2.5 Regional Forecast Impacts from GEMS Observations

Joseph G. Dreher* and John Manobianco
ENSCO, Inc., Melbourne, FL

1. INTRODUCTION

Technological advancements in Micro Electro Mechanical Systems (MEMS) and nanotechnology have inspired a concept for a revolutionary observing system called Global Environmental Micro Sensors (GEMS). The system features a wireless network of in situ, buoyant airborne probes that can monitor all regions of the Earth with unprecedented spatial and temporal resolution. The probes will be designed to remain suspended in the atmosphere for hours to days and take measurements of temperature, humidity, pressure, and wind velocity that are commonly used as dependent variables in numerical weather prediction (NWP) models. As a result, it will not be necessary to develop complex algorithms for assimilating such data into research or operational models.

This paper provides a discussion of the system used to simulate dispersion of and observations collected by an ensemble of probes, highlights a possible deployment scenario and describes a series of Observing System Simulation Experiments (OSSEs) performed to assess the impacts of simulated GEMS data on regional weather forecasts.

2. SIMULATION SYSTEM

The Advanced Regional Prediction System (ARPS; Xue et al. 2000; Xue et al. 2001) coupled with a Lagrangian particle model (LPM) is used to simulate the dispersion of observations collected by an ensemble of probes. The ARPS is a complete, fully automated, stand-alone system designed to forecast explicitly storm- and regional-scale weather phenomena. It includes a data ingest, quality control, and objective analysis package known as ADAS (ARPS Data Analysis System; Brewster 1996), a prediction model, and a post-processing package.

Probe dispersion is simulated using the LPM embedded within ARPS. The probes are assumed to be passive tracers moving independent of one another and transported by the wind. The LPM tracks the location of each probe based on three-dimensional wind components and updates probe position using the resolvable-scale components of wind velocity directly from the ARPS model, as well as turbulent velocity fluctuations. The turbulent velocity fluctuations are estimated from a subgrid scale (SGS) turbulence parameterization (Mellor and Yamada 1980) similar to the SGS turbulence scheme of Deardorff (1980) used in the ARPS model. A parameterization scheme for wet deposition or precipitation scavenging is included in the LPM to simulate the impact of frozen and liquid precipitation on probe trajectory and possible washout (Seinfield and Pandis 1998). The scavenging process is parameterized in the LPM by accounting for two effects. First, rain or snow is assumed to wet the probe shell and increase its mass. Second, rain drops are assumed to impart their momentum to the probe upon impact.

3. DEPLOYMENT SCENARIO

A large number ($>10^6$) of simulated probes can be deployed at any time during the model integration, and at any latitude, longitude, and altitude within the three-dimensional ARPS domain. The LPM provides accurate position information because the velocity variables are updated every model time step by interpolating to the actual probe locations.

Two 30-day periods from June and December 2001 were selected to study the dispersion characteristics of GEMS under differing weather regimes. The summer case was chosen to assess probe dispersion during a weather pattern with relatively weak large-scale flow, and the winter case was selected to analyze probe dispersion with strong jet streams and progressive large-scale features. The differences in simulated probe dispersion were expected to be substantial depending on the prevailing weather patterns, so it was important to study the dispersion patterns under two widely varying weather scenarios.

A strategy was developed to deploy positively buoyant probes that ascend upward through the atmosphere. For this scenario, simulated probes were released from surface weather station sites around the northern hemisphere and ascended to a level of neutral buoyancy that depends on probe mass. This scenario examines the impact of probes remaining neutrally buoyant throughout 30-day simulations versus becoming negatively buoyant and falling out of the air gradually. Depictions of the resulting probe distribution for both June and December 2001, over the northern hemisphere after 14 days are shown in Figure 1.

* *Corresponding author address*: Joseph G. Dreher, ENSCO, Inc., 4849 N. Wickham Rd., Melbourne, FL 32940. e-mail: dreher.joe@ensco.com

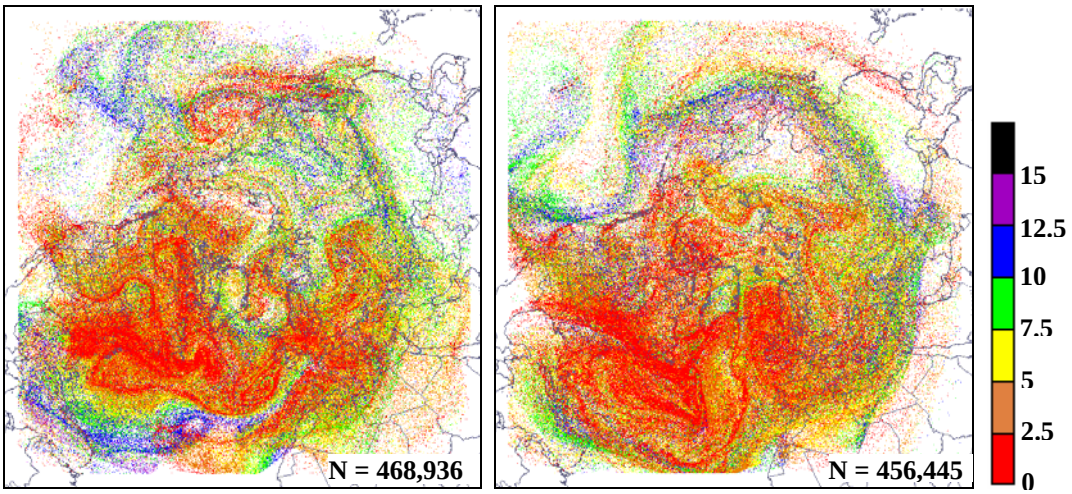

Figure 1. Probe positions for the hemispheric ARPS simulations at (a) 0000 UTC 15 June 2001, and (b) 0000 UTC 15 December 2001, 14 days after model initialization time. The probe altitude (km) is denoted by the colors according to the key provided and total number of probes is given by N.

4. REGIONAL OSSEs

OSSEs are used to assess the impact of probe measurements on weather analyses and forecasts following Atlas (1997) and Lord (1997). The model used for OSSEs is the Pennsylvania State University (PSU)/National Center for Atmospheric Research (NCAR) Fifth-generation Mesoscale Model (MM5; Grell et al. 1995). The MM5 is configured in such a manner as to generate a significantly different solution from a nature simulation to approximate the differences between a state-of-the-art model and the real atmosphere (Atlas 1997). The OSSE methodology consists of three steps:

- **Nature simulations.** These forecast runs are considered "truth" and the trajectories of all simulated probes are tracked and extracted. In addition to simulated GEMS data. all surface, rawinsonde. and aircraft data are extracted from these model simulations as well. The ARPS model is used for the nature simulations.
- **Conventional simulations (Cnv)**. Simulated surface, rawinsonde, and aircraft observations are intermittently assimilated into the MM5 at specified times.
- **Conventional & GEMS simulations (CnvGEMS)**. In addition to conventional data, simulated GEMS data are intermittently assimilated into the MM5 at specified times.

4.1 Nature Runs

Two ARPS 50-km hemispheric nature runs (domain A, Figure 2) were initialized using Aviation Model (AVN) re-analysis fields (1° x 1°) from 0000 UTC 1 June 2001 and 0000 UTC 1 December 2001, respectively, and run for 30 days to simulate large-scale dispersion of GEMS probes. The AVN grids were also used to provide lateral boundary conditions at 12-h intervals throughout each model run (Kalnay et al. 1996). A one-way nested 15-km domain covering a large portion of the United States and Canada (domain B, Figure 2) was initialized at 0000 UTC 10 June 2001 and 0000 UTC 10 December 2001, respectively, and run 10 days. For each 15-km ARPS simulation, lateral boundary conditions were supplied by the ARPS 50-km simulation at 3-h intervals. Simulated measurements from conventional networks and GEMS probes were extracted at 3-h intervals during each 15-km ARPS simulation.

