INVESTIGATING THE PERFORMANCE OF CONVECTION-ALLOWING HINDCAST SIMULATIONS DURING THE NORTH AMERICAN MONSOON WITH AND WITHOUT GPS-PWV DATA ASSIMILATION

By

James M. Moker, Jr.

A Dissertation Submitted to the Faculty of the

DEPARTMENT OF HYDROLOGY AND ATMOSPHERIC SCIENCES

In Partial Fulfillment of the Requirements
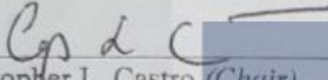
For the Degree of

DOCTOR OF PHILOSOPHY
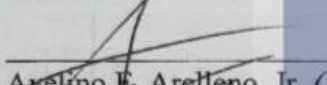
In the Graduate College

THE UNIVERSITY OF ARIZONA

2019

THE UNIVERSITY OF ARIZONA
GRADUATE COLLEGE

As members of the Dissertation Committee, we certify that we have read the dissertation prepared by James M. Moker, Jr. titled **Investigating the performance of convection-allowing hindcast simulations during the North American Monsoon with and without GPS-PWV data assimilation** and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy.
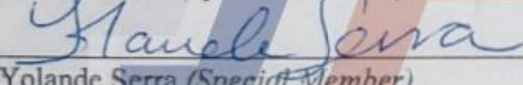
_____ Date: 4/15/2019
Christopher L. Castro *(Chair)*

_____ Date: 4/15/2019
Avelino F. Arelleno, Jr. *(Member)*

_____ Date: 4/15/2019
Thomas Galarneau *(Member)*
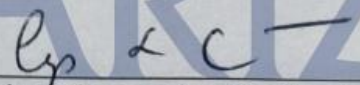
_____ Date: 4/15/2019
Richard Bennett *(Member)*

_____ Date: 4/15/2019
David Adams *(Special Member)*

_____ Date: 4/15/2019
Yolande Serra *(Special Member)*

Final approval and acceptance of this dissertation is contingent upon the candidate's submission of the final copies of the dissertation to the Graduate College.

I hereby certify that I have read this dissertation prepared under my direction and recommend that it be accepted as fulfilling the dissertation requirement.

_____ Date: 4/15/2019
Dissertation Director: Christopher L. Castro

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

ABSTRACT

The North American monsoon (NAM) is responsible for summer severe weather in the Southwest U.S. and northwest Mexico and its associated rainfall contributes the highest percentage of yearly precipitation to this region. Short-term convection-allowing model forecasts have shown difficultly in replicating the diurnal cycle of NAM convective precipitation. Generally, convection initiating over the Sierra Madre Occidental (SMO) mountain range in the early afternoon may later organize into mesoscale convective systems (MCSs) that propagate west towards the lower elevations and Gulf of California (GOC). MCSs account for the greatest proportion of all NAM-related precipitation. In my dissertation, I investigate the performance of daily short-term WRF hindcasts on moisture and precipitation with and without the assimilation of precipitable water vapor (PWV) measurements from Global Positioning System (GPS) ground receivers in the NAM GPS Transect Experiment 2013.

In Chapter 2.1, I investigate the overall performance of daily hindcasts during the 2013 monsoon season without data assimilation. The 2.5-km convection-allowing hindcasts consistently display a moist bias in their initial conditions compared to GPS-PWV observations; this leads to diurnal convection beginning 3-6 hours earlier than observations. Because the precipitation forecast skill varies with the proximity of an inverted trough (IV), I compare the days when an IV is present ("strongly forced") to days when an IV is not present ("weakly forced"). I find that strongly forced days display higher precipitation forecast skill than weakly forced days especially in the slopes of the northern SMO west of the crest that is associated with MCSs. In a case study spanning 8-9 July 2013, when nearly identical MCSs evolved over northern Sonora on consecutive days, the MCS is poorly simulated on the first day (weakly forced) when the IV is east of the SMO while a salient MCS is simulated on the second day (strongly forced) when that IV is over the SMO. I find a greater ensemble-based sensitivity to the initial specification of PWV for the weakly forced day when compared to the strongly forced day. Therefore, GPS-PWV data assimilation has the potential to benefit weakly forced days the most.

In Chapter 2.2, for the weakly forced day (8 July 2013), I explore the impact of ensemble data assimilation of GPS-PWV observations to the model fields and the

hindcast simulation of an MCS that occurred 9-15 hours after forecast initialization. I find that GPS-PWV DA improves location and intensity of the MCS. For all experiments, the GPS-PWV DA reduces the PWV root-mean-square-error to within the GPS-PWV observation error of 1-2 mm at initialization and reduces the initial wet bias. Although there is a short "memory" of these adjustments in that the PWV RMSE across the sites rises quickly and approaches the RMSE of the non-DA experiment after 2 hours, this is due to advective effects near the GOC in that the adjustments move quickly away or toward the sites. Assimilating GPS-PWV observations lowers the moisture (water vapor mixing ratio) error in the lower atmosphere to where is it within the instrument error aboard the radiosonde. From my sensitivity analyses, I conclude that increasing the covariance localization cutoff radius improves the MCS when adjusting all state variables but degrades it when adjusting only thermodynamic variables. Also, I recommend assimilating a mean hourly observation ("superobbing") rather than individual 5-min observations as it allows for more stable adjustments. Lastly, I note that having a 12-h spin-up improves the MCS simulation because the initial conditions have a chance to make their way to the convection-allowing grid before GPS-PWV DA adjusts the MCS towards the observation.

In a region of complex terrain that suffers from unreliable observations and poor convective forecasts during the NAM, I have shown utility from GPS-PWV observations in a) the diagnosis of wet model bias, b) the improvement of the initial conditions via convection-allowing ensemble data assimilation, and c) the improvement of MCS simulation. The results of this dissertation point to a need for more observations in the vertical and a deeper understanding of sensitivities of atmospheric variables to one another, such as what can be gained with with a network of ground-based lidars that continuously monitor the boundary layer.

CHAPTER 1: INTRODUCTION

1.1 BACKGROUND OF THE NORTH AMERICAN MONSOON (NAM)

Beginning in July and lasting through September, the North American monsoon (NAM), with increased moisture and instability, is associated with an influx of convective precipitation to northwest Mexico and the adjacent southwest U.S. Relative to annual precipitation, the NAM provides 60-80% in northwestern Mexico and 40% in the southwest United States (Douglas et al. 1993; Adams and Comrie 1997). Severe weather from the NAM is one of this binational region's principal natural hazards. During "burst" periods when organized convection is more favored (e.g., Carleton 1986, Douglas et al. 1993; Adams and Comrie 1997), mesoscale convective systems (MCSs) are the dominant mechanism for severe weather (McCollum et al. 1995; Lang et al. 2007; Newman and Johnson 2012; Rowe et al. 2012).

1.2 DIURNAL CYCLE OF NAM CONVECTION AND THE MCS

Convective development in the NAM is linked to the diurnal cycle of heating of complex terrain and associated mountain-valley circulations (e.g., Diem and Brown 2003) in both the southwest U.S. (Raymond and Wilkening 1980; Damiani et al. 2008) and northwest Mexico (Gochis et al. 2007; Lang et al. 2007; Nesbitt et al. 2008). Convection begins as air-mass type thunderstorms that form over the Sierra Madre Occidental (SMO), a prominent northwest-southeast-oriented mountain range in northwest Mexico, in the early to mid-afternoon (Janowiak et al. 2007). Then, depending on the synoptic environment, these storms may organize into an MCS that propagates off the SMO in the direction of the upper-level steering flow that is usually westward toward lower elevations and the Gulf of California (GOC). Convective propagation occurs through successive outflow boundaries from the cold pools of leading convective lines that mechanically lift moist and unstable air thus initiating new convection (Corfidi 2003). MCSs can be sustained into the early morning of the following day (Lang et al. 2007; Rotunno et al. 1988) with MCS-related precipitation approaching the GOC about

12 hours later than the time of maximum daytime convection in the SMO to the east (Johnson et al. 2007; Lang et al. 2007; Zuidema et al. 2007; Nesbitt et al. 2008). MCSs account for the majority of the rainfall west of the SMO (Castro et al. 2007; Newman and Johnson 2012) in the NAM core region (a rectangle bounded by 24°-30° N and 112°-106° W) (Higgins et al. 2006).

1.3 INGREDIENTS FOR NAM CONVECTION

In order for NAM convection to occur, atmospheric instability and moisture are necessary (Johnson et al. 2007; Becker and Berbery 2008; Adams and Souza 2009) that are characterized by convective available potential energy (CAPE; Moncrieff and Miller 1976) and precipitable water vapor (PWV; Moore et al. 2015) metrics, respectively. For convection to organize and propagate into an MCS, from a synoptic perspective, some vertical wind shear is also necessary. The presence of a transient upper tropospheric inverted trough (IV; Pytlak et al. 2005) facilitates convective organization (resembling an MCS) by increasing instability and dynamical forcing (Douglas and Englehart 2007; Bieda et al. 2009; Finch and Johnson 2010; Castro et al. 2007; Newman and Johnson 2012; Lahmers et al. 2016) and enhancing mid-level flow and vertical wind shear (Finch and Johnson 2010) during NAM "burst" periods (Carleton 1986; Carleton et al. 1990; Adams and Souza 2009). In fact, easterly vertical wind shear is associated with one of the leading modes of NAM precipitation (Seastrand et al. 2014).

Along with atmospheric instability and vertical wind shear, moisture is also necessary for NAM convection. In the southwest U.S., surges of low-level tropical moisture in the GOC northwestward ("gulf surges"; Douglas and Leal 2003; Rogers and Johnson 2007) are important for the development of convection (Hales 1972; Brenner 1974; Fuller and Stensrud 2000; Higgins et al. 2004), although I find there is low correlation of gulf surges to convection in northwest Mexico because of high values in that locale that are nearly constant during the NAM (Moker et al. 2018). "Major" gulf surges are usually triggered by the passage of a low-pressure disturbance near the GOC mouth, such as a tropical cyclone (TC) or tropical easterly wave (TEW) (Fuller and Stensrud 2000; Douglas and Leal 2003; Higgins and Shi 2005), and traverse the entire

length of the GOC over a period of several days (Zehnder 2004, Serra et al 2016). Confined to the northern GOC, "minor" gulf surges, in contrast, are triggered by the convective outflow boundaries of decaying MCSs and last several hours (Fuller and Stensrud 2000; Lang et al. 2007). "Minor" gulf surges can also be indirectly caused by IVs because of their enhanced support of MCS development (Douglas and Leal 2003; Lang et al. 2007).

1.4 MOTIVATION

With my experience of working at the National Weather Service, my research interest lies on the manifold between research and operations in the world of short-term weather forecasting and analysis. The research conducted in my dissertation is motivated by operational forecast models poorly capturing the diurnal cycle of convection in the NAM as well as MCSs (discussed in Section 1.6) for the following documented reasons: 1) grid-spacing too coarse to depict phenomena related to convection on the meso-γ-scale and 2) the lack of quality moisture observations in the high terrain of the SMO where convection initiates. Kursinski et al. (2008) showed that a 5% decrease or increase in initial PWV significantly changes the amount of convective precipitation over northwest Mexico therefore implying that NAM convective forecasts are sensitive to the initial specification of PWV. Past studies have shown the improvement of short-term forecasts of convection with ensemble methods (e.g., Torn 2010; Ancell et al. 2011; Suarez et al. 2012; Meng and Zhang 2007). With the correction of the initial moisture specification in a convection-allowing forecast model (via ensemble data assimilation) from information from a novel in-situ GPS-PWV observational dataset that addresses the need for high temporal moisture observations in the SMO (discussed in Section 1.5), I seek to improve the accuracy of short-term forecasts of the location and timing of NAM-related convection for early warning of life and property.

1.5 LACK OF RELIABLE OBSERVATIONS IN NORTHWEST MEXICO

The NAM Experiment (NAME) (Higgins et al. 2006) in 2004 advanced our knowledge of convective processes in the NAM core region (see Higgins and Gochis 2007) and provided a valuable dataset for assessing model representation of NAM convection. However, a noted major weakness in the NAME observational network experimental design was the lack of measurements of both the diurnal cycle of the boundary layer and moisture fluxes at higher elevations away from the coast (Higgins and Gochis 2007), a critical region for NAM convective initiation. This lack of observational data strongly motivated the NAM GPS Transect Experiment 2013 (hereafter, "Transect 2013") (Adams et al. 2014; Serra et al. 2016) that included ten GPS meteorological stations in 3 transects that were installed in northwest Mexico. One of the transects were to capture the evolution of PWV during convection initiation in the SMO. It also motivated the North American Monsoon GPS Hydrometeorological Network 2017 under Consortium for Arizona – Mexico Arid Environment (CAZMEX 2017) (Risanto et al. 2019), a field campaign in 2017 with 20 GPS-PWV stations.

1.6 MODEL REPRESENTATION OF NAM CONVECTION

The accuracy of numerical weather prediction (NWP) forecasts during the NAM depends on how well a given model forecast system can represent the physical processes that drive thunderstorm development and the key regional synoptic-scale features that facilitate convective organization. Meso-α-scale (100-1000 km) to meso-β-scale (10-100 km) operational models are inadequate to explicitly represent convection; they generally have a poor representation of the terrain-forced diurnal cycle of convection (e.g., Collier and Zhang 2007; Lee et al. 2007) for example occurring during the NAM as described in analyses of radar data (Lang et al. 2007; Nesbitt et al. 2008). Operational forecast models can easily resolve features at the meso-α-scale therefore representing features important to the NAM like the longwave atmospheric circulation pattern over western North America during the warm season (e.g., monsoon ridge positioning) and IVs. Regional models used at the meso-β-scale tend to overestimate monsoon precipitation in mountainous regions and underestimate precipitation associated with organized propagating convection (Castro et al. 2012; Bukovsky et al. 2013). For example, a

commonly used numerical atmospheric model, the Weather Research and Forecasting (WRF) model, was unable to simulate a MCS during an Intensive Observing Period (IOP) of NAME with a 10-km grid spacing (Cassell et al. 2015), consistent with other studies during NAME that attempt to simulate organized convection using different regional atmospheric models and/or model domain and parameterization configurations (Li et al. 2008; Newman and Johnson 2012).

To explicitly represent (i.e., simulate precipitation directly from the model cloud microphysics scheme without cumulus parameterization) storm-scale structures like squall lines and outflow boundaries (e.g., Li et al. 2008), the use of convection-allowing atmospheric modeling at the meso-γ-scale (i.e., model grid spacing on the order of 1-4 km) is necessary. MCSs cannot be resolved by larger-domain operational forecast models (Gutzler et al. 2009) since they are dependent upon antecedent meso-γ scale meteorological features that would not be resolved by the model. A convection-allowing grid is therefore required to reasonably represent NAM convection including MCSs (e.g., Cassell et al. 2015).

As a result of poorly-understood processes involved with the initiation and growth of deep convection over complex terrain in the NAM core region, both operational and modeling studies using convection-allowing grids have difficulty simulating the timing and subsequent propagation of deep convection (e.g., Li et al. 2008; Castro et al. 2012; Pearson et al. 2014). For example, Janowiak et al. (2007) found that the peak of simulated diurnal precipitation in the NAM maximized 3-6 hours earlier than observations during NAME. I found a similar result of the 3-6-hour precipitation time lag in my first manuscript that characterizes the model performance of NAM convection in the 2013 season without the use of data assimilation (Moker et al. 2018).

1.7 DATA ASSIMILATION RESEARCH TESTBED

In my dissertation, I use the ensemble adjustment Kalman Filter (EAKF; Anderson 2001) within the Data Assimilation Research Testbed (DART; Anderson et al. 2009), which is a community DA software that allows for the application of ensemble algorithms in operational forecasting.

For an observed variable (e.g., PWV), weighted by the error characteristics of the ensemble representation of the observation and the instrument and representativeness of the observation itself, the EAKF first shifts the background (or "prior") ensemble so it becomes the same mean as the analysis (or "posterior") and then the ensemble linearly contracts around the analysis mean resulting in a standard deviation equal to that of the analysis. Once arriving at the posterior values of the model-equivalent observed variables, the increments of each component of the prior state vector are computed from the observation increments via linear regression with error covariance localization.

Anderson (2003) applies a "local least-squares" framework to the EAKF where each ensemble member at each model grid space is updated for each observation. To assimilate an GPS-PWV observation, I follow these steps derived from Anderson (2003) that are available in the EAKF within DART:

1) Start with an ensemble of the background state of the atmosphere (prior).

2) Get an ensemble sample of the GPS-PWV. Since GPS-PWV is not a WRF variable, I must use an observational operator on the state variable for each member i: $PWV_i^f = h(x_i^f)$. The RHS is represented by a bilinear interpolation of the 4 closest model grids and then a summation of the water vapor (from the water vapor mixing ratio) in the column.

3) Get the observation from the instrument PWV^o and the observation error distribution $(\sigma^o)^2$.

4) Inflate the prior variances $(\sigma^f)^2$ to maintain model spread because of the limited sample size and model error (Hamill et al. 2001; Anderson 2001; Whitaker and Hamill 2002). This can be done multiplicative (Anderson 2001) or additive (Hamill and Whitaker 2005), but the use of a relaxation can limit excessive ensemble spread in data-sparse regions caused by the previous two techniques (Zhang et al. 2004). In my dissertation, I use a newer technique that uses adaptive inflation that evolves in time and varies in space (Anderson 2009).

5) Calculate the analysis ensemble mean of the model-equivalent GPS-PWV observation ($\overline{PWV^a}$) (eq. 1) and then for each ensemble member i compute the analysis model-equivalent GPS-PWV observation (PWV_i^a) (eq. 2) and observation increment ΔPWV_i (eq. 3).

$$\overline{PWV^a} = \left(\frac{1}{\frac{1}{(\sigma^f)^2} + \frac{1}{(\sigma^o)^2}}\right)\left(\frac{\overline{PWV^f}}{(\sigma^f)^2} + \frac{PWV^o}{(\sigma^o)^2}\right) \qquad (1)$$

$$PWV_i^a = \sqrt{\frac{(\sigma^o)^2}{(\sigma^f)^2 + (\sigma^o)^2}}\left(PWV_i^f - \overline{PWV^f}\right) + \overline{PWV^a} \qquad (2)$$

$$\Delta PWV_i = PWV_i^a - PWV_i^f \qquad (3)$$

6) Linearly regress the observation increments ΔPWV_i onto state variable increments Δx_i for each ensemble member (eq. 4) and add to the prior state ensemble sample x_i^f to get the analysis value for the state variable x_i^a (eq. 5). The state variable can be water mixing ratio or U wind, for example.

$$\Delta x_i = \frac{cov\left(x_i^f, PWV_i^f\right)}{var\left(PWV_i^f\right)}\Delta PWV_i \qquad (4)$$

$$x_i^a = x_i^f + \Delta x_i \qquad (5)$$

7) Continue until all observations are processed within a time window and state variables are updated to the analysis. In my case, I use GPS-PWV observations within 30 minutes of the observation time. Then, advance the ensemble members to the next time observations are available via the non-linear forward model (WRF-ARW).

To limit filter divergence from spurious correlations far away from the observation (Houtekamer and Mitchell 2001; Hamill et al. 2001) which could render the EAKF useless, covariance localization must be applied. I use the Gaspari-Cohn (Gaspari and Cohn 1999) 3-D covariance localization function, which is based on a 5th order polynomial based on a cutoff distance relative to the GPS-PWV observation site location. Appendix C expands on this ensemble DA algorithm.

1.8 OBJECTIVES

I conducted two studies that make up the main structure of my dissertation. In the first study, I assess the baseline performance of the hindcasts (retrospective forecasts) from downscaled operational models, 32-km North American Mesoscale Model and the

0.25-degree Global Forecast System (GFS), to a convection-allowing grid during Transect 2013. This study is attached as Appendix A and is summarized in Chapter 2.1 with the following objectives:

1) Determine the current initialization error and handling of the representation of NAM convection (e.g., diurnal cycle and MCSs) in the short-term from operational forecast models downscaled onto a convection-allowing grid (WRF-ARW).

2) Determine the role of NAM-modulating features such as gulf surges, TEWs, and IVs in convective precipitation in the southwest US versus northwest Mexico.

3) Determine the difference in forecast performance of convection on days when there is a transient upper-level feature (i.e., an IV) present that enhances upward vertical motion, instability, and wind shear to days when there is none.

4) Determine the sensitivity of initial PWV to forecast moisture and precipitation.

5) Determine the most appropriate satellite-derived precipitation dataset that is closest to "truth" in verifying NAM precipitation in complex terrain.

In the second study, I investigate the utility of the assimilation of GPS-PWV observations into the initial fields and forecasts of a "weakly forced" day whose forecast rainfall and moisture fields were shown to have a high sensitivity to the initial specification of PWV in the western SMO foothills. This study is attached as Appendix B and summarized in Chapter 2.2 with the following objectives:

1) Develop an algorithm to assimilate GPS-PWV observations from Transect 2013 into the WRF-ARW initial model fields of short-range convective simulations.

2) Determine the impact of GPS-PWV observations on the initial model fields and forecasts especially of the development of an MCS several hours after initialization.

3) From an engineering perspective and physical considerations, investigate the impact of increasing covariance localization radius, localizing the type of WRF state variable to adjust in the algorithm, spinning up the model, doubling the number of ensemble members, and assimilating a mean observation (i.e., "super-obbing") instead of individually.

CHAPTER 2: PRESENT STUDY

2.1 CONVECTIVE-PERMITTING HINDCAST SIMULATIONS DURING THE NORTH AMERICAN MONSOON GPS TRANSECT EXPERIMENT 2013: ESTABLISHING BASELINE MODEL PERFORMANCE WITHOUT DATA ASSIMILATION

In this study, in days coinciding with the deployment of the GPS-PWV stations in northwest Mexico during Transect 2013 (from 26 Jun through 12 Sep 2013), I run 24-h convection-allowing WRF-ARW hindcasts (retrospective forecasts) in two deterministic simulations, WRF-GFS and WRF-NAM, which are downscaled from the operational Global Forecast System model and North American Mesoscale model, respectively. I initialize the runs at 12z with periodic updates to the lateral boundary conditions every 6 h. The Rapid Refresh (RAP) model is used to initialize the soil moisture and temperature because it has a finer spatial resolution compared to the other operational models.

I find a consistent moist bias in the initial specification of PWV relative to the GPS-PWV observations with WRF-NAM having the higher bias. The high moisture bias in the initialization of the simulations leads to relatively high moisture biases in the western slopes of the northern SMO prior to 0000 UTC leading to convection beginning along the SMO 3-6 hours early in the diurnal cycle relative to TRMM precipitation. This time shift error in representing the diurnal cycle of convection is consistent with previous studies. The high PWV biases decrease with time within the diurnal cycle.

Gulf surges did not have noticeable impact the development of MCSs and related convection in northwest Mexico. No TEWs or TCs directly impacted the NAM.

I classify days based on the presence of an IV; days when an IV located in the NAM core region is "strongly forced" while all other days with appreciable rainfall are "weakly forced". I find that strongly forced days display higher modeled precipitation forecast skill than weakly forced days on the slopes of the northern SMO away from the crest especially towards the west.

A case study spanning 8-9 July 2013 illustrates consecutive days when nearly identical MCSs evolve over northern Sonora. The MCS is poorly simulated on the weakly forced day (8 July 2013) when an IV is east of the SMO (near Big Bend, TX), but

a salient MCS is simulated on the strongly forced day (9 July 2013) when an IV is further west over the central SMO in the core NAM region.

There is greater ensemble-based sensitivity of the initial specification of PWV to forecast moisture and precipitation in the weakly forced day relative to the strongly forced day. For weakly-forced days, one could conclude that a more accurate specification of local moisture conditions from the assimilation of GPS-PWV observations would have a greater impact on the model simulation of organized MCS-type convection.

2.2 UTILITY OF GPS-DERIVED PWV MEASUREMENTS IN CONSTRAINING HIGH-RESOLUTION WRF FORECASTS OVER THE NORTH AMERICAN MONSOON REGION

In this study, I investigate the utility of assimilating GPS-PWV observations from the Transect 2013 sites into WRF-ARW and its impact on simulating the MCS that reached peak intensity at 2300-0200 LT on 8-9 Jul 2013. This was the weakly forced day in the case study in Chapter 2.1 (Moker et al. 2018) that was shown high sensitivity of forecast rainfall and moisture to initial PWV specification in the western SMO foothills. I use the ensemble adjustment Kalman filter (EAKF) in the Data Assimilation Research Testbed (DART) software with a 6-hour spin-up, 6.5 1-hour DA cycles and 20 ensemble members. I used a Gaspari-Cohn covariance localization cutoff of 0.07 radians for each observation, which equates to a horizontal-equivalent distance of ~450 km and vertical-equivalent distance of 3.5 km. With this cutoff, the mean function value of 0.5 across all sites produces a contour with an area of ~150,000 km^2, a magnitude that is the similar to the cloud shield of an MCS at maximum intensity (Maddox 1980). The correlation of PWV to moisture in the vertical from soundings in the NAM region have the highest values ~3.5 km from the surface.

A 24-h deterministic forecast is initialized from the ensemble-mean analysis in what is referred to as the *assim* experiment. In addition to the *assim* experiment, the other control experiments are a deterministic forecast initialized at 12z (*cold_start* experiment*)* and a 12-h ensemble spin-up with a deterministic forecast initialized by the ensemble

mean (*no_assim* experiment). From an engineering standpoint, I also investigate the sensitivity of different localizations lengths and adjusted variable localization among other impacts on the initial condition and simulated MCS 9-15 hours after forecast initialization time.

Assimilating GPS-PWV observations improves the simulation of the MCS. I use the definition from Maddox (1980) of an MCS that has the following criteria:

1) A cloud shield colder than -32 °C that has an area of at least 100,000 km^2.

2) An interior cloud shield colder than -52 °C that has an area of at least 50,000 km^2.

3) The cloud shield has an eccentricity of at least 0.7 at the time of maximum extent.

4) The size conditions in 1 and 2 must persist for at least 6 hours.

Because I found that the simulated MCSs were much warmer than the observed, we evaluate hindcast performance by using the percentage of area < -32 °C of the simulated cloud shield that matches the < -32 °C region of the observed MCS cloud shield at the times the MCS criteria above are met. I use the model state at 6z Jul 9 to assess the MCS as it is the time when the model MCS is at the end of its peak intensity and the observed MCS is at the beginning of its peak intensity. First, compared to a hindcast beginning at 0500 LT (*cold_start* experiment) (34% during observed maximum intensity [at, e.g., 2300 LT]), there is improved simulation of the MCS when the hindcast initializes 12 h earlier (42% for ensemble spin-up [*no_assim* experiment] and 37% for deterministic spin-up [*warm_start* experiment]) because, as a first order, the meso-γ-scale flow has time to establish in the convection-allowing grid. Then, the assimilation of GPS-PWV observations (*assim* experiment) nudges the location and broadens the coverage of the MCS towards the observation for further improvement of the simulated MCS (50%).

