges at the southern end of the valley.
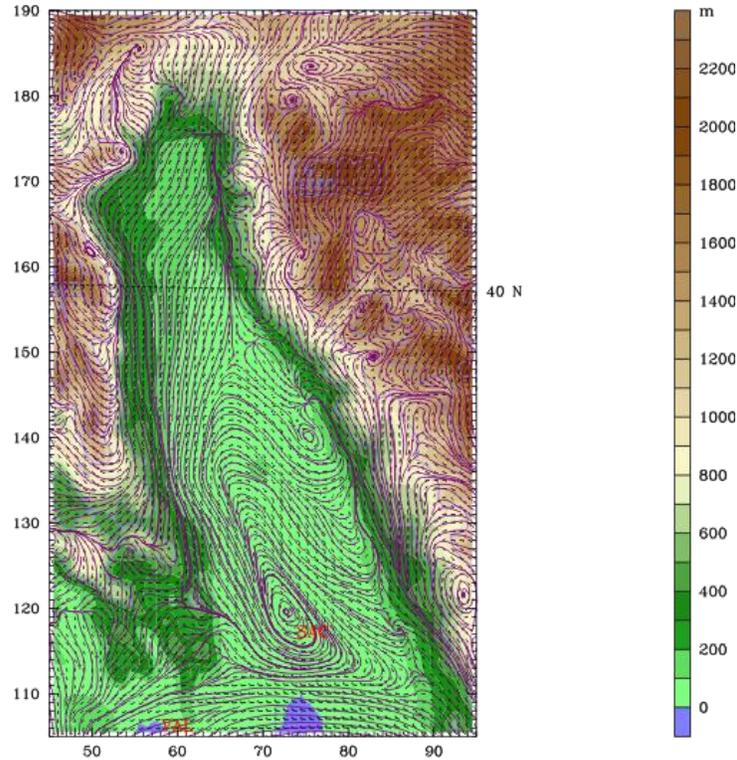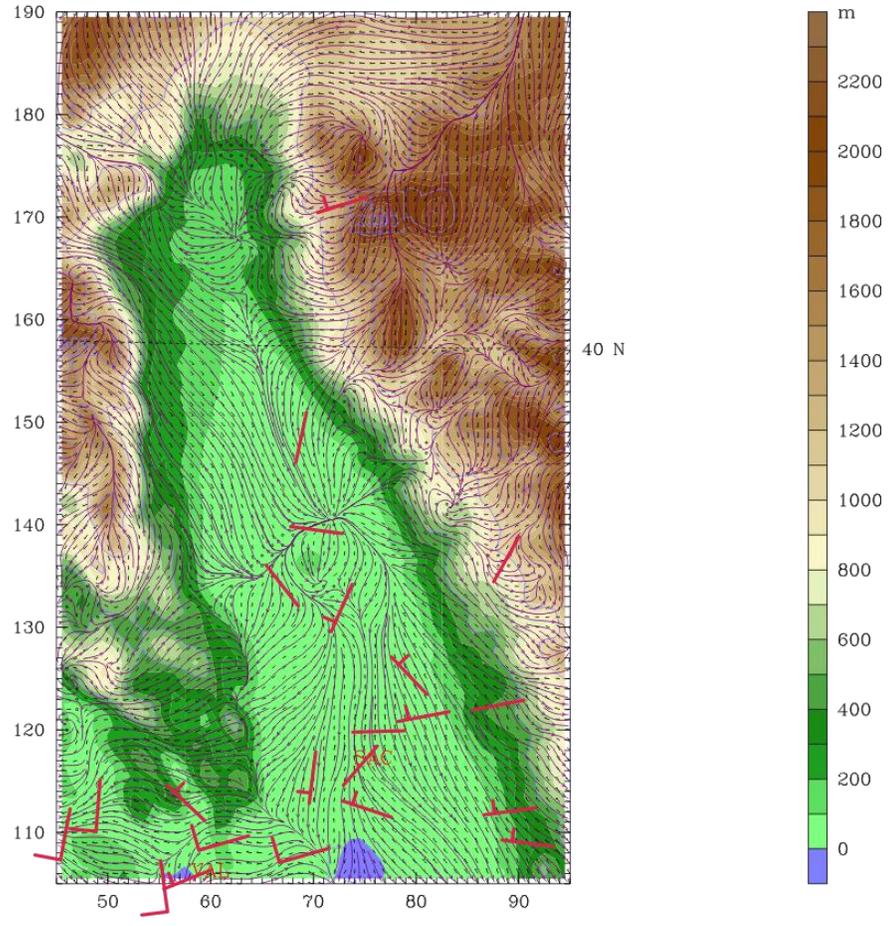
a)

Dataset: d03 RIP: 2000-07 29-03 NO FDDA d03 wi   Init: 1200 UTC Sat 29 Jul 00
Fcst: 24.00 h   Valid: 1200 UTC Sun 30 Jul 00 (0400 PST Sun 30 Jul 00)
Terrain height AMSL
Terrain height AMSL
Horizontal wind vectors   at k-index = 42
Horizontal wind streamlines   at k-index = 42

BARB VECTORS: FULL BARB = 10 kts
CONTOURS: UNITS=m LOW= 0.0000   HIGH= 3500.0   INTERVAL= 500.00

86

b)

Figure 5-42: Simulated winds in the San Joaquin Valley region at 1200 UTC 30 July 2000. a) NOFDDA. b) MFDDA5.

Six hours later, at 1800 UTC 30 Dec 2000, 30 hours into the simulation, the daytime pattern had developed in both simulations (Figure 5-43). Both simulations were similar, except that the winds in the MFDDA5 simulation turned 180 degrees from northerly to southerly near Fresno before they went up the slope of the Sierra Nevada Mountains, which agreed with observations. Winds in the NOFDDA simulation did not agree well with observations. The run produced flows that exited the San Joaquin Valley through passes and over the mountains in such areas as the Pacheco Pass (southwest of Modesto) and the Tehachapi Mountains (at the southern end of the San Joaquin Valley). Overall for the entire 5-day simulation, both simulations captured the northerly mean flow along the axis of the valley, with the MFDDA5 simulation agreeing better with the observed winds.

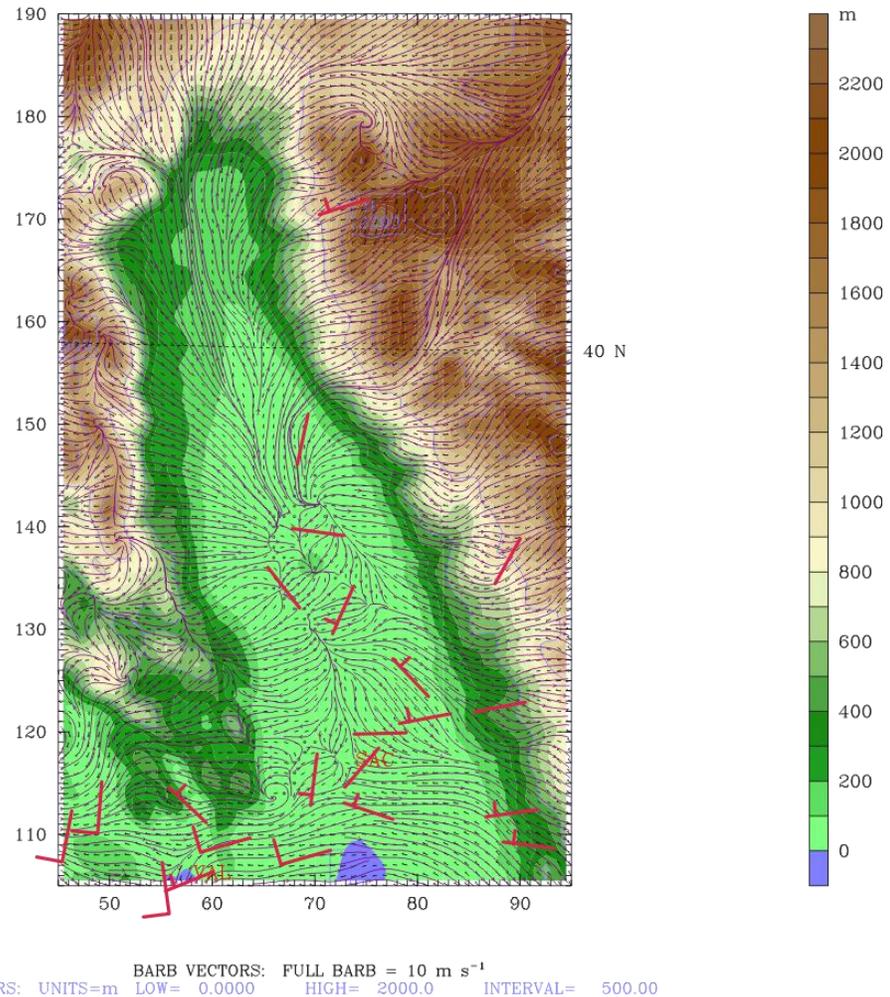Dataset: d03 RIP: 2000-07 29-03 NO FDDA d03 wi   Init: 1200 UTC Sat 29 Jul 00
Fcst: 30.00 h   Valid: 1800 UTC Sun 30 Jul 00 (1000 PST Sun 30 Jul 00)
Terrain height AMSL
Terrain height AMSL
Horizontal wind vectors   at k-index = 50
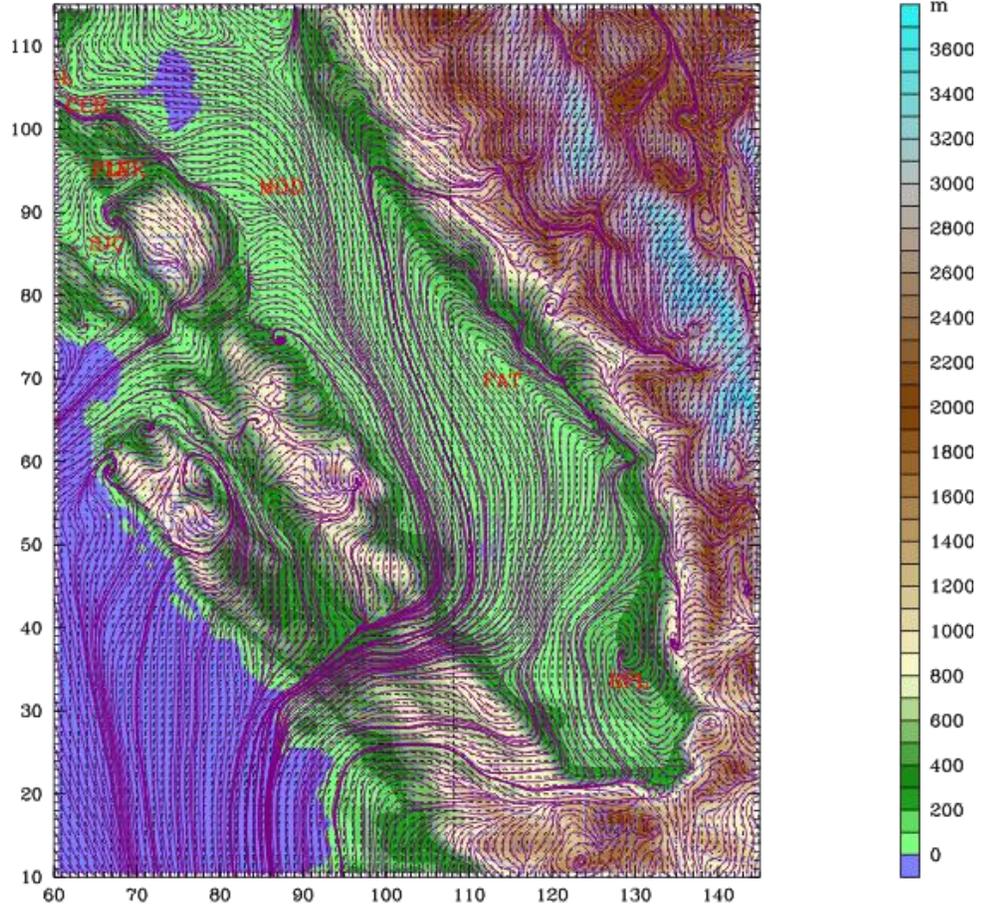Horizontal wind streamlines   at k-index = 50