To simulate measurements obtained from GEMS probes, a tri-linear interpolation algorithm within the LPM was used to extract values of temperature, humidity, u and v wind components, and pressure at locations throughout the model integration. A random component to represent instrument error was added to the simulated observations in order to address questions regarding instrument accuracy.

Only conventional in situ surface, upper air, and aircraft observations were extracted from the ARPS simulations. Remote sensing observations were not used in any of the experiments. To simulate measurements obtained from surface and rawinsonde instrumentation, 3D position data were used to extract measurements from the ARPS simulations. The simulated rawinsonde data were only extracted at 12-h

intervals (0000 UTC and 1200 UTC) similar to the observation frequency of the current operational rawinsonde network. Each simulated rawinsonde observation contained 26 levels of data in order to emulate the significant and mandatory levels reported by current rawinsonde measurements.

To simulate observations from the Aircraft Communications Addressing and Reporting System (ACARS), both time and position interpolation were used to extract measurements from the ARPS simulations using actual ACARS flight positions obtained from the Forecast Systems Laboratory (FSL; Dr. John Smart, personal communication). The FSL ACARS flight position data were obtained for a typical 24-h period and used to approximate the positions during the period of interest. In addition, simulated moisture data were not included in the ACARS data suite since the operational system generally does not include such data. In a similar manner to the simulated GEMS observations, random error was added to all the simulated conventional data.

4.2 Regional Assimilation Runs

The MM5 simulations were configured to provide sufficiently different forecasts from the nature runs by degrading the horizontal grid spacing to 60 km and 30 km for the hemispheric and regional assimilation runs, respectively. The MM5 60-km runs were initialized by interpolating the ARPS 50-km nature solutions to the MM5 grids at 0000 UTC 10 June and 0000 UTC 10 December 2001. The MM5 60-km grid covers approximately the same area as the 50-km ARPS domain A in Figure 2. As with the ARPS nature runs, AVN re-analysis fields (1° x 1°) supplied lateral boundary conditions at 12-h intervals throughout each model run.

The regional MM5 30-km simulations, covering approximately the ARPS domain B in Figure 2 were initialized at 0000 UTC 11 June and 0000 UTC 11 December 2001. The MM5 30-km simulations were conducted in a one-way nested configuration from the MM5-60-km simulations and run until 0000 UTC 18 June and 0000 UTC 18 December 2001. By integrating the MM5 60-km simulations for 1 day prior to the initialization of the regional runs, the ARPS and MM5 solutions diverge over the assimilation domain due to the inherent disparities between the NWP models and different horizontal resolutions. These differences between the nature and assimilation run are designed to approximate the typical differences between a "state of the art" NWP model and the real atmosphere (Atlas 1997).

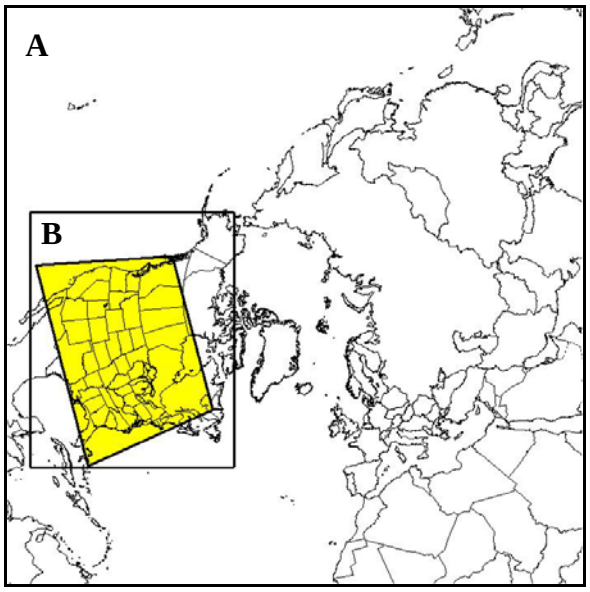

Figure 2. Grid configuration for the ARPS nature and MM5 OSSE simulations. Grid A represents the outer ARPS 50-km and MM5 60-km domains, while grid B denotes the ARPS 15-km and MM5 30-km domains, respectively. The yellow-shaded box represents the area of objective verification statistics described in section 5.

Simulated conventional and/or GEMS data obtained from the ARPS 15-km simulations were intermittently assimilated into the MM5 at 3-h intervals throughout each run. For both 30-km MM5 experiments, the 60-km MM5 simulations supplied the lateral boundary conditions. The ARPS nature and MM5 OSSE methodology is summarized in Figure 3.

Data were assimilated into the MM5 using an intermittent DA technique similar to Rogers et al. (1996) and Manobianco (2002) as depicted in Figure 4. This technique incorporated simulated data from the ARPS model runs into the MM5 integration by using a two successive-scan Cressman scheme with quality-control checks that subsequently adjusted the analyses from the first guess towards the observations at 3-h intervals. Each 3-h background field contains information from the previous observations through the analysis and forecasts of MM5. This cycle was repeated every 3-h throughout the 7-day forecast periods from 11-18 June and December 2001.

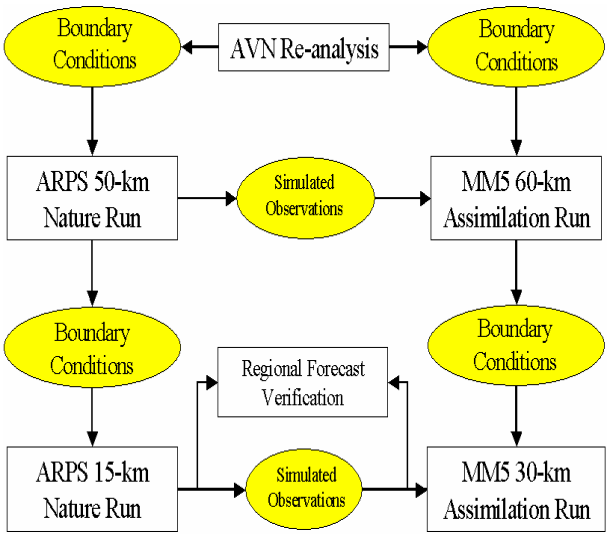

Figure 3. Flowchart for the nature simulations and OSSEs.

In order to mimic a regional operational forecast cycle, 48-h forecasts were generated at 6-h intervals during the intermittent DA cycle for both the June and December 2001 OSSEs following Weygandt et al. (2004). A total of 29 forecasts were conducted for each OSSE scenario. A summary of the dates and duration of the regional OSSE forecasts is presented in Table 1. Weygandt et al. (2004) allowed for a 48-h period of model adjustment before extracting observations because they used different models for the global and regional assimilation runs. No such spin-up period was needed for the regional OSSEs in this study because the MM5 was used for both the hemispheric and regional assimilation runs.

Since information from the regional lateral boundaries propagates through the regional simulations especially at later forecast times (Warner et al. 1997), the regional domains were chosen as large as computationally practical. Furthermore, simulated conventional data (rawinsonde, surface, and aircraft) were assimilated into each MM5 60-km run at 12-h intervals to provide better initial and boundary conditions.