The DA algorithm reduces the PWV RMSE from 3 mm at the beginning of the DA cycling period to 0.25-1 mm at the final analysis for all experiments across the GPS-PWV sites. This falls within the GPS observation error of 1-2 mm. There is a short memory for these adjustments in that the PWV RMSE rises quickly becoming in line with the non-DA experiments by hour 2. Advection of the moisture adjustments,

20

specifically near the GOC, can explain the short "memory". The adjustments indeed have an impact on MCS formation 6+ hours later.

CHIH was the only location that had an upper air site collocated with a GPS-PWV site. Assimilating GPS-PWV reduces moisture error at and below ~700 hPa to within the instrument error on the radiosonde, but the information contained in ensemble-based covariances were not able to recover the winds from statistical connections to the change in PWV. At MAZT, the adjustments to the atmospheric moisture profile were negligible as the closest GPS-PWV location was ~150 km away.

Increasing the covariance localization cutoff when adjusting all variables (*all_vars* experiments) improves the simulated MCS (from 31% for 0.03 radians to ~50% for 0.07 and 0.10 radians). Increasing the covariance localization cutoff when adjusting only thermodynamic variables (*no_winds* experiments) slightly improves the simulated MCS from 0.03 radians (36%) to 0.07 radians (42%) but degrades the simulated MCS at 0.10 radians (25%). Assimilating a mean hourly observation (*super_ob* experiment) instead of individual 5-min observation allows for smaller increments and more stability during the DA cycles. The resulting MCS is 42% matched which is similar to the *no_assim* experiment.

This study has shown the need for vertical information as it is a source of uncertainty when assimilating an integrated quantity such as GPS-PWV. A field campaign with a network of lidars that can continuously observe atmospheric variables a few km from the ground can aid in assessing statistical connections of quantities of moisture and winds in three dimensions.

REFERENCES

Adams, D. K., and A. C. Comrie, 1997: The North American Monsoon. *Bull. Am. Meteorol. Soc.*, **78**, 2197–2213, doi:10.1175/1520-0477(1997)078<2197:TNAM>2.0.CO;2.

Adams, D. K. and E. P. Souza, 2009: CAPE and convective events over the southwest U.S. during the North American monsoon. *Monthly Weather Review*, **137**, 83–98.

Adams, D. K., C. Minjarez, Y. Serra, A. Quintanar, L. Alatorre, A. Granados, E. Vázquez, and J. Braun, 2014: Mexican GPS tracks convection from North American monsoon. Eos, Trans. *Amer. Geophys. Union*, 95, **61**, doi:10.1002/2014EO070001.

Ancell, B. C., C. F. Mass, and G. J. Hakim, 2011: Evaluation of Surface Analyses and Forecasts with a Multiscale Ensemble Kalman Filter in Regions of Complex Terrain. *Mon. Weather Rev.*, **139**, 2008–2024, doi:10.1175/2010mwr3612.1.

Anderson, J., T. Hoar, K. Raeder, H. Liu, N. Collins, R. Torn, and A. Arellano, 2009: The data assimilation research testbed a community facility. *Bull. Am. Meteorol. Soc.*, **90**, 1283–1296, doi:10.1175/2009BAMS2618.1.

Anderson, J., 2001: An Ensemble Adjustment Kalman Filter for Data Assimilation. *Mon. Weather Rev.*, **129**, 2884–2903, doi:10.1175/1520-0493(2001)129<2884:AEAKFF>2.0.CO;2. http://journals.ametsoc.org/doi/abs/10.1175/1520-0493%282001%29129%3C2884%3AAEAKFF%3E2.0.CO%3B2.

Anderson, J., 2003: A Local Least Squares Framework for Ensemble Filtering. *Mon. Weather Rev.*, **131**, 634–642, doi:10.1175/1520-0493(2003)131<0634:ALLSFF>2.0.CO;2. http://journals.ametsoc.org/doi/abs/10.1175/1520-0493%282003%29131%3C0634%3AALLSFF%3E2.0.CO%3B2.

Anderson, J., 2009: Spatially and temporally varying adaptive covariance inflation for ensemble filters. *Tellus, Ser. A Dyn. Meteorol. Oceanogr.*, **61 A**, 72–83, doi:10.1111/j.1600-0870.2008.00361.x.

Becker, E. J., and E. H. Berbery, 2008: The Diurnal Cycle of Precipitation over the North American Monsoon Region during the NAME 2004 Field Campaign. *J. Climate*, **21**,

Bieda, S. W., C. L. Castro, S. L. Mullen, A. C. Comrie, and E. Pytlak, 2009: The relationship of transient upper-level troughs to variability of the North American monsoon system. *J. Clim.*, **22**, 4213–4227, doi:10.1175/2009JCLI2487.1.

Brenner, I. S., 1974: A Surge of Maritime Tropical Air-Gulf of California to the Southwestern United States. *Mon. Wea. Rev.*, **102**(5), 375–389, doi:10.1175/1520-0493(1974)102<0375:ASOMTA>2.0.CO;2.

Bukovsky, M., D. J. Gochis, and L. O. Mearns, 2013: Towards Assessing NARCCAP Regional Climate Model Credibility for the North American Monsoon: Current Climate Simulations. J. Climate, 26, 8802–8826.

Carleton, A. M., 1986: Synoptic-dynamic character of "bursts" and "breaks" in the southwest U.S. summer precipitation singularity. *J. Climatol.*, **6**, 605–623.

Carleton, A. M., D. A. Carpenter, and P. J. Weber, 1990: Mechanisms of interannual variability of the southwest United States summer rainfall maximum. *J. Climate*, **3**, 999–1015.

Cassell W. W., C. L. Castro, T. M. Luong, and Q. Xiao, 2015: Simulating organized convection during the 2004 North American Monsoon Experiment and its sensitivity to the specification of initial conditions. Submitted to Monthly Weather Review.

Castro, C. L., R. A. Pielke Sr., and J. O. Adegoke, 2007: Investigation of the summer climate of the contiguous United States and Mexico using the Regional Atmospheric Modeling System (RAMS). Part I: Model climatology (1950–2002). *J. Climate*, **20,** 89–110.

Castro, C. L., H. I. Chang, F. Dominguez, C. Carrillo, J. K. Schemm, and H. M. H. Juang, 2012: Can a regional climate model improve the ability to forecast the North American monsoon? *J. Clim.*, **25**, 8212–8237, doi:10.1175/JCLI-D-11-00441.1.

Collier, J. C., and G. J. Zhang, 2007: Effects of increased horizontal resolution on simulation of the North American monsoon in the NCAR CAM3: An evaluation based on surface, satellite, and reanalysis data. *J Climate*, **20**, 1843–1861, doi:10.1175/JCL14099.1.

Corfidi S. F., 2003: Cold Pools and MCS Propagation: Forecasting the Motion of Downwind-Developing MCSs. Wea. Forecasting, 18, 997–1017.

Damiani, R., and Coauthors, 2008: The cumulus, photogrammetric, in situ, and doppler observations experiment of 2006. *Bull. Amer. Meteor. Soc.*, **89,** 57–73.

Diem J.E., and D.P. Brown, 2003: Anthropogenic impacts on summer precipitation in central Arizona, U.S.A., The Professional Geographer, 55, 343-355.

Douglas, M. W., R. A. Maddox, K. Howard, and S. Reyes, 1993: The Mexican monsoon. *J. Clim.*, **6**, 1665–1677, doi:10.1175/1520-0442(1993)006<1665:TMM>2.0.CO;2.

Douglas, M. W., and J. C. Leal, 2003: Summertime Surges over the Gulf of California: Aspects of Their Climatology, Mean Structure, and Evolution from Radiosonde, NCEP Reanalysis, and Rainfall Data. *Wea. and Forecasting*, **18**, 55–74, doi:10.1175/1520-0434(2003)018<0055:SSOTGO>2.0.CO;2.

Douglas, A. V., and P. J. Englehart, 2007: A climatological perspective of transient synoptic features during NAME 2004. *J. Climate*, **20**, 1947-1954.

Finch, Z. O., and R. H. Johnson, 2010: Observational Analysis of an Upper-Level Inverted Trough during the North American Monsoon Experiment. *Mon. Wea. Rev.*, **138**, 3540-3555.

Fuller, R. D., and D. J. Stensrud, 2000: The relationship between tropical easterly waves and surges over the Gulf of California during the North American Monsoon. *Mon. Wea. Rev.*, **128,** 2983–2989.

Gaspari, G., and S. E. Cohn, 1999: Construction of correlation functions in two and three dimensions. *Q. J. R. Meteorol. Soc.*, **125**, 723–757, doi:10.1256/smsqj.55416.

Gochis, D. J., C. J. Watts, J. Garatuza-Payan, and J. Cesar-Rodriguez, 2007: Spatial and temporal patterns of precipitation intensity as observed by the NAME Event Rain gauge Network from 2002 to 2004. *J. Clim.*, **20**, 1734–1750, doi:10.1175/JCLI4092.1.

Gutzler, D.S., L.N. Long, J. Schemm, S.B. Roy, M. Bosilovich, J.C. Collier, M. Kanamitsu, P. Kelly, D. Lawrence, M. Lee, R.L. Sánchez, B. Mapes, K. Mo, A. Nunes, E.A. Ritchie, J. Roads, S. Schubert, H. Wei, and G.J. Zhang, 2009: Simulations of the 2004 North American Monsoon: NAMAP2. J. Climate, 22, 6716–6740, https://doi.org/10.1175/2009JCLI3138.1

Hales, J., 1972: Surges of Maritime Tropical Air Northward Over Gulf of California. *Mon. Wea. Rev.*, **100**, 298–306, doi: 10.1175/1520-0493(1972)100<0298:SOMTAN>2.3.CO;2.

Hamill, T. M., and J. S. Whitaker, 2005: Accounting for the Error due to Unresolved Scales in Ensemble Data Assimilation: A Comparison of Different Approaches. *Mon. Weather Rev.*, **133**, 3132–3147, doi:10.1175/MWR3020.1. https://doi.org/10.1175/MWR3020.1.

Hamill, T. M., J. S. Whitaker, and C. Snyder, 2001: Distance-Dependent Filtering of Background Error Covariance Estimates in an Ensemble Kalman Filter. *Mon. Weather Rev.*, **129**, 2276–2790, doi:10.1080/00365519009091567.

Higgins, R.W. and W. Shi, 2005: Relationships between Gulf of California Moisture Surges and Tropical Cyclones in the Eastern Pacific Basin. J. Climate, 18, 4601–4620, https://doi.org/10.1175/JCLI3551.1

Higgins, R., W. Shi, and C. Hain, 2004: Relationships between Gulf of California moisture surges and precipitation in the southwestern United States. *J. Climate*, **17**, 2983–2997.

Higgins, W., and D. Gochis, 2007: Synthesis of Results from the North American Monsoon Experiment (NAME) Process Study. *J. Clim.*, **20**, 1601–1607, doi:10.1175/JCLI4081.1.

Higgins, W., and Coauthors, 2006: The NAME 2004 field campaign and modeling strategy. *Bull. Am. Meteorol. Soc.*, **87**, 79–94, doi:10.1175/BAMS-87-1-79.

Houtekamer, P. L., and H. L. Mitchell, 2001: A Sequential Ensemble Kalman Filter for Atmospheric Data Assimilation. *Mon. Weather Rev.*, **129**, 123–137, doi:10.1175/1520-0493(2001)129<0123:ASEKFF>2.0.CO;2. http://journals.ametsoc.org/doi/abs/10.1175/1520-0493%282001%29129%3C0123%3AASEKFF%3E2.0.CO%3B2.

Janowiak, J. E., V. J. Dagostaro, V. E. Kousky, and R. J. Joyce, 2007: An examination of precipitation in observations and model forecasts during NAME with emphasis on the diurnal cycle. *J. Clim.*, **20**, 1680–1692, doi:10.1175/JCLI4084.1.

Johnson, R. H., P. E. Ciesielski, B. D. McNoldy, P. J. Rogers, and R. K. Taft, 2007: Multiscales variability of the flow during the North American Monsoon experiment. *J. Clim.*, **20**, 1628–1648, doi:10.1175/JCLI4087.1.

Kursinski, E. R., D. K. Adams, and M. Leuthold, 2008: GPS observations of precipitable water and implications for the predictability of precipitation during the North American monsoon. *CLIVAR Exchanges*, No. 45, International CLIVAR Project Office, Southampton, United Kingdom, 14, 19–21.

Lahmers, T. M., C. L. Castro, D. K. Adams, Y. L. Serra, J. J. Brost, and T. Luong, 2016: Long-term changes in the climatology of transient inverted troughs over the North American monsoon region and their effects on precipitation. *J. Climate*, *29*, 6027-6064.

Lang, T. J., D. A. Ahijevych, S. W. Nesbitt, R. E. Carbone, S. A. Rutledge, and R. Cifelli, 2007: Radar-observed characteristics of precipitating systems during NAME 2004. *J. Clim.*, **20**, 1713–1733, doi:10.1175/JCLI4082.1.

Lee, M.-I., and Coauthors, 2007: Sensitivity to horizontal resolution in the AGCM simulations of warm season diurnal cycle of precipitation over the United States and northern Mexico. *J Climate*, **20**, 1862–1881, doi:10.1175/JCL14090.1.

Li, J., S. Sorooshian, W. Higgins, X. Gao, B. Imam, and K. Hsu, 2008: Influence of spatial resolution on diurnal variability during the north American monsoon. *J. Clim.*, **21**, 3967–3988, doi:10.1175/2008JCLI2022.1.

Maddox, R. A., 1980: Mesoscale convective complexes. *Bull. Amer. Meteor. Soc.*, **61**, 1374–1387, doi:https://doi.org/10.1175/1520-0477(1980)061<1374:MCC>2.0.CO;2.

McCollum, D., R. Maddox, and K. Howard, 1995: Case study of a severe mesoscale convective system in central Arizona. *Wea. Forecasting*, **10**, 643–665.

Meng, Z., and F. Zhang, 2007: Tests of an Ensemble Kalman Filter for Mesoscale and Regional-Scale Data Assimilation. Part II: Imperfect Model Experiments. *Mon. Weather Rev.*, **135**, 1403–1423, doi:10.1175/MWR3352.1. http://journals.ametsoc.org/doi/abs/10.1175/MWR3352.1.

Moker, J. M., C. L. Castro, A. F. Arellano, Y. L. Serra, and D. K. Adams, 2018: Convective-permitting hindcast simulations during the North American Monsoon GPS Transect Experiment 2013: Establishing baseline model performance without data assimilation. *J. Appl. Meteorol. Climatol.*, **57**, 1683–1710, doi:10.1175/JAMC-D-17-0136.1.

Moker, J. M., A. F. Arellano, C. L. Castro, Y. L. Serra, D. K. Adams, and C. B. Risanto, 2019: Utility of GPS-derived PWV Measurements in Constraining High-Resoluation WRF Forecasts Over the North American Monsoon Region. *In preparation.*

Moore, A., I. Small, S. Gutman, Y. Bock, J. Dumas, P. Fang, J. Haase, M. Jackson, and J. Laber, 2015: National Weather Service Forecasters Use GPS Precipitable Water Vapor for Enhanced Situational Awareness during the Southern California Summer Monsoon. *Bull. Amer. Meteor. Soc.,* **96**, 1867–1877, doi:10.1175/BAMS-D-14-00095.1.

Moncrieff, M. W., and M. J. Miller, 1976: The dynamics and simulation of tropical cumulonimbus and squall lines. *Q. J. R. Meteorol. Soc.*, **102**, 373–394, doi:10.1002/qj.49710243208.

Nesbitt, S. W., D. J. Gochis, and T. J. Lang, 2008: The Diurnal Cycle of Clouds and Precipitation along the Sierra Madre Occidental Observed during NAME-2004: Implications for Warm Season Precipitation Estimation in Complex Terrain. *J. Hydrometeorol.*, **9**, 728–743, doi:10.1175/2008JHM939.1. http://journals.ametsoc.org/doi/abs/10.1175/2008JHM939.1.

Newman, A., and R. H. Johnson, 2012: Mechanisms for Precipitation Enhancement in a North American Monsoon Upper-Tropospheric Trough. *J. Atmos. Sci.*, **69**, 1775–

1792, doi:10.1175/JAS-D-11-0223.1. http://journals.ametsoc.org/doi/abs/10.1175/JAS-D-11-0223.1.

Pearson, K. J., G. M. S. Lister, C. E. Birch, R. P. Allan, R. J. Hogan, and S. J. Woolnough, 2014: Modelling the diurnal cycle of tropical convection across the "grey zone." *Q. J. R. Meteorol. Soc.*, **140**, 491–499, doi:10.1002/qj.2145.

Pytlak, E., M. Goering, and A. Bennett, 2005: Upper tropospheric troughs and their interaction with the North American monsoon. *19th Conf. on Hydrology*, San Diego, CA, Amer. Meteor. Soc., P2.3, https://ams.confex.com/ams/pdfpapers/85393.pdf.

Raymond, D. J., and M. H. Wilkening, 1980: Mountain induced convection under fair weather conditions. *J. Atmos. Sci.*, **37,** 2693–2706.

Rogers, P. J., and R. H. Johnson, 2007: Analysis of the 13–14 July Gulf Surge Event during the 2004 North American Monsoon Experiment. *Mon. Wea. Rev.*, **135**, 3098–3117, doi:10.1175/MWR3450.1.

Risanto, C. B., C. L. Castro, J. M. Moker, A F. Arellano, D. K. Adams, L. M. Fierro, C. M. Minjarez-Sosa, 2019: Evaluating Forecast Skill of Moisture from Convective-Permitting WRF-ARW Model during 2017 North American Monsoon Season. *In preparation.*

Rotunno, R., J. B. Klemp, and M. L. Weisman, 1988: A theory for strong, long-lived squall lines. J. Atmos. Sci., 45, 463–485.

Rowe, A. K., S. A. Rutledge, and T. J. Lang, 2012: Investigation of Microphysical Processes Occurring in Organized Convection during NAME. *Mon. Wea. Rev.*, **140**, 2168–2187, doi:10.1175/MWR-D-11-00124.1.

Seastrand, S., Y. Serra, C. Castro, and E. Ritchie, 2015: The dominant synoptic-scale modes of North American monsoon precipitation. *Int. J. Climatol.*, **35**, 2019–2032, doi:10.1002/joc.4104. http://doi.wiley.com/10.1002/joc.4104.

Serra, Y. L., D. K. Adams, C. Minjarez-Sosa, J. M. Moker, A. F. Arellano, C. L. Castro, A. I. Quintanar, L. Alatorre, A. Granados, G. Vazquez, K. Holub, and C. C. DeMets, 2016, 2016: The north American monsoon GPS transect experiment 2013. *Bull. Am. Meteorol. Soc.*, **97**, 2103–2115, doi:10.1175/BAMS-D-14-00250.1.

Suarez, A., H. D. Reeves, D. Wheatley, and M. Coniglio, 2012: Comparison of Ensemble Kalman Filter–Based Forecasts to Traditional Ensemble and Deterministic Forecasts for a Case Study of Banded Snow. *Weather Forecast.*, **27**, 85–105, doi:10.1175/waf-d-11-00030.1.

Torn, R. D., 2010: Performance of a Mesoscale Ensemble Kalman Filter (EnKF) during the NOAA High-Resolution Hurricane Test. *Mon. Weather Rev.*, **138**, 4375–4392, doi:10.1175/2010mwr3361.1.

Whitaker, J. S., and T. M. Hamill, 2002: Ensemble Data Assimilation without Perturbed Observations. *Mon. Weather Rev.*, **130**, 1913–1924, doi:10.1175/1520-0493(2002)130<1913:EDAWPO>2.0.CO;2. http://journals.ametsoc.org/doi/abs/10.1175/1520-0493%282002%29130%3C1913%3AEDAWPO%3E2.0.CO%3B2.

Zehnder, J. A., 2004: Dynamic mechanisms of the gulf surge. *J Geophys Res-Atmos*, **109**, doi:10.1029/2004JD004616.

Zhang, F., C. Snyder, and J. Sun, 2004: Impacts of Initial Estimate and Observation Availability on Convective-Scale Data Assimilation with an Ensemble Kalman Filter. *Mon. Weather Rev.*, **132**, 1238–1253, doi:10.1175/1520-0493(2004)132<1238:IOIEAO>2.0.CO;2. https://doi.org/10.1175/1520-0493(2004)132%3C1238:IOIEAO%3E2.0.CO.

Zuidema, P., C. Fairall, L. M. Hartten, J. E. Hare, and D. Wolfe, 2007: On the air–sea interaction at the mouth of the Gulf of California. *J. Climate*, **20,** 1649–1661.

APPENDIX A: CONVECTIVE-PERMITTING HINDCAST SIMULATIONS DURING THE NORTH AMERICAN MONSOON GPS TRANSECT EXPERIMENT 2013: ESTABLISHING BASELINE MODEL PERFORMANCE WITHOUT DATA ASSIMILATION

(Published in *Journal of Applied Meteorology and Climatology*)

James M. Moker Jr.[1], Christopher L. Castro[1], Yolande L. Serra[2], Avelino Arellano Jr.[1], and David K. Adams[3]

[1] *Department of Hydrology and Atmospheric Sciences, University of Arizona, Tucson, Arizona*
[2] *Joint Institute for Study of Atmosphere and Ocean, University of Washington, Seattle, Washington*
[3] *Centro de Ciencias de la Atmósfera, Universidad Nacional Autónoma de México (UNAM), México, D.F.*

Corresponding author address: James M. Moker Jr., Department of Hydrology and Atmospheric Sciences, Harshbarger Building, Room 214A, 133 E. James E. Rogers Way, Tucson, AZ 85721-0081
E-mail: jmoker@email.arizona.edu

Abstract

During the North American monsoon global positioning system (GPS) Transect Experiment 2013, daily convective-permitting WRF simulations are performed in northwestern Mexico and the southern Arizona border region using the operational Global Forecast System (GFS) and North American Mesoscale Forecast System (NAM) models as lateral boundary forcing and initial conditions. Compared to GPS precipitable water vapor (PWV), the WRF simulations display a consistent moist bias in the initial specification of PWV leading to convection beginning 3–6 h early. Given appreciable observed rainfall, days are classified as strongly and weakly forced based only on the presence of an inverted trough (IV); gulf surges did not noticeably impact the development of mesoscale convective systems (MCSs) and related convection in northwestern Mexico. Strongly forced days display higher modeled precipitation forecast skill than weakly forced days in the slopes of the northern Sierra Madre Occidental (SMO) away from the crest, especially toward the west where MCSs account for the greatest proportion of all monsoon-related precipitation. A case study spanning 8–10 July 2013 illustrates two consecutive days when nearly identical MCSs evolved over northern Sonora. Although a salient MCS is simulated on the strongly forced day (9–10 July 2013) when an IV is approaching the core monsoon region, a simulated MCS is basically nonexistent on the weakly forced day (8–9 July 2013) when the IV is farther away. The greater sensitivity to the initial specification of PWV in the weakly forced day suggests that assimilation of GPS-derived PWV for these types of days may be of greatest value in improving model precipitation forecasts.

1. Introduction

Severe thunderstorms are one of the principal natural hazards in the southwestern United States and adjacent northwestern Mexico, occurring predominantly from July to mid-September and especially during "burst" periods when organized convection is more favored (e.g., Carleton 1986; Douglas et al. 1993; Adams and Comrie 1997). Monsoon thunderstorms can cause damage and hazards from blowing dust from strong outflows, flash flooding from torrential rainfall, and power outages from lightning strikes (McCollum et al. 1995; Gochis et al. 2007; Magirl et al. 2007; Griffiths et al. 2009). With respect to annual precipitation, convective precipitation related to the North American monsoon accounts for 60%–80% in northwestern Mexico and 40% in southwest United States (Douglas et al. 1993). The ability to quickly and accurately forecast the location and timing of monsoon thunderstorms is critical for the timely issuance of official watches and warnings in this binational region.

The accuracy of numerical weather prediction (NWP; acronyms used in this paper are listed in appendix B) forecasts during the monsoon depends on how well a given model forecast system can deterministically represent thunderstorm development and the key regional synoptic-scale features that facilitate convective organization. Thunderstorm development is strongly tied to the diurnal cycle of convection over complex terrain in both the southwestern United States (Raymond and Wilkening 1980; Damiani et al. 2008) and northwestern Mexico (Gochis et al. 2007; Lang et al. 2007; Nesbitt et al. 2008) with mesoscale convective systems (MCSs) being the dominant mechanism for severe weather (McCollum et al. 1995; Lang et al. 2007; Newman and Johnson 2012; Rowe et al. 2012). Precipitation from MCSs approaches the Gulf of California (GoC) about 12 h

later than the time of maximum diurnal convection in the SMO to the east (Johnson et al. 2007; Lang et al. 2007; Zuidema et al. 2007; Nesbitt et al. 2008). In the North American monsoon core region, defined as the region bounded by 24°–30°N and 112°–106°W (Higgins et al. 2006), MCSs account for the majority of the rainfall that falls west of the SMO (Castro et al. 2007; Newman and Johnson 2012).

A necessary condition for the development of monsoon convection is a favorable thermodynamic environment in terms of atmospheric instability and moisture (Johnson et al. 2007; Becker and Berbery 2008; Adams and Souza 2009). These criteria have been traditionally characterized by convective available potential energy (CAPE; Moncrieff and Miller 1976) and precipitable water vapor (PWV; Moore et al. 2015) metrics. However, during monsoon "burst" periods (Carleton 1986; Carleton et al. 1990; Adams and Souza 2009), transient upper-tropospheric inverted troughs (IVs; Pytlak et al. 2005) facilitate convective organization, likely through both increasing instability and dynamical forcing (Douglas and Englehart 2007; Bieda et al. 2009; Finch and Johnson 2010). For example, Finch and Johnson (2010) argue that IVs modulate and intensify midlevel flow and shear leading to favorable conditions for convective organization in northwestern Mexico. From a climatological perspective, easterly vertical wind shear is also associated with one of the leading modes of monsoon precipitation (Seastrand et al. 2015). Climatologically, IVs are associated with an increase in organized propagating (MCS-like) convection in the North American monsoon core region (Bieda et al. 2009; Lahmers et al. 2016). Assessing exactly how the presence or absence of IVs impacts model forecast skill is a major objective of this study.

Surges of low-level tropical moisture (generally below 600 hPa) up the GoC

(Douglas and Leal 2003; Rogers and Johnson 2007), hereinafter referred to as "gulf surges," are also important for the development of convection in the southwest United States (Hales 1972; Brenner 1974; Fuller and Stensrud 2000; Higgins et al. 2004), although our study brings into question their relevance for low-level moisture south of this region in northwest Mexico. "Major" gulf surges are often triggered by the passage of a low-pressure disturbance near the mouth of the gulf such as a tropical cyclone (TC) or tropical easterly wave (TEW; Fuller and Stensrud 2000; Douglas and Leal 2003; Higgins and Shi 2005), and traverse the entire length of the gulf over a period of several days (Zehnder 2004; Serra et al. 2016). TCs can also make direct landfall into northwest Mexico and the southwest United States, although this typically occurs in October at the end of the monsoon season (Wood and Ritchie 2013). "Minor" gulf surges triggered by the convective outflow boundaries of decaying MCSs can last several hours and are confined to the northern GoC (Hales 1972; Fuller and Stensrud 2000; Lang et al. 2007). IVs can also be an indirect cause of minor gulf surges in their support of MCS development (Douglas and Leal 2003; Lang et al. 2007).