BARB VECTORS: FULL BARB = 10 m s⁻¹

a)   CONTOURS: UNITS=m LOW= 0.0000   HIGH= 3500.0   INTERVAL=   500.00
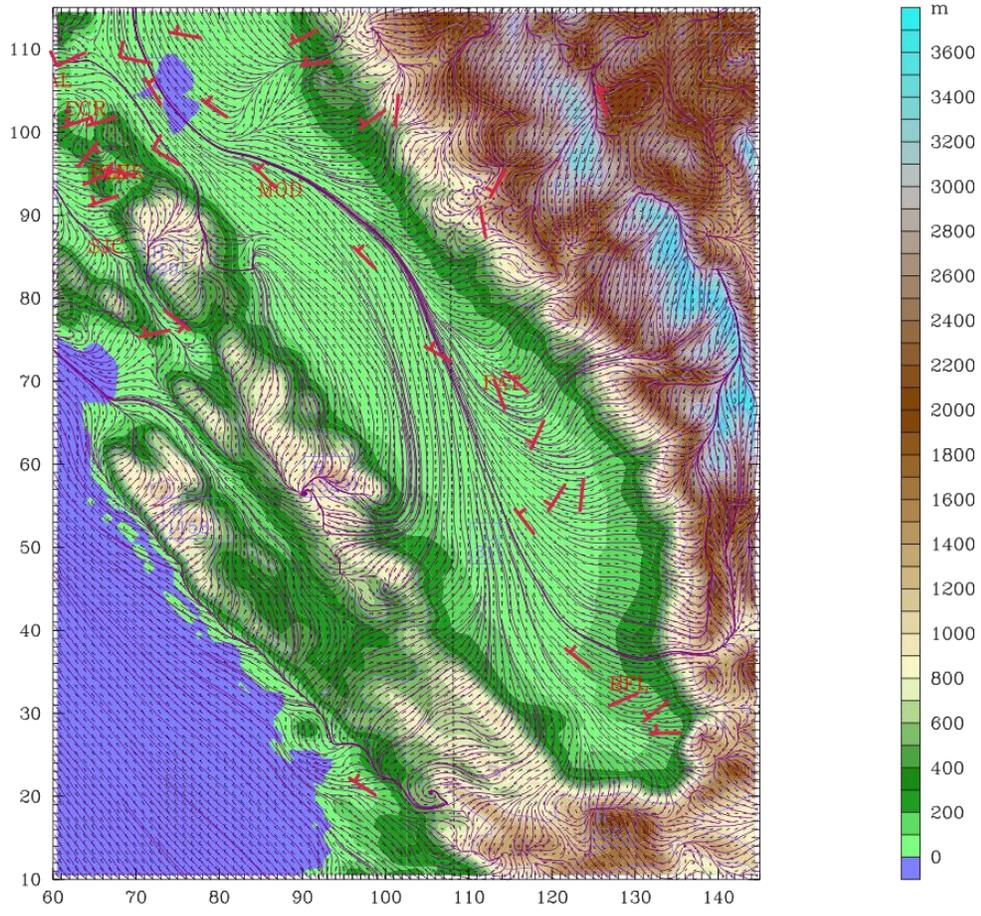
Dataset: d03 RIP: 2000-07 29-03 MFDDA5 d03 win   Init: 1200 UTC Sat 29 Jul 00
Fcst: 30.00 h   Valid: 1800 UTC Sun 30 Jul 00 (1000 PST Sun 30 Jul 00)
Terrain height AMSL
Terrain height AMSL
Horizontal wind vectors   at k-index = 50
Horizontal wind streamlines   at k-index = 50

b)

Figure 5-43: Simulated winds in the San Joaquin Valley region at 1200 UTC 30 July 2000. a) NOFDDA. b) MFDDA5.

5.3.2.4 Time Series Analysis

The modeled and observed time series results of temperature, wind speed, and wind direction for selected observation sites were produced. In the Bay Area region at Livermore (KLVK), the model did an excellent job of simulating the diurnal temperature cycle, in agreement with the observations, as the temperature errors were small (Figure 5-44a). Both model runs tended to be slightly cooler than the observations during the day by 2-3 °C, but MFDDA5 agreed better with the observations at night. The observations showed that a diurnal cycle also existed for wind speed (Figure 5-44b), which the model did a reasonable job of simulating. The fastest speeds occurred during the day, which both methods underpredicted by about 2 m/s. However, MFDDA5 did a better job of simulating the quick changes in wind speeds by the hour than NOFDDA. Many of the wind direction observations were either missing

or removed due to an excessive amount of calm winds; therefore, there were large gaps of the observed wind direction time series plot (Figure 5-44c). The observations that were available showed the winds were mostly westerly in direction throughout the study period, which both methods used in the model did a reasonable job of simulating.

At Concord (KCCR), the model again did a good job of simulating the diurnal temperature cycle (Figure 5-45a). Similar to Livermore, the model was slightly cooler than the observations during the day and slightly warmer than the observations at night for the NOFDDA simulation. The MFDDA5 run partially corrected the biases at night and, thus, was in slightly better agreement with observations. The observations showed a diurnal cycle for the wind speeds as well (Figure 5-45b), but the model did not simulate this cycle as well with either method, unlike at Livermore. The wind speeds in the simulations tend to remain mostly between 2 m s^{-1} and 5 m s^{-1}, while observed values are as large as ~6 m s^{-1} for day 1 and day 2. The MFDDA5 run did agree well with the observations between 105 hours and 112 hours. The model simulated a diurnal cycle for the wind direction (Figure 5-45c), but the observations that were available did not show this cycle. The observations showed that the winds were mostly northwesterly in direction for the entire period, but the model produced southwesterly winds at night and northwesterly winds during the day.

At Pleasanton, the model simulated the diurnal temperature cycle present in the observations (Figure 5-46a). The observations showed that the temperature increased as time progressed, another feature both simulations showed. Similar to Livermore and Concord, the MFDDA5 simulation agreed better with the observations during the night than the NOFDDA run. The observations showed a diurnal cycle for wind speed, which the model did a reasonable job simulating (Figure 5-46b). MFDDA5 did a better job with the quick changes in wind speed than NOFDDA. The observations showed that the winds were mostly westerly throughout the entire period, which both methods modeled well (Figure 5-46c). The observations showed that during the early morning hours the winds were easterly for a before becoming westerly again. The NOFDDA run did a better job of simulating this pattern than the MFDDA5simualtion. Notice that this obverted pattern did not appear at Livermore, spatially very close to Pleasanton. It was possible that the difference was caused by the proximity of Pleasanton to the western end of the valley and the west-east oriented Dublin Canyon, while the Livermore site was in the middle of the valley away from many locally produced terrain effects. The terrain effects only lasted a few hours each day before the bulk westerly flow began to take over. Unfortunately, the MFFDDA5 simulation did not perform as well as the NOFDDA run.

In the Sacramento Valley at Sacramento (KSAC), the model did a good job of simulating the diurnal temperature cycle, in good agreement with observations (Figure 5-47a). During the first three days the model was slightly cooler than the observations during the day for both methods. Observations showed that large wind speed changes occurred nearly every hour during the study period. The MFDDA5 simulation did a slightly better job in reproducing this pattern (Figure 5-47b). Since these oscillations also existed in the NOFDDA simulation, it was possible these were the signals of the transient motions (e.g. gravity waves) that passed though the site. For wind direction, both models simulated a diurnal pattern for the first three days where the winds were westerly during the day and southerly during the night (Figure 5-47c). The observations that were available showed that a diurnal pattern may have existed, but too many hours were missing to declare which of the methods did a better job.

In the San Joaquin Valley, at Fresno (KFAT), the model simulated the diurnal temperature cycle well, but was cooler than the observations throughout the entire study period

by several degrees for both methods (Figure 5-48a). The observations showed rapid changes in wind speeds every hour, but the model did an inadequate job simulating them (Figure 5-48b), unlike at Sacramento. The model also tended to have slower winds than the observations. The modeled developed a diurnal pattern for wind direction, with northeasterly winds during the night (which represented downslope flow from the Sierra Nevada Mountains) and westerly winds during the day (which represented flow up the mountains) (Figure 5-48c). The MFDDA5 simulation was in better agreement with the observations.

At Bakersfield (KBFL), the model simulated the diurnal temperature cycle, but was several degrees cooler than the observations throughout the entire study period for both methods (Figure 5-49a). The observations showed a diurnal cycle for the wind speeds as well (Figure 5-49b), which the model did a good job of simulating. However, the modeled maximum speeds were has slower winds than observations. The observations showed a diurnal pattern in the wind direction, which the model again simulated fairly well (Figure 5-49c).