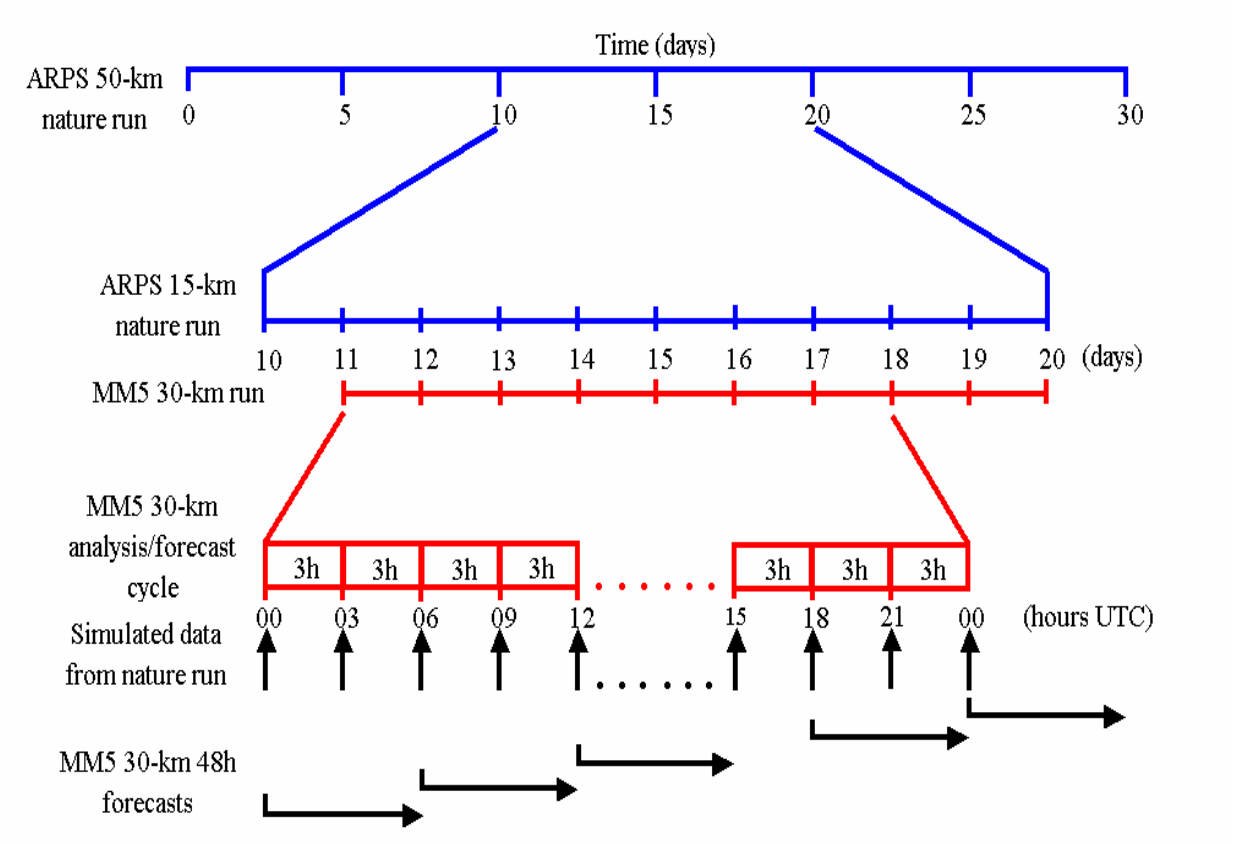

Figure 4. Schematic of the regional OSSEs timeline and data assimilation methodology.

Table 1. Summary of the regional OSSE forecasts forth the June and December 2001 experiments.

Simulation	Dates	Duration	Experiment
Regional nature run – ARPS 15-km simulations	10-20 June and December	10 days	ARPS regional forecast
	10-11 June and December	1 day	ARPS 15-km spin-up
	11-18 June and December	7 days	Simulated surface and aircraft observations extracted at 3-h intervals (simulated rawinsonde extracted at 12-h intervals)
Regional assimilation run – MM5 30-km simulations	11-18 June and December	7 days	Intermittent DA cycle with 3-h update cycle using ARPS simulated observations
	11-18 June and December	7 days	Generation of 48-h forecasts at 6-intervals
	11-20 June and December	9 days	Verification of MM5 forecasts against ARPS nature simulations

One of the limitations in the current study is that simulated satellite observations were not extracted from ARPS and assimilated into MM5 on either the hemispheric or regional domains. Since the regional model runs were conducted primarily over land regions, excluding satellite data in the OSSEs has less of an impact than if such data were used over oceanic regions where in situ data are sparse. In fact, Zapotocny et al. (2005a, b) demonstrated that, except for cloud track winds, satellite data have much less impact on 12- to 24-h forecasts over data dense regions of the U.S where rawinsonde data are most prevalent. A future set of OSSEs using a global modeling system with capabilities to simulate a full suite of in situ and remotely sensed data is needed to study this issue in greater detail and mitigate the impact of lateral boundary conditions inherent with regional modeling.

4.3 Conventional OSSEs (Cnv)

The Cnv OSSEs include only in situ data from simulated conventional networks. The Cnv simulations serve as a reference against which the experiments are compared, since no simulated GEMS observations were assimilated.

4.4 Conventional & GEMS (CnvGEMS)

In addition to simulated conventional data, the CnvGEMS OSSEs include simulated data obtained from the GEMS surface deployment scenario. All simulated data (conventional and GEMS) were assimilated into the MM5 30-km OSSEs at 3-h intervals.

5. REGIONAL OSSE VERIFICATION

In order to verify the regional forecasts with and without GEMS data, the ARPS nature runs were interpolated to a grid identical to that of the MM5 30-km simulations (following Hamill and Colucci 1997). Objective verification of the OSSEs was then accomplished by calculating gridded bias and root mean square (RMS) errors of temperature, dew point, and vector wind over a sub-domain centered on much of the U.S. (yellow shaded box in Figure 2). Vector wind verification was chosen to summarize the results from both the u and v wind components using a single statistic. These aggregate statistics provide a straightforward comparison between the Cnv and CnvGEMS OSSEs. The use of more extensive objective and subjective verification strategies (e.g. Nutter and Manobianco 1999; Ebert and McBride 2000; Case et al. 2002) was beyond the scope of the present study.

If Φ represents a predicted variable from the benchmark simulation or OSSEs, then forecast error is defined as:

$$\Phi' = \Phi_{exp} - \Phi_{nat}, \qquad (1)$$

where the subscripts *exp* and *nat* denote the experiment (OSSE) and nature quantities, respectively. The bias represents the average model error of the benchmark or OSSEs, and is computed as:

$$\text{Bias} = \frac{1}{N}\sum_{i=1}^{N} \Phi', \qquad (2)$$

where N represents the total number of grid points (171 x 203) times the number of forecasts (29) at any given height in the atmosphere. The RMS error is calculated as:

$$\text{RMS Error} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\Phi')^2} \quad . \qquad (3)$$

The statistics were computed from gridded forecast errors fields where the ARPS nature simulations represented the observed fields and the MM5 experiments represented the forecasted fields. For data impact comparisons between the CnvGEMS and Cnv OSSEs, the forecast impact was normalized by computing a percentage forecast improvement of the experiment compared to the control forecast as follows:

$$\%\ \text{Improvement} = 100 \times \frac{\text{CNTL - EXP}}{\text{CNTL}} \quad , \qquad (4)$$

where CNTL is the control RMS error and EXP is the experiment RMS error. Positive (negative) values indicate improved (worsened) impact of the assimilated data on the forecasts. In this case, the CNTL RMS error represents the Cnv OSSEs with the entire suite of simulated conventional observations. On the other hand, the EXP RMS error represents the CnvGEMS OSSEs with simulated conventional and GEMS observations.

A total of 9 OSSEs were conducted and the details of each experiment are summarized in Table 2 along with the ARPS nature simulations. Each OSSE, with the exception of sensitivity Exp. 3, assumed perfect observations with no instrument errors. In Exp. 3, random errors for the simulated probes were added to each variable, based on typical errors for MEMS sensors (Kristofer S. J. Pister, personal communication). In addition, random errors were added to each variable for the simulated conventional observations. Exps. 6 and 7 reduced the total number of probes to 10% and 1%, respectively. Data thinning

Table 2. Summary of the simulations and regional OSSE experiments for June 2001 and December 2001. For each experiment, the variables assimilated into the OSSE (if applicable) are provided, along with a description of experiment. Experiment descriptions are only given for the regional ARPS 15-km and MM5 30-km simulations, respectively. Sensitivity experiments 3-8 were conducted only for June 2001.