Operational forecast models typically resolve features at the meso-α scale (100–1000 km) therefore representing features important to the monsoon like the longwave atmospheric circulation pattern over western North America during the warm season (e.g., monsoon ridge positioning) and transient synoptic features (e.g., IVs). However, to explicitly represent storm-scale structures like squall lines and outflow boundaries (e.g., Li et al. 2008), the use of convective-permitting atmospheric modeling at the meso-γ scale (i.e., model grid spacing on the order of 1–4 km without cumulus parameterization) is necessary. Since MCSs are dependent upon antecedent meso-γ-scale features, they

cannot be resolved by large-domain operational forecast models that focus on the meso-α scale and larger (Gutzler et al. 2009).

The North American Monsoon Experiment (NAME; Higgins et al. 2006) in 2004 advanced our knowledge of convective processes in the North American monsoon core region (see Higgins and Gochis 2007) and provided a valuable dataset for assessing model representation of monsoon convection. However, a noted major weakness in the NAME observational network experimental design was the lack of measurements of both the diurnal cycle of the boundary layer and moisture fluxes at higher elevations away from the coast (Higgins and Gochis 2007), a critical region for monsoon convective initiation. This lack of measurements strongly motivated the North American Monsoon GPS Transect Experiment 2013 (Transect 2013) (Adams et al. 2014; Serra et al. 2016) that included 10 GPS meteorological stations that were installed in northwest Mexico to capture the evolution of PWV during convection initiation through organization and propagation from the highest elevations of the SMO to the coastal plains of the GoC.

In our study, we use a convective-permitting model to downscale operational forecasts over northwest Mexico during Transect 2013. Using a similar model, Kursinski et al. (2008a) showed that a 5% change in initial PWV values (within the analysis error) significantly changes the amount of convective precipitation over northwestern Mexico, therefore implying that monsoon convective forecasts are sensitive to the initial specification of PWV. Here, we report on the results of validating our forecasts against satellite rainfall and surface observations from the Transect 2013 dataset. A follow-on study will document the results of the assimilation of GPS PWV into the model forecasts.

2. Data and model description

a) GPS PWV from the Transect 2013 experiment

In the previous section, we introduced Transect 2013 where 10 GPS meteorological sensors were deployed across northwest Mexico within the North American monsoon core region during summer 2013. These GPS sensors are indicated by black dots in the bottom panel of Fig. 1 and were set up into three strategically placed transects (Serra et al. 2016) whose configuration and monitoring purposes are summarized in Table 1. The longer-term SuomiNet GPS stations (http://www.suominet.ucar.edu), principally located within the United States at the time of this study, are indicated by triangles the bottom panel of Fig. 1. Gulf surges identified by the coastal transect have the advantage of being based on full-tropospheric moisture and are consequently less subject to localized land surface effects (e.g., surface dewpoint). Additionally, the advection of moisture along the gulf can be monitored to estimate the speed and extent of the surge, which, as discussed above, has important consequences for the location of convective outbreaks mainly in the southwest United States. The SMO transect, whose terrain cross section is highlighted in Fig. 2, offers a first-time look at in situ high-frequency PWV evolution during convective initiation in the higher elevations and thereby filling the gap in the NAME dataset in this locale that was noted by Higgins and Gochis (2007).

GPS meteorological sensors also provide standard meteorological variables including precipitation at 1-min temporal resolution. GPS PWV is high frequency (~5 min) and all weather, which makes it particularly advantageous for rapidly evolving cloudy and rainy conditions. For example, the strong upswing in PWV prior to deep

convective events in association with water vapor convergence (Kursinski et al. 2008a,b; Adams et al. 2011, 2013, 2015) permits the use of the time rate of change of PWV for identifying the events as well as a proxy for their intensity. Global Navigation and Satellite Systems (GNSS)-Inferred Positioning System and Orbit Analysis Simulation Software (GIPSY-OASIS; https://gipsy-oasis.jpl.nasa.gov/) is used to obtain PWV from the GPS signal at 5-min temporal resolution. For more information on the derivation of PWV from the GPS signal, see Bevis et al. (1992). One of the 10 stations, RAYN, failed in mid-July and was excluded from the analysis.

b) Gridded precipitation dataset

The gridded rainfall dataset that we use for verification purposes in our northwest Mexico domain is the Tropical Rainfall Measuring Mission (TRMM) Multisatellite Precipitation Analysis (TMPA) 3B42, version 7, dataset (referred to hereinafter as TRMM; Huffman et al. 2007) based on its past performance that compared it three other datasets with similar subdaily temporal resolutions capable of resolving the diurnal cycle of convective precipitation: Precipitation Estimation from Remotely Sensed Information Using Artificial Neural Networks (PERSIANN; Sorooshian et al. 2000), Global Satellite Mapping of Precipitation (GSMaP; Okamoto et al. 2005; Kubota et al. 2007; Aonashi et al. 2009; Ushio et al. 2009), and the Climate Prediction Center (CPC) morphing technique (CMORPH; Joyce et al. 2004) datasets. These datasets have unique ways of incorporating infrared geostationary satellite data, polar-orbiting satellite microwave data, and/or gauge observations, as well as additional analysis methods to determine quantitative precipitation estimates (QPE) that are explained in detail in their respective references in Table 2. Stage IV is a combined WSR-88D radar and gauge-based gridded

36

rainfall product within the United States that Chen et al. (2015) consider the "benchmark for validating other radar- or satellite-based QPE products" (p. 4445). They show that Stage IV and TRMM display similar spatial precipitation patterns over the United States and state that their results "cast a vote of confidence for the satellite QPE algorithm" of TRMM to be a reference for developers of a QPE algorithm in the Global Precipitation Measurement (GPM). Stillman et al. (2016) examine TRMM, CMORPH, and PERSIANN satellite precipitation datasets across the Walnut Gulch Experimental Watershed (WGEW), a 150 km^2 desert watershed in southeast Arizona just to the north of the SMO, and find that TRMM performs the best and PERSIANN the worst when compared to observations during the warm season. Tian et al. (2010) compared TRMM, GSMaP, CMORPH, and PERSIANN. In the western United States, in contrast to TRMM, they show that GSMaP, CMORPH, and PERSIANN generally miss the heaviest rain rates (greater than 40 mm day^{-1}) and overestimate precipitation, particularly in the warm season.

c) Other observational and atmospheric reanalysis data sources

Data sources used to identify synoptic features such as IVs, TEWs, and tropical cyclones (TCs) as well as to calculate gulf surges outside of the coastal transect are described in Table 3.

d) WRF-ARW Model configuration and hindcast simulations

The Advanced Research version of the Weather Research and Forecasting (WRF-ARW; Skamarock et al. 2008) Model, version 3.4.1, is used for retrospective daily convective simulations (referred to as hindcasts) for the duration of Transect 2013. The WRF-ARW configuration is based on the real-time quasi-operational model at the

University of Arizona within the Department of Hydrology and Atmospheric Sciences (UA HAS), hereinafter referred to as UA-WRF (http://www.atmo.arizona.edu/?section=weather&id=wrf), whose configuration was also the basis for a study that dynamically downscaled regional climate models (Luong et al. 2017). Our model uses three nested domains (d01, d02, and d03), as described in Table 4, that feature 27 vertical levels with a terrain-following hydrostatic pressure coordinate that is a traditional sigma coordinate (Skamarock et al. 2005). The innermost domain (d03) closely corresponds to the NAME Tier I region (which itself encompasses the monsoon core region) and since it has a 2.5-km horizontal grid spacing, it explicitly resolves convection at the meso-γ scale. The coarser domains of d01 and d02 (30- and 10-km horizontal grid spacing, respectively) employ the Kain–Fritsch cumulus parameterization scheme (Kain 2004). Other model physics applied to all domains is shown in Table 5.

The hindcasts are executed daily during the Transect 2013 period for 79 days (from 26 June through 12 September 2013) in two sets of simulations that have different initial conditions and 6-hourly lateral boundary conditions: 1) the NAM model (32-km horizontal grid spacing) and 2) GFS model (0.25° horizontal grid spacing). To capture the diurnal cycle of convection, the hindcasts are initialized at 1200 UTC (0500 LT) and run for 24 h. The Rapid Refresh (RAP) model is used to initialize the soil moisture and temperature because these surface data have finer spatial resolution as compared to the NAM and GFS models.

3. Analysis methods

a) Assessment of synoptic features

Following Douglas and Englehart (2007), who summarized transient features of

the 2004 monsoon season in NAME, we track IVs, gulf surges, TEWs, and tropical

cyclones (TCs) during the Transect 2013 period. These features are then used to

categorize days by the synoptic forcing conditions to assess their relative impact on

forecast skill within the hindcasts.

1) IDENTIFICATION OF TRANSIENT SYNOPTIC FEATURES

We subjectively determine an IV, based on Bieda et al. (2009), by an area of

enhanced relative vorticity that moves westward along the southern periphery of the

monsoon high and is depicted by a swirl pattern in the GOES water vapor infrared

channel. Additionally, North American Regional Reanalysis (NARR)-A 300-hPa

geopotential height and wind fields are also examined to confirm the IV presence. We

note IVs that approach the NAME Tier I region (monsoon core).

We objectively identify TEWs from the ERA-Interim reanalysis and GFS analysis

using a vorticity-tracking algorithm (Hodges 1994, 1995) that was previously shown to

be effective in tracking TEWs across the tropical Atlantic and east Pacific (e.g.,

Thorncroft and Hodges 2001; Serra et al. 2010). We identify a TEW as a vertically

averaged relative vorticity feature over the 850–600-hPa layer that exceeds $+5 \times 10^{-6}$ s^{-1}

with the following requirements: 1) the vorticity feature persists for at least 2 days, 2) its

track has a length of at least 1000 km, and 3) its track passes within 500 km of the mouth

of the GoC. Similar TEW tracks were found in the ERA-Interim and GFS vorticity fields

(not shown). If a TEW track coincided with a TC identified in the "best track" database,

then that track is categorized as a TC (not a TEW). Note: Two TEWs that occurred were

not initially found by the objective vorticity tracking algorithm and were not included.

They were later verified by both GOES infrared imagery and NWS area forecast

discussions.

Gulf surges are identified using hourly Transect 2013 GPS PWV at the three coastal transect stations (MOCH, KINO, and PSCO) and hourly dewpoint temperature at Yuma, Arizona (KNYL), located north of the GoC. In the coastal transect, per Serra et al. (2016), we first calculate a 24-h moving average of PWV to smooth the data. Then, we calculate a percent increase from the minimum PWV in a previous 24-h period to the maximum PWV in the current 24-h period (ending at 1200 UTC to match the hindcasts). The thresholds used to identify a gulf surge are specific to each of the three GoC sites and are based on the minimum PWV percent increase for each site out of all three gulf surges in Fig. 4 of Serra et al. (2016): 11% at MOCH, 19% at KINO, and 28% at PSCO. North of the GoC, after smoothing hourly dewpoint temperatures at KNYL via a 24-h moving average, we identify gulf surges with two criteria: 1) a minimum dewpoint increase of 4°C in 2 consecutive 24-h periods that is based on the three gulf surges in Fig. 4 of Serra et al. (2016), and 2) a dewpoint temperature of at least 18°C (~64°F) in the current 24-h period that is partially based on the NWS method that uses this daily mean dewpoint temperature as one of its two threshold criteria (NWS Tucson 2015, personal communication).

2) CRITERIA FOR IDENTIFYING STRONGLY-FORCED AND WEAKLY-FORCED DAYS

As a first pass, we classify days as strongly or weakly forced based on the presence of one or more of the synoptic features discussed in the previous section known from the literature to facilitate convective organization in the North American monsoon region based on Douglas and Englehart (2007). We choose days when organized convection propagates off the high terrain of the SMO toward the west over the lower

elevations of Sonora and the border region of southern Arizona, as documented in the 2013 Monsoon Weather Discussions led by UA HAS graduate students and part of the Transect 2013 field campaign (https://monsoonwx2013.wordpress.com). To disregard days with monsoon breaks and light convective activity, we use TRMM to identify days when precipitation fell along the northern SMO crest and to the north and west toward the lower elevations of northwest Mexico and adjacent extreme southern Arizona. Only days with 24-h accumulations of at least 20 mm in at least five grid points in a region bounded by 26.0°–32.5°N and 114°–107.25°W (except for the GoC and Baja Peninsula) are considered.

b) Statistical performance metrics for WRF Model simulations

Given the sensitivity of model rainfall to initial specification of PWV found by Kursinski et al. (2008a), we compare model PWV for the sets of strongly and weakly forced days to observed GPS PWV. These results are then used to interpret analysis of hindcast rainfall against TRMM observations.

1) EVALUATION OF MODELED PWV

For direct comparisons of GPS-derived PWV and WRF-simulated PWV, the time and location are first matched with instantaneous PWV values extracted directly from the Transect 2013 and SuomiNet GPS datasets every third hour (i.e., 1200, 1500, 1800, …, 1200 UTC). Then, the WRF-simulated PWV is mapped to each GPS site using vertical integration and inverse-distance squared weighting schemes that are described in appendix A. Mean bias and root-mean-square error (RMSE) are computed between GPS observations and WRF-simulated (model) values. Here, mean bias is defined as the model mean minus the observation mean. In this convention, a negative (positive) mean

bias indicates a dry (moist) model bias relative to the observations. RMSE is defined as the square root of the mean squared differences between the observed and modeled values (Wilks 1995). A paired observation-model two-tailed t test is used to determine the level of statistical significance of the differences of the mean biases.

2) EVALUATION OF MODELED RAINFALL

To compare TRMM precipitation with the d03 WRF modeled precipitation, the 3-h TRMM mean precipitation rate is first converted to 3-h accumulations to match the time interval of the output of the WRF hindcasts. Then, WRF precipitation is scaled up from its 2.5-km horizontal resolution to 0.25° to match that of TRMM using the Earth System Modeling Framework (ESMF) "conserve" function within NCL. For each 3-h forecast period, modeled precipitation (WRF) is subtracted from observed precipitation (TRMM) to produce the bias maps for the combination of both strongly and weakly forced days.

Precipitation forecast skill is evaluated by first completing a 2×2 forecast contingency matrix (Table 6) and then calculating the critical success index (CSI; Donaldson et al. 1975), probability of detection (POD), and false-alarm ratio (FAR) at each grid point for a subset of days. Mason (1989) and Schaefer (1990) have shown that CSI is a biased metric that is dependent on the number of observed events. To address this bias, POD and FAR were added to the analysis. The metrics range from 0 to 1. CSI and POD have a perfect (zero) forecast skill of 1 (0). FAR has a perfect (zero) forecast skill of 0 (1). CSI is defined as a ratio of hits (A) to observed events ($A + C$) and false alarms (B):

$$\text{CSI} = \frac{A}{(A + B + C)} \quad (1)$$

POD is defined as the ratio of hits (A) to observed events ($A + C$):

$$POD = \frac{A}{(A + C)} \quad (2)$$

FAR is defined as the ratio of false alarms (B) to total forecasts ($A+B$):

$$FAR = \frac{B}{(A + B)} \quad (3)$$

We define a precipitation event at a grid point that has at least 2.5 mm (10 mm) of accumulation in the 6-h (daily) periods and at least 3 observed events in both the strongly and weakly forced days subsets. We use a neighborhood verification technique that considers modeled events in ±2 grid points where each grid point is assigned a weighted average of the difference of the metric between the strongly and weakly forced subsets of days. The metric is undefined for a grid point if any of the subsets have 1) less than 3 observed events or 2) no modeled events in all neighborhood grid points. For each grid point, a two-tailed statistical local significance test (p value < 0.10) was established via 1000 permutations in a Monte Carlo resampling method. Only grid points that containing at least 900 unique values are used. Finally, the statistical field significance is obtained in a method similar to Livezey and Chen (1983) using the same permutation method as a local test but with the resampling of the maps. The 900th value (90%) of the histogram is the critical value for statistical field significance. Pattern correlations are computed between the common grid points of the subdaily maps and daily maps.

    3) FORWARD MODEL SENSITIVITY EXPERIMENTS

    To demonstrate the potential impact of constraining the moisture fields in WRF, we conduct a suite of analyses on the forward sensitivity of WRF PWV and rainfall to the initial specification of PWV at GPS transect sites that is described in appendix A.

4. Model performance for strong and weak days

a) Overview of 2013 monsoon season

Synoptic forcing mechanisms affecting convective precipitation coinciding with the Transect 2013 period are shown in Fig. 3. IVs were the most prevalent features with 28 days impacted by 12 events. This was followed by 26 days being impacted by 6 TCs. Two TEW tracks (impacting 4 days) that came within 500 km of the mouth of the GoC were not associated with TCs. The remaining TEWs were actually TCs, or in one instance developed into a TC, and thus were not included in the TEW count (Fig. 3). The tropical activity in the eastern North Pacific in 2013 was above average with 18 TCs occurring versus a seasonal average of 15 for this basin from 1971 to 2009 (http://www.nhc.noaa.gov/, accessed 2 December 2016). No TCs directly impacted the land areas in northwest Mexico and adjacent southwest United States.

Sixteen gulf surges were identified in the coastal transect and KNYL with 5 surges initiated by TCs. One TEW initiated an additional gulf surge. Gulf surges were identified at MOCH (10), KINO (12), PSCO (11), and KNYL (7). In the coastal transect, an adjacent upstream station that recorded a gulf surge on the same or following day was considered part of the same event. Because surface dewpoint temperature data could lag or lead a PWV signal, we allowed gulf surges identified at KNYL an error of ±1 day to be considered part of the same event. Adams and Comrie (1997) differentiated between "major" and "minor" surges where "major" surges initiate near the GoC mouth and traverse the entire GoC while "minor" surges initiate partway up the GoC from cool MCS outflow and run the balance of the GoC. Out of the 16 total gulf surges, 7 were "major" in that they were observed at all 3 coastal transect sites (MOCH, KINO, and PSCO) and 4

44

of these also being observed at KNYL. Five of the gulf surges were "minor" if they were observed at PSCO or KNYL, but did not include all coastal transect sites. A third category of "partial" was defined for 4 gulf surges that did not reach PSCO.

Considering only days when appreciable rainfall was observed in northwest Mexico and adjacent southwest United States (i.e., monsoon active periods), we identify 22 strongly forced and 41 weakly forced days (Table 7). Since no consistent relationship is observed between the presence or absence of a gulf surge and a strongly forced or weakly forced day, we conclude that a strongly forced day solely requires the presence of an IV and, conversely, a weakly forced day requires the absence of an IV. The lack of a strong relationship between gulf surges and day classification suggests that MCSs initiating in northwest Mexico are not dependent upon gulf surges for their development. This is in contrast to convection that occurs in Arizona where gulf surges play a more integral role as indicated by previous literature (e.g., Hales 1972; Brenner 1974; Fuller and Stensrud 2000; Higgins et al. 2004). Comparing the mean upper-tropospheric winds at 300 hPa from NARR data suggests that enhanced easterly winds, and thus higher vertical wind shear, were present over the northern SMO on the strongly forced days than on weakly forced days (not shown). We hypothesize that the hindcasts will exhibit better performance of representing observed precipitation for the strongly forced days because an IV would be present within the convective-permitting domain of the WRF simulations (core monsoon region). In contrast, relatively poor model forecast performance is expected on weakly forced days when there are no obvious dynamic forcing mechanisms present to facilitate convective organization.

b) Model diurnal cycle in PWV

Relatively larger moist biases and RMSEs are found at GPS sites located in Mexico than in the United States (e.g., MULT) in both the WRF-NAM (Fig. 4) and WRF-GFS (Fig. 5) hindcasts. The largest errors occur at the sites located on the western slope of the SMO (e.g., MULT), where diurnally generated convection transitions to more organized (MCS type) convection during the late afternoon. Thus, the SMO transect has captured weaknesses in the forecast model PWV in the same regions that were noted in the NAME 2004 field campaign dataset (Higgins and Gochis 2007). We later show that these weaknesses in the model PWV fields may also result in rainfall forecast errors. The moist bias is greatest overall in the WRF-NAM hindcasts contributing to a higher number of GPS sites that display statistically significant positive differences in PWV at model initialization. In contrast, the smaller PWV biases in the WRF-GFS hindcasts at model initialization result in fewer GPS sites having statistically significant differences.

We investigate the model initialization of the WRF-NAM hindcasts in greater detail since they display the highest PWV biases and RMSEs (Fig. 4). Nine out of 15 GPS sites show significant positive PWV biases greater than 2 mm generally west of the SMO crest toward lower elevation. Of the three sites that have a significant positive PWV bias of 0–2 mm, one is located near Phoenix in central Arizona (SA31), another is in extreme southeast Arizona (AZCO), while the final one is located in a relatively dry area east of the SMO crest (CUAH). The highest RMSE values are found on the western SMO slope and foothills with 5–7 mm at MULT, ONVS, and BGTO. The lowest RMSE (< 2 mm) is located at CUAH east of the SMO crest. RMSEs of 2–3 mm are located in southern Arizona at the same sites that feature the lowest model bias (AZCO and SA31).

As a rule, both the WRF-NAM and WRF-GFS hindcasts (Figs. 4 and 5) show a decrease in bias with time within the diurnal cycle (wet model biases decrease and even become biased dry). MULT is the only SMO transect site that maintains a statistically significant moist model bias throughout the diurnal cycle in both sets of hindcasts. In the WRF-NAM hindcasts (Fig. 4), PSCO is the only site that retains a statistically significant positive PWV bias, increasing from +2 to +4 mm at model initialization to +4 to +6 mm at 0600 UTC and lasting until the end of the diurnal cycle. The increased moist bias trend with time at PSCO, also in the WRF-GFS hindcasts, is the exception to the overall behavior of the GPS sites in that they typically show increased dry bias (decreased moist bias) throughout the diurnal cycle. The number of sites with statistically significant biases decrease with time correspondingly as the wet biases decrease in the context of the WRF-NAM hindcasts. While statistically significant dry biases of −2 to 0 mm at BASC (higher elevation) are not observed until 0900 UTC and through the end of the diurnal cycle, statistically significant moist biases of +2 to +4 mm are observed at ONVS (lower elevation) at model initialization (1200 UTC) and 1500 UTC. In the WRF-GFS hindcasts (Fig. 5), BASC is the only site with a statistically significant dry bias (−2 to 0 mm) at model initialization and does not become statistically significant again until 0600 UTC through 0900 UTC with values of −2 to 0 mm. The bias then becomes drier with values of −4 to −2 mm at the end of the diurnal cycle. Around the same time, lower-elevation ONVS has a statistically significant dry bias of −4 to −2 mm from 0900 UTC until 1200 UTC.

c) Model diurnal cycle in rainfall

The spatially averaged RMSE of daily precipitation accumulation between the

WRF hindcasts and each satellite-based rainfall product all fall within the same order of magnitude (Table 8). In additional to being supported by the previous studies mentioned in section 2b, the choice of TRMM as the primary source of hindcast rainfall verification is also justified by the fact that it has the lowest RMSE with 8.7 mm day^{-1} (7.8 mm day^{-1}) for WRF-NAM (WRF-GFS). The highest RMSE is observed in PERSIANN with 9.9 mm day^{-1} (9.2 mm day^{-1}) for WRF-NAM (WRF-GFS). For each satellite-based product, the wetter WRF-NAM has a higher RMSE than its respective WRF-GFS RMSE.

In Figs. 6 and 7, we verify the mean diurnal cycle of WRF precipitation that is produced explicitly from the cloud microphysics in d03 (top row) against TRMM (middle row) and display the spatial patterns of the biases (bottom row) in addition to the overall spatially averaged values across the domain (Table 9) for the period of Transect 2013. We normalize the 3- and 24-h hindcast precipitation accumulations into a mean hourly rainfall rate. Because there are no statistically significant differences in spatially averaged rainfall biases between the strongly and weakly forced days, we present the combined set of days where blue (red) areas indicate a wetter (dryer) model bias.

The model-simulated precipitation biases are strongly tied to the evolution of the diurnal cycle of convection. In the WRF-NAM hindcasts (Fig. 6), the precipitation analysis of the diurnal cycle, whose overall positive (moist) PWV bias was mentioned in the previous subsection, correspondingly shows a larger positive (wet) model rainfall bias compared to the WRF-GFS (Fig. 7) as seen through the 2100–0000 UTC period. WRF-NAM hindcast rainfall begins earlier than the WRF-GFS as seen in the 1200–1500 UTC and 1500–1800 UTC periods (Figs. 6 and 7, left two panels in the top rows). This contributes to the WRF-NAM spatially averaged rainfall model bias both maximizing 3 h

earlier and being greater in magnitude when compared to the WRF-GFS (+0.12 mm h^{-1} at 1800–2100 UTC for WRF-NAM versus +0.06 mm h^{-1} at 2100–0000 UTC for WRF-GFS). While the spatially averaged TRMM maximizes during the 0000–0300 UTC period (0.18 mm h^{-1}), the spatially averaged modeled precipitation maximizes 3 h earlier during the 2100–0000 UTC period with 0.28 mm h^{-1} (0.24 mm h^{-1}) for WRF-NAM (WRF-GFS) (Table 9). At the time of maximum precipitation bias (1800–2100 UTC for WRF-NAM and 2100–0000 UTC for WRF-GFS), the SMO Transect 2013 sites ONVS, MULT, and BASC (located near the center of each map) coincide with the largest gridpoint-based positive precipitation biases. The spatially averaged precipitation bias is approximately zero by the 0000–0300 UTC period before minimizing in the 0600–0900 UTC period (−0.07 mm h^{-1} for WRF-NAM and −0.06 mm h^{-1} for WRF-GFS) and remaining negative (dry) for the remainder of the diurnal cycle. The dry precipitation bias region that begins to appear during the 0000–0300 UTC period is most pronounced just to the north of the ONVS and MULT stations with a smaller area farther south that are both on the western slope of the SMO. The transition from relatively large positive precipitation biases prior to 0000 UTC to relatively large negative precipitation biases after 0300 UTC suggests that WRF is challenged overall in representing the evolution of organized convection in northwest Mexico with respect to timing. A wet model bias (more pronounced in WRF-NAM) is observed in the western slopes of SMO south of the mountain transect sites in the 24-h period. This bias is not seen across the region of MCS activity on this large time scale because of the canceling effect of the forecast timing error. In summary, regardless of the source of boundary forcing, WRF consistently initiates the convection on the crest of the SMO too early and underestimates the

propagating, more MCS-type precipitation that occurs in the afternoon and early evening hours. Our WRF-ARW hindcast results using a convective-permitting grid spacing are broadly similar to the earlier findings of Lee et al. (2007) and Collier and Zhang (2007), who also examined model simulations of the diurnal cycle for the monsoon, although at coarser spatial scales.

d) Evaluation of model performance for rainfall

We use the CSI, POD, and FAR metrics to evaluate the hindcasts for both strongly forced and weakly forced subsets of days (Figs. 8 and 9). We display the differences of the forecast metrics between the two subsets of days for all but the first 6-h period because convection is limited 1200–1800 UTC (see Fig. 6, row 2). We also perform this analysis for the entire 24-h forecast period (1200–1200 UTC). Blue (red) areas indicate a higher forecast skill for strongly (weakly) forced days. The rank of field statistical significance (pattern correlation of the subdaily grid to the daily grid) is displayed in the lower left (upper right) of each map with both statistics listed in Table 10.