Overall the model did a good job of simulating the diurnal patterns of temperature, wind speed, and wind directions. The model tended to be cooler than the observations during the day and warmer than the observations at night, except in the San Joaquin Valley, where the model was cooler throughout the entire five days. The temperature biases were likely due to a number of factors. The model may not have properly simulated characteristics of the desert land areas within the San Joaquin Valley, which would affect the temperature fluxes over the area. The effects of fog which would keep temperatures cooler may also have been underpredicted. The model also tended to simulate slower winds than the observations, but the MFDDA5 method did a better job of simulating the rapid changes in wind speed that often occurred hourly.
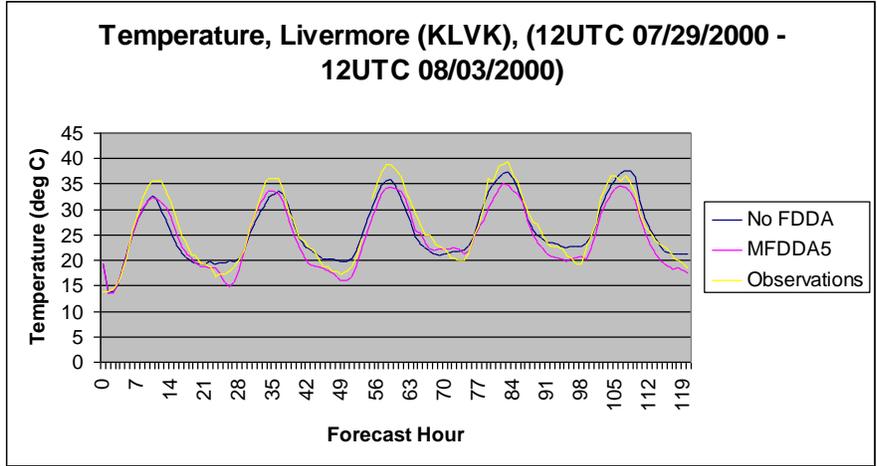
a)
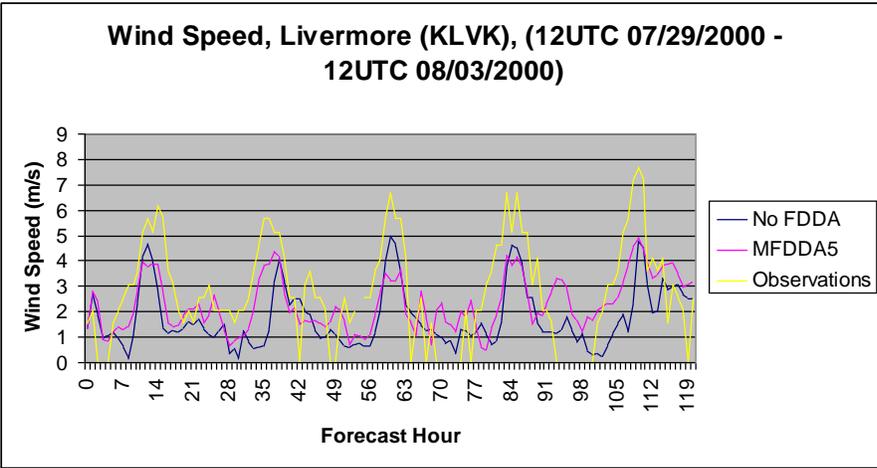
b)
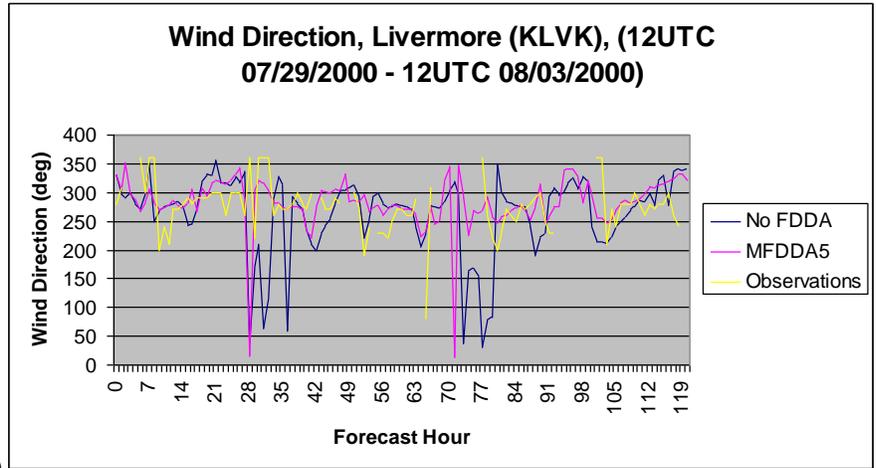
c)