Simulations	**Variables Assimilated**	**Experiment Description**
Nature	N/A	ARPS 10-day regional simulations
Cnv OSSEs (June and December 2001)	T, p, T_d, u, v*	Simulated surface, rawinsonde, and aircraft observations assimilated into MM5
CnvGEMS OSSEs (June and December 2001)	T, p, T_d, u, v	Same as Cnv OSSE, except that in addition to conventional data, simulated GEMS data are assimilated into MM5
Sensitivity Experiment 1 (June 2001)	T, p, u, v	Same as CnvGEMS OSSEs, but exclude T_d
Sensitivity Experiment 2 (June 2001)	T, p, T_d	Same as CnvGEMS OSSEs, but exclude u, v (winds)
Sensitivity Experiment 3 (June 2001)	T, p, T_d, u, v	Same as CnvGEMS OSSEs, but include random probe and conventional observation errors
Sensitivity Experiment 4 (June 2001)	T, p, T_d, u, v	Same as CnvGEMS OSSEs, but include precipitation scavenging of probes
Sensitivity Experiment 5 (June 2001)	T, p, T_d, u, v	Same as CnvGEMS OSSEs, but use a 6-h intermittent data assimilation cycle
Sensitivity Experiment 6 (June 2001)	T, p, T_d, u, v	Same as CnvGEMS OSSEs, but use only 10% of GEMS data
Sensitivity Experiment 7 (June 2001)	T, p, T_d, u, v	Same as CnvGEMS OSSEs, but use only 1% of GEMS data
*T = temperature, p = pressure, T_d = Dew point, u = u-wind component, v = v-wind component.		

was performed by excluding probes randomly without replacement throughout the assimilation domain to reduce the effective resolution of the assimilated data. The random thinning scheme was designed to emulate changing the deployment strategy to release probes from fewer surface stations. Once a probe was randomly excluded, the data from that probe were not used at any subsequent assimilation times.

6. REGIONAL OSSE RESULTS

Overall, both the June and December 2001 OSSEs demonstrate that the assimilation of simulated GEMS observations extracted from the nature run improved the predicted primary variables over the Cnv experiments. To diagnose the distribution of impacts throughout the entire troposphere, vertical profiles of RMS errors as a function of forecast hour (0 h, 12 h and 24 h) were also plotted. Vertical profiles of percent improvement for RMS errors were generated to display the impact of assimilating GEMS data relative to runs with only conventional data. The percent improvement profiles were further stratified by forecast initialization times with (0000 and 1200 UTC) and without (0600 and 1800 UTC) standard rawinsonde data. The impact of GEMS data was expected to be greater at the non-rawinsonde initialization times when there is substantially less conventional in situ data above the surface.

The Cnv OSSEs included only simulated conventional in situ observations and all sampled meteorological variables were assimilated during both the June and December 2001 time periods. The CnvGEMS OSSEs were similar but included both simulated conventional and GEMS data. The Cnv OSSE experiments were designed to emulate an operational regional forecast assimilation system and serve as a point of reference to benchmark the other experiments. For the June and December 2001 CnvGEMS OSSEs, the RMS errors were generally smaller which indicates a significant positive forecast improvement when compared to the Cnv OSSEs.

6.1 June 2001

The vertical profiles of RMS errors for the CnvGEMS forecasts indicate large error reductions for each variable as benchmarked against the Cnv simulation (Figure 5). The largest RMS error differences occur at 0 h in the mid and upper troposphere when GEMS data have the most impact by improving the analyses used as initial conditions in the subsequent 48-h forecasts (Figure 5a, d, g). In fact, the CnvGEMS RMS errors of vector wind in the upper troposphere at 0 h are 2 m s^{-1} lower than the Cnv forecasts (Figure 5g). However, those effects decrease with time to ~1 m s^{-1} at 12 h and 24 h (Figure 5h, i). It is important to note that the smaller RMS error differences below 900-hPa for each variable are likely due to the positive impact of the conventional surface data (Figure 5). In this case, there are ~1,000 near-surface GEMS observations in addition to 2,337 conventional surface stations.

The vertical profiles generally show a 5-10% larger improvement in the mid and upper troposphere for the non-rawinsonde (06, 18 UTC) versus rawinsonde (00, 12 UTC) initialization times when simulated GEMS observations are competing with more conventional observations (Figure 6). Overall, the CnvGEMS forecasts at 0 h show improvements of greater than 30% for each variable throughout the depth of the troposphere (Figure 6a, d, g). The percent improvement for vector wind reaches a maximum of 60% at 0 h for the non-rawinsonde forecasts above 600-hPa (Figure 6g). The percent improvements for each variable at 12 h and 24 h are 15-20% smaller than at 0 h (Figure 6b-i); however, even for the 24-h dew point forecasts, the improvement reaches a maximum of 35% above 250-hPa (Figure 6f). The larger percent improvements in dew point above 250-hPa are likely due to the absence of moisture data in the simulated ACARS data; whereas the improvements in temperature and vector wind are likely smaller because ACARS already provides these data.

6.2 December 2001

The vertical profiles of RMS errors from December 2001 show substantial differences between the Cnv and CnvGEMS OSSEs throughout the troposphere, especially at the 0-h and 12-h forecasts (Figure 7). The largest differences occurred at 0 h when the RMS errors of vector wind from CnvGEMS in the upper troposphere are 2 m s^{-1} smaller than those from the Cnv forecasts (Figure 7g). However, the 24-h forecasts from December 2001 show 1 m s^{-1} smaller vector wind error differences between Cnv and CnvGEMS in the mid and upper troposphere than those from June 2001 (compare Figure 7i and Figure 5i). This result is consistent with the point made by Warner et al. (1997) that the lateral boundary impacts would be larger for stronger (December) versus weaker (June) flow regimes. Essentially, the forecast improvements from assimilating GEMS data diminish more rapidly during stronger flow regimes due to the nature and OSSE grid configurations.

As in June 2001, the largest percent improvements for December 2001 occur for non-rawinsonde initialization times (Figure 8). Notable in the dew point percent improvement statistics is the larger forecast impact at non-rawinsonde initialization times on the order of 75% at 0-h, 50% at 12 h, and 40% at 24 h

around 200 hPa (Figure 8d-f). Similar to June 2001, the large percent improvements around 200 hPa are likely due to the fact that simulated ACARS observations do not include moisture data. Therefore, GEMS data are making larger impacts (10-20% larger than Cnv) on forecasted dew points at these levels (Figure 8d-f). The slight forecast degradation above 200 hPa at 12 h and 24 h for temperature and vector wind is likely due to the impact from lateral boundary conditions affecting the verification domain (Figure 8b, c, h, i). No such forecast degradations were evident in the June 2001 case, which indicates that the stronger flow regime in December may have contributed to lateral boundary condition contamination in the regional OSSE methodology.

6.3 Sensitivity Analysis

For brevity, the results for June 2001 sensitivity experiments are summarized without accompanying figures in the sections that follow:

Experiment 1 – No Dewpoint

Exp. 1 included the same GEMS data as CnvGEMS but the dewpoint variable was excluded from the data assimilation cycle. The most significant impact of excluding dewpoint data was that dewpoint RMS errors were very similar to Cnv dewpoint errors at all forecast times and levels. There was a slight increase in the vector wind RMS errors at all forecast times and throughout the troposphere. Excluding dewpoint data had no significant impact on the temperature RMS errors.