The CSI, POD, and FAR differences display similar spatial patterns within each time period with the POD and FAR having a greater range of difference than the CSI. In the 24-h period, there is a common pattern of greatest CSI and POD and least FAR for the strongly forced days in an area across the high terrain just west of the SMO crest (around BASC) and into lower elevations of Sonora (Figs. 8 and 9, column 4). This area is where convection initiates over the highest terrain and propagates generally toward the west in late afternoon and continues toward the GoC into the evening and early morning as MCSs. Between the strongly and weakly forced days, the patterns of the differences of

CSI and POD are field statistically significant across both the WRF-NAM and WRF-GFS

hindcasts, while the FAR differences are field statistically significant only in the WRF-

GFS during this time.

In the 1800–0000 UTC period, the spatial patterns of the metric differences show

increased forecast skill for the strongly forced days across the western SMO slope in the

BASC–MULT–ONVS region. In the WRF-GFS hindcasts, the spatial patterns display

field statistical significance and have the highest map correlations with the 24-h period

(0.55, 0.50, and 0.55 for CSI, POD, and FAR, respectively). In contrast, in the WRF-

NAM hindcasts, only the CSI difference patterns are field statistically significant while

the POD difference pattern is the highest correlated metric to the 24-h period (0.41).

In the 0000–0600 UTC period, we find the lowest ranks of field statistical

significance (highest rank is 74.7%) indicating that, when the precipitation biases

between the WRF and TRMM minimize (Figs. 6 and 7, bottom row) as TRMM

precipitation transitions to the lower terrain, the differences in forecast skill has the

highest likelihood of occurring by chance. The CSI and FAR difference patterns in the

WRF-NAM at this time have the highest correlations to the 24-h pattern (0.43 and 0.42,

respectively).

In the 0600–1200 UTC period, WRF precipitation is confined to the eastern SMO

slope (CUAH–CHIH), near the immediate coast, and over the GoC (Figs. 6 and 7, top

row). The differences of the metrics between the strongly and weakly forced days in the

lowest terrain (<1000 m) of Sonora is minimal (±0.1). The CSI and FAR difference

patterns are field statistically significant in both the WRF-NAM and WRF-GFS, while

the POD difference patterns are field statistically significant only in the WRF-NAM. In

the WRF-NAM hindcasts, the strongly forced days have highest forecast verification metrics south of MOCH–BGTO (tropical related) along the coast and in the northern SMO and east slope (BASC–CUAH–CHIH). The WRF-GFS hindcasts show the highest skill in the strongly forced days across eastern SMO slope in the CSI and POD difference maps.

In the 1800–0000 UTC and 24-h periods, the strongly forced days display higher precipitation forecast skill than the weakly forced days in all or some of the three metrics that are field statistically significant. This includes areas across west of the SMO crest, western SMO foothills, and toward the GoC where MCSs mature and decay and supports the idea that the WRF simulations in the weakly forced days are more challenged to capture the mature stages of MCSs once they propagate off the western slopes of the SMO. This finding is consistent with our hypothesis that MCS development in WRF will tend to more preferentially occur when an IV is nearby and conversely WRF is more challenged to forecast MCS development in the absence of an IV.

There is some agreement, in terms of spatial patterns, in the rainfall verification metrics among the four satellite-based precipitation datasets. This is showcased in the differences in daily (24 h) forecast verification metrics between the strongly forced days and weakly forced days for WRF-NAM (WRF-GFS) in Fig. 10 (Fig. 11) using TRMM, CMORPH, PERSIANN, and GSMaP. Additionally, CSI (CSI and FAR) differences in the WRF-NAM (WRF-GFS) have field statistical significance across all products.

5. Case study of 8–10 July 2013

a) Synoptic overview

An IV approached from the east and across the SMO in conjunction with the

second major gulf surge of the season triggered by TC Erick on 8 July 2013. Based on

our classification criteria, the first day (8–9 July) is considered weakly forced while the

second day (9–10 July) is considered strongly forced. GOES water vapor IR imagery

depicted a mature MCS cloud shield each evening at approximately 0600 UTC in similar

locations over northern Sonora and the immediate border region of southern Arizona

(Fig. 12, top panel). Pressure and winds on the 2.0 potential vorticity unit (PVU) surface

indicate the location of the IV as shown by a PV anomaly at 1800 UTC from the 1200

UTC GFS 6-h forecast (Fig. 12, bottom panel). Rising motion from the upward tilting of

isentropes ahead of the PV anomaly was closer to the area of the MCS development on

the second day.

At 1200 UTC 8 July, an IV was located south of Big Bend, Texas, and was

moving west into the Mexican state of Chihuahua as shown on the dynamic tropopause

map (Fig. 12, bottom left). The lowest pressure indicated by the 6-h GFS forecast of the

associated PV anomaly was between 280 and 300 hPa. GOES infrared imagery showed

convective clouds near the southern tip of the Baja California Peninsula in association

with TC Erick, which triggered a gulf surge at the mouth of the GoC earlier in the day

that was observed by all 3 transect sites and Yuma. As shown in the 3-hourly GOES IR

and TRMM accumulations (Fig. 13, rows 1 and 2), thunderstorms developed along the

Sonora–Chihuahua border along the northern SMO crest at approximately 2100 UTC 8

July. By 0000 UTC 9 July, an MCS began to form near ONVS–MULT–BASC. Between

0300 and 0600 UTC, rainfall maximized in intensity with a large 3-h accumulation swath

of 20 mm north of ONVS. A mature MCS evident by a vast cloud shield over Sonora at

0600 UTC began to dissipate by 0900 UTC. While the synoptic lift from the IV was not

likely playing a role in the development of the MCS because of its distance, the 20–25-kt

(1 kt = 0.51 m s−1) east winds above 200 mb over the northern SMO likely created

favorable vertical wind shear for the formation and maintenance of the MCS.

After lingering convection and debris clouds at 1500 UTC (Fig. 14, top row), the

following strongly forced day is similar in terms of the MCS formation and location. It is

classified as strongly forced because at 1200 UTC 9 July the IV was located on the

eastern border of the Tier I region (core) and therefore had more influence than on the

previous day despite the PV anomaly being weaker (one grid point with 260–280 hPa).

Its location directly over the northern SMO at 1800 UTC (Fig. 12, bottom right) aided in

upward motion and wind shear for the MCS that would form just to the west. TRMM

precipitation (Fig. 14, row 2) showed that diurnal convection began approximately three

hours earlier than the previous day at 1800 UTC near CUAH just east of the SMO crest

in association with the IV. By 2100 UTC, convection in the northern SMO was in a

similar region to the previous evening, but with greater spatial extent. Between 0300 and

0600 UTC 10 July, the 3-h rainfall accumulation maximized at the same time as the day

before but with a larger 20-mm swath that is also shifted north and west at a location

around MOCH and toward the northeast into southern Arizona. By 0600 UTC, the MCS

cloud shield reached maximum coverage similar to the previous day. A small 20-mm

swath is apparent in the next 3-h period near the border region, but after 0900 UTC the

MCS cloud shield quickly dissipated. Only lingering showers were present between

PSCO and Yuma at 1200 UTC.

b) Evaluation of model hindcast precipitation

Here, we illustrate the importance of synoptic forcing to MCS development in

existing models and WRF. The 24-h rainfall accumulations across all four satellite-based rainfall products display roughly similar rainfall patterns of the MCSs that had occurred each evening and night, but the magnitudes were different (Fig. 15). TRMM (PERSIANN) gives the lowest (highest) rainfall accumulations for both days. For 9 July 2013 (Fig. 15, bottom row), the large 90-mm isohyet in PERSIANN (and to a less extent, CMORPH) is likely an overestimate that may have been contaminated by the high IR brightness from the massive anvil. PERSIANN is also the only product that does not include gauge bias correction.

Figure 16 displays the 24-h precipitation accumulation from the model providing the lateral boundary conditions (columns 2 and 4), their respective WRF simulations (columns 3 and 5), and TRMM (column 1) for both the weakly forced day (top row) and strongly forced day (bottom row). The coarse resolution of the parent models requires cumulus parameterization, as they cannot represent precipitation explicitly as is apparent in the precipitation accumulation pattern. TRMM shows MCS-related precipitation of upward of 20–50 mm (yellow, orange, and red) over a large portion of northern Sonora on the weakly forced day. While the NAM model shows a closed 10-mm isohyet north of the MCS maximum precipitation area, the GFS model shows a small closed 20-mm isohyet on the western edge of this region.

The respective WRF-NAM shows a semblance of the NAM model precipitation pattern in northern Sonora, but the WRF-GFS has does not replicate the GFS model precipitation area in central Sonora. The weakly forced day displayed poor WRF performance in terms of MCS location and timing (Fig. 13, rows 3–4). A large area of rainfall with accumulations of greater than 20 mm in WRF-NAM begins 1800–2100

UTC just west of the crest in the southern SMO around and south of BGTO. TRMM at the same time showed accumulations of no more than 6 mm. This swath expanded into the northern SMO toward MULT by 2100–0000 UTC. The WRF-GFS demonstrates a similar pattern, but with less intensity. At 0300–0600 UTC, the WRF-NAM indicated convection south of ONVS–MULT–BASC, while TRMM showed the MCS occurring north of there. By 0600–0900 UTC, the WRF-GFS shows an area of rainfall in the same region as the WRF-NAM 3 h earlier. Both parent models and their respective WRF simulations show precipitation along the SMO from central Sonora south that is not observed in the TRMM. For the strongly forced day, a larger area of 20–50-mm accumulation is shown in central Sonora close to the Arizona border region from the MCS. WRF precipitation developed rapidly prior to 2100 UTC from the central SMO into southern Arizona and continued through 0000 UTC near the Arizona border near the area of TRMM rainfall, although it was 3–6 h too early (Fig. 14, rows 3–4). At 2100–0000 UTC, a large rainfall swath greater than 20 mm was shown at lower elevations west of MULT in both WRF-NAM and WRF-GFS for the strongly forced day as the model advances the precipitation off the SMO too quickly. By 0000–0300 UTC, the 20-mm swath was located near KINO close to the actual MCS location. After this time, the WRF-GFS performed relatively well in simulating the MCS near its actual location through 0900 UTC with 3-h swaths of rainfall greater than 20 mm that matched the timing of TRMM. In contrast to the weakly forced day, the rainfall patterns of the parent models and their respective WRF simulation are similar in nature and cover the general location of the MCS observed in TRMM. Similar to the weakly forced day, there is still there is an overforecast of rainfall in the WRF simulations (especially WRF-NAM) south

of the MCS locations along the SMO, reflecting the erroneous early initiation of convective precipitation and positive modeled precipitation biases early in the day, as discussed above in reference to Figs. 6 and 7.

6. Discussion

In both WRF-NAM and WRF-GFS hindcasts, regardless of the strong or weak classification criteria, relatively large positive biases in modeled simulated PWV occur on the western slopes of the northern SMO during the first part of the day (prior to 0000 UTC). These moist biases likely contribute to an erroneous model representation of the diurnal cycle of convection that initiates 3–6 h earlier than observed in TRMM; simulated convective rainfall is too high prior to 0000 UTC and too low after 0300 UTC. MCS-type convection is more likely to occur in the latter period. These same types of systematic errors in the modeled representation of the diurnal cycle of convection have also been noted in prior studies modeling North American monsoon precipitation using a coarser resolution with parameterized convection.

Given the relative poor performance of the WRF Model for organized MCS-type convection during the weakly forced days without the appreciable influence of an IV, we hypothesize that the convective-permitting WRF Model simulations would potentially benefit the most from the assimilation of GPS-PWV observations in these types of days. This hypothesis is well supported by an ensemble-based model sensitivity analysis of the 8–10 July 2013 case study that included a weakly forced day followed by a strongly forced day. The sensitivity of model-simulated PWV and rainfall relative to the initial PWV specification at ONVS for the weakly forced day (8–9 July 2013) and strongly forced day (9–10 July 2013) is shown in Fig. 17 with the methodological approach

57

described in appendix A. The ONVS Transect 2013 site was chosen in this analysis for the following reasons: 1) its location west of the crest of the SMO approximately corresponds to where the transition to organized MCS-type convection occurs during these particular days, 2) it exhibits the largest relative biases and RMSE in modeled simulated PWV within the SMO transect, and 3) its location at model initialization is far enough upstream to be a source of moisture feeding the convection that would commence along the SMO crest approximately 6 h later owing to diurnal mountain/valley-breeze circulation. Other sites closer to the SMO crest (i.e., MULT and BASC) do not appear to have a large PWV "memory" indicated by the lower sensitivities there (not shown). At the model initialization and 12-h forecast times (left two panels of Fig. 17, respectively), we find clear higher sensitivity of the modeled simulated PWV to specification of PWV initial conditions at this site for the weakly forced day than for the strongly forced day. This results in relatively high sensitivity of the model-simulated rainfall in northern Sonora to the initial PWV specification at ONVS (Fig. 17, top right) and in the same general location where the MCS is observed to occur via GOES IR imagery (Fig. 13, row 1) and TRMM precipitation (Fig. 13, row 2). Similar patterns of PWV and precipitation sensitivity can also be found at higher-elevation stations along the SMO transect, like MULT and CHIH (not shown), though these stations did not exhibit as strong a sensitivity to forecast PWV and rainfall as ONVS.

These sample results strongly suggest that the WRF Model simulation results would be more sensitive to the initial specification of local PWV west of the SMO crest during the weakly forced days. In other words, as compared to days when an IV is present in this region, a more accurate specification of local moisture conditions during

the weakly forced days would have a greater relative impact on the model simulation of organized MCS-type convection. Therefore, we hypothesize that assimilation of GPS-derived PWV may be of greatest value to improve the WRF precipitation forecasts during the types of days when the obvious synoptic-scale forcing mechanisms to facilitate organized convection are absent. As mentioned in our prior work of Serra et al. (2016), a subsequent study will focus on the assimilating the GPS PWV observations into the WRF-GFS configuration considered here to formally assess this hypothesis.

7. Conclusions

Daily WRF-ARW hindcasts of monsoon convection are performed using the forcing data for initial and boundary conditions from the operational GFS and NAM model during the period of Transect 2013. Both WRF-GFS and WRF-NAM hindcasts display a consistent moist bias in the initial specification of PWV when compared to GPS-derived PWV with the WRF-NAM being the wetter of the two. We classify days by their level of synoptic forcing; strongly and weakly forced days are differentiated by the clear presence of an IV. The presence or absence of a gulf surge is not found to alter the WRF hindcasts for MCS development over northwest Mexico in contrast to MCS development in southern Arizona where previous literature indicates a high dependence on antecedent gulf surges. Model forecast precipitation skill is evaluated using the objective CSI, POD, and FAR metrics for the days when appreciable NAM precipitation in observed by TRMM. The strongly forced days initiated over the SMO displayed notably higher precipitation forecast skill than the weakly forced days, especially for areas west of the SMO crest. Therefore, strongly forced days appear to improve the model's ability to deterministically forecast more organized, propagating MCS-type

convection that accounts for a greater proportion of the monsoon precipitation west of the SMO crest toward the GoC. The 8–10 July 2013 case study is a clear illustration of this point. During these two days, nearly identical MCSs evolved in the same area in northern Sonora in terms of precipitation amounts and spatial extent of the cloud shield. However, WRF reasonably simulates the MCS only on the day classified as strongly forced, or when an IV is near the initiation region at the highest elevations of the SMO (9–10 July 2013), but not on the weakly forced day (8–9 July 2013) when the IV is located farther east. Relative to the specification of initial modeled PWV at the ONVS Transect 2013 site near the western SMO foothills, we find higher sensitivity to the initial PWV field, 12-h forecast PWV (0000 UTC), and 12-h forecast 12-h rainfall accumulation (0000–1200 UTC) for the weakly forced day than for the strongly forced day. Because ONVS shows higher sensitivity than higher-elevation sites (e.g., MULT and BASC), we hypothesize that the initial specification of PWV west of the SMO crest is crucial for improved convective-permitting forecasts especially for MCS-type precipitation. Furthermore, these simulations would potentially benefit the most from the assimilation of GPS-PWV observations on days when synoptic-scale forcing mechanisms (e.g., IVs) that facilitate organized convection are absent. To formally assess this hypothesis, we introduced a subsequent study focusing on assimilating the GPS PWV observations into the WRF configuration in our prior work (Serra et al. 2016).

Acknowledgements

References

Adams, D. K., and A. C. Comrie, 1997: The North American Monsoon. *Bull. Amer. Meteor. Soc.*, **78**(10), 2197–2213, doi:10.1175/1520-0477(1997)078%3C2197:TNAM%3E2.0.CO;2.

Adams, D. K. and E. P. Souza, 2009: CAPE and convective events over the southwest U.S. during the North American monsoon. *Monthly Weather Review*, **137**, 83–98.

Adams, D. K., and Coauthors, 2011: A dense GNSS meteorological network for observing deep convection in the Amazon, *Atmos. Sci. Let.*, **12**, 207–212, doi:10.1002/asl.312.

Adams, D. K., S. I. Gutman, K. L. Holub, and D. S. Pereira, 2013: GNSS observations of deep convective time scales in the Amazon. *Geophys. Res. Lett.*, **40**, 2818–2823, doi:10.1002/grl.50573.

Adams, D. K., C. Minjarez, Y. Serra, A. Quintanar, L. Alatorre, A. Granados, E.

Vázquez, and J. Braun, 2014: Mexican GPS tracks convection from North American monsoon. Eos, Trans. *Amer. Geophys. Union*, 95, **61**, doi:10.1002/2014EO070001.

Adams, D. K., and Coauthors, 2015: The Amazon Dense GNSS Meteorological Network: A New Approach for Examining Water Vapor and Deep Convection Interactions in the Tropics. *Bull. Amer. Meteor. Soc.*, **96**, 2151–2165, doi:10.1175/BAMS-D-13-00171.1.

Aonashi, K., and Coauthors, 2009: GSMaP passive microwave precipitation retrieval algorithm: Algorithm description and validation. *J. Meteor. Soc. Japan*, **87A**, 119–136.

Becker, E. J., and E. H. Berbery, 2008: The Diurnal Cycle of Precipitation over the North American Monsoon Region during the NAME 2004 Field Campaign. *J. Climate*, **21**,

Bevis, M., S. Businger, T. A. Herring, C. Rocken, R. A. Anthes, and R. H. Ware, 1992: GPS Meteorology: Remote Sensing of Atmospheric Water Vapor Using the Global Positioning System. *J. Geophys. Res*, **97**, 15787–15, 801, doi:10.1029/92JD01517.

Bieda, S. W., III, C. L. Castro, S. L. Mullen, A. C. Comrie, and E. Pytlak, 2009: The Relationship of Transient Upper-Level Troughs to Variability of the North American Monsoon System. *J. Climate*, **22**, 4213–4227, doi:10.1175/2009JCLI2487.1.

Brenner, I. S., 1974: A Surge of Maritime Tropical Air-Gulf of California to the Southwestern United States. *Mon. Wea. Rev.*, **102**(5), 375–389, doi:10.1175/1520-

0493(1974)102<0375:ASOMTA>2.0.CO;2.

Carleton, A. M., 1986: Synoptic-dynamic character of "bursts" and "breaks" in the
southwest U.S. summer precipitation singularity. *J. Climatol.*, **6**, 605–623.

Carleton, A. M., D. A. Carpenter, and P. J. Weber, 1990: Mechanisms of interannual
variability of the southwest United States summer rainfall maximum. *J. Climate*,
3, 999–1015.

Castro, C. L., R. A. Pielke Sr., and J. O. Adegoke, 2007: Investigation of the summer
climate of the contiguous United States and Mexico using the Regional
Atmospheric Modeling System (RAMS). Part I: Model climatology (1950–
2002). *J. Climate*, **20,** 89–110.

Chen, S., and Coauthors, 2015: Intercomparison of precipitation estimates from WSR-
88D radar and TRMM measurement over continental United States. *IEEE Trans.
Geosci. Remote Sens.*, **53**, 4444–4456,
https://doi.org/10.1109/TGRS.2015.2399307.

Chou, M. D., and M. J. Suarez, 1999: A solar radiation parameterization for atmospheric
studies. *NASA Tech. Memo.* 104606, **15**, 40 pp.

Chou, M. D., M. J. Suarez, X. Z. Liang, and M. M. H. Yan, 2001: A thermal infrared
radiation parameterization for atmospheric studies. *NASA Tech. Memo.*, 104606,
19, 68 pp.

Collier, J. C., and G. J. Zhang, 2007: Effects of increased horizontal resolution on
simulation of the North American monsoon in the NCAR CAM3: An evaluation
based on surface, satellite, and reanalysis data. *J Climate*, **20**, 1843–1861,
doi:10.1175/JCL14099.1.

Damiani, R., and Coauthors, 2008: The cumulus, photogrammetric, in situ, and doppler
observations experiment of 2006. *Bull. Amer. Meteor. Soc.*, **89,** 57–73.

Dee, D., and Coauthors, 2011a: The ERA-Interim reanalysis: Configuration and
performance of the data assimilation system. *Quart. J. Roy. Meteor.
Soc.*, **137**, 553–597.

Donaldson, R. J., R. M. Dyer and M. J. Krauss, 1975: An objective evaluator of
techniques for predicting severe weather events. Preprints: 9th Conf. Severe Local
Storms, Norman, Oklahoma. Amer. Meteor. Soc., 321-326.

Douglas, M. W., R. A. Maddox, K. Howard, and S. Reyes, 1993: The Mexican
Monsoon. *J. Climate*, **6,** 1665–1677.

Douglas, M. W., and J. C. Leal, 2003: Summertime Surges over the Gulf of California:
Aspects of Their Climatology, Mean Structure, and Evolution from Radiosonde,
NCEP Reanalysis, and Rainfall Data. *Wea. and Forecasting*, **18**, 55–74,
doi:10.1175/1520-0434(2003)018<0055:SSOTGO>2.0.CO;2.

Douglas, A. V., and P. J. Englehart, 2007: A climatological perspective of transient
synoptic features during NAME 2004. *J. Climate*, **20**, 1947-1954.

Finch, Z. O., and R. H. Johnson, 2010: Observational Analysis of an Upper-Level
Inverted Trough during the North American Monsoon Experiment. *Mon. Wea.
Rev*., **138**, 3540-3555.

Fuller, R. D., and D. J. Stensrud, 2000: The relationship between tropical easterly waves
and surges over the Gulf of California during the North American Monsoon. *Mon.
Wea. Rev.*, **128,** 2983–2989.

Gochis, D. J., C. J. Watts, J. Garatuza-Payan, and J. Cesar-Rodriguez, 2007: Spatial and

temporal patterns of precipitation intensity as observed by the NAME Event Rain Gauge Network from 2002 to 2004. *J. Climate*, **20**, 1734–1750.

Griffiths, P. G., C. S. Magirl, R. H. Webb, E. Pytlak, P. A. Troch, and S. W. Lyon, 2009: Spatial distribution and frequency of precipitation during an extreme event: July 2006 mesoscale convective complexes and floods in southeastern Arizona. *Water Resour. Res.*, **45**, W07419, doi:10.1029/2008WR007380.

Gutzler, D. S., and Co-authors, 2009: Simulations of the 2004 North American Monsoon: NAMAP2. *J. Climate*, **22**, 6716-6740.

Hales, J., 1972: Surges of Maritime Tropical Air Northward Over Gulf of California. *Mon. Wea. Rev.*, **100**, 298–306, doi: 10.1175/1520-0493(1972)100<0298:SOMTAN>2.3.CO;2.

Higgins, W., and W. Shi, 2005: Relationships between Gulf of California moisture surges and tropical cyclones in the eastern Pacific basin. *J. Climate*, **18**, 4601–4620, https://doi.org/10.1175/JCLI3551.1.

Higgins, W., and D. Gochis, 2007: Synthesis of results from the North American Monsoon Experiment (NAME) process study. *J. Climate*, **20**, 1601–1607, doi:10.1175/JCLI4081.1

Higgins, R., W. Shi, and C. Hain, 2004: Relationships between Gulf of California moisture surges and precipitation in the southwestern United States. *J. Climate*, **17**, 2983–2997.

Higgins, R.W. and Coauthors, 2006: The NAME 2004 field campaign and modeling strategy. *Bull. Amer. Meteor. Soc.*, **87**, 79-94.

Hodges, K. I., 1994: A general method for tracking analysis and its application to

meteorological data. *Mon. Wea. Rev.*, **122**, 2573–2586.

Hodges, K. I., 1995: Feature tracking on the unit sphere. *Mon. Wea. Rev.*, **123,** 3458–3465.

Hong, S.–Y., and J.–O. J. Lim, 2006: The WRF single–moment 6–class microphysics scheme (WSM6). *J. Korean Meteor. Soc.*, **42**, 129–151.

Hong, S.–Y., Y. Noh, J. Dudhia, 2006: A new vertical diffusion package with an explicit treatment of entrainment processes. *Mon. Wea. Rev.*, **134**, 2318–2341, doi: 10.1175/MWR3199.1.

Huffman, G. J., R. F. Adler, D. T. Bolvin, G. Gu, E. J. Nelkin, K. P. Bowman, E. F. Stocker, and D. B. Wolff, 2007: The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. *J. Hydrometeor.*, **8,** 38–55.

Iacono, M. J., J. S. Delamere, E. J. Mlawer, M. W. Shephard, S. A. Clough, and W. D. Collins, 2008: Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models. *J. Geophys. Res.*, **113**, D13103, https://doi.org/10.1029/2008JD009944

Johnson, R., P. Ciesielski, and B. McNoldy, 2007: Multiscale variability of the flow during the North American Monsoon Experiment. *J. Climate*, **20**, 1628–1648, doi: 10.1175/JAS-D-13-065.1.

Joyce, R. J., J. E. Janowiak, P. A. Arkin, and P. Xie, 2004: CMORPH: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. *J. Hydrometeor.*, **5**, 487–503, doi: 10.1175/1525-7541(2004)005,0487:CAMTPG.2.0.CO;2.

Kain, J. S., 2004: The Kain–Fritsch convective parameterization: An update. *J. Appl. Meteor.*, **43**, 170–181, doi: 10.1175/1520-0450(2004)043<0170:TKCPAU>2.0.CO;2.

Kubota, T., and Coauthors, 2007: Global precipitation map using satellite borne microwave radiometers by the GSMaP Project: Production and validation. *IEEE Trans. Geosci. Remote Sens.*, **45**, 2259–2275.

Kursinski, E. R., D. K. Adams, and M. Leuthold, 2008a: GPS observations of precipitable water and implications for the predictability of precipitation during the North American Monsoon. *CLIVAR Exchanges*, **13**, 14.