Figure 5-44: Time series plots for Livermore in the Bay Area region. a) Temperature. b) Wind Speed. c) Wind Direction.
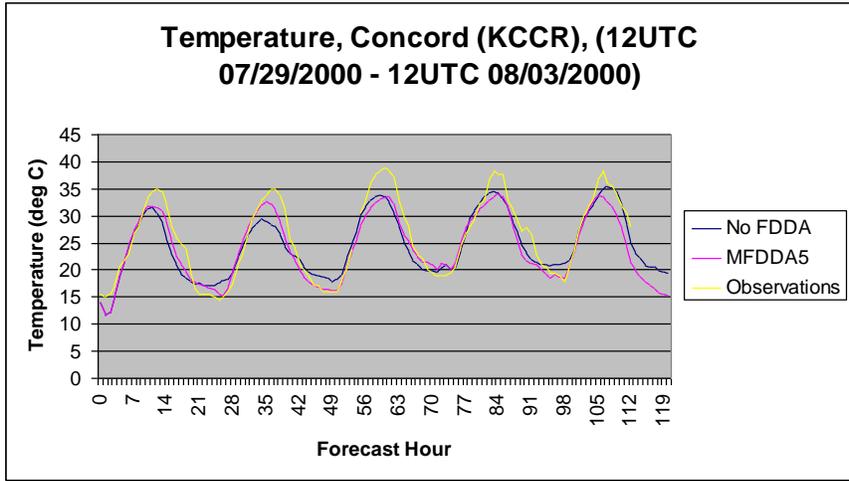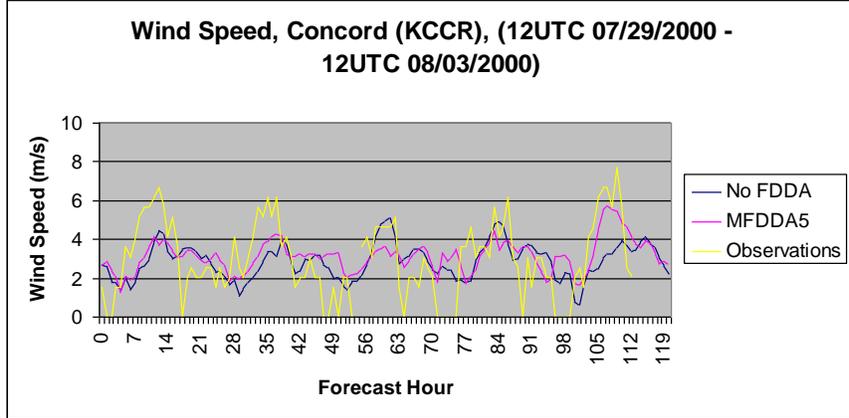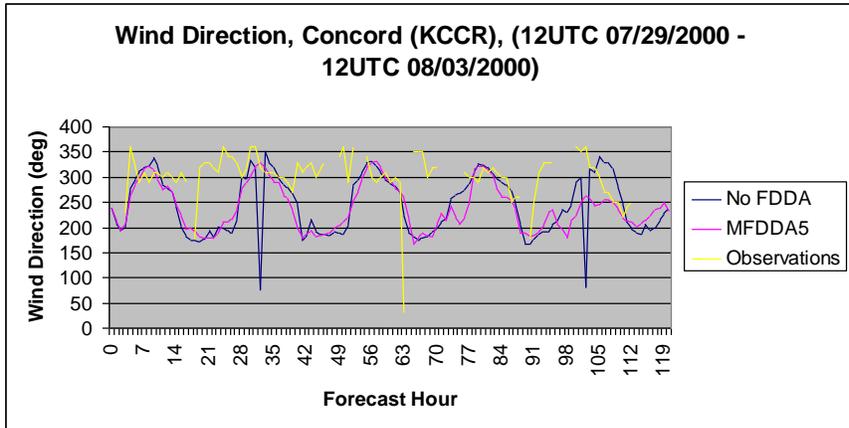
Figure 5-45: Time series plots for Concord in the Bay Area region. a) Temperature. b) Wind Speed. c) Wind Direction.
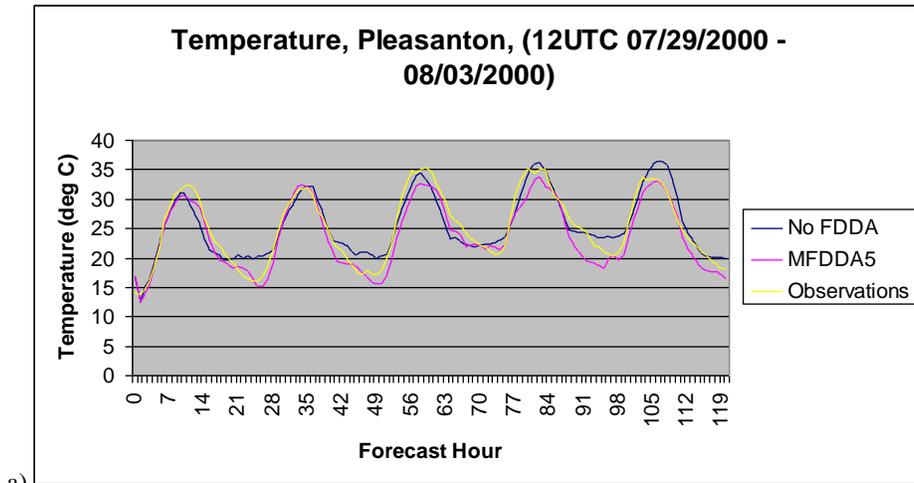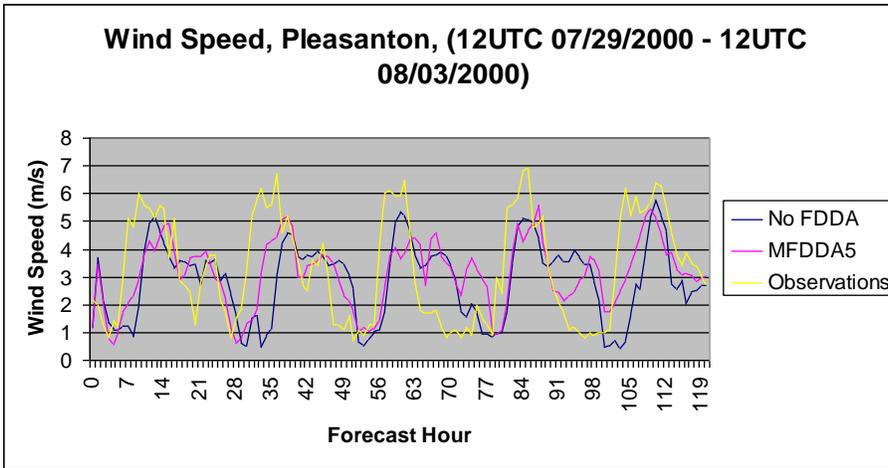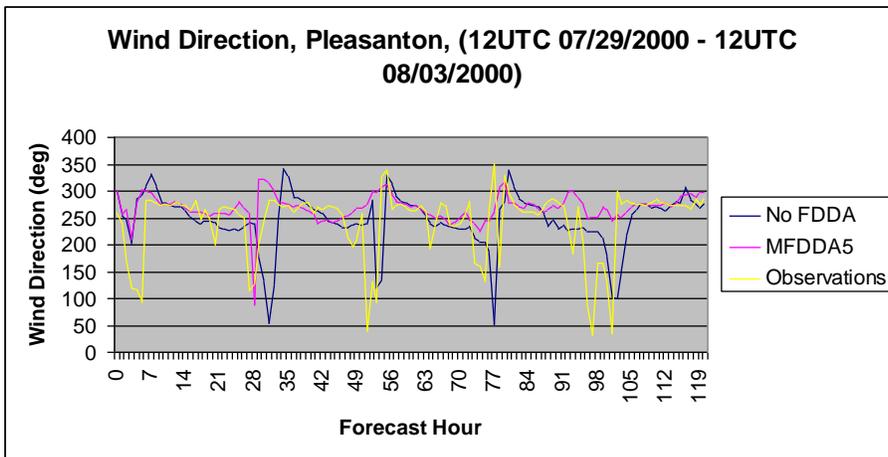
Figure 5-46: Time series plots for Pleasanton in the Bay Area region. a) Temperature. b) Wind Speed. c) Wind Direction.
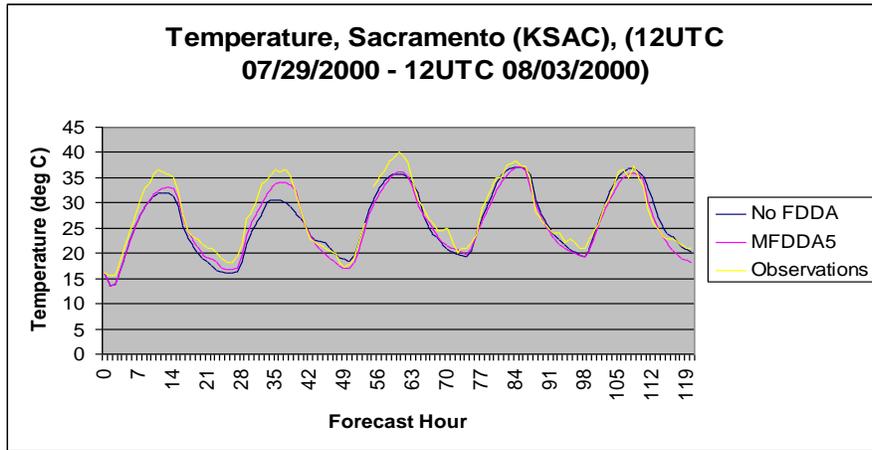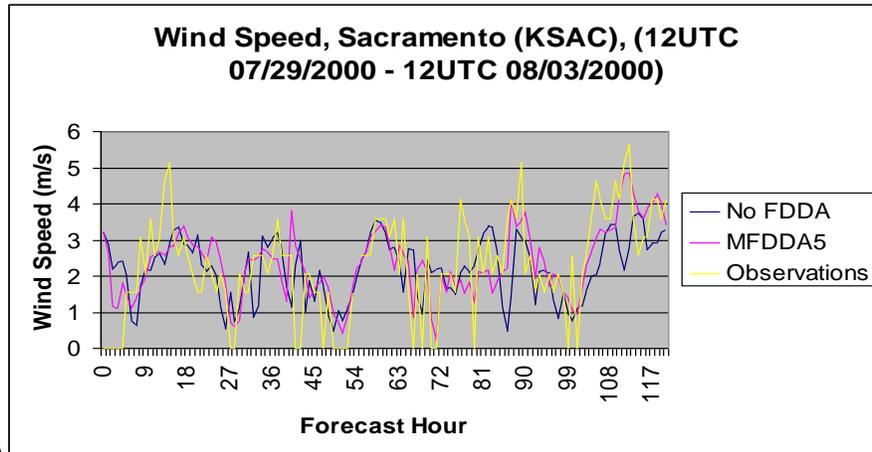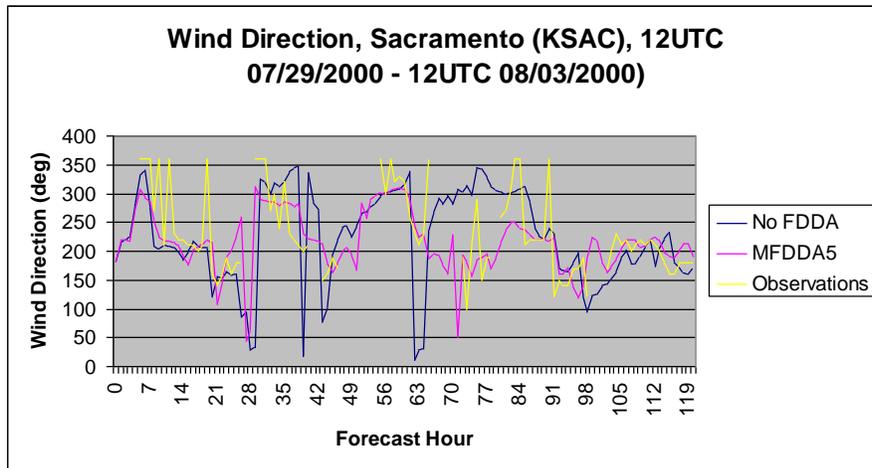
Figure 5-47: Time series plots for Sacramento in the Sacramento Valley. a) Temperature. b) Wind Speed. c) Wind Direction.
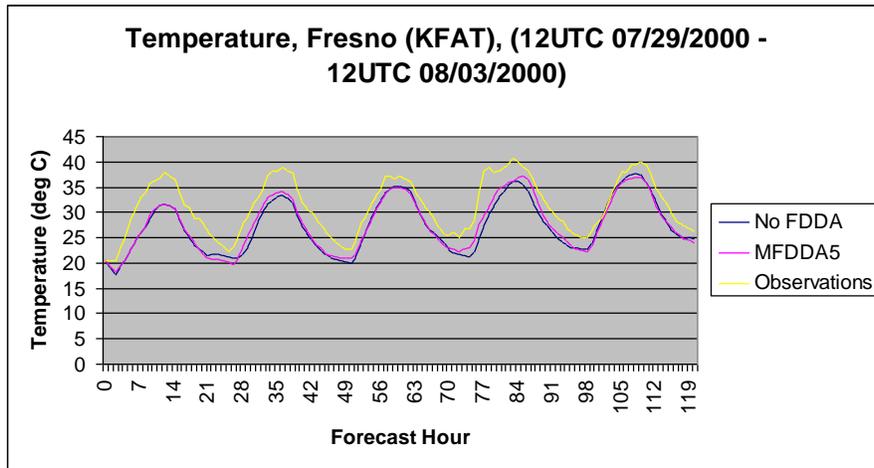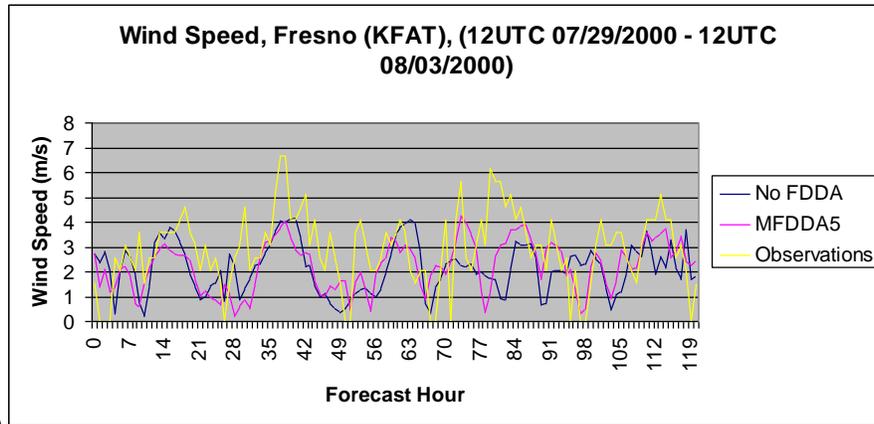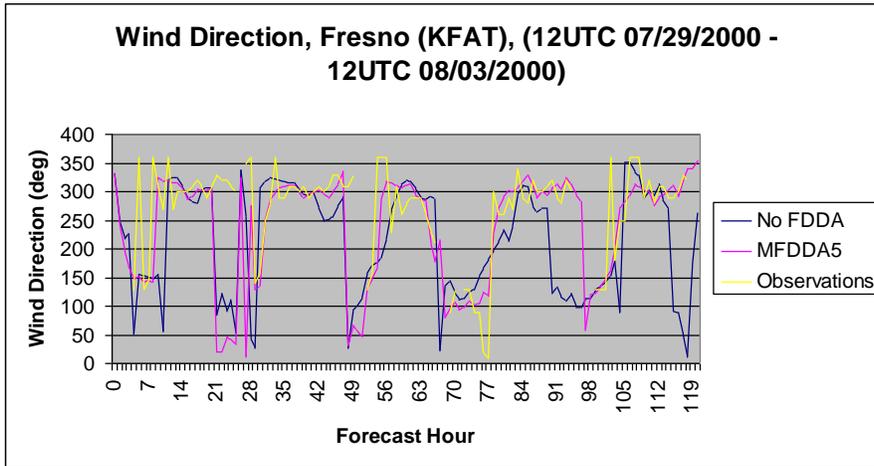
Figure 5-48: Time series plots for Fresno in the San Joaquin Valley. a) Temperature. b) Wind Speed. c) Wind Direction.
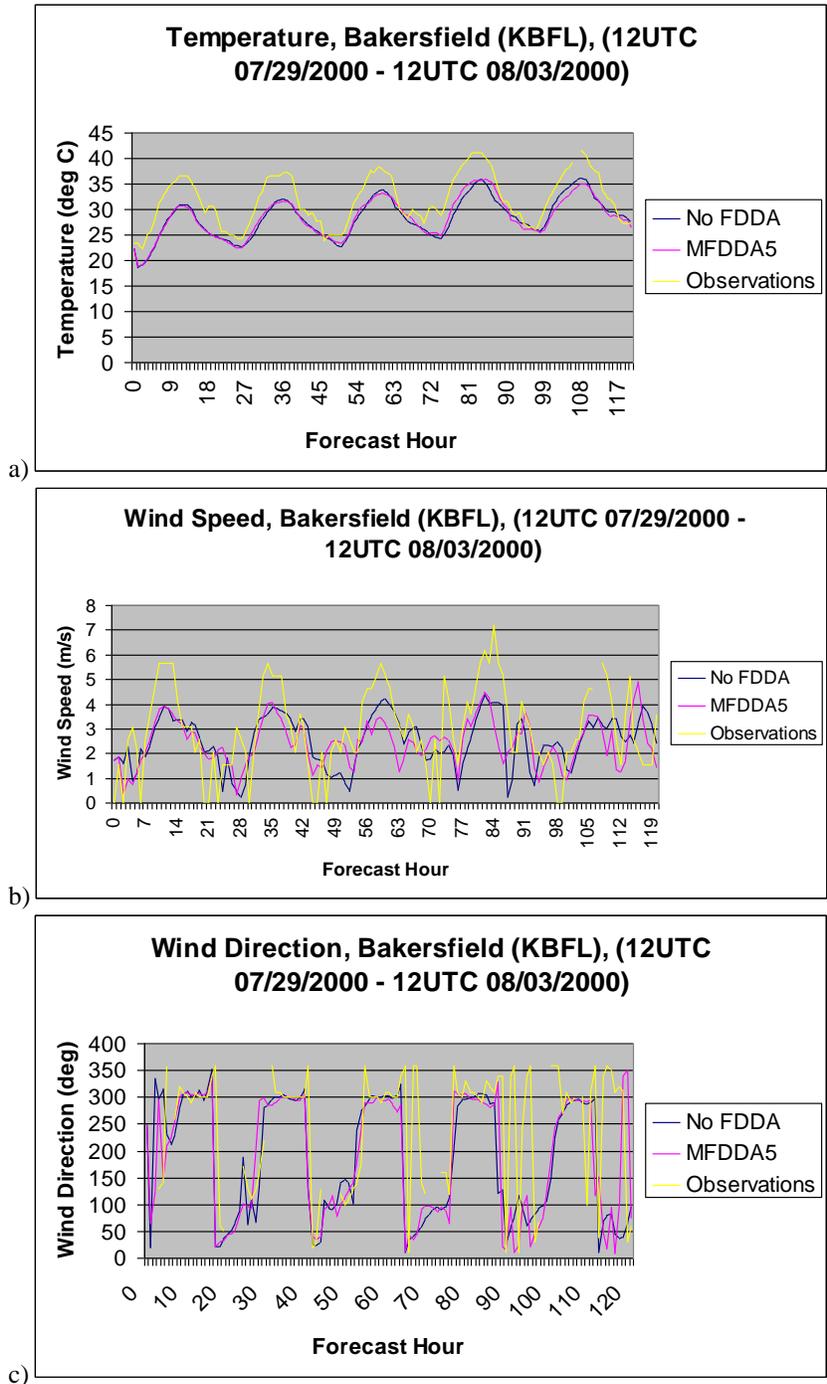
Figure 5-49: Time series plots for Bakersfield in the San Joaquin Valley. a) Temperature. b) Wind Speed. c) Wind Direction.