Experiment 2 – No Wind

Exp. 2 included the same GEMS data as in the CnvGEMS simulation, but both the u and v components of the wind were withheld from the data assimilation cycle. The magnitude of the vector wind errors degrades to that of the Cnv forecasts, which indicates that the full suite of GEMS data have substantial impact on the wind forecasts. Both the temperature and dew point RMS errors increase by 0.5 K and 1 K, respectively, below 400 hPa, for the 12-h and 24-forecasts. In fact, the temperature and dew point RMS errors at 24 h approach the magnitudes of the RMS errors from the Cnv forecasts at 24 h. This result indicates the importance of assimilating wind data in order to obtain more accurate forecasts of both temperature and dew point.

Experiment 3 – Instrument Errors

Exp. 3 included the same GEMS and conventional data as CnvGEMS, but with random observational errors. Introducing errors caused little degradation in the temperature, dewpoint and vector wind forecasts when compared with CnvGEMS.

Experiment 4 – Precipitation Scavenging

For Exp. 4 the simulated GEMS data were extracted from an ARPS nature simulation where probe precipitation and ice scavenging was activated in the LPM. Precipitation scavenging made very little difference in the temperature and dewpoint errors, except for slight degradation of the temperature RMS errors by approximately 0.25 K below 400-hPa for the 24-h forecasts. The largest differences were in the vector wind RMS errors with degradation at all levels and forecast times when precipitation scavenging was activated. Vector wind RMS errors were approximately 0.25 m s^{-1} higher at all levels for the 0-h and 12-h forecasts when comparing the RMS errors to CnvGEMS. However, for the 24-h forecasts, the Exp. 4 RMS errors approach the magnitude of those from the Cnv forecasts at 24-h.

Experiment 5 – 6-h DA Frequency

Exp. 5 included the same probes and conventional data as in CnvGEMS, but all simulated data were assimilated at 6-h instead of 3-h intervals. This experiment was designed to test the sensitivity of data assimilation frequency. By assimilating the data less often at 6-h intervals, the RMS errors for all variables did not show degradation for the 0- and 12-forecasts when compared to CnvGEMS.

Experiment 6 - 10% of probe data

Excluding 90% of the probes from the DA cycle did not substantially degrade the 48-h forecasts of temperature, dewpoint and vector wind when comparing the RMS errors to the full data set used for CnvGEMS.

Experiment 7 – 1% of probe data

Excluding 99% of the probes from the DA cycle and subsequent 48-h forecasts substantially degraded the forecasts of temperature, dewpoint and vector wind when comparing the RMS errors to the full data set used for CnvGEMS. For all variables and forecast times the errors approach the magnitudes of Cnv errors, especially for the 12-h and 24-h forecasts. Excluding 99% of the GEMS data provided relatively little forecast improvement over the Cnv simulation.

Exps. 6 and 7 were designed to test the sensitivity of the data impact to the number of probes and effective resolution of the assimilated data using the same deployment strategy. The mean nearest neighbor distances of probes in a selected 50-hPa layer (475-525 hPa) from the CnvGEMS (100%), Exp.6 (10%), and Exp. 7 (1%) OSSEs are plotted in Figure 9. The ±25 hPa layer was chosen to highlight the average spacing

of the observations used by the MM5 objective analysis at a given pressure level. Note that the average probe spacing within the 50-hPa layer decreases from ~25 km for the full dataset to ~90 km for 10% of the probes and greater than 300 km for 1% of the probes at day 5 (0000 UTC 15 June 2001; Figure 9).

Previous studies focusing on objective analysis suggest that the optimum observation spacing is 2 times the model grid spacing (Koch et al. 1983). Since the MM5 regional forecasts were run at 30-km grid spacing, it was not advantageous to have 100% of the probe data with an average spacing of ~25 km because that value is substantially smaller than twice the grid spacing. In effect, 100% of the probe data over-samples the scales of motion that can be resolved using a model with 30-km grid spacing. Exp. 6, that included data from only 10% of probes, produced only slight degradations in RMS errors of temperature, dew point and vector wind when the compared with the full data set. This result is consistent with the fact that 10% of probe data yields an average probe spacing of ~90 km that is closer to but still larger than twice the model grid spacing (60 km). However, the RMS errors from Exp. 7 including only 1% of the probe data approach those of Cnv. For that experiment, the probe spacing of ~300 km is much greater than 2 times the model grid spacing and closer to the average spacing of conventional upper air observations which explains why the results are similar to the Cnv forecasts.

7. SUMMARY AND CONCLUSIONS

A series of OSSEs were conducted to assess the impact of assimilating GEMS data on regional weather forecasts compared with assimilating only current conventional in situ surface, upper air, and aircraft observations. The regional OSSEs were conducted for 10-day periods from the June and December 2001 hemispheric cases used to study probe dispersion. The OSSEs were configured to mimic an operational regional forecast cycle by running 48-h forecasts with and without simulated GEMS data from 29 consecutive initialization times at 3-h intervals between days 10 through 18 of the 30-day hemispheric simulation periods.

The OSSEs demonstrated that the addition of simulated GEMS observations had a significant impact on improving the bias and RMS errors in temperature, dew point, and vector wind forecast compared with the conventional simulations. The improvements to the regional forecast errors exceed 50% especially for the 0- to 12-h forecasts over an already data-rich region. The large improvements in the early forecast period reflect the fact that GEMS data have the most impact on improving the model initial conditions even when using an intermittent data assimilation cycle and simplistic objective analysis scheme. Overall, the forecasts impacts were generally similar for both the June and December OSSEs. Based on this result, data impacts did not depend much on the large-scale prevailing weather patterns that were quite different between the June and December 2001 cases. The only exception was that the impacts of the GEMS data for the December 2001 case decrease faster with forecast hour as stronger flow regimes allow the lateral boundary conditions to affect the interior of the domain more rapidly.

A number of sensitivity experiments were conducted including data thinning that used the same deployment strategy but decreased the number of probes in the network and increased the distance between adjacent nodes. The data thinning OSSE produced very similar forecast impacts using only 10% of GEMS probe data included in the full simulation. Based on the results obtained from these regional OSSEs, further work dealing with GEMS simulated probes is planned. Experiments are currently underway to validate the OSSEs by comparing the forecast impact to assimilating real observations (following Weygandt et al. 2004).

8. ACKNOWLEDGEMENTS

This work was supported by the Universities Space Research Association's NASA Institute for Advanced Concepts.

9. REFERENCES

Atlas, R., 1997: Atmospheric observations and experiments to assess their usefulness in data assimilation. *J. Royal Meteor. Soc. Japan*, **75**, 111-130.

Brewster, K., 1996: Application of a Bratseth analysis scheme including Doppler radar data. Preprints, *15th Conf. on Weather Analysis and Forecasting*, Amer. Meteor. Soc., Norfolk, VA, 92-95.

Case, J. L., J. Manobianco, A. V. Dianic, M. M. Wheeler, D. E. Harms, and C. R. Parks, 2002: Verification of high-resolution RAMS forecasts over east-central Florida during the 1999 and 2000 summer months. *Wea. Forecasting*, **17**, 1133-1151.

Deardorff, J. W., 1980: Stratocumulus-capped mixed layers derived from a three-dimensional model. *Bound.-Layer Meteor*., **7**, 199-226.

Ebert, E., and J. L. McBride, 2000: Verification of precipitation in weather systems: Determination of systematic errors. *J. Hydrol.*, **239**, 179-202.

Grell, G., J. Dudhia, and D. Satuffer, 1995: A description of the Fifth-Generation Penn State/NCAR mesoscale model (MM5). NCAR / TN-398 + STR [Available on line at http://www .mmm.ucar.edu/mm5/mm5-home.html].

Hamill, T., M., and S., J., Colucci, 1997: Verification of Eta–RSM Short-Range Ensemble Forecasts. *Mon. Wea. Review*, **125**, 1312–1327.