Kursinski, E. R., R. A. Bennett, D. Gochis, S. I. Gutman, K. L. Holub, R. Mastaler, C. Minjarez Sosa, I. Minjarez Sosa, and T. van Hove, 2008b): Water vapor and surface observations in northwestern Mexico during the 2004 NAME Enhanced Observing Period, *Geophys. Res. Lett.*, **35**, L03815, doi:10.1029/2007GL031404.

Lahmers, T. M., C. L. Castro, D. K. Adams, Y. L. Serra, J. J. Brost, and T. Luong, 2016: Long-term changes in the climatology of transient inverted troughs over the North American monsoon region and their effects on precipitation. *J. Climate*, *29*, 6027-6064.

Lang, T., D. Ahijevych, S. Nesbitt, R. E. Carbone, S. A. Rutledge, and R. Cifelli, 2007: Radar-observed characteristics of precipitating systems during NAME 2004. *J. Climate*, **20**, 1713–1733, doi:10.1175/JCLI4082.1.

Lee, M.-I., and Coauthors, 2007: Sensitivity to horizontal resolution in the AGCM simulations of warm season diurnal cycle of precipitation over the United States and northern Mexico. *J Climate*, **20**, 1862–1881, doi:10.1175/JCL14090.1.

Li, J., S. Sorooshian, W. Higgins, X. Gao, B. Imam, and K. Hsu, 2008: Influence of Spatial Resolution on Diurnal Variability during the North American Monsoon. *J. Climate*, **21**, 3967–3988, doi:10.1175/2008JCLI2022.1.

Livezey, R. E., and W. Y. Chen, 1983: Statistical significance and its determination by Monte Carlo techniques. *Mon. Wea. Rev*., **111**, 46–59, doi:10.1175/1520-0493(1983)111,0046:SFSAID.2.0.CO;2

Luong, T.M., C.L. Castro, H. Chang, T. Lahmers, D.K. Adams, and C.A. Ochoa-Moya, 2017: The More Extreme Nature of North American Monsoon Precipitation in the Southwestern United States as Revealed by a Historical Climatology of Simulated Severe Weather Events. *J. Appl. Meteor. Climatol.*, **56**, 2509–2529, https://doi.org/10.1175/JAMC-D-16-0358.1

Magirl, C. S., and Coauthors, 2007: Impact of recent extreme Arizona storms. Eos, *Trans Amer. Geophys. Union*, **88**, 191-193.

Mason, I., 1989: Dependence of the critical success index on sample climate and threshold probability. *Aust. Meteor. Mag.,* **37**, 75-81.

McCollum, D., R. Maddox, and K. Howard, 1995: Case study of a severe mesoscale convective system in central Arizona. *Wea. Forecasting*, **10**, 643–665.

Mesinger, F., and Coauthors, 2006: North American Regional Reanalysis. *Bull. Amer. Meteor. Soc.*, **87**, 343–360, doi:10.1175/BAMS-87-3-343.

Moncrieff, M., and M. Miller, 1976: The dynamics and simulation of tropical cumulonimbus and squall lines. *Quart. J. Roy. Meteor. Soc.*, **102,** 373–394.

Moore, A., I. Small, S. Gutman, Y. Bock, J. Dumas, P. Fang, J. Haase, M. Jackson, and J. Laber, 2015: National Weather Service Forecasters Use GPS Precipitable Water

Vapor for Enhanced Situational Awareness during the Southern California Summer Monsoon. *Bull. Amer. Meteor. Soc.,* **96**, 1867–1877, doi:10.1175/BAMS-D-14-00095.1.

Nesbitt, S. W., D. J. Gochis, and T. J. Lang, 2008: The Diurnal Cycle of Clouds and Precipitation along the Sierra Madre Occidental Observed during NAME-2004: Implications for Warm Season Precipitation Estimation in Complex Terrain. *J. Hydrometeorol.*, **9**, 728–743, doi:10.1175/2008JHM939.1.

Newman, A., and R. H. Johnson, 2012: Mechanisms for Precipitation Enhancement in a North American Monsoon Upper-Tropospheric Trough. *J. Atmos. Sci.*, **69**, 1775–1792, doi:10.1175/JAS-D-11-0223.1.

Okamoto, K. I., T. Ushio, T. Iguchi, N. Takahashi, and K. Iwanami, 2005: The global satellite mapping of precipitation (GSMaP) project. *Proc. 25th Int. Symp. on Geoscience and Remote Sensing,* Seoul, South Korea, IEEE, 3414–3416.

Pytlak, E., M. Goering, and A. Bennett, 2005: Upper Tropospheric Troughs and Their Interaction with the North American Monsoon. *19th Conf. on Hydrology*, San Diego, CA, Amer. Meteor. Soc., 1–5. https://ams.confex.com/ams/pdfpapers/85393.pdf.

Raymond, D. J., and M. H. Wilkening, 1980: Mountain induced convection under fair weather conditions. *J. Atmos. Sci.*, **37,** 2693–2706.

Rogers, P. J., and R. H. Johnson, 2007: Analysis of the 13–14 July Gulf Surge Event during the 2004 North American Monsoon Experiment. *Mon. Wea. Rev.*, **135**, 3098–3117, doi:10.1175/MWR3450.1.

Rowe, A. K., S. A. Rutledge, and T. J. Lang, 2012: Investigation of Microphysical

Processes Occurring in Organized Convection during NAME. *Mon. Wea. Rev.*, **140**, 2168–2187, doi:10.1175/MWR-D-11-00124.1.

Schaefer, J. T., 1990: The critical success index as an indicator of warning skill. *Wea. Forecasting*, **5**, 570–575, https://doi.org/10.1175/1520-0434(1990)005<0570:TCSIAA>2.0.CO;2.

Seastrand, S., Y. Serra, C. L. Castro, and E. A. Ritchie, 2014: The dominant synoptic-scale modes of North American monsoon precipitation. *Int. J. Climatol.*, **35**(8), 2019–2032, doi:10.1002/joc.4104.

Serra, Y. L., G. N. Kiladis, and K. I. Hodges, 2010: Tracking and mean structure of easterly waves over the Intra-Americas Sea. *J. Climate*, **23**, 4823–4840

Serra, Y. L., D. K. Adams, C. Minjarez-Sosa, J. M. Moker, A. F. Arellano, C. L. Castro, A. I. Quintanar, L. Alatorre, A. Granados, G. Vazquez, K. Holub, and C. C. DeMets, 2016: The North American Monsoon GPS Transect Experiment 2013. *Bull. Amer. Meteor. Soc.*, **97**, 2103–2115, doi:10.1175/BAMS-D-14-00250.1.

Skamarock, W. C., J. B. Klemp, J. Dudhia, D. O. Gill, D. M. Barker, W. Wang, and J. G. Powers, 2005: A description of the Advanced Research WRF version 2. NCAR Tech. Note NCAR/TN-468+STR, 88 pp

Skamarock, W. C., J. B. Klemp, J. Dudhia, D. O. Gill, D. M. Barker, X.-Y. Huang, W. Wang, and J. G. Powers, 2008: A Description of the Advanced Research WRF Version 3. *NCAR Technical Note No. 475*, pp. 125.

Sorooshian, S., K. Hsu, X. Gao, H. V. Gupta, B. Imam, and D. Braithwaite, 2000: Evaluation of PERSIANN system satellite-based estimates of tropical rainfall. *Bull. Amer. Meteor. Soc.*, **81**, 2035–2046, doi:10.1175/1520-

0477(2000)081,2035: EOPSSE.2.3.CO;2.

Stillman, S., X. Zeng, and M.G. Bosilovich, 2016: Evaluation of 22 Precipitation and 23
 Soil Moisture Products over a Semiarid Area in Southeastern Arizona. *J.
 Hydrometeor.*, **17**, 211–230, https://doi.org/10.1175/JHM-D-15-0007.1

Tewari, M., F. Chen, W. Wang, J. Dudhia, M. A. LeMone, K. Mitchell, M. Ek, G.
 Gayno, J. Wegiel, and R. H. Cuenca, 2004: Implementation and verification of the
 unified NOAH land surface model in the WRF model. *20th conference on
 weather analysis and forecasting/16th conference on numerical weather
 prediction*, pp. 11–15.

Thorncroft, C. D., and K. Hodges, 2001: African easterly wave variability and its
 relationship to Atlantic tropical cyclone activity. *J. Climate*, **14**, 1166–1179,
 doi:10.1175/1520-0442(2001)014<1166:AEWVAI>2.0.CO;2.

Tian, Y., C. D. Peters-Lidard, R. F. Adler, T. Kubota, and T. Ushio, 2010: Evaluation of
 GSMaP precipitation estimates over the contiguous United States. *J.
 Hydrometeor.*, **11**, 566–574, https://doi.org/10.1175/2009JHM1190.1.

Torn, R. D., and G. J. Hakim, 2008: Ensemble-based sensitivity analysis. *Mon. Wea.
 Rev.*, **136** (2), 663-677, doi:10.1175/2007MWR2132.1.

Ushio, T., and Coauthors, 2009: A Kalman filter approach to the Global Satellite
 Mapping of Precipitation (GSMaP) from combined passive microwave and
 infrared radiometric data. *J. Meteor. Soc. Japan*, **87A**, 137–151.

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

Wood, K. M. and E. A. Ritchie, 2013: An updated climatology of tropical cyclone

impacts on the southwestern United States. *Mon. Weather Rev.* **141**: 4322–4336.

Zehnder, J. A., 2004: Dynamic mechanisms of the gulf surge. *J Geophys Res-Atmos*, **109**, doi:10.1029/2004JD004616.

Zuidema, P., C. Fairall, L. M. Hartten, J. E. Hare, and D. Wolfe, 2007: On the air–sea interaction at the mouth of the Gulf of California. *J. Climate*, **20,** 1649–1661.

Tables

Table 1: List of station names, locations, and elevations for each station in each transect within Transect 2013 along with its purpose.

	Station	Lat (°N)	Lon (°W)	Elevation (m MSL)	Purpose
SMO Transect	KINO	28.8149	111.9287	7	To observe the development of MCSs in the core region at high temporal resolution from convective initiation at the northern SMO crest (BASC-CUAH) to organization and propagation along the western SMO slope/foothills (MULT-ONVS) and towards the GoC (KINO)
	ONVS	28.4602	109.5288	189	
	MULT	28.6356	108.7595	1550	
	BASC	28.2035	108.2098	1999	
	CUAH	28.4079	106.8922	2058	
	CHIH	28.6224	106.1006	1463	
	Station	**Lat (°N)**	**Lon (°W)**	**Elevation (m MSL)**	**Purpose**
Coastal Transect	MOCH	25.7815	109.0264	15	To observe the propagation of gulf surges along the GoC from south to north (MOCH-KINO-PSCO)
	KINO	28.8149	111.9287	7	
	PSCO	31.3004	113.5483	53	
	Station	**Lat (°N)**	**Lon (°W)**	**Elevation (m MSL)**	**Purpose**
Southern Transect	MOCH	25.7815	109.0264	15	To observe the strong precipitation gradient between the central GoC (MOCH) and the central SMO foothills (BGTO)
	BGTO	25.3625	107.5511	207	

Table 2: Satellite-based precipitation datasets available during the period of study.

Product	Source	Time Resolution	Spatial Resolution	References
CMORPH	NOAA	3-hourly	0.25°	Joyce et al. (2004)
GSMaP	JAXA	Hourly	0.10°	Okamoto et al. (2005) Kubota et al. (2007) Aonashi et al. (2009) Ushio et al. (2009)
PERSIANN	UCI	3-hourly	0.25°	Sorooshian et al. (2000)
TRMM	NASA	3-hourly	0.25°	Huffman et al. (2007)

Table 3: Descriptions of other observational and atmospheric reanalysis datasets used in this study.

Dataset	Source	Purpose
Surface meteorological data from Naval Air Station Yuma (KNYL)	Mesowest (http://mesowest.utah.edu)	To calculate a gulf surge in addition to the coastal transect
0.25° 3-hourly TMPA (Huffman et al. 2007)	NASA's Goddard Earth Sciences Data and Information Services Center (https://mirador.gsfc.nasa.gov/)	Observed precipitation to compare with hindcasts
NOAA's 4-km GOES-East satellite imagery (water vapor IR channel)	Iowa Environmental Mesonet (http://mesonet.agron.iastate.edu/archive/data)	To subjectively depict IVs
NARR dataset (29 vertical levels, 32-km horizontal resolution, 3-h temporal resolution) (Mesinger et al. 2006)	NOAA Operational Model Archive and Distribution System (NOMADS) (https://nomads.ncdc.noaa.gov/data/narr/)	A secondary tool to visualize IVs
ERA-Interim dataset (60 vertical levels, ~70-km horizontal resolution, 3-h temporal resolution) (Dee et al. 2011)	ECMWF Public Datasets (http://apps.ecmwf.int/datasets)	To track TEWs
"Past track seasonal maps" in the Eastern Pacific for 2013 based on the HURDAT "best track" database	NOAA's National Hurricane Center (http://www.nhc.noaa.gov/data)	To identify TCs

Table 4: Horizontal resolutions of the 3 nested WRF-ARW domains along with the cumulus parameterization scheme, if used.

	d01	d02	d03
Horizontal Resolution	30 km	10 km	2.5 km
Cumulus Parameterization	Kain-Fritsch	Kain-Fritsch	none

Table 5: Listing of the physics schemes that are used in our WRF-ARW configuration and applied to all domains.

Category	Scheme	Reference
Microphysics	WRF single-moment 6-class	Hong and Lim (2006)
Planetary Boundary Layer	Yonsei University	Hong et al. (2006)
Longwave Radiation	Rapid Radiative Transfer Model	Iacona et al. (2008)
Shortwave Radiation	Goddard	Chou and Suarez (1999); Chou et al. (2001)
Land Surface Model	Unified Noah	Tewari et al. (2004)

Table 6: A 2×2 forecast contingency table as input for the CSI, POD, and FAR metrics to calculate precipitation forecast skill for each grid point.

		Observed?	
		Yes	No
Forecast?	Yes	A (hits)	B (false alarms)
	No	C (misses)	D (correct negatives)

Table 7: List of strongly forced days and weakly forced days based on the presence of an inverted trough and appreciable TRMM rainfall in northwest Mexico. Because no lateral boundary conditions were unavailable for 7 Jul (marked with asterisk), a weakly forced day, that day was omitted from the analysis. With the exception of 7 Jul, there are 22 strongly forced days and 40 weakly forced days during Transect 2013.

Strongly forced days	Weakly forced days	
9 Jul	26 Jun	5 Aug
10 Jul	30 Jun	9 Aug
12 Jul	1 Jul	10 Aug
13 Jul	2 Jul	11 Aug
17 Jul	3 Jul	12 Aug
18 Jul	4 Jul	13 Aug
22 Jul	5 Jul	14 Aug
31 Jul	7 Jul*	18 Aug
1 Aug	8 Jul	19 Aug
2 Aug	11 Jul	23 Aug
15 Aug	15 Jul	24 Aug
16 Aug	16 Jul	28 Aug
20 Aug	23 Jul	1 Sep
21 Aug	24 Jul	6 Sep
22 Aug	25 Jul	7 Sep
25 Aug	26 Jul	8 Sep
26 Aug	27 Jul	9 Sep
27 Aug	28 Jul	10 Sep
30 Aug	29 Jul	12 Sep
31 Aug	30 Jul	
3 Sep	3 Aug	
5 Sep	4 Aug	

Table 8: Grid-based RMSE mean of daily rainfall accumulation (mm day^{-1}) from each satellite-based precipitation product and each hindcast simulation for the duration of Transect 2013. Values are listed in ascending order.

Product	WRF-NAM	WRF-GFS
TRMM	8.7	7.8
CMORPH	9.3	8.6
GSMaP	9.6	8.8
PERSIANN	9.9	9.2

Table 9: Mean grid precipitation rate for TRMM, WRF-NAM, WRF-GFS, and model bias (WRF minus TRMM) for the diurnal cycle (mm hr^{-1}) for 3-h and 24-h intervals of the hindcasts during the Transect 2013 period. Highest hourly mean precipitation rate and model biases are bolded.

Time (UTC)	TRMM	WRF-NAM	WRF-NAM bias	WRF-GFS	WRF-GFS bias
1200-1500	0.05	0.07	+0.02	0.02	-0.03
1500-1800	0.04	0.13	+0.09	0.05	+0.01
1800-2100	0.09	0.21	**+0.12**	0.14	+0.05
2100-0000	0.16	**0.28**	+0.10	**0.24**	**+0.06**
0000-0300	**0.18**	0.21	+0.01	0.20	-0.01
0300-0600	0.15	0.11	-0.06	0.11	-0.06
0600-0900	0.12	0.07	-0.07	0.07	-0.06
0900-1200	0.08	0.05	-0.04	0.05	-0.04
1200-1200	0.11	0.14	+0.02	0.11	-0.01

Table 10: For both the WRF-NAM (top) and WRF-GFS (bottom) hindcasts, field significance values of the differences in forecast metrics between the two subsets of strongly and weakly forced days are listed. Field significance levels that are greater than 90% are in bold and those between 80% and 90% are in italics. Also, pattern correlations between the subdaily and daily grids are shown with the greatest pattern correlation in bold.

WRF-NAM	$CSI_{strong} - CSI_{weak}$		$POD_{strong} - POD_{weak}$		$FAR_{strong} - FAR_{weak}$	
Time (UTC)	Field Significance Level	Pattern Correlation to Daily	Field Significance Level	Pattern Correlation to Daily	Field Significance Level	Pattern Correlation to Daily
1800-0000	**97.4%**	+0.39	81.3%	**+0.41**	46.4%	+0.38
0000-0600	28.6%	**+0.43**	0.3%	+0.30	74.7%	**+0.42**
0600-1200	**>99%**	+0.40	**94.1%**	+0.26	**>99%**	+0.10
1200-1200	**>99%**		**98.1%**		56.1%	

WRF-GFS	$CSI_{strong} - CSI_{weak}$		$POD_{strong} - POD_{weak}$		$FAR_{strong} - FAR_{weak}$	
Time (UTC)	Field Significance Level	Pattern Correlation to Daily	Field Significance Level	Pattern Correlation to Daily	Field Significance Level	Pattern Correlation to Daily
1800-0000	**>99%**	**+0.55**	**98.5%**	**+0.50**	**94.8%**	**+0.55**
0000-0600	17.4%	+0.45	2.3%	+0.43	55.5%	+0.45
0600-1200	**>99%**	+0.45	*82.6%*	+0.19	**>99%**	+0.40
1200-1200	**98.7%**		**93.4%**		**93.6%**	