5.4   Optimal Radius of Influence

Observational nudging used a radius of influence (RIN) to determine which individual observations had an influence on the model outputs at each grid point. If the RIN was too small,

97

fewer observations would be used to smooth out the model errors. RINs too big would use values from observations sites that were very far from a grid point, possibly creating even larger errors and render the obs nudging useless. The RIN was normally chosen based on previous modeling studies. For this study error correlation, calculations were made to objectively determine an optimal RIN for use during obs nudging.

Ten observation sites were chosen at various locations within the 4-km domain. The temperature innovation values from the observations, and from the NOFDDA model simulation at those sites, were used to calculate the correlation coefficients between two sites. Equation 5.1 shows the correlation calculation:

$$ r = \frac{\sum_{i=1}^{N}\left((x_i - \bar{x})(y_i - \bar{y})\right)}{\sqrt{\sum_{i=1}^{N}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{N}(y_i - \bar{y})^2}} \tag{5.1} $$

where r was the correlation coefficient, N was the number of times error values were taken, x_i was the error (i.e. innovation) between the observation and the model at the first station at the ith time, x bar was the mean innovation error over N times at the first station, y_i was the error between the observation and the model at the second station at the ith time, and y bar was the mean innovation error over N times at the second time. The innovation was the observation value minus the model value. Figure 5-50 shows the chart of the correlation coefficients vs. the horizontal distance between various pairs of observation sites. The red line indicates a correlation coefficient of 0.5, chosen as the threshold value for determining which coefficient values from a pair of sites were considered good. If a coefficient value was above the threshold then the corresponding horizontal distance was considered as a possible value to use as the RIN during obs nudging. Thirteen correlation coefficients were shown to provide a good distribution of the horizontal distances and to provide a good distribution for of high and low coefficients. For this experiment the largest horizontal distance that had a correlation coefficient above 0.5 without having low values in this range was 105 km, but the largest correlation was at 63 km. Therefore, this distance could be considered as the largest possible value that can be used as the RIN and still provide an effective data assimilation. This experiment actually validated the selection of the RIN used for obs nudging in the FDDA experiments made during this study, which were 50 km at the surface and 100 km above the surface.
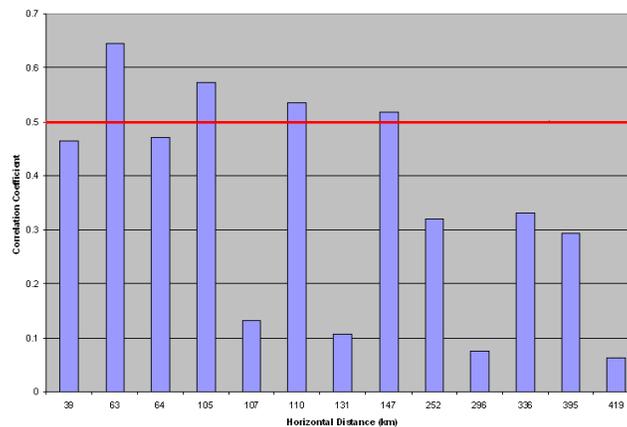
Figure 5-50: Correlation coefficient vs. horizontal distance between various pairs of observation sites. The red line indicates a correlation coefficient of 0.5.

6   DISCUSSIONS

The objective of this research was to assist the BAAQMD transition from its current MM5 modeling system to the newly developed WRF-ARW modeling system. To achieve this goal, various approaches to find the optimal WRF-ARW model configuration for the region were taken. Model performances were evaluated over the Bay Area for selected cases where MM5/CMAQ and MM5/CAMx systems were used for previous air quality studies. One major concern at the beginning of this project was that since WRF-ARW FDDA capabilities were still under development (through a different project with the Defense Threat Reduction Agency (DTRA)), performance was an unknown, especially when applied in the Bay Area where a large data void region existed upstream over the Pacific Ocean.

The investigation began with setting up a baseline configuration before FDDA was applied. This was achieved by testing and evaluating different radiation schemes and LSMs that were used within the WRF model. Using the determined baseline configuration, FDDA capabilities that were recently implemented into WRF-ARW were tested and evaluated. To achieve the goal, both subjective and objective verification methods were used. The focus was on evaluating the model-simulated RH, temperature, wind speeds, and wind directions, critically important to the dispersion of air pollutants. An optimal model configuration for the SFBA region was achieved, although more case studies will be required to further verify the conclusions. The major finding from this research was that the WRF-ARW modeling system with the optimal physics options found is this research was able to reproduce the weather features in the SFBA. The inclusion of FDDA further improved the model solutions as well as the error statistics. However, to safely transition from the MM5 to WRF, systematic benchmark comparisons between the two modeling systems are needed, as indicated in the next section.

7   CONCLUSIONS

The two atmospheric radiation schemes that were tested were the Rapid Radiative Transfer Model (RRTM) and the Rapid Radiative Transfer Model for Global Climate Models (RRTMG). It was found that both radiative schemes produced similar results for all verification fields (i.e. wind speed, wind direction, temperature and water vapor mixing ratio), with a slight degradation shown in some fields in the RRTMG experiment. Considering the fact that most air-pollution cases were under weakly-forced fair weather conditions, WRF-ARW with varying atmospheric radiation schemes would be less likely produce dramatically different solutions. Therefore, all of the numerical experiments used the RRTM/Dudhia radiation scheme, for both the winter and summer cases.

The four land surface models tested were the simple 5-layer thermal diffusion scheme, the Noah LSM, the Rapid Update Cycle (RUC) LSM, and the Pleim-Xiu (PX) LSM. For the winter case the PX LSM produced the smallest errors over most domains for most fields and therefore was chosen as the LSM for the model simulations. For the summer case no particular LSM performed better than the others; therefore the PX LSM was once again chosen for this case because it was tested and used in air quality applications by the EPA.