Kalnay, E., and Coauthors, 1996: The NCEP/NCAR 40-year reanalysis project. *Bull. Amer. Meteor. Soc*., **77**, 437-471.

Lord, S. J., E. Kalnay, R. Daley, G. D. Emmitt, and R. Atlas, 1997: Using OSSEs in the design of the future generation of integrated observing systems. Preprints, *First Symposium on Integrated Observing Systems*, Amer. Meteor. Soc., Long Beach, CA, 45-47.

Manobianco, J., 2002: Global Environmental MEMS Sensors (GEMS): A Revolutionary Observing System for the 21st Century, Phase I Final Report. [Available online at http://www.niac. usra.edu/studies/].

Mellor, G. L., and T. Yamada, 1982: Development of a turbulence closure model for geophysical fluid problems. *Rev. Geophys. Space Phys*., **20**, 851-875.

Nutter, P. A., and J. Manobianco, 1999: Evaluation of the 29-km Eta model. Part I: Objective verification at three selected stations. *Wea. Forecasting*, **14,** 5-17.

Rogers, E., T. L. Black, D. G. Deaven, G. J. DiMego, Q. Zhao, M. Baldwin, N. W. Junker, and Y. Lin, 1996: Changes to the operational "early" eta analysis/forecast system at the National Centers for Environmental Prediction. *Wea. Forecasting*, **11,** 391-413.

Seinfeld, J. H., and S. N. Pandis, 1998: *Atmospheric Chemistry and Physics – From Air Pollution to Climate Change*, John Wiley and Sons Inc., New York, 1326 pp.

Warner, T.T., Peterson, R.A. and Treadon, R.E. (1997): A tutorial on lateral boundary conditions as a basic and potentially serious limitation to regional numerical weather prediction. *Bull. Amer. Meteor. Soc*., **78**, 2599–2617.

Weygandt, S.S., and coauthors, 2004: Potential forecast impacts from space-based lidar winds: Regional observing system simulation experiments. *8th Symp. Int. Obs. and Assim. Systems for Atm. Oceans and Land Surf.,* Seattle, Amer. Meteor. Soc.

Xue, M., K. K. Droegemeier, and V. Wong, 2000: The Advanced Regional Prediction System (ARPS) — A multi-scale nonhydrostatic atmospheric simulation and prediction model. Part I: Model dynamics and verification. *Meteor. Atmos. Phys.*, **75**, 161-193.

_____, and Coauthors, 2001: The Advanced Regional Prediction System (ARPS) — A multiscale nonhydrostatic atmospheric simulation and prediction tool. Part II: Model physics and applications. *Meteor. Atmos. Phys*., **76**, 143-165.

Zapotocny, T. H., W. P. Menzel, J. A. Jung, and J. P. Nelson III, 2005a: A four-season impact study of rawinsonde, GOES, and POES data in the eta data assimilation system. Part I: The total contribution. *Wea. Forecasting*, **20**, 161-177. **11,** 391-413.

_____, T. H., W. P. Menzel, J. A. Jung, and J. P. Nelson III, 2005b: A four-season impact study of rawinsonde, GOES, and POES data in the eta data assimilation system. Part II: Contribution of the components. *Wea. Forecasting*, **20**, 178-198.

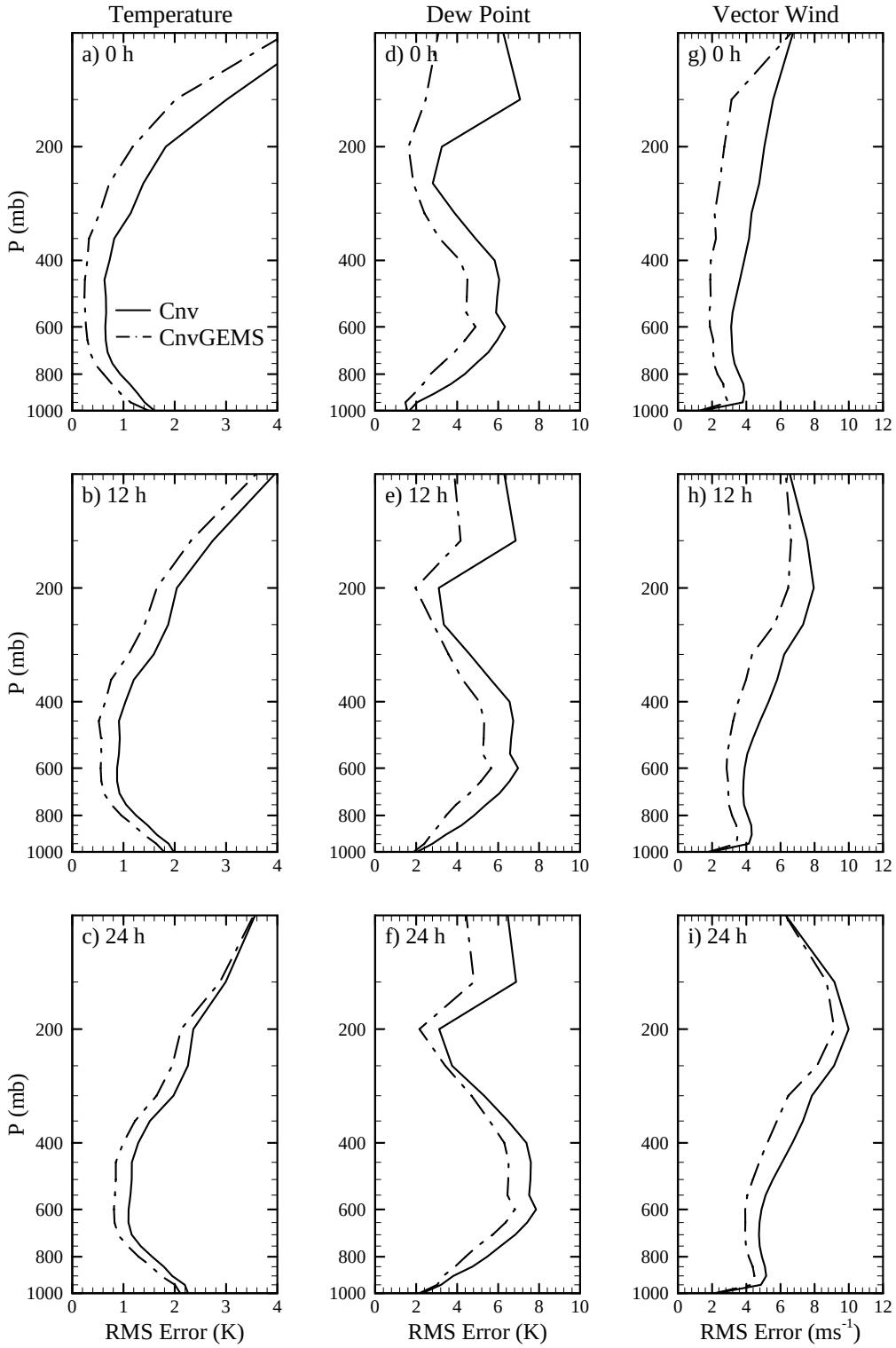

Figure 5. Vertical profiles of the temperature (a-c), dew point (d-f), and vector wind (g-i) root mean square (RMS) error for the Cnv (solid lines) and CnvGEMS (dot-dashed lines) OSSE forecasts from June 2001. Data are presented for the 0-h (a, d, g), 12-h (b, e, h), and 24-h (c, f, i) forecasts. Statistics were computed over the OSSE verification domain shown in Figure 2.