Figures
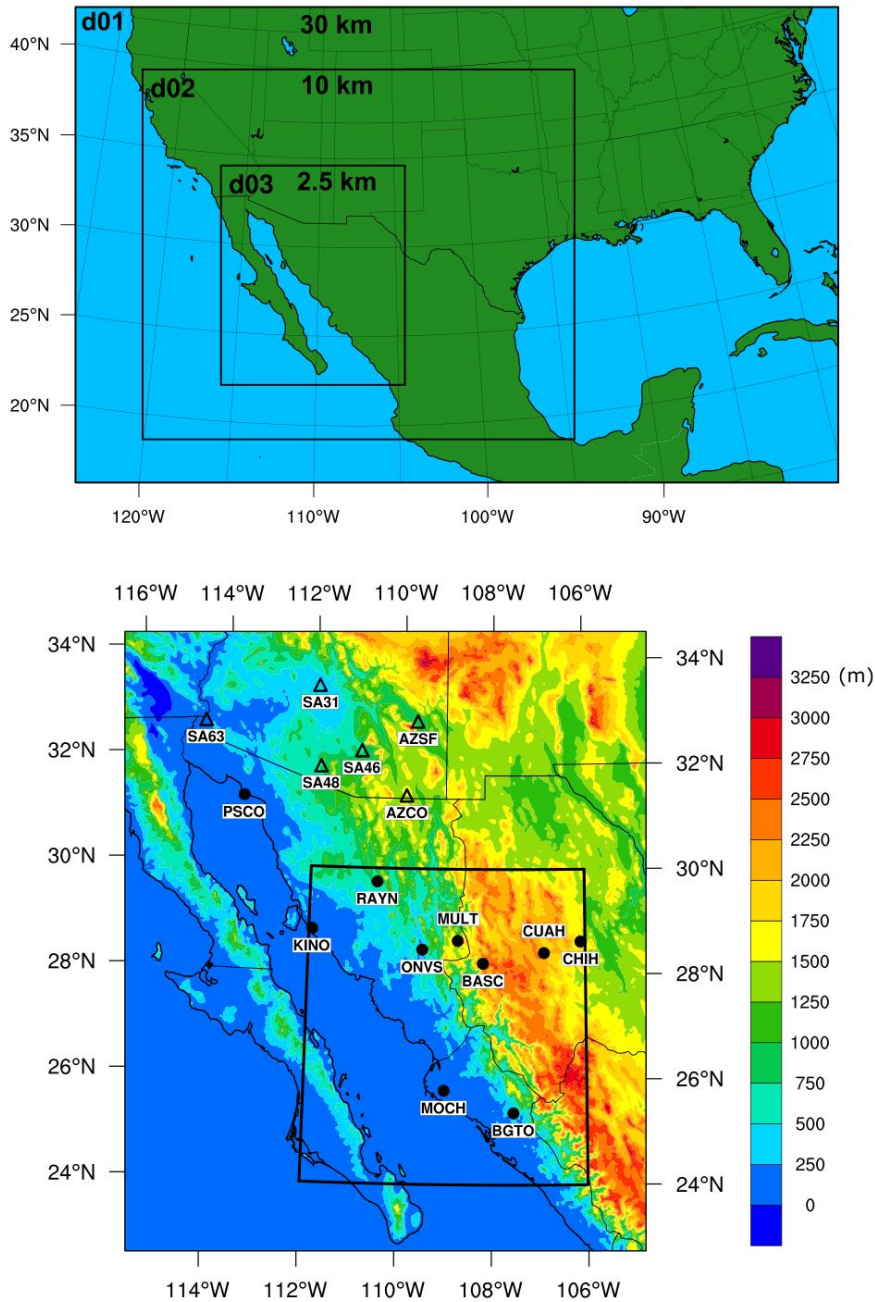
Figure 1: (top) The WRF-ARW nested domain configuration in our hindcast setup. Domain d01 has 159 (west–east) × 99 (south–north) grid points with a horizontal spacing of 30 km, domain d02 has 270 × 231 grid points with a horizontal spacing of 10 km, and the innermost domain d03 has a 460 × 548 grid points with a horizontal spacing of 2.5 km. (bottom) Domain d03 with locations of the Transect 2013 (circles) and SuomiNet (triangles) stations as well as terrain shaded every 250 m. The SMO transect (KINO–CHIH) is situated within the northern third of the North American monsoon core region outlined in black and bounded by 24°–30°N and 112°–106°W.
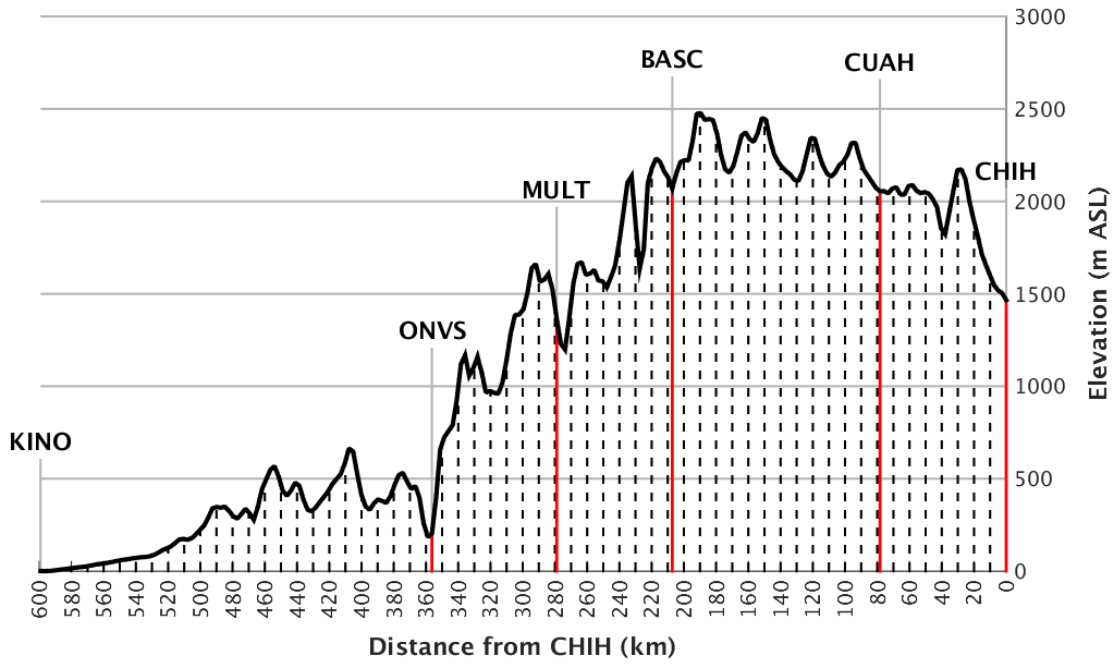
Figure 2: A west-to-east cross section of the mountainous terrain (m ASL) shows the locations of the SMO transect stations indicated by red solid vertical lines. Distance relative to the easternmost station CHIH is indicated by black dashed lines along the abscissa.
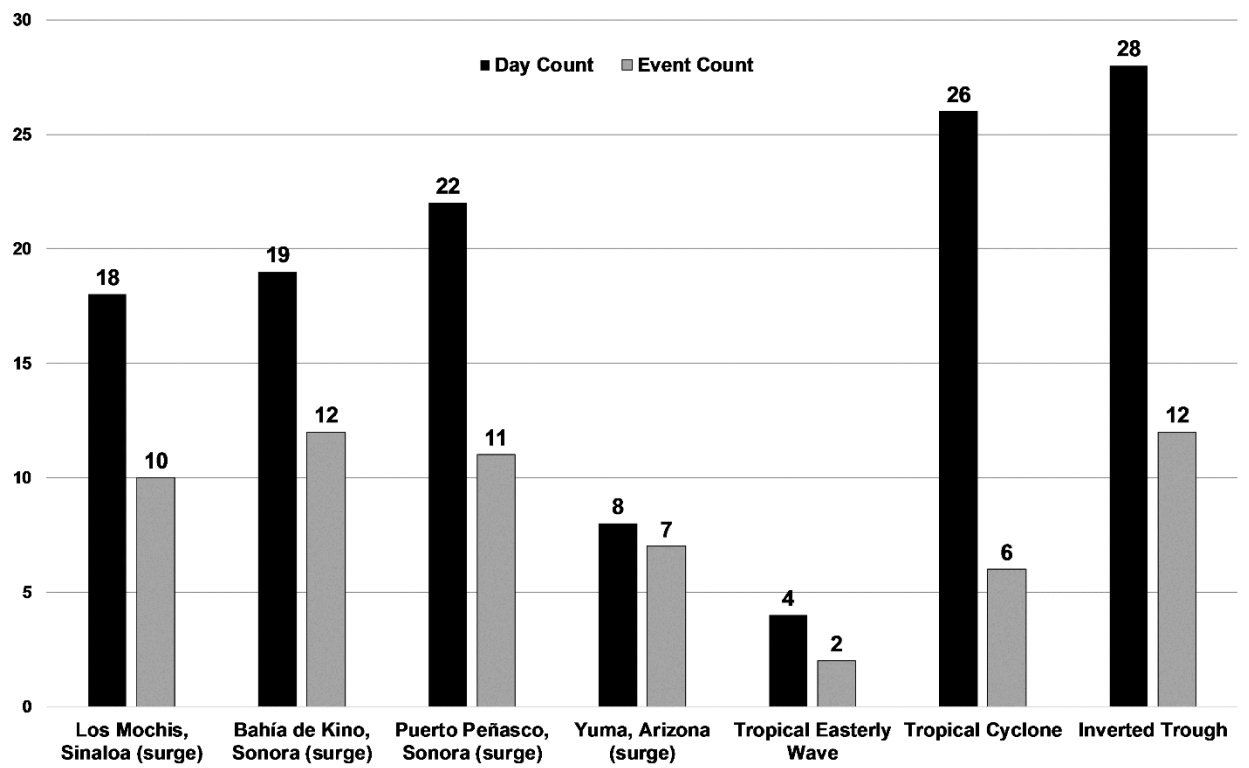
Figure 3: Counts of synoptic forcing mechanisms during the 2013 season (27 Jun–9 Sep) that have been demonstrated to impact North American monsoon convection in northwest Mexico and adjacent southwest United States. The first four cities indicate sites recording gulf surges. The event count (gray) shows the actual number of forcing mechanisms, while the day count (black) is the number of days that were affected by that particular forcing mechanism.
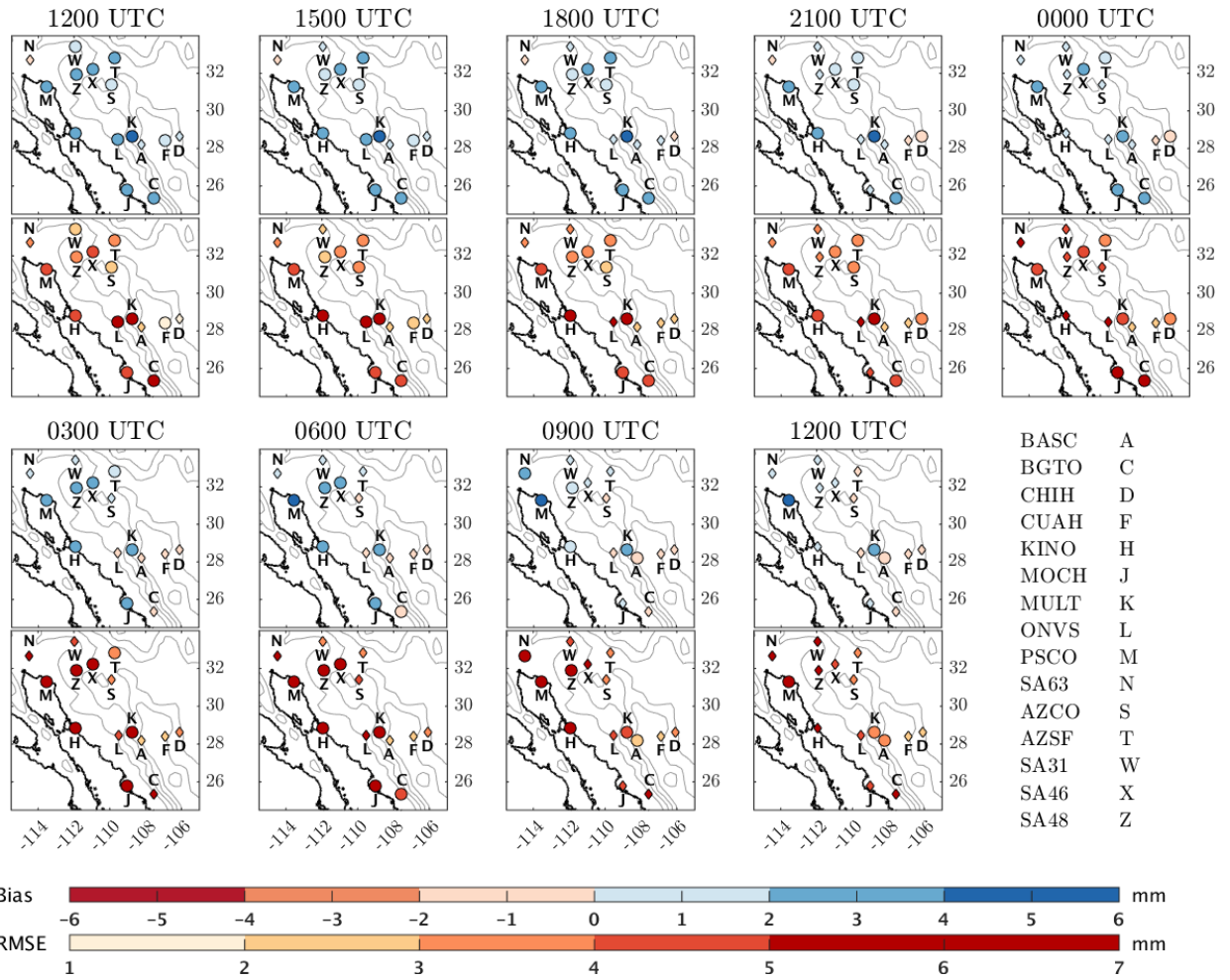
Figure 4: Mean 3-hourly model bias (top plot of each hour) and RMSE (bottom plot of each hour) of PWV and for all WRF-NAM hindcasts relative to GPS-derived PWV across northwest Mexico and adjacent southwest United States (Transect 2013 and selected SuomiNet stations) within the diurnal cycle. Circles indicate stations with model biases that are statistically significant (p value < 0.01). Terrain is contoured every 500 m (gray lines).
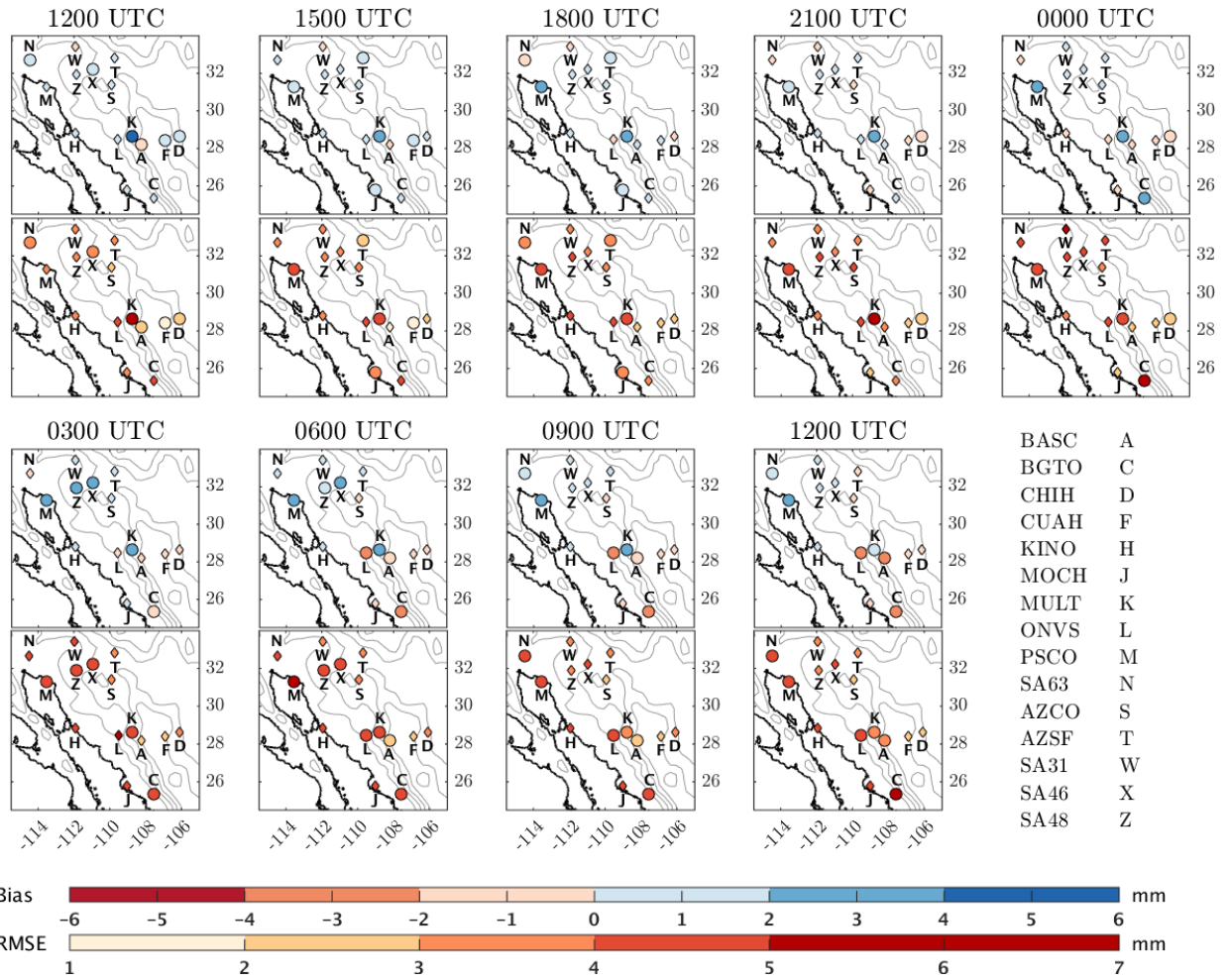
Figure 5: As in Fig. 4, but for the WRF-GFS hindcasts.
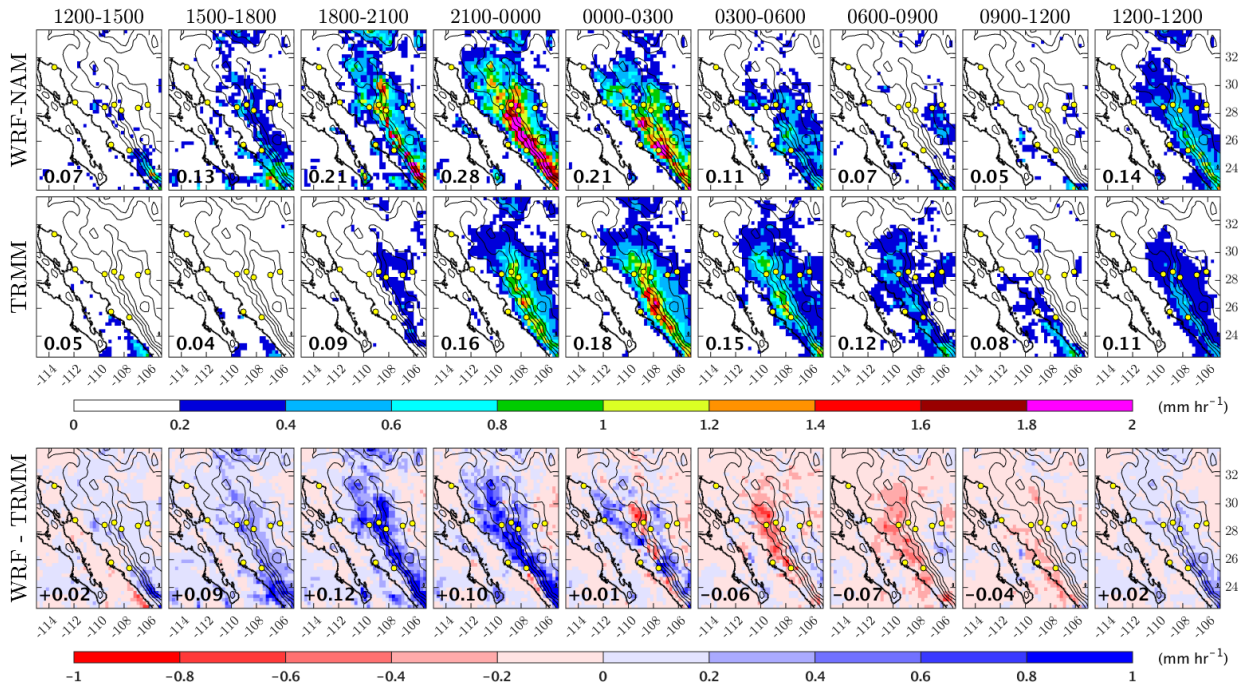
Figure 6: (top) Mean hourly rainfall rate of WRF-NAM hindcasts, (middle) TRMM precipitation analysis, and (bottom) model bias (WRF minus TRMM) for the combined set of strongly and weakly forced days within diurnal cycle for 3-hourly intervals (columns 1–8) and 24 h (column 9). The WRF hindcasts were scaled up to the TRMM 0.25° grid using the "conserve" method of the ESMF software function within NCL. Mean values across all grid boxes for rainfall rate (top two rows) and biases (bottom row) are shown in the bottom-left corner of each map. Blue (red) pixels indicate a wet (dry) model bias. Yellow circles indicate the locations of a subset of the Transect 2013. Terrain is contoured every 500 m (black lines).
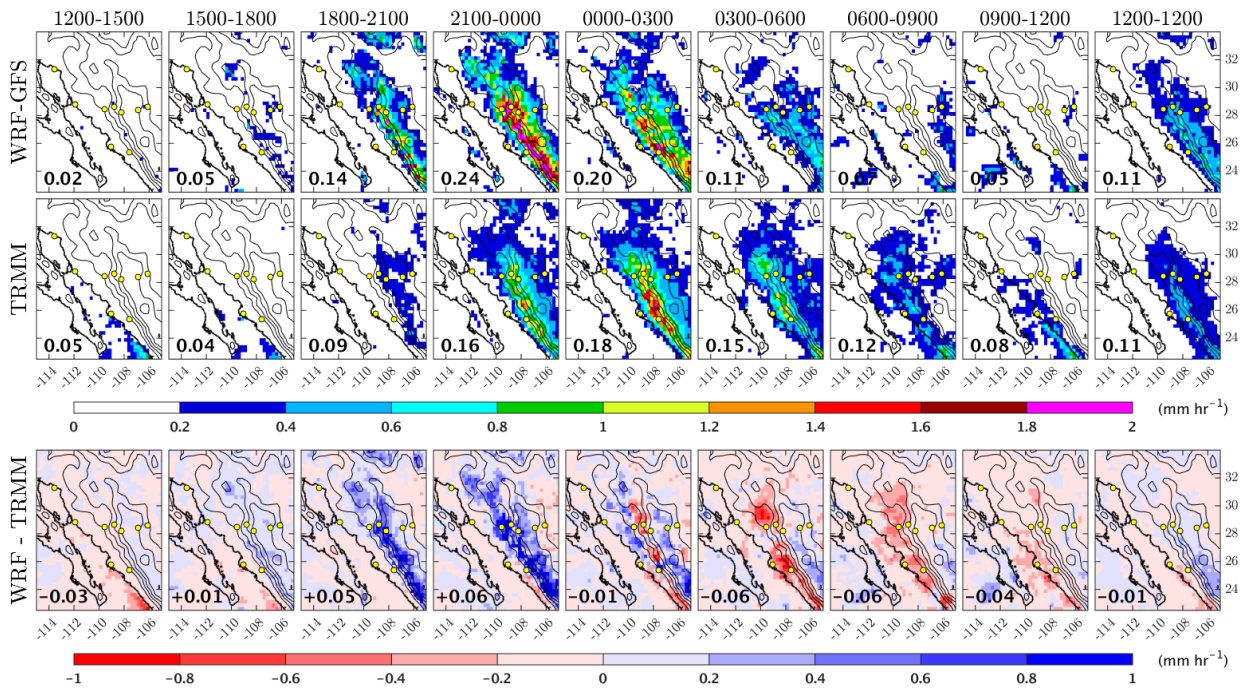
Figure 7: As in Fig. 6, but for the WRF-GFS hindcasts.
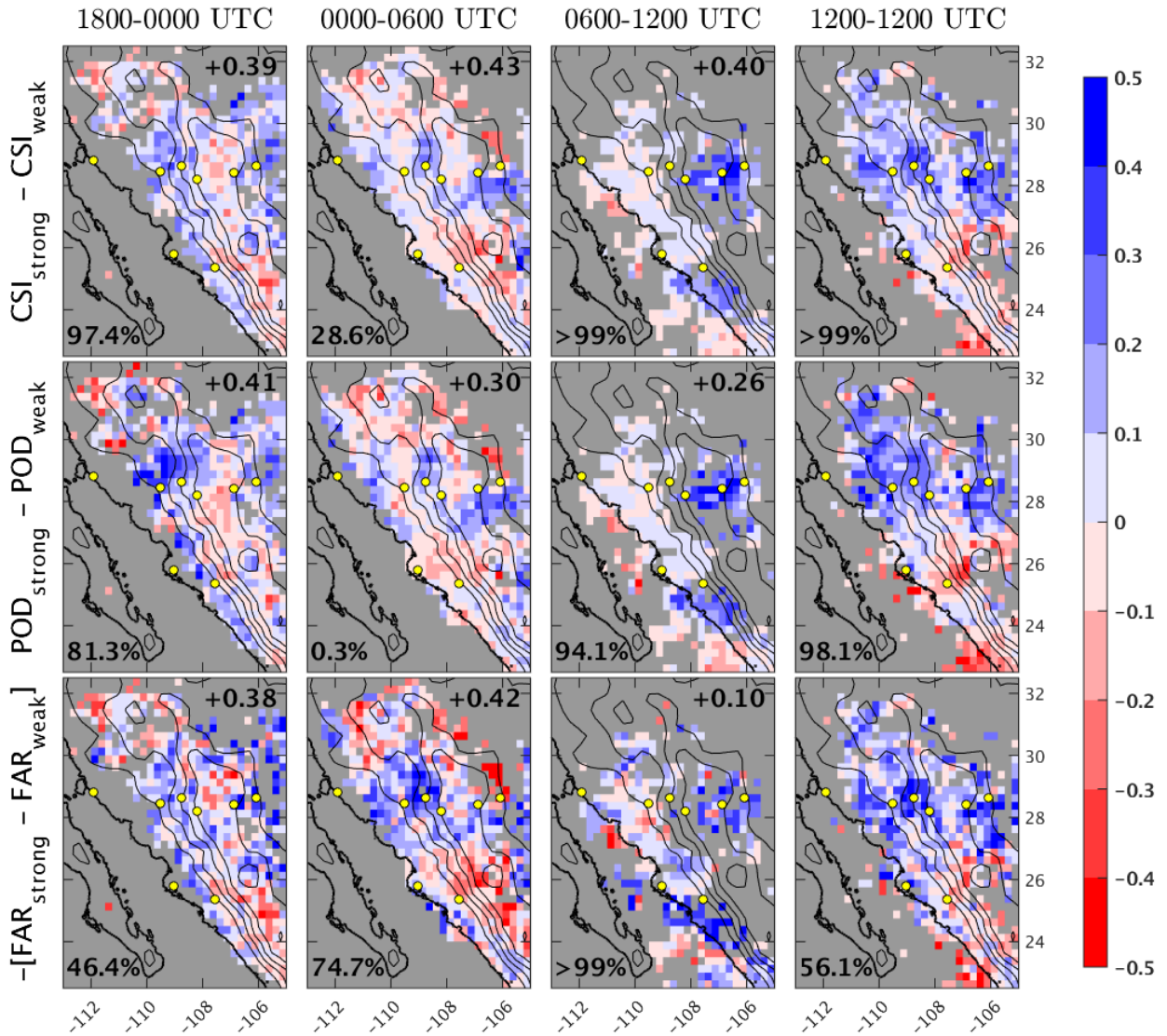
Figure 8: Differences between the strongly and weakly forced days of grid-based (top) CSI, (middle) POD, and (bottom) FAR forecast verification metrics for WRF-NAM hindcasts (modeled) and TRMM (observed) rainfall for the 1800–0000 UTC (column 1), 0000–0600 UTC (column 2), 0600–1200 UTC (column 3), and 1200–1200 UTC (daily) (column 4) periods. Increased forecast skill is shown in red (blue) for the strongly (weakly) forced days. Gray pixels indicate where metrics could not be computed. See text for metrics description and local significance methodology. The statistical field significance rank of each map is displayed in the bottom-left corner of each map and was obtained via 1000 permutations in a Monte Carlo resampling technique. The pattern correlation of metric between the subdaily (columns 1–3) and daily (column 4) forecast metric is given in the top-right corner of each subdaily map. Yellow circles indicate the locations of a subset of Transect 2013. Terrain is contoured every 500 m (black lines).
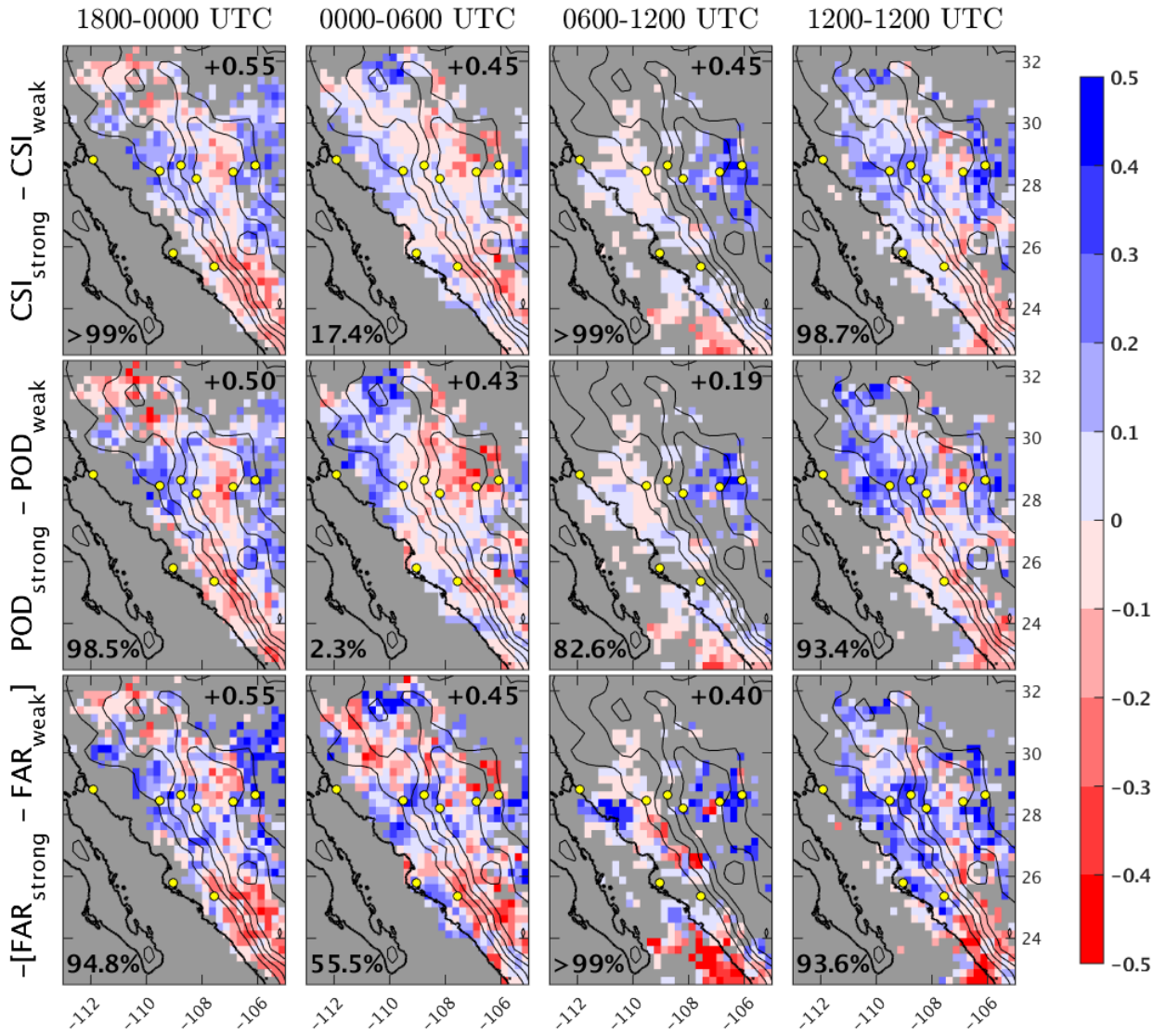
Figure 9: As in Fig. 8, but forecast metrics calculated for the WRF-GFS hindcasts.
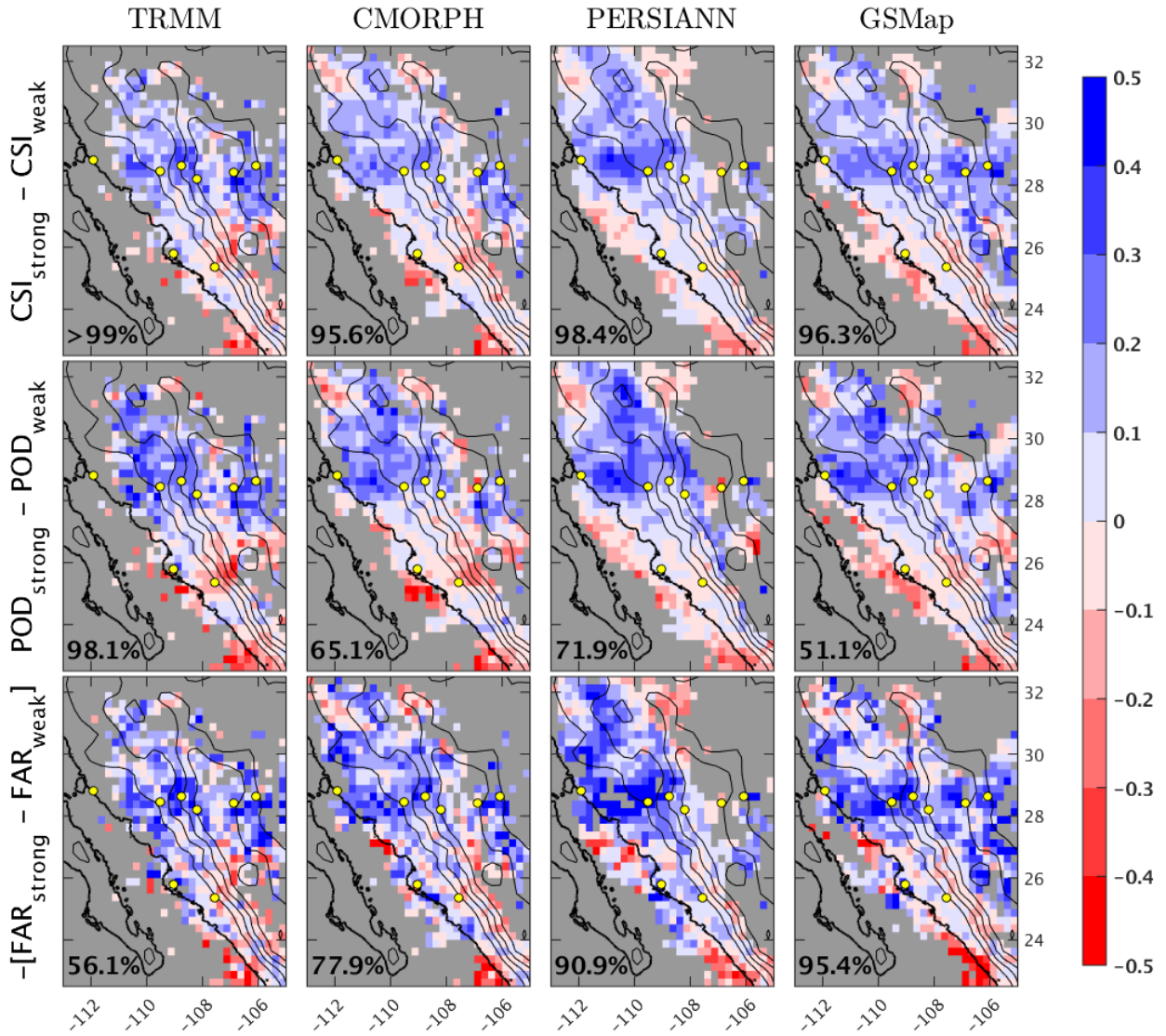
Figure 10: Differences between the strongly and weakly forced days of grid-based daily rainfall forecast verification metrics for WRF-NAM hindcasts and each satellite-based precipitation product of TRMM (column 1), CMORPH (column 2), PERSIANN (column 3), and GSMaP (column 4). The metrics include (top) CSI, (middle) POD, and (bottom) FAR. See Fig. 8 caption for description and text for complete methodology.