A number of FDDA strategies were used in the model simulations. 3D analysis nudging was used above the surface on the 36-km and 12-km domains, while surface analysis nudging of wind fields was also used in other simulations. Observational nudging was used on all three domains for the winter case simulations, and only on the 12-km and 4-km domains in the summer case simulations. A multiscale FDDA strategy was also used that combined both analysis nudging and obs nudging.

Statistical analysis showed that using FDDA reduced the model errors compared to the NOFDDA simulation that did not use data assimilation in the winter and summer cases. The simulations that used the multiscale FDDA strategy produced the smallest errors of all the experiments for all four fields, even when moisture and temperature data were not assimilated at the surface. The MFDDA2SP simulation performed the best for the winter case, while the MFDDA5 simulation was the best for the summer case. The FDDA simulations performed better than the NOFDDA simulations when they were compared on the subregion level within the 4-km domain, but the improvements in the errors were somewhat smaller for the Bay Area region. The FDDA simulations also performed better than the NOFDDA simulation when they were compared against an independent dataset.

For the summer case, a sensitivity study was conducted to show the effect obs nudging on the 36-km coarse domain had on the 4-km WRF-ARW solutions. As expected, with and without obs nudging on the36-km domain, the 4-k WRF-ARW solutions were similar. The effect of Oakland radiosonde was also evaluated. The 4-km domain WRF-ARW solutions with and without assimilating Oakland radiosonde were compared statistically, and found no difference in model statistics, although subjective comparisons were recommend for a future study.

Subjective analysis showed that even without data assimilation the model could represent the humidity, temperature, and the winds reasonably well for both cases. On the large scale, the locations of WRF-simulated highs and lows on the 36-km domain were well placed and consistent with observations. The large scale wind patterns were also found to be in line with observations. Using the multiscale FDDA strategy improved wind patterns the most. Observed troughs and ridges were found to be simulated with a good degree of accuracy.

On the 4-km domain, even without data assimilation, the model simulated mesoscale wind patterns reasonably well for both cases. The simulations with and without FDDA were able to develop upslope and downslope flows that occurred along the mountains that surround the Central Valley, as well as resolve the mesoscale eddies that developed within the valley. The multiscale FDDA simulations were superior to the NOFDDA runs in developing mesoscale eddies seen in observations.

In subregions that were within the 4-km domain, the effect of FDDA was not as large for both cases in the Bay Area region. Both the NOFDDA and the multiscale FDDA simulations simulated the winds over the region with very few differences. In the Sacramento Valley and the San Joaquin Valley, MFDDA performed better than NOFDDA in the development and placement of the Fresno Eddy and the Schultz Eddy during the summer case. Both simulations simulated the upslope and downslope flows along the mountains surrounding both valleys, but MFDDA reproduced the southerly flow in the Sacramento Valley better in the summer than NOFDDA, which developed a divergence zone in the middle of the valley that was not observed.

To help evaluate WRF-ARW performance in each subregion, times series at several selected sites were analyzed. It was found that in the Bay Area WRF-ARW showed a warms bias at night for the winter PM case (e.g. at San Jose, Livermore, and Vallejo), and a cold bias during daytime for the summer ozone case (e.g. Livermore, Concord), with Pleasanton as the only exception. FDDA, in general, tried to correct these temperature biases. For the surface winds, the effect of FDDA was not as evident. For example, at Pleasanton, NOFDDA performed better in reproducing the morning hour southeasterly flow induced by the terrain. In Sacramento, WRF-ARW, especially when the FDDA was applied, was able to reproduce observed conditions fairly well, with transient waves present in both model results and observations. In the San Joaquin Valley, for the winter PM case, WRF-ARW showed cold biases at Fresno and warm biases at Bakersfield as large as 10 degrees. For the summer ozone case, WRF-ARW showed a colder bias as large as 6-7 degrees. The temperature biases were likely due to a number of factors that included the model possibly not properly simulating the characteristics of the desert land within the San Joaquin Valley or the presence of fog within the valley, which could keep observed temperatures cooler. The use of FDDA did not completely resolve this issue. For surface winds, the WRF-ARW was able to reproduce the observed up-and downslope motions and the diurnal variations. The multiscale FDDA simulations agreed better with the observations than the NOFDDA runs.

An optimal RIN was objectively determined by calculating the correlation coefficients of the innovations between two observation stations at different distance separations at 4-km. The results of the correlation calculation showed that a correlation coefficient greater than 0.5 was found for two observation sites that were approximately 105 km apart from each other, with the largest correlation at 63 km. These values were consistent with the 50-km RIN used at the surface on the 4-km domain, and the 100-km RIN used near the surface.

7.1 Recommendations for BAAQMD

Based on this research, the following WRF modeling configurations for the BAAQMD are recommended:

1) The RRTM radiation scheme was tested and found to perform equal to or better than the RRTMG scheme, thus it is recommended for use in BAAQMD's WRF configuration.

There is no need to switch to the new RTMG scheme that was originated from the GCMs.

2) The Pleim-Xiu land surface model was found to be the best LSM in this research, thus, it should be used in the BAAQMD's WRF configuration.

3) A Multiscale FDDA approach that combined analysis nudging on the coarser grids and obs nudging on the finer grids was found to be more effective in reducing model errors than using analysis nudging and obs nudging alone. Based on the WRF configuration used in this research, nudging for mass fields is turned off within the PBL in both analysis nudging (including surface analysis nudging) and obs nudging.

4) It was not important whether or not obs nudging was used on the 36-km coarse domain, since the sensitivity study showed nearly identical results for the 4-km inner domain with or without obs nudging on the 36-km domain. This was expected because the lateral boundary effect from the 36-km domain on the 4-km domain was minimal and obs nudging was used on the 12-km domain.

5) Use of the special data in FDDA has added value.

6) Further recommendations include the future research areas listed below.

7) Since the 40-km Eta analysis is close to the 36-km coarse grid resolution used in this research, the BAAQMD should consider using the 12-km domain as the coarse grid and add a 1-km grid centered over the Bay Area.

7.2 Future Research

There are a number of areas in this investigation that can be improved in the future research:

1) **WRF-MM5 comparison:** Although this research has shown that WRF with use of FDDA can reduce model errors as expected, systematic benchmark comparisons between the two modeling system are still needed before a decision is made to transition from MM5 to WRF. The goal of the comparison is to show if the WRF model performs significantly better than MM5 using the same study cases and with the same input and model configurations.

2) **Improving the quality of the observations:** During this study it was found that many observed wind speeds had values that showed unrealistic high-frequency oscillations between a zero and a non-zero value, especially in light wind scenarios (e.g. San Jose, Livermore, Sacramento, Modesto, Fresno and Bakersfield for the winter PM case, and Concord, Sacramento, Bakersfield for the summer ozone case). It might be possible that the observing instruments were not properly designed or calibrated to accurately measure the low wind speeds. These zero wind speeds may adversely affect the FDDA results and the objective verification. This issue should be further investigated.

103

3) **Improving the observation quality control (QC) procedure**: The BAAQMD found that redundant observations could pass through the obs-processing software and exist in the obs nudging input. For example, two observations that have identical (or nearly identical) location information are currently not treated as the duplicates because they likely come from different observation systems, thus they carry different identifiers. Although the WRF obs nudging code has some level of QC that can remove the duplicates, the code may not be designed to handle the special case like this. Since BAAQMD is rich with observations from various observing systems, it is critical to address these QC issues and implement effective algorithms to improve the obs processing procedures.

4) **Improving WRF FDDA strategies/parameters when using higher resolution modeling**: Due to the coastal region and complex terrain over the BAAQMD modeling domain, especially over the Bay Area subregion, improved model resolution could bring the benefit of improved representation of topography and land surface characteristics. However, this would also bring challenges to the existing FDDA strategies. These challenges include dealing with representativeness of observations over the coastal and complex terrain regions. Further investigation of FDDA over complex terrain (e.g. determination of the optimal influence distance) and its application in high-resolution modeling is critical to the Bay Area and a subject for future research.

5) **Improving existing FDDA techniques:** The current nudging-based FDDA techniques have limitations. For example, the nudging parameters (e.g. nudging strength, RIN, and time window) are specified based on experience. Optimal nudging parameters can be determined by using the hybrid nudging EnKF techniques that have been under development for last several years at Penn State (Lei et al. 2011a, 2011b and 2011c). In the hybrid techniques, the nudging parameters are determined using the background error covariances based on an ensemble of model simulations.   The flow-dependent information can be used to determine the nudging parameters.   This technique is particularly attractive to coastal and complex terrain regions like California.

6) **Improving the land surface representation and LSM initialization**: As indicated earlier, the temperature biases found in the time series analyses (e.g. Fresno and Bakersfield) were likely caused by the inaccurate representation of the land surface characteristics within the model. Although a land surface model is being used (i.e. the Pleim-Xiu suite of physics), the soil initialization processes in the model could still be inadequate over the region. Further improvement of the land surface modeling includes use of a better land surface data (e.g. satellite observed vegetation fraction).

REFERENCES

Bao, J-W., S. A. Michelson, P. O. G. Persson, I. V. Djalalova, and J. M. Wilczak, 2008: Observed and WRF-Simulated Low-Level Winds in a High-Ozone Episode during the Central California Ozone Study. *J. Appl. Meteor. Climatol.,* **47**, 2372-2394.

Beaver, S., 2008: Cluster Analysis of Air Quality Data for CCOS Ozone Study, Final Report. ARB contract #06-1CCOS, Nov. 25, 2008. 319 pp.

Benjamin, S. O. and Nelson L. Seaman, 1985: A Simple Scheme for Objective Analysis in Curved Flow. *Mon. Wea. Rev.*, **113**, 1184-1198.

Byun, D., and K.L. Schere, 2006: Review of the Governing Equations, Computational Algorithms, and Other Components of the Models-3 Community Multiscale Air Quality (CMAQ) Modeling System. *Applied Mechanics Reviews*, **59**, 51-77.

Chen, F., and J. Dudhia, 2001: Coupling an Advanced Land Surface–Hydrology Model with the Penn State–NCAR MM5 Modeling System. Part I: Model Implementation and Sensitivity. *Mon. Wea. Rev.*, **129**, 569–585.

Deng, A., N. L. Seaman, G. K. Hunter, and D. R. Stauffer, 2004: Evaluation of Interregional Transport Using the MM5-SCIPUFF System. *J. Appl. Meteor.*, **43**, 1864-1886.