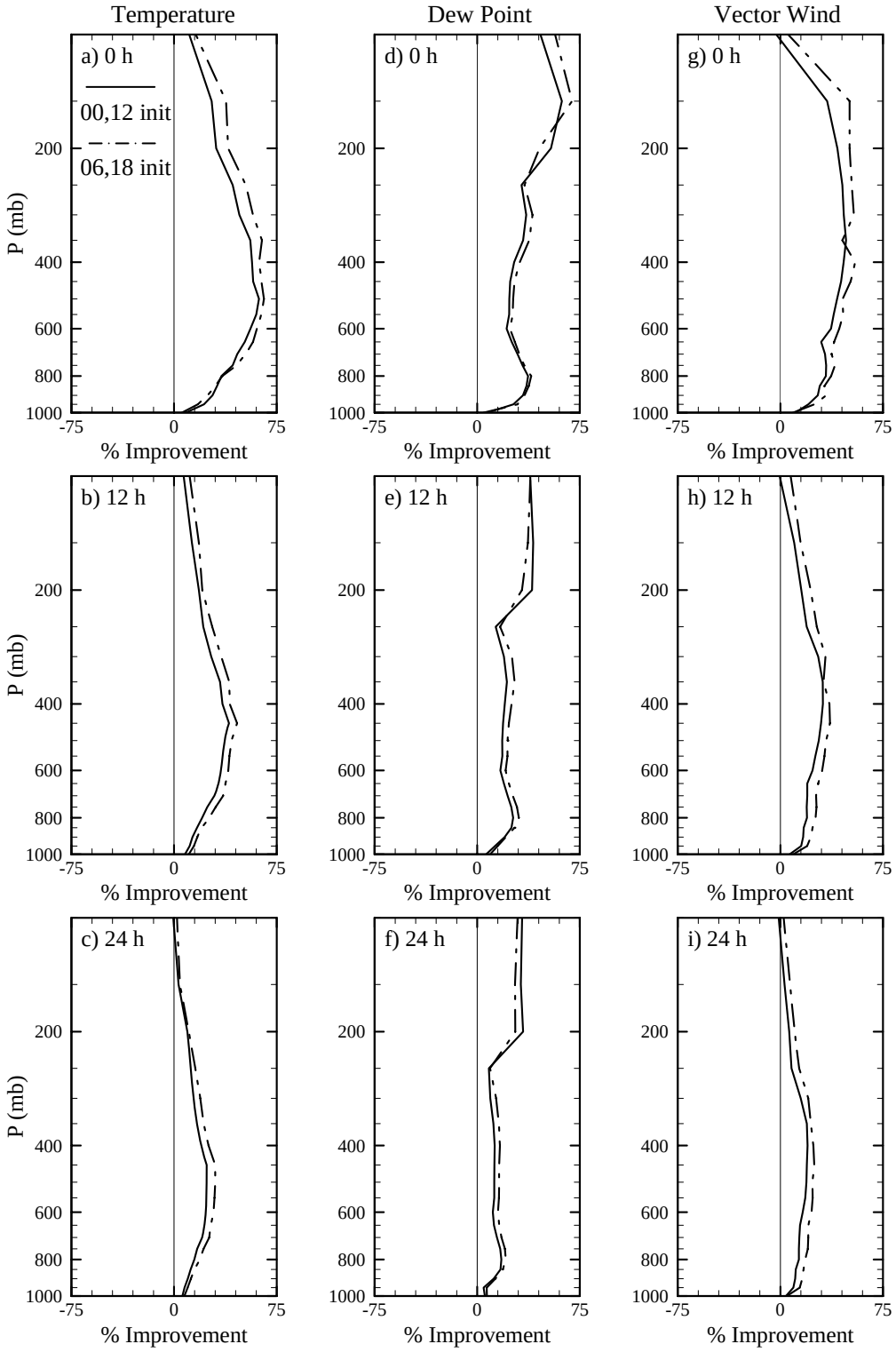

Figure 6. Vertical profiles of the temperature (a-c), dew point (d-f), and vector wind (g-i) percent improvement for the GEMS OSSE forecasts from June 2001 shown for rawinsonde (solid lines) and non-rawinsonde initialization (dot-dashed lines) times. Data are presented for the 0-h (a, d, g), 12-h (b, e, h), and 24-h (c, f, i) forecasts. Statistics were computed over the OSSE verification domain shown in Figure 2.

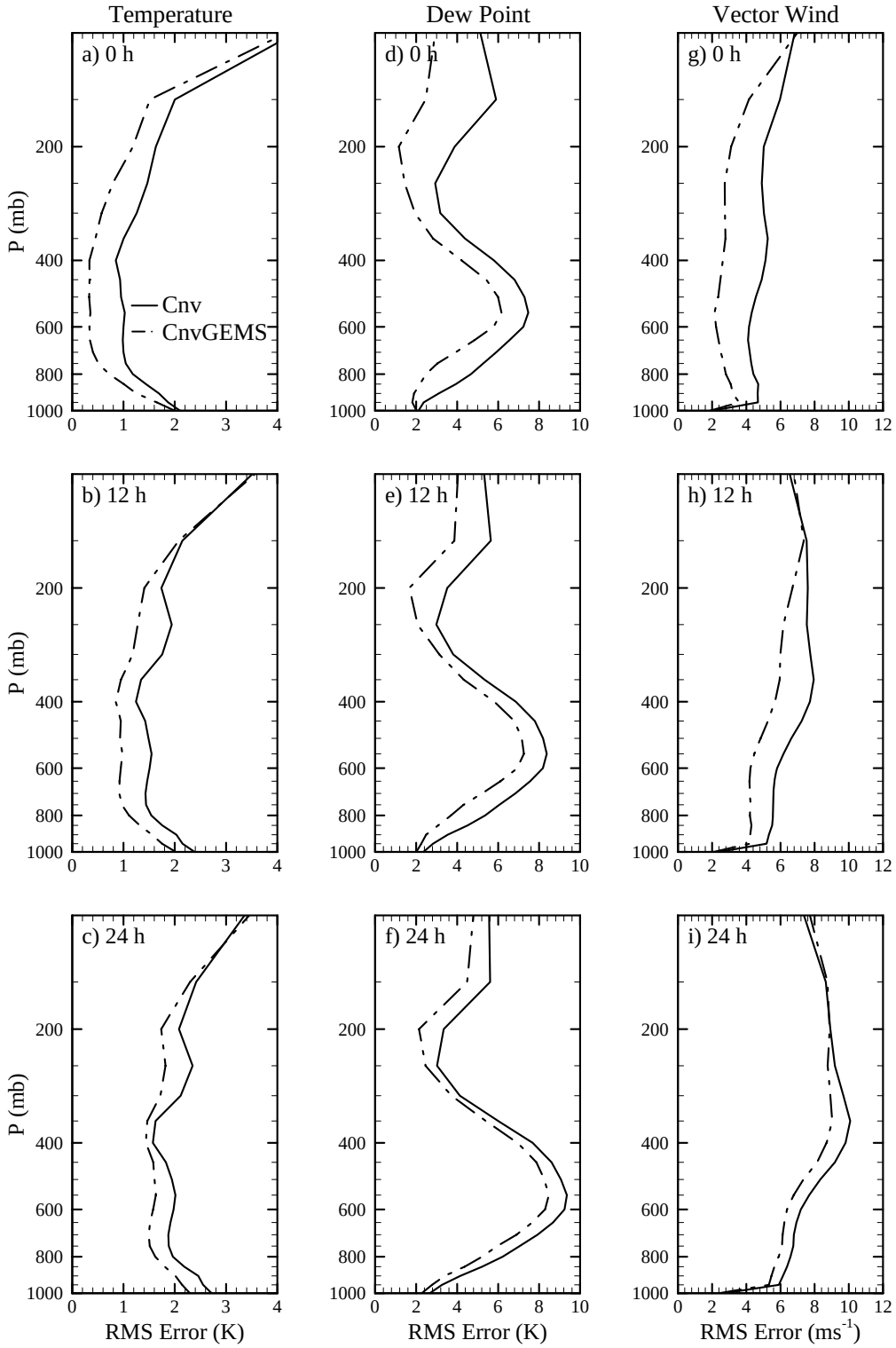

Figure 7. Vertical profiles of the temperature (a-c), dew point (d-f), and vector wind (g-i) root mean square (RMS) error for the Cnv (solid lines) and CnvGEMS (dot-dashed lines) OSSE forecasts from December 2001. Data are presented for the 0-h (a, d, g), 12-h (b, e, h), and 24-h (c, f, i) forecasts. Statistics were computed over the OSSE verification domain shown in Figure 2.