92
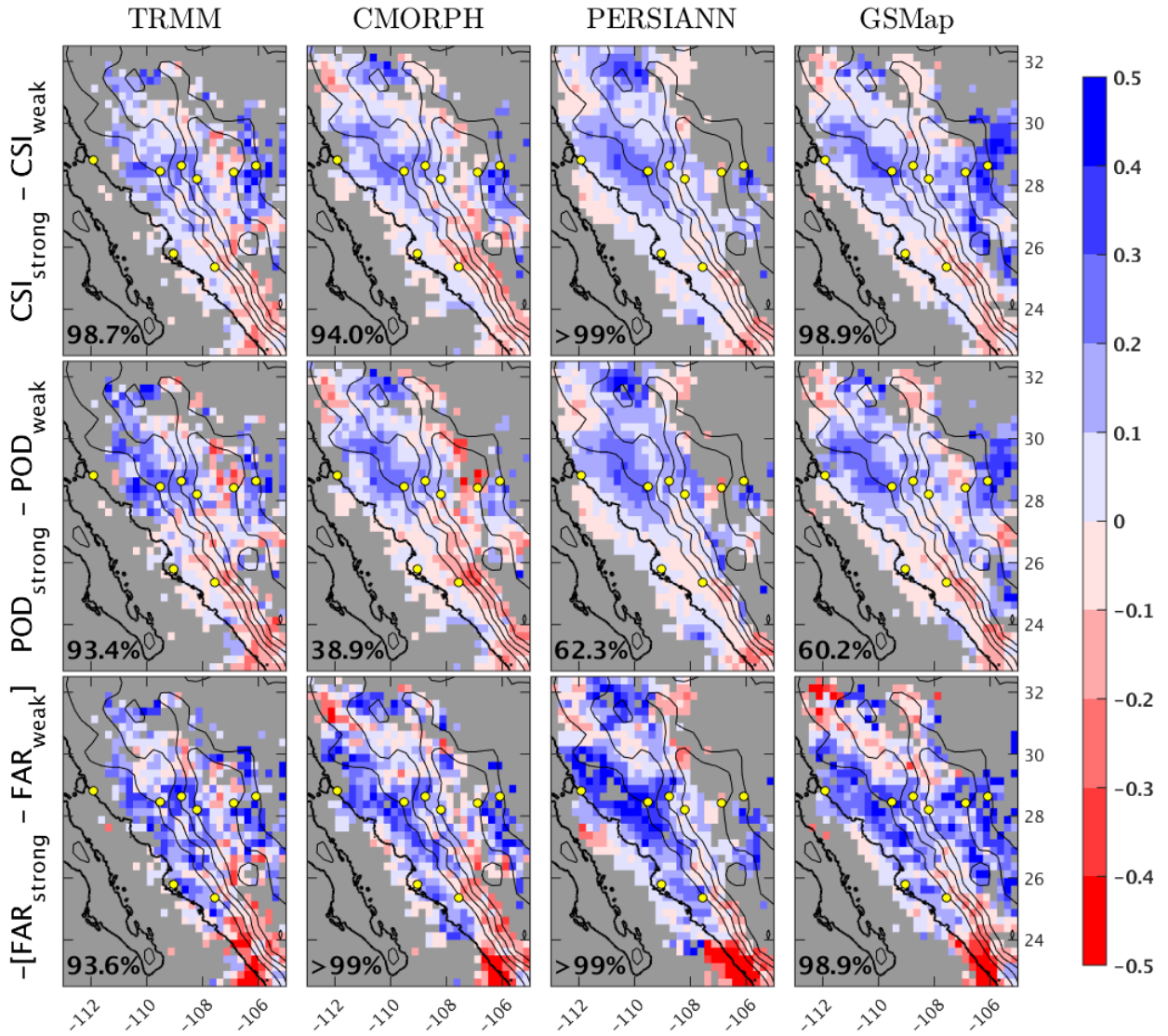
Figure 11: As in Fig. 10, but forecast metrics calculated for the WRF-GFS hindcasts.
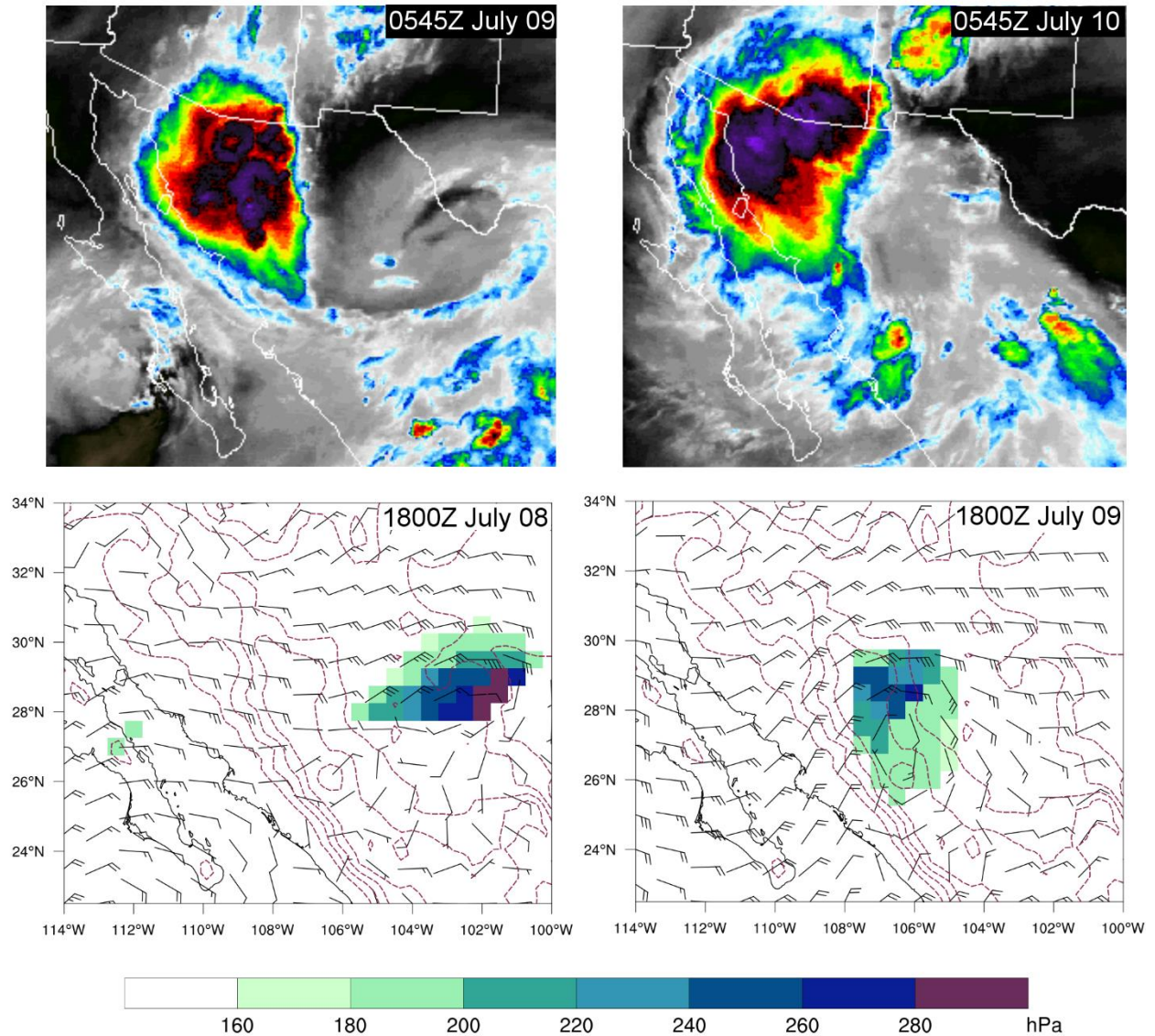
Figure 12: (bottom) Pressure (shaded every 20 hPa below 160 hPa) and winds (kt) are displayed on the 2.0 PVU surface (dynamic tropopause) for the 6-h 1200 UTC GFS forecast valid at 1800 UTC with terrain contoured every 500 m (purple dashed lines). (top) Approximately 12 h later (0545 UTC), the observed water vapor IR channel from GOES satellite. These maps are shown for a (left) weakly forced and (right) strongly forced day composing the 8–10 Jul 2013 case study.
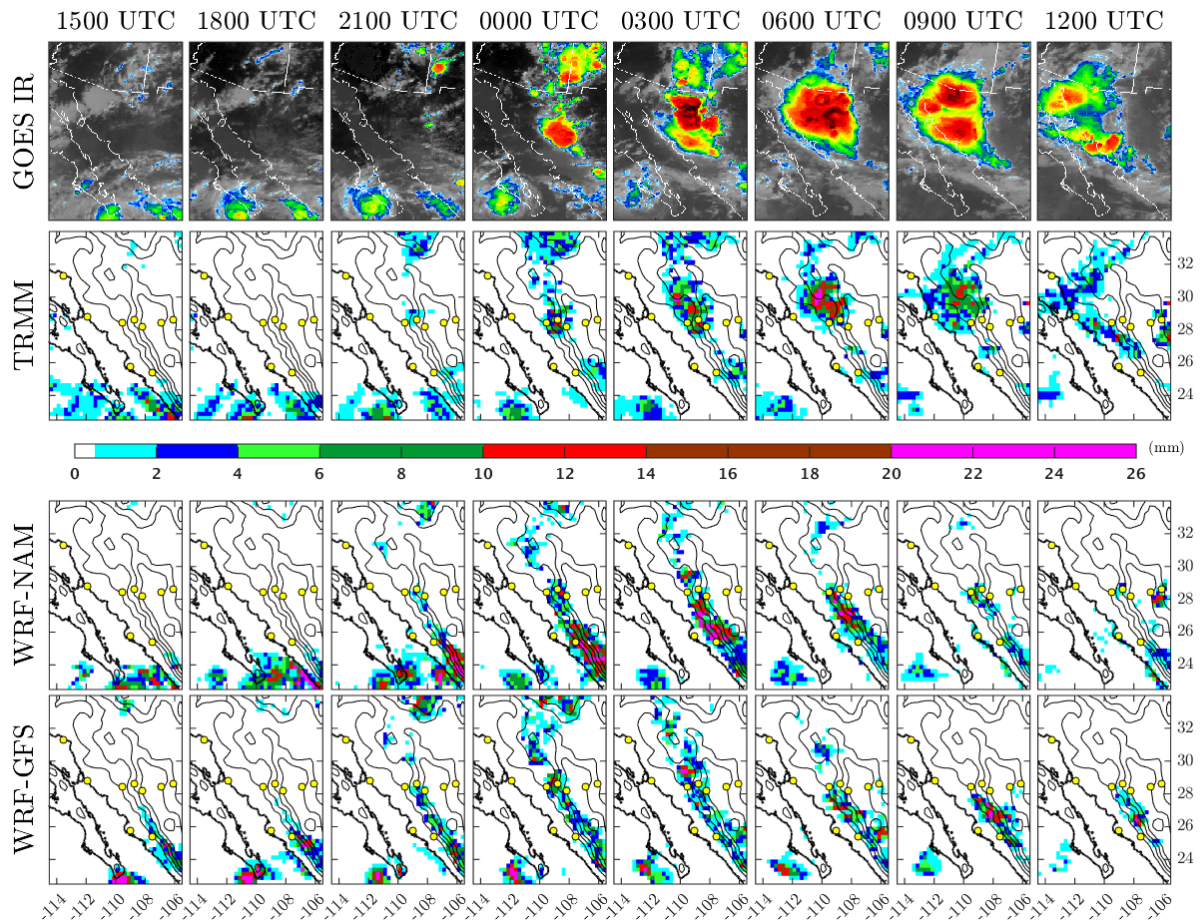
Figure 13: 3-hourly GOES IR (11-*μm* channel) imagery (row 1), 3-h rainfall accumulation of TRMM (row 2), WRF-NAM hindcast (row 3), and WRF-GFS hindcast (row 4) for the weakly forced day of 8 Jul 2013. Times are in UTC and indicate the time of the GOES IR image and the ending time of the 3-h rainfall accumulation. Yellow circles indicate the locations of a subset of Transect 2013. Terrain is contoured every 500 m (black lines).
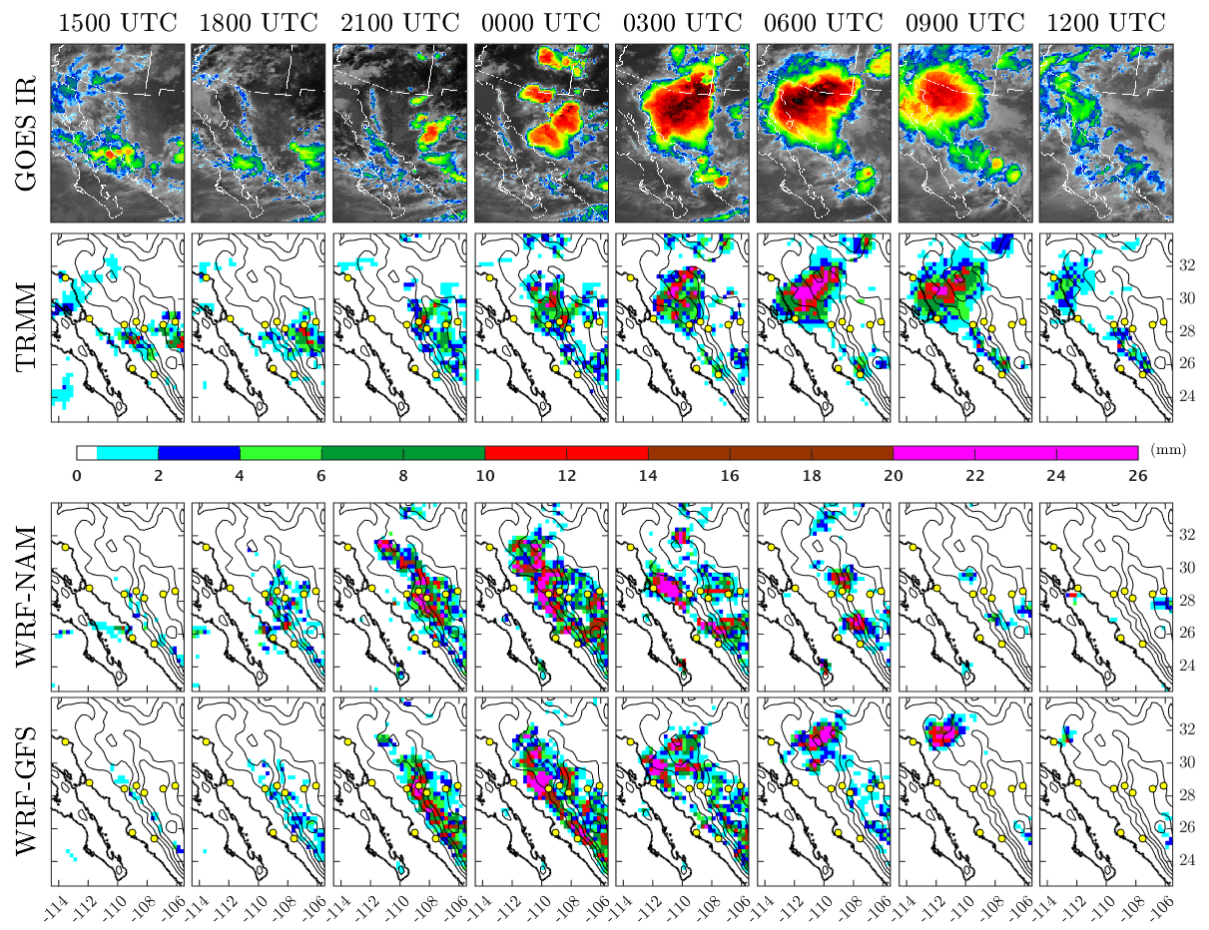
Figure 14: As in Fig. 13, but for the strongly forced day of 9 Jul 2013.
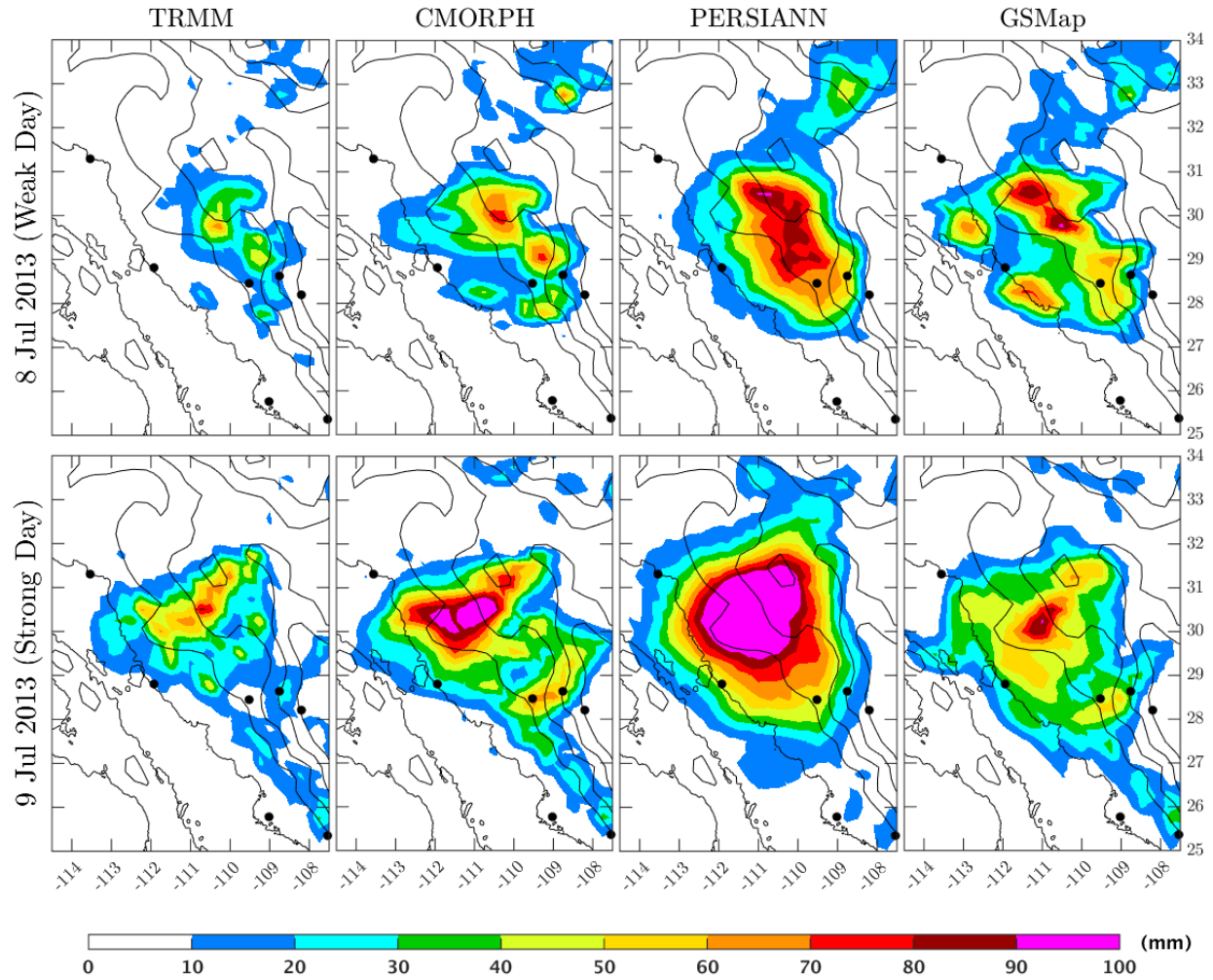
Figure 15: For the (top) weakly forced (1200 UTC 8 Jul 2013 to 1200 UTC 9 Jul 2013) and (bottom) strongly forced (1200 UTC 9 Jul 2013 to 1200 UTC 10 Jul 2013) day, 24-h rainfall accumulations are shown for the satellite-based precipitation products of TRMM (column 1), CMORPH (column 2), PERSIANN (column 3), and GSMaP (column 4). Block dots indicate the locations of a subset of Transect 2013. Terrain is contoured every 500 m (black lines).
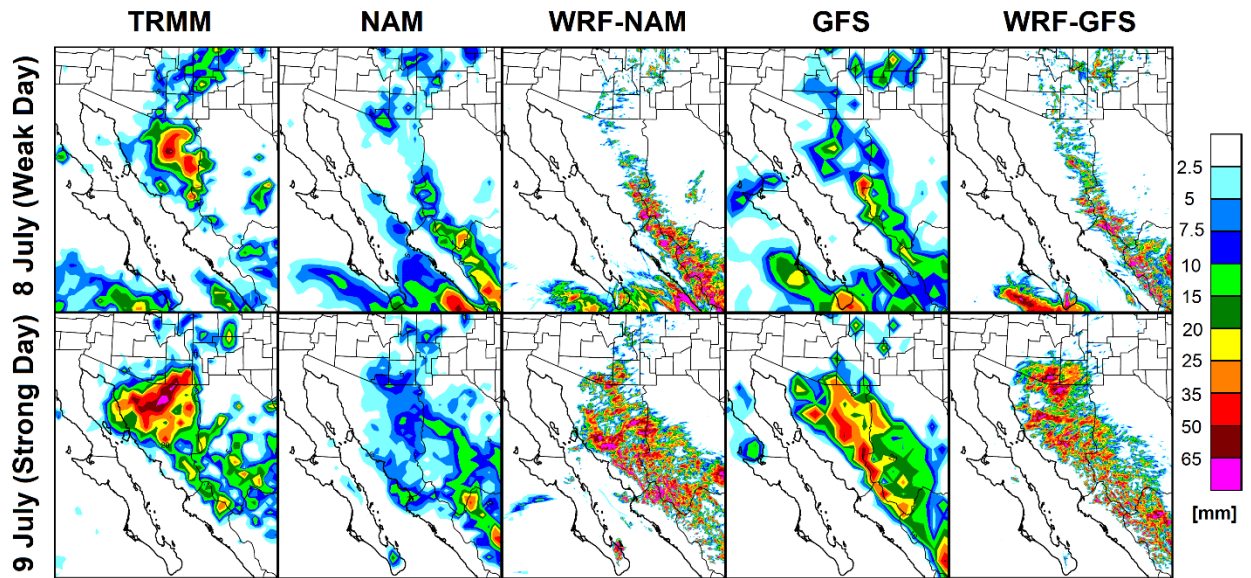
Figure 16: For the (top) weakly forced (1200 UTC 8 Jul 2013 to 1200 UTC 9 Jul 2013) and (bottom) strongly forced (1200 UTC 9 Jul 2013 to 1200 UTC 10 Jul 2013) day, 24-h rainfall accumulations are shown for TRMM (column 1), NAM (column 2), WRF-NAM (column 3), GFS (column 4), and WRF-GFS (column 5).
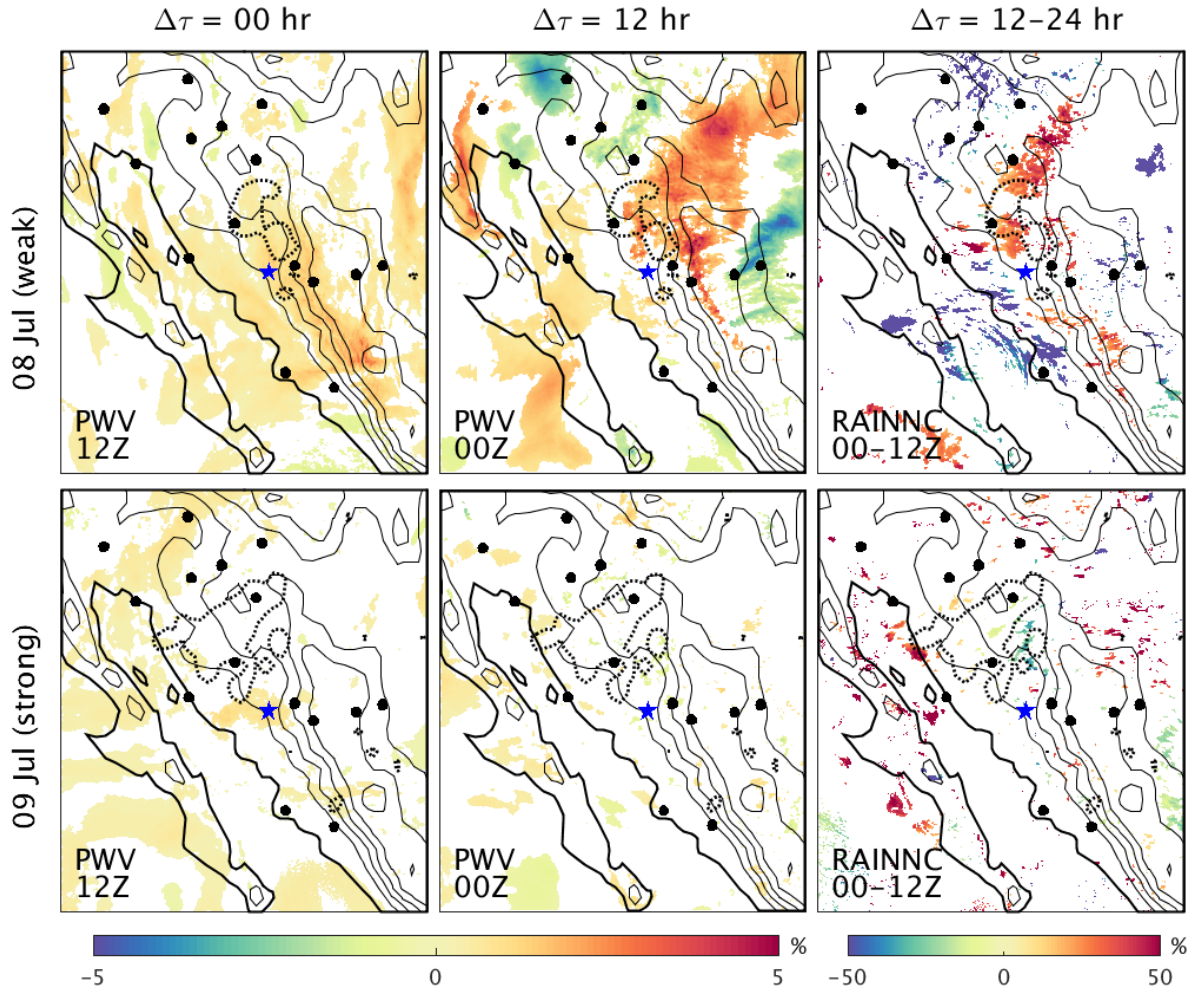
Figure 17: Relative sensitivity of WRF PWV (columns 1 and 2) and grid-scale rainfall (RAINNC; in domain d03) (column 3) at time $t + \Delta\tau$ to changes in the initial condition of PWV at ONVS GPS ground site (marked as a blue star, lon: 109.53°W, lat: 28.46°N). The sensitivities of ONVS PWV to WRF PWV (% PWV per % PWV) at $\Delta\tau = 0$ and $\Delta\tau = 12$ (columns 1 and 2) and ONVS PWV to RAINNC (% PWV per % mm) at $\Delta\tau = 12$ to 24 (column 3) are presented for both the (top) weakly forced (1200 UTC 8 Jul 2013 to 1200 UTC 9 Jul 2013) and (bottom) strongly forced (1200 UTC 9 Jul 2013 to 1200 UTC 10 Jul 2013) day. Note: only statistically significant correlations (p value < 0.05) are shown. The 24-h 25-mm TRMM isohyet is plotted by a dotted line. See text for calculation of ensemble–based sensitivities. Terrain is contoured every 500 m (black lines).