Deng, A., D. R. Stauffer, B. J. Gaudet, J. Dudhia, J. Hacker, C. Bruyere, W. Wu, F. Vandenberghe, Y. Liu, and A. Bourgeois, 2009: Update on WRF-ARW End-to-End Multi-scale FDDA System. *10th Annual WRF Users' Workshop*, Boulder, CO, June 23, 14 pp.

Dudhia, J., 1989: Numerical Study of Convection Observed during the Winter Monsoon Experiment Using a Mesoscale Two-Dimensional Model. *J. Atmos. Sci.*, **46**, 3077-3107.

Grell, G. A., J. Dudhia, and D. R. Stauffer, 1994: A Description of the Fifth-Generation Penn State/NCAR Mesoscale Model (MM5). NCAR Tech. Note NCAR/TN-398+STR, 128 pp.

Hong, S. Y., J. Dudhia, and S. H. Chen, 2004: A Revised Approach to Ice Microphysical Processes for the Bulk Parameterization of Clouds and Precipitation. *Mon. Wea. Rev.*, **132**, 103-120.

Iacono, M., J. Delamere, E. Mlawer, M. Shephard, S. Clough, and W. Collins, 2008: Radiative Forcing by Long-lived Greenhouse Gases: Calculations with the AER Radiative Transfer Models*, J. Geo-phys. Res*., **113**, D13103, doi:10.1029/2008JD009944.

Janjic, Z. I., 1996: The Surface Layer in the NCEP Eta Model, *Eleventh Conference on Numerical Weather Prediction*, Norfolk, VA, 19–23 August; Amer. Meteor. Soc., Boston, MA, 354–355.

Janjic, Z. I., 2002: Nonsingular Implementation of the Mellor–Yamada Level 2.5 Scheme in the NCEP Meso Model, NCEP Office Note, No. 437, 61, 107 pp.

Kain, J. S., 2004: The Kain–Fritsch Convective Parameterization: An Update. *J. Appl. Meteor.*, **43**, 170–181.

Kain, J. S. and M. Fritsch, 1990: A One-Dimensional Entraining/Detraining Plume Model and Its Application in Convective Parameterization. *J. Atmos. Sci.*, **47**, 2784-2802.

Kumar, N. and F. W. Lurmann, 1997: Peer Review of ENVIRON's Ozone Source Apportionment Technology and the CAMx Air Quality Model. REVISED FINAL REPORT STI-996203-1732-RFR. June, 1997.

Lei, L., D. R. Stauffer, S. E. Haupt, and G. S. Young, 2011a: A hybrid nudging-ensemble Kalman filter approach to data assimilation. Part I: Application in the Lorenz system. Submitted to *Tellus*.

Lei, L., D. R. Stauffer, and A. Deng, 2011b: A hybrid nudging-ensemble Kalman filter approach to data assimilation. Part II: Application in a shallow water model. Submitted to *Tellus*.

Lei, L., D. R. Stauffer, and A. Deng, 2011c: A hybrid nudging-ensemble Kalman filter approach to data assimilation in WRF/DART. In preparation.

Lin, Y. L. and I. C. Jao, 1995: A numerical study of flow circulations in the Central Valley of California and formation mechanisms of the Fresno eddy. *Mon. Wea. Rev.*, **123**, 3227-3239.

Mlawer, E. J., S. J. Taubman, P. D. Brown, M. J. Iacono, and S. A. Clough, 1997: Radiative Transfer for Inhomogeneous Atmosphere: RRTM, A Validated Correlated-k Model for the Longwave. *J. Geophys. Res.*, **102 (D14),** 16663–16682.

Otte, T. L., 2008a: The Impact of nudging in the meteorological model for retrospective air quality simulations. Part I: Evaluation against national observation networks. *J. Appl. Meteor. Climatol.*, **47**, 1853-1867.

Otte, T. L., 2008b: The Impact of nudging in the meteorological model for retrospective air quality simulations. Part II: Evaluating collocated meteorological and air quality observations. *J. Appl. Meteor. Climatol.*, **47**, 1868-1887.

Pleim, J. E. and A. Xiu, 1995: Development and Testing of a Surface Flux and Planetary Boundary Layer Model for Application in Mesoscale Models. *J. Appl. Meteor.*, **34**, 16–32.

Pleim, J. E. and A. Xiu, 2003: Development of a Land Surface Model. Part II: Data Assimilation. *J. Appl. Meteor.*, **42**, 1811-1822.

Pleim, J. E. and R. Gilliam, 2009: An Indirect Data Assimilation Scheme for Deep Soil Temperature in the Pleim–Xiu Land Surface Model. *J. Appl. Meteor. Climatol*., **48**, 1362–1376.

Seaman, N. L., D. R. Stauffer, and A. M. Lario-Gibbs, 1995: A Multiscale Four-Dimensional Data Assimilation System Applied in the San Joaquin Valley during SARMAP. Part I: Modeling Design and Basic Performance Characteristics. *J. Appl. Meteor.*, **34**, 1739-1761.

Skamarock, W. C., J. B. Klemp, J. Dudhia, D. O. Gill, D. M. Barker, M. G. Duda, X.-Y. Huang, W. Wang, and J. G. Powers, 2008: A Description of the Advanced Research WRF Version 3. NCAR Tech. Note NCAR/TN-475+STR, 113 pp.

Smirnova, T. G., J. M. Brown, and S. G. Benjamin, 1997: Performance of Different Soil Model Configurations in Simulating Ground Surface Temperature and Surface Fluxes. *Mon. Wea. Rev.*, **125**, 1870–1884.

Smirnova, T. G., J. M. Brown, S. G. Benjamin, and D. Kim, 2000: Parameterization of Cold season Processes in the MAPS Land-Surface Scheme. *J. Geophys. Res.*, **105 (D3),** 4077–4086.

Soong, S. T., P. T. Martien, C. L. Archer, S. Tanrikulu, J. W. Bao, J. M. Wilczak, S. A. Michelson, Y. Jia, and C. Emery, 2006. COMPARISON OF WRF/CAMx AND MM5/CAMx SIMULATIONS FOR AN OZONE EPISODE IN CALIFORNIA. *14th Joint Conference on the Applications of Air Pollution Meteorology with A&WMA*, Jan 30-Feb 2, 2006, Atlanta.

Stauffer, D. R. and N. L. Seaman, 1990: Use of Four-Dimensional Data Assimilation in a Limited-Area Mesoscale Model. Part I: Experiments with Synoptic-Scale Data. *Mon. Wea. Rev.*, **118**, 1250-1277.

Stauffer, D. R. and N. L. Seaman, 1994: Multiscale Four-Dimensional Data Assimilation. *J. Appl. Meteor.*, **33**, 416-434.

Tanrikulu, S., D. R. Stauffer, N. L. Seaman, and A. J. Ranzieri, 2000: A Field-Coherence Technique for Meteorological Field-Program Design for Air Quality Studies. Part II: Evaluation in the San Joaquin Valley. *J. Appl. Meteor.*, **39**, 317-334.

Xiu, A. and J. E. Pleim, 2001: Development of a Land Surface Model. Part I: Application in a Mesoscale Meteorological Model. *J. Appl. Meteor.*, **40**, 192–209.

APPENDIX I: Objective Verification

The MAE is used to measure how close the model values are compared to the observed values. The MAE is defined as:

$$MAE = \frac{\sum_{i=1}^{N} |x_{\text{mod}} - x_{obs}|}{N}$$ (A.1)

The ME measures the bias of the model values compared to the observed values. The ME is defined as:

$$ME = \frac{\sum_{i=1}^{N} (x_{\text{mod}} - x_{obs})}{N}$$ (A.2)

To calculate the wind direction statistics, both the model values and the observed values for the u and the v wind components are taken from the pairs files.

$$Wind\_Dir = a\tan 2\left(\frac{v}{u}\right) * \frac{180}{\pi}$$ (A.3)

where $a$tan2 is a function that accounts for the signs of the u and the v wind components in order to place the angle in the correct quadrant and π is approximated to a value of 3.1416 for this equation. Using equation A.3 and subtracting the result from 270, both the model wind direction and the observed wind direction are calculated in degrees at each observation location. Subtracting from 270 gives the correct heading for the wind direction based on meteorological convention (the direction the wind is from).

APPENDIX II: Project Specific Codes

1) WRF

The WRF code used in this research was WRFV3.2.1 release. In addition to the new obs nudging capabilities (described in section 2) that have now become standard WRF obs nudging codes, the WRF model used for the summer ozone case period includes two of the recent Penn State modifications in WRF obs nudging capability. The first modification allows WRF obs nudging to use an MM5 method (Stauffer and Seaman 1994) to define the horizontal radius of influence especially important in complex terrain. The WRF default method can produce adverse effects of the observations' influence over complex terrain. Comparison of the WRF solutions using the two different methods indicated that the MM5 method has a slight advantage (not shown). The second modification allows the lowest sounding level to be treated like the regular surface observations in terms of how surface obs are spread horizontally. These modifications have been delivered to BAAQMD. There has been an intention to implement some of these capabilities into a future version of WRF through NCAR.

2) OBSGRID

The WRF objective analysis software, OBSGRID, used in this research was the 22 April, 2011 build. OBSGRID was still under development at the time when this project started. Some of the new functionalities were made available to this project through a different research project in collaboration with NCAR. One of the important modifications to OBSGRID is to allow the MM5-RAWINS type of the Cressman option, which automatically assigns the scale factors for each analysis cycle by explicitly setting the radius of influence to zero. For this study the scale factors that were automatically assigned by the OBSGRID program were 15 (or 540 km), 11, 8, and 6. These modifications are now available for public use. In order to help the subregion analysis, we modified OBSGRID to allow graphic display of the observations in a user-specified subregion. This capability is not available to public, but has been delivered to BAAQMD.

3) Special Obs Processing Software

To process the special obs and convert them into the OBSGRID-compatible format, several software modules for processing special surface wind obs, surface AQS obs, and special wind profiler data were created. These codes have been delivered to BAAQMD.

APPENDIX III: Effects of Assimilating Special Wind Profiler Observations

For the winter case, an additional experiment was conducted to evaluate the effetcts of assimilating the special wind profiler data from 18 profile stations for the period 12 UTC, 16 Dec. 2000 to 12 UTC, 21 Dec. 2000. These profiler data included wind speed, wind direction and site location (lat/lon) in one single file that was downloaded from the BAAQMD FTP site. PSU created software to convert the data to OBSGRID-input format.