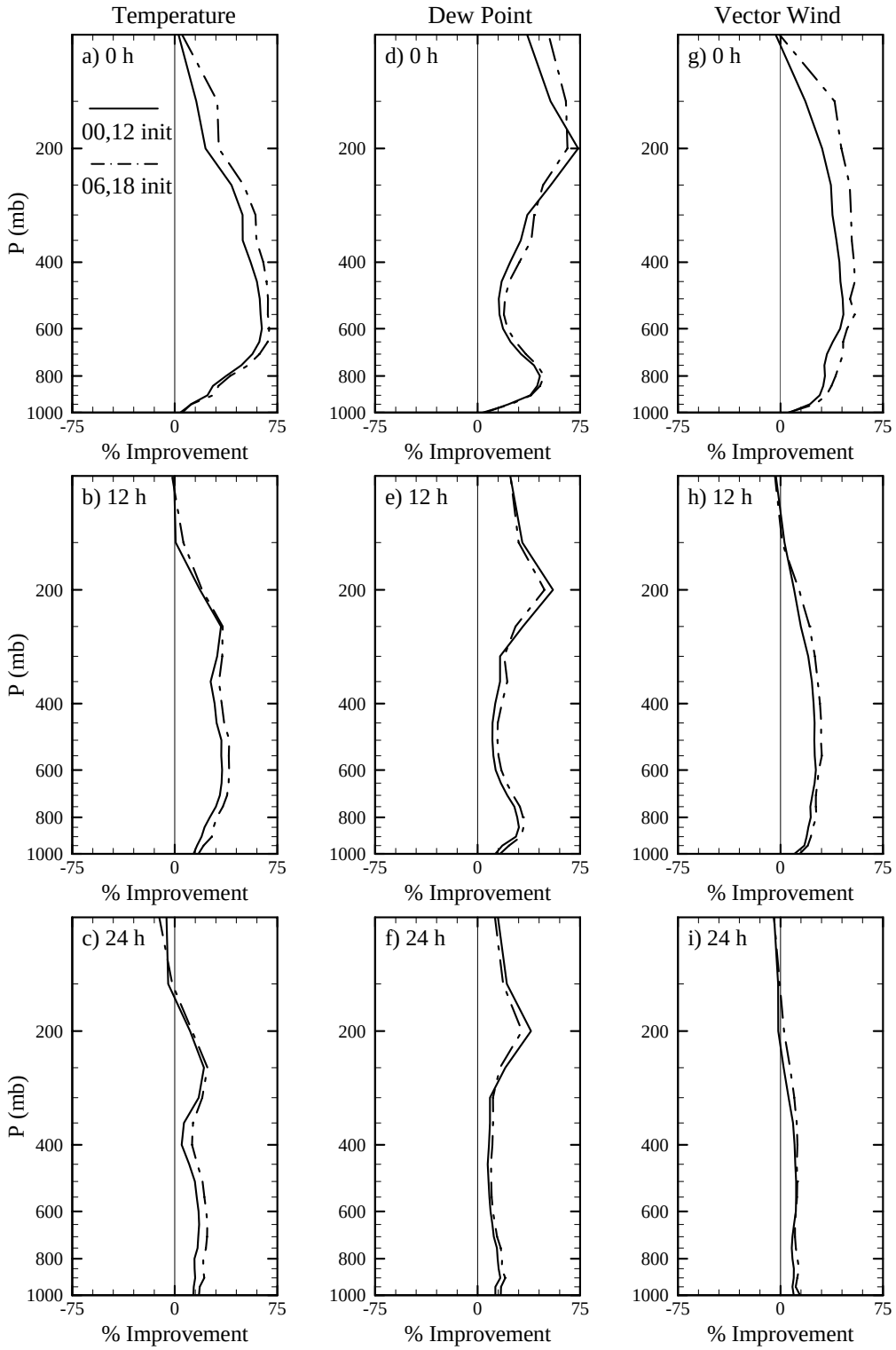

Figure 8. Vertical profiles of the temperature (a-c), dew point (d-f), and vector wind (g-i) percent improvement for the GEMS OSSE forecasts from December 2001 shown for rawinsonde (solid lines) and non-rawinsonde initialization (dot-dashed lines) times. Data are presented for the 0-h (a, d, g), 12-h (b, e, h), and 24-h (c, f, i) forecasts. Statistics were computed over the OSSE verification domain shown in Figure 2.

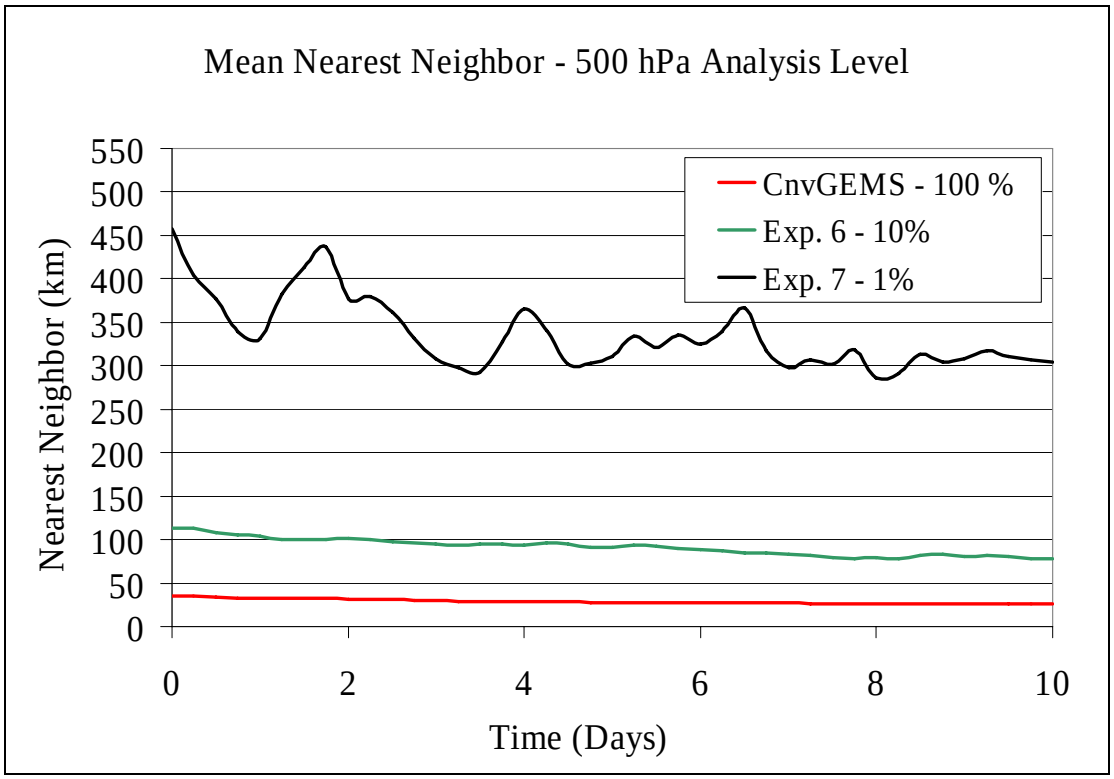

Figure 9. Mean nearest neighbor (NN) distances for 100%, 10% and 1% of the probe data at the 500-hPa analysis level over the ARPS 15-km domain shown in Figure 2.