Two obs nudging experiments were compared: 1) OFDDA2_HR − obs nudging only using the standard WMO data on all grids using the higher resolution version of UNGRIB for improved QC, same as OFDDA described in section 4; and 2) OFDDA2_spec − same as OFDDA2_HR except that the special wind profiler data were also included in obs nudging. FDDA parameters used in these two experiments are shown in Table A-1. Nudging of the wind field was applied through all model layers, but nudging for the mass field was only allowed above the model-simulated PBL. No nudging was allowed for surface mass field observations. A time window of two hours was used for upper air observations, with a reduced window of one hour at the surface (SFCFACT=0.5).   A reduced radius of influence from 100-km (SFCFACR=0.67 used in both OFDDA2_HR and OFDDA2_spec) for surface data was also used.

Table A-1: FDDA parameters

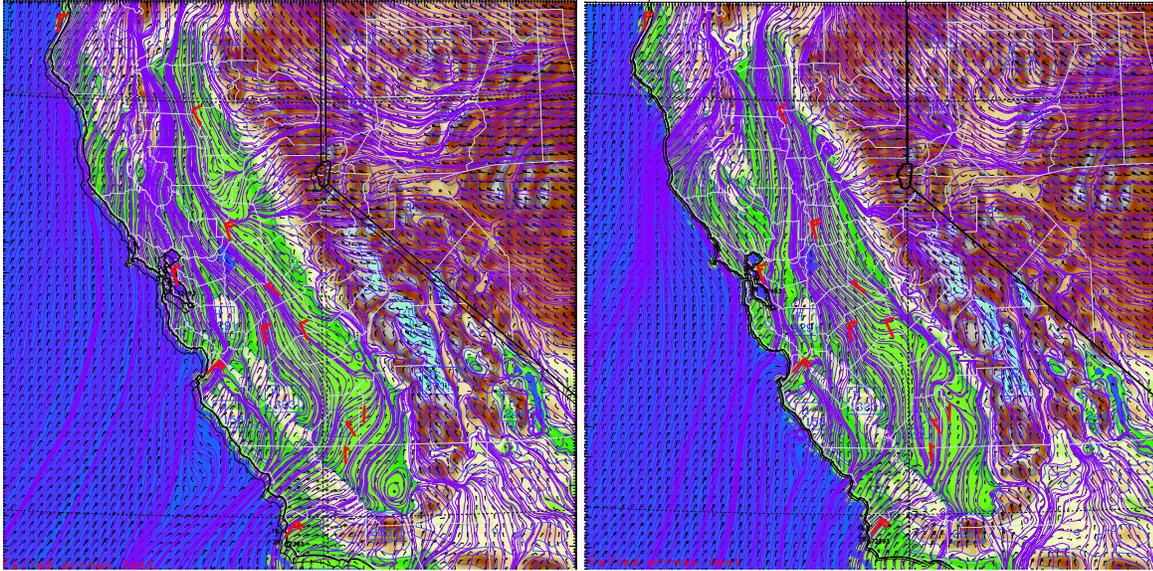|	OBS Nudging		
	36km	12km	4km
G (1/sec)	0.0004	0.0004	0.0004
3-D wind field	Nudging all layers	Nudging all layers	Nudging all layers
3-D mass field	Nudging above PBL	Nudging above PBL	Nudging above PBL
Sfc wind field	Used within PBL	Used within PBL	Used within PBL
Sfc mass field	Not used	Not used	Not used
RINXY (km)	150	100	100
TWINDO (hr)	2	2	2
dt (sec)	180	60	20
IONF	2	4	10

Figures A-1a and A-1b show the model-simulated winds and streamlines at the k=42 sigma level (equivalent to ~250m AGL at Fresno, CA) overlaid with the special profiler data at the 988-mb level (also equivalent to ~250 m AGL at Fresno, CA), for OFDDA2_HR and OFDDA2_spec, respectively. Comparing figures A-1a and A-1b demonstrates the added value

of using special wind profiler in addition to assimilating the two WMO radiosondes on the 4-km domain. The simulated winds using only WMO observations in the San Joaquin Valley (Figure A-1a) are in the opposite direction (southerly) from the observations (northerly), which is corrected by assimilating the special wind profiler data (Figure A-1b). Without the special data, the model also has trouble simulating the winds along the coast due to the low density of standard WMO observations (e.g. the model-simulated wind directions over the Monterey Bay have large errors). The flow patterns over the San Joaquin Valley are corrected by assimilating observations. The OFDDA2_spec experiment simulates the northerly flow along the coast near the northwest corner of the domain while the OFDDA2_HR fails to simulate it and instead has southerly flow throughout the region. The OFDDA2_spec experiment maintains southerly flow throughout the valley.
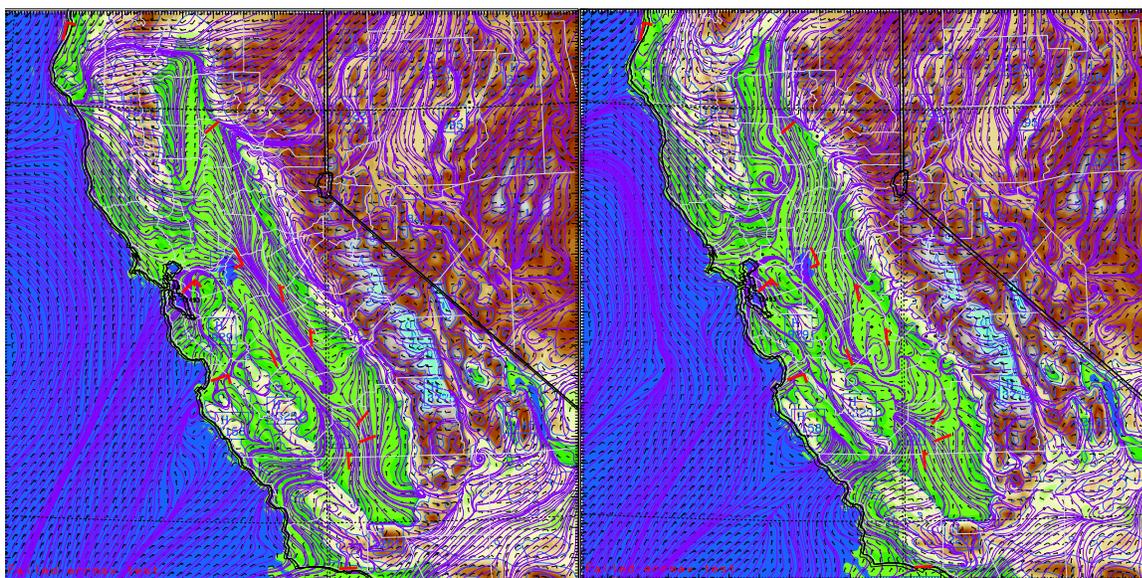

Figure A-1: WRF-simulated daytime winds at 988 mb level at 00 UTC (16 PST), 17 Dec. 2000, over the entire 4-km domain, overlaid with the special wind profiler observations at the same time. a) Obs nudging with WMO obs only. b) Obs nudging with WMO and the special wind profiler obs.

Figures A-2a and A-2b are similar to figures A-1a and A-1b except they are for the nighttime conditions. The FDDA experiments, with and without special observation assimilation, simulate northerly flow through the entire Central Valley beginning with drainage flow off the mountains in the northern Sacramento Valley and ending with the flow going over the Coastal Ranges and offshore in the southern San Joaquin Valley. Without special data assimilation, a large vortex develops south of Bakersfield in the southern San Joaquin Valley, which does not agree well with some of the observations in the southern valley, while with special data assimilation there is better agreement with the observations. With special data assimilation (Figure 2-Ab) the large vortex in the southern San Joaquin Valley was not simulated but a smaller vortex is starting to form near Fresno, possibly a sign of the Fresno Eddy. With special data assimilation another eddy forms near Sacramento that is not as evident in figure A-2a where special data was not assimilated. Over the northeastern side of the Bay Area region the model-simulated winds in figure A-2b are in better agreement with the obs than in figure A-2a.

111

Figure A-2: WRF-simulated nighttime winds at 988 mb level at 12 UTC (04 PST), 17 Dec. 2000, over the entire 4-km domain, overlaid with the special wind profiler observations at the same time. a) Obs nudging with WMO obs only. b) Obs nudging with WMO and the special wind profiler obs.

Figure A-3 compares the upper air statistical scores between experiments OFDDA2_HR and OFFDA2_spec. This comparison demonstrates the added value of assimilating the special wind profiler data. Consistent with the subjective verification using mesoscale analyses, the added value of assimilating the special wind profiler data is clear. The statistical scores indicate significant improvement for all domains with the largest error reduction in wind fields on the 4-km grid.
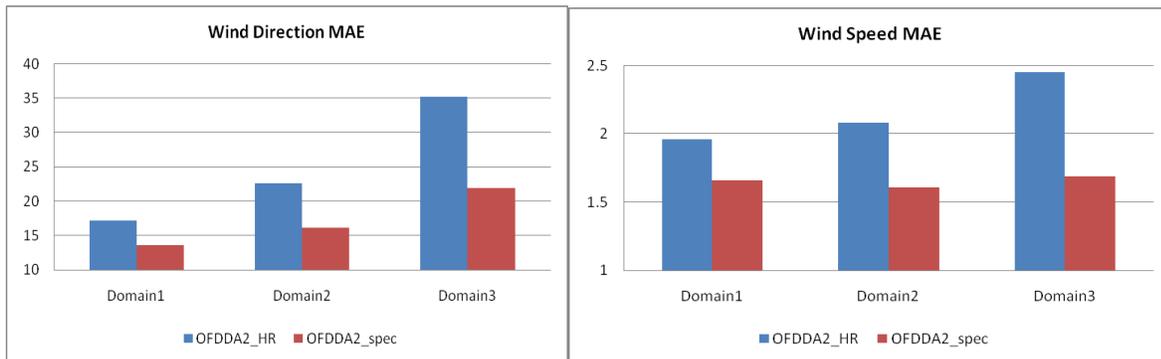
Figure A-3: Mean absolute error (MAE) of the WRF-simulated upper air a) wind direction, and b) wind speed, for all three grids, averaged over the entire 5-day period, for experiments OFDDA2_HR and OFDDA_spec.