APPENDIX A

Expanded Analysis Methods

a) Calculation of PWV in WRF

WRF PWV is calculated as a discrete summation across all model levels for each grid point (i,j) in the d03 domain:

$$\text{PWV}_{i,j} = \sum_{k=1}^{\text{nlev}} \left(\frac{p_k}{R_d \, T_{vk}} \right) (\text{QVAPOR}_k) \, (Z_{k+0.5} - Z_{k-0.5}), \quad \text{(A1)}$$

where p_k is pressure (Pa), T_{vk} is virtual temperature (K), and QVAPOR_k is the water vapor mixing ratio (kg water vapor/kg dry air), and all are associated with the mass point at model level k. The geopotential height (m) respective upper and lower bounds on the staggered grid at model level k are $Z_{k+0.5}$ and $Z_{k-0.5}$ (m) are. R_d is the dry air gas constant, and nlev is 27, the number of model levels. The variables used in this PWV calculation are first interpolated to the location of the GPS station using an inverse-distance weighting of the four closest corresponding model grid points.

b) Forward sensitivity analysis methodology

Before calculating the sensitivities, we spin up the WRF simulation for 12 h to get the appropriate variance of state variables in d03. We first add small perturbations on WRF meteorological fields (in the outermost domain: d01) based on the NCEP GFS error covariance (cv3) using the WRF Data Assimilation (WRFDA) tool. Then, we integrate WRF-GFS for 12 h to propagate these perturbations to the innermost domain (d03). The resulting 40-member ensemble in d03 composes our WRF initial conditions. We sample the model-equivalent GPS PWV for each ensemble member, iens, at a site (refer to section a in this appendix), isite ($\text{PWV}^{\text{mod}}_{\text{isite},t=0,\text{iens}}$), using these initial conditions, that is,

$$\text{PWV}^{\text{mod}}_{\text{isite},t=0,\text{iens}} = \mathbf{H}\mathbf{x}^{\text{mod}}_{\text{iens}} \quad (A2)$$

where $\mathbf{x}^{\text{mod}}_{\text{iens}}$ is the QVAPOR 3D field at $t=0$ for ensemble member iens and $\mathbf{H}$ is the

observation operator that maps the model moisture field to site-specific PWV as would

be observed from a GPS ground site [see Eq. (A1)]. This operator bilinearly interpolates

QVAPOR to a site location and vertically integrates the interpolated profile to get PWV.

Thus, we have a 40-element vector, $y^{\text{mod}}_{\text{isite},t=0}$, for each GPS site of our study domain.

Starting with the ensemble d03 initial conditions, we generate an ensemble of

hourly WRF 2.5-km forecasts with $\Delta\tau$ out to 24 h. For each grid point (i,j) of the model

forecast $(t + \Delta\tau)$, we can calculate the relative local sensitivity of WRF Model variable

$X_{i,j,t+\Delta\tau}$ to PWV initial conditions at a particular site $\text{PWV}_{\text{isite},t=0}$ using the statistics

from the ensemble forecasts expressed as

$$\left(\frac{\Delta X_{i,j,t+\Delta\tau}}{\Delta\text{PWV}_{\text{isite},t=0}}\right) = \frac{\text{cov}(X_{i,j,t+\Delta\tau}, \text{PWV}^{\text{mod}}_{\text{isite},t=0})}{\text{var}(\text{PWV}^{\text{mod}}_{\text{isite},t=0})} \times \frac{\langle\text{PWV}^{\text{mod}}_{\text{isite},t=0}\rangle}{\langle X_{i,j,t+\Delta\tau}\rangle}. \quad (A3)$$

Here, $X_{i,j,t+\Delta\tau}$ represents the ensemble WRF forecast of PWV or rainfall at grid point

(i,j). In principle, $X_{t+\Delta\tau}$ can be any 3D model forecast variable. The $\langle\cdot\rangle$, $\text{cov}(\cdot)$ and

$\text{var}(\cdot)$ notations represent the expected value (ensemble mean in this case), covariance,

and variance across the ensemble, respectively. The PWV at a model grid point is

calculated in the same way as Eq. (A2) with $\mathbf{H}$ now corresponding to just the integration

of the QVAPOR profile without interpolation. The sensitivity in Eq. (A3) (which is a

linear regression estimate) can be interpreted as the local linear sensitivity of the model to

changes in PWV initial conditions. This is analogous to the linear sensitivity component

of the Kalman gain in an ensemble Kalman filter. A more general concept of ensemble-

based sensitivity analysis has been discussed in detail by Torn and Hakim (2008).

APPENDIX B

Acronyms

CMORPH	CPC morphing technique
CSI	Critical success index
ECMWF	European Centre for Medium-Range Weather Forecasts
ERA	ECMWF Re-Analysis
ESMF	Earth System Modeling Framework
FAR	False-aarm ratio
GFS	Global Forecast System
GoC	Gulf of California
GOES	Geostationary Operational Environmental Satellite
GPM	Global Precipitation Measurement
GPS	Global positioning system
GSMaP	Global Satellite Mapping of Precipitation
HAS	Department of Hydrology and Atmospheric Sciences
IV	Inverted trough
JAXA	Japan Aerospace Exploration Agency
LT	Local time
MCS	Mesoscale convective system
NAM	North American Mesoscale Forecast System
NAME	North American Monsoon Experiment
NARR	North American Regional Reanalysis
NASA	National Aeronautics and Space Administration
NCL	NCAR Command Language
NOAA	National Oceanic and Atmospheric Administration
NWP	Numerical weather prediction
NWS	National Weather Service
PERSIANN	Precipitation Estimation from Remotely Sensed Information Using Artificial Neural Networks
POD	Probability of detection
PV	Potential vorticity
PWV	Precipitable water vapor
RMSE	root-mean-square error
SMO	Sierra Madre Occidental
TC	Tropical cyclone
TEW	Tropical easterly wave
TMPA	TRMM Multisatelite Precipitation Algorithm
TRMM	Tropical Rainfall Measuring Mission
TS	Tropical storm
UA	University of Arizona
UCI	University of California, Irvine

UTC	Universal coordinated time
WRF-ARW	Advanced Research version of the Weather Research and Forecasting Model
WSR-88D	Weather Surveillance Radar-1988 Doppler

APPENDIX B: UTILITY OF GPS-DERIVED PWV MEASUREMENTS IN CONSTRAINING HIGH-RESOLUTION WRF FORECASTS OVER THE NORTH AMERICAN MONSOON REGION

James M. Moker Jr.[1*], Avelino F. Arellano, Jr.[1], Christopher L. Castro[1], Yolande L. Serra[2], David K. Adams[3], and C. Bayu Risanto[1]

[1] *Department of Hydrology and Atmospheric Sciences, University of Arizona, Tucson, Arizona*
[2] *Joint Institute for Study of Atmosphere and Ocean, University of Washington, Seattle, Washington*
[3] *Centro de Ciencias de la Atmósfera, Universidad Nacional Autónoma de México (UNAM), México, D.F.*

Manuscript draft for submission to *Monthly Weather Review*

Corresponding author address: James M. Moker, Jr., Harshbarger Bldg., Rm. 214A, Dept. of Hydrology and Atmospheric Sciences, PO Box 210011, 133 E. James E. Rogers Way, Tucson, AZ 85721-0011
E-mail: jmoker@email.arizona.edu

Abstract

In northwest Mexico on 8 July 2013, early afternoon convection in the Sierra Madre Occidental (SMO) mountain range propagated north and west and organized into a Mesoscale Convective System (MCS) by early evening in the lower elevations of Sonora and the northern Gulf of California (GOC) reaching peak intensity at 2300-0200 LT. The main objective of this study is to investigate the utility, including potential improved simulation of the MCS, of assimilating observations from ground-based GPS-PWV stations from the North American Monsoon Transect Experiment 2013 into convection-allowing Advanced Weather Research and Forecasting Model (WRF-ARW) retrospective forecasts (hindcasts).

We use an ensemble adjustment Kalman filter (EAKF) in the Data Assimilation Research Testbed (DART) software with a 6-h spin-up, 6.5 1-h DA cycles, Gaspari-Cohn 3-D covariance localization, and 20 ensemble members. Then, we initialize a 24-h deterministic hindcast from the ensemble-mean analysis at 0500 LT (1200 UTC). Using an engineering approach for an observation network novel to this region, we run a series of sensitivity experiments investigating, for example, the impact of different covariance localization cutoffs and types of adjusted variables within the ensemble DA algorithm.

For all experiments, across the GPS-PWV stations, the DA algorithm reduces the PWV RMSE from 3 mm at the prior at the beginning of the DA cycling period to 0.25-1 mm at forecast initialization. This falls within the GPS-PWV observation error of 1-2 mm. There is a short memory at the 9 GPS-PWV stations for these adjustments in that the PWV RMSE rises quickly and becomes similar to the non-DA experiment by hour 2. Especially with the GPS-PWV stations near the GOC, this short "memory" is partially

105

attributed to moisture adjustments in non-uniform regions being advected north. Assimilating GPS-PWV observations reduces water vapor mixing ratio error near the surface to within the radiosonde humidity instrument error. The assimilation also increases Convective Available Potential Energy (CAPE) in the region of the MCS.

Assimilating GPS-PWV observations improves the simulation of the MCS. To evaluate hindcast performance, we use the percentage of area < -32 °C of the simulated cloud shield that matches the < -32 °C region of the observed MCS cloud shield at times spatial MCS criteria are met. First, compared to a hindcast beginning at 0500 LT (*cold_start* experiment) (34% during observed maximum intensity [at, e.g., 2300 LT]), there is improved simulation of the MCS when the hindcast initializes 12 h earlier (42% for ensemble spin-up [*no_assim* experiment] and 37% for deterministic spin-up [*warm_start* experiment]) because, as a first order, the meso-γ-scale flow has time to establish in the convection-allowing grid. Then, the assimilation of GPS-PWV observations (*assim* experiment) nudges the location and broadens the coverage of the MCS towards the observation for further improvement of the simulated MCS (50%).

Increasing the covariance localization cutoff when adjusting all variables (*all_vars* experiments) improves the simulated MCS (from 31% for 0.03 radians to ~50% for 0.07 and 0.10 radians). Increasing the covariance localization cutoff when adjusting only thermodynamic variables (*no_winds* experiments) slightly improves the simulated MCS from 0.03 radians (36%) to 0.07 radians (42%) but degrades the simulated MCS at 0.10 radians (25%).

1. Introduction

a) Importance of NAM rainfall and NWP

The North American Monsoon (NAM) is responsible for approximately 50-70% of the annual precipitation in northwest Mexico and the Southwest US (Douglas et al. 1993; Adams and Comrie 1997). The NAM begins in southern Mexico in May and expands north and westward along the Sierra Madre Occidental (SMO) mountain range in northwest Mexico and southwest U.S. in June and July. The convective rainfall at the northern periphery of the NAM in southwest U.S. peaks in August before declining in late September. In both Mexico and the U.S., the NAM has significant impacts on water supply and demand, severe weather, extreme heat, drought, and wildfire during the warm season (Ray et al. 2007).

The North American Monsoon Experiment (NAME; Higgins et al. 2006; Higgins and Gochis 2007) was a multi-agency field campaign that took place in the summer of 2004 whose goal was to determine the limits of predictability of summertime precipitation of the NAM through both an enhanced observational campaign and follow-up numerical weather prediction (NWP) studies. The Sierra Madre Occidental (SMO) mountain range has been shown to have a primary role in the spatial and temporal evolution of convective rainfall throughout the diurnal cycle (e.g., Gochis et al. 2004,2007; Lang et al. 2007). Sea breezes from the Gulf of California (GOC) also contribute to moist upslope flow that aids in this convection (Johnson et al. 2007). The diurnal cycle of convection begins with individual thunderstorms forming along the crest of the SMO and its western slope in the early afternoon (Adams and Comrie 1997; Nesbitt et al. 2008). Then, these individual cells may organize into a mesoscale

convective system (MCS) that propagates generally westward towards lower elevations and the GOC throughout the evening and early morning hours (Lang et al. 2007; Nesbitt et al. 2008). A peak in seasonal observed late-evening/early-morning precipitation across lower elevations is associated with these MCSs (Lang et al. 2007). Although not necessary, an upper-level inverted trough (IV; Adams and Comrie 1997; Pytlak et al. 2005; Bieda et al. 2009; Newman and Johnson 2012; Seastrand et al. 2015), a transient synoptic scale feature, can aid in the formation of an MCS. NAME-related research has shown that meteorological processes, such as the ones contributing to MCSs, occur on the meso-γ and meso-β scales (1-100 km). Regional modeling studies have demonstrated that a convection-allowing modeling system is necessary to effectively represent MCSs in this region (Gao et al. 2007; Li et al. 2008).

b) Lack of in-situ observations in the NAM region

NAME-related research has shown that the lack of in-situ observations at higher elevations of the SMO is a source of uncertainty in characterizing the environmental conditions of the location in which convective initiation occurs along with moisture transport from lower elevations (Higgins and Gochis 2007). In-situ meteorological data in Mexico are sparser and less reliable than in the U.S.; this leads to poor initial condition specification in the forecast model in Mexico irrespective of the model's spatial resolution. High-density observational networks that have a high temporal frequency would allow for better conceptualization of the initiation and upscale growth of convection over northwest Mexico. While there have been efforts to restore the radar and radiosonde networks throughout the country (Zavaleta and Vargas 2012), these low-density observations are costly. In addition, the radar is subject to partial signal blockage

from nearby complex terrain (e.g., Minjarez-Sosa et al. 2012). To address the need for high-quality moisture observations near the SMO, the North American Monsoon GPS Transect Experiment 2013 (Transect 2013; Serra et al. 2016) was deployed in northwestern Mexico during the 2013 NAM season. Transect 2013 involved nine Global Positioning System precipitable water vapor (GPS-PWV) ground-based stations that monitored the diurnal cycle of the moisture in this region and are used in the current study (described in Section 2a).

c) Errors in simulating the timing and propagation of NAM convection

Mesoscale processes associated with the initiation and growth of deep convection, particularly over the highest elevations of the SMO, and the sources of water vapor relevant to these processes, are poorly understood. As a result, both operational and high-resolution models have difficulty simulating the timing and subsequent propagation of deep convection over the SMO (e.g., Li et al. 2008; Castro et al. 2012; Pearson et al. 2014). Janowiak et al. (2007) showed that the peak in the numeral modeling of diurnal precipitation in the NAM is maximized 3-6 hours earlier than observations during NAME. A similar temporal shift was reported by Moker et al. (2018) near the SMO during the 2013 NAM season for situations when a precursor IV was absent ("weakly forced" day). In our study, expanding on the work of Moker et al. (2018), we investigate the impact of assimilating GPS-PWV observations into high-resolution model forecasts of organized convective events such as the MCSs.

d) DA studies in the NAM region

Data assimilation (DA) studies in the NAM region is limited to Mo et al. (2007). With the assimilation of soundings during NAME, they found that the impacts are

regional and concentrated over the core monsoon region with only small differences in the upper-level circulation features such as the jet stream and large-scale waves. At lower levels, they found that the differences depend on the assimilation system. In addition to the soundings improving short-range forecasts over the core monsoon region, they also generally improve the analyses over the areas where the assimilation system has the largest uncertainties and errors.

e) Advantages of ensemble DA

The ensemble Kalman Filter (EnKF) uses a flow-dependent and multivariate (via ensemble) estimation of forecast error covariance, which permits correction of forecast errors in state variables such as wind, temperature, and moisture. The error reduction is done in a more consistent way relative to older approaches like the 3D-VAR (e.g., Meng and Zhang 2008). An EnKF system permits adjustments of the model grid based the sensitivity to the data being assimilated. Demonstrating this concept, Liu et al. (2012), in their study, assimilate satellite PWV data via the EnKF and find that it improves the short-range hindcast of a tropical storm, particularly the early genesis and later intensification. The moisture in the column is adjusted based on the sensitivity to the PWV observation rather than adjusting the entire moisture profile based on fixed and predetermined correlation lengths, as is the case for optical interpolation (OI) and 3D-VAR approaches. The EnKF has been shown to improve forecasts and reduce uncertainty for large-scale model prediction systems (e.g. Whitaker and Hamill 2002) and short-term analysis and prediction of MCSs (e.g. Torn 2010; Ancell et al. 2011; Suarez et al. 2011; Meng and Zhang 2007).

f) Previous studies on HiRes WRF modeling in the NAM

Kursinski et al. (2008) show that precipitation in high-resolution modeling studies over the SMO is sensitive to PWV initialization. In hindcasts of NAME Intensive Observing Periods (IOPs) using a WRF adjoint modeling system, Cassell and Castro (2011) show that the development of MCSs in Sonora and Arizona is sensitive to the initial specification of atmospheric water vapor where convection initiates in the SMO.

During the 2013 season, Moker et al. (2018) showed the that diurnal cycle of convection in hindcast simulations begins and ends 3-6 hours too early compared to satellite (TRMM) rainfall observations. This is reflected in a high moisture bias at model initialization they found compared to the GPS-PWV observations. Also, on days when synoptic forcing is weak (no IV), the precipitation forecast was poorer compared to strong days (with an IV) in areas west of the SMO into the lower elevations of Sonora where MCSs tend to propagate in the early evening and overnight hours. Finally, the sensitivity to rainfall and PWV several hours after initialization to initial PWV at a lower elevation site was higher in the weakly-forced day than the strongly-forced day. This would indicate that there would be greater change in the forecast for the weakly-forced day with PWV assimilation.

g) Previous studies involving PWV DA

The value of PWV from radiosondes and ground-based GPS PWV (including zenith total delay (ZTD) and slant wet delay (SWD)) for improving moisture and precipitation forecasts has been demonstrated in several studies covering a variety of terrain across the U.S., Europe, and Japan.

1) NORTH AMERICAN STUDIES

Over the Central US, Kuo et al. (1993) pioneered the assimilation of PWV into a weather prediction model. They assimilate radiosonde-based PWV into a mesoscale model by relaxing the predicted value toward the observed value while maintaining the 3D structure of the moisture field. They find improved the short-range precipitation forecasts as a result. Using the same assimilation method, Kuo et al. (1996) assimilates GPS-based PWV using 4D-VAR with improved moisture analysis. However, they use a sounding for a reference of the vertical structure of moisture. Guo et al. (2000) assimilates GPS-PWV using 4D-VAR and succeeds in reproducing the observed precipitation pattern associated with a squall line. The vertical structure of the moisture could only be recovered with the addition of wind profiler data. Rainfall structure improved from the assimilation of rainfall and surface dewpoint. Ha et al. (2003) assimilates SWD into a 4D-VAR modeling system and finds that it improves cold front location and strength resulting in a more accurate short-range precipitation forecast. Compared to assimilating PWV, assimilating SWD resolves more moisture information in between stations.

Also, in the Central US, Gutman et al. (2004) and a companion study (Smith et al. 2007) find improvements to lower (below 500 hPa) tropospheric humidity in the RUC (now RAP) model forecast system, even though this region already benefits from a dense network of surface observations. Both studies rely on the RUC configured using OI for PWV (Smith et al. 2007). Smith et al. (2007) report that PWV has the greatest impact on 3- and 6-h forecasts but continues to have a noticeable impact on 9- and 12-h forecasts. In addition, the densification of the US GPS network over the course of the experiment period (18 in 1999 to ~275 in 2004) had a positive impact on the relative humidity,

convective available potential energy (CAPE) (e.g., Moncrieff and Miller 1976), and PWV (Gutman et al. 2004; Smith et al. 2007) forecast errors. Smith et al. (2007) indicate that their results ultimately depend on the DA technique; better results are expected with dimensional variational techniques that allow for the direct assimilation of ZTD (thereby avoiding the retrieval of PWV).

2) EUROPEAN STUDIES

In a European study, Poli et al. (2007) examine the impact of the assimilation of ZTD observations into the 4D-VAR Météo-France global forecasting system (Action de Recherche Petite Echelle Grande Echelle [ARPEGE]). Their results indicate that these data positively constrain the synoptic patterns up to 4 days during all seasons over France including the improvement of the 12- to 36-h forecasts of rainfall patterns during spring and summer.

Yan et al. (2009) assimilated ZTD observations from the Convective and Orographically-induced Precipitation Study (COPS) campaign, which examined the impact of a dense installation of GPS observations in northeast France and southwest Germany on high-resolution precipitation forecasts, into the Météo-France Applications of Research to Operations at Mesoscale (AROME) regional non-hydrostatic model with a 2.5 km horizontal resolution and 3-h 3D-VAR data assimilation cycles using the Poli et al. (2007) scheme. Their results suggest that while the COPS ZTD observations provide additional improvements to short-range forecasts of summertime convective rainfall compared to assimilation of the European operational GPS network alone, most of the improved skill results from the assimilation of ZTD observations from the operational network that was already in place. Like in the Gutman et al. (2004) and Smith et al.

113

(2007) studies, improvements in precipitation forecasts result primarily through modification of the lower tropospheric moisture profile in their simulations. Yan et al. (2009) also note that for their COPS case studies the convection improved in terms of both the timing and amounts.

3) JAPANESE STUDIES

Seko et al. (2011) conducted an assimilation experiment of GEONET-derived PWV by using a local ensemble transform Kalman filter (LETKF; Hunt et al. 2007) method based on the Japan Meteorological Agency Non-Hydrostatic Model (JMANHM; Saito et al. 2007).

Seko et al. (2013) developed a two-way nested NHM-LETKF system and investigated the synergistic effects of simultaneous assimilation of the Doppler radar radial wind velocity and water vapor data observed by GEONET (i.e., PWV and SWD, which is the accumulated water vapor amount along the ray path of a GPS signal). They succeeded in increasing the number of ensemble forecasts that reproduced localized heavy rainfall by assimilating the GPS data and the Doppler radar data.

Oigawa et al. (2018) is the first study that assimilates PWV data with a horizontal resolution of less than 10 km and uses a small horizontal localization radius for PWV data over a precipitating region.

h) Objective

On 8 July 2013, early-afternoon convection formed in the SMO in northwest Mexico that propagated and organized into an MCS by early evening as it moved west and north towards lower elevations of Sonora and the GOC, peaking in intensity at 2300-0200 LT. Moker et al. (2018) showed that the downscaled operational models were

114

unable to simulate this MCS. The main objective of the current study is to investigate the utility of assimilation of GPS-PWV observations from Transect 2013 into convection-allowing WRF-ARW hindcasts on this day, including any potential improvement of the simulated MCS. To address this objective, we develop a configuration of the EAKF in DART and run a series of sensitivity tests.

This manuscript is broken down as follows: we describe the data and methods in Section 2, describe the sensitivity and impact assimilation experiments in Section 3, report the results in Section 4, discuss and make implications of the results in Section 5, and finally summarize and make recommendations in Section 6.

2. Data and methods

In this section, we discuss the data and methods involved in our study and executing the objective in Section 1h. This section is organized into the following subsections: GPS-PWV observation and Transect 2013 network description (Subsection a), model physics and domain description (Subsection b), data assimilation scheme (Subsection c), experimental design (Subsection d), and forecast verification methods (Subsection e). Refer to Table 1 for more information on the datasets used.

a) GPS-PWV data

PWV data in this study is gathered from the nine GPS-PWV stations in Transect 2013 (Serra et al. 2016; Adams et al. 2014) that were deployed across northwest Mexico from late July to middle September 2013. Transect 2013 was comprised of three transects (Fig. 1b) whose purpose was to observe: 1) convective initiation and diurnal cycle in the SMO in Sonora and western Chihuahua in 6 stations that addresses a study gap documented in NAME (Higgins and Gochis 2007), 2) gulf surge progression in the GOC

in 3 coastal stations in Sonora and Sinaloa, and 3) atmospheric moisture behavior between two close sites with a high precipitation gradient in northern Sinaloa. An all-weather ground-based GPS-met station can measure PWV with a high temporal resolution (~1-min). PWV is derived from the GPS signal delay which is proportional to the vertically integrated amount of water vapor in the air (Bevis et al. 1992). Technical details on GPS-to-PWV conversion methods are found in Bevis et al. (1992). GPS-PWV has an error of a few percent that translates to ~1-2 mm (Bevis et al. 1992; Rocken et al. 1993; Duan et al. 1996; Wolfe and Gutman 2000; Adams et al. 2011). For Transect 2013 data, Global Navigation and Satellite Systems (GNSS)-Inferred Positioning System and Orbit Analysis Simulation Software (GIPSY-OASIS; https://gipsy-oasis.jpl.nasa.gov) is employed to obtain PWV from the GPS signal at a 5-min temporal resolution. This procedure uses satellite information greater than 10° from the horizon resulting in a cone that represents an observation with a 10-15 km radius (Serra et al. 2016).

b) Model description

We use version 3.4.1 of the Advanced Research Weather Research and Forecasting (WRF-ARW; Skamarock et al. 2008) model with 27 vertical levels, three nested domains (Fig. 1a; Table 2); and a 50-hPa top to produce the experimental hindcast simulations. Previous NAME research had concluded that a convective-allowing horizontal grid spacing is required for adequate representation of the diurnal cycles of processes in the NAME Tier I region. The innermost domain (d03) closely corresponds to the NAME Tier I region and has a 2.5-km horizontal grid spacing that indeed allows for explicit convection. The coarser domains (d01 and d02) employ the Kain–Fritsch cumulus parameterization scheme (Kain 2004). Except for using adaptive time stepping,

116

this model configuration is identical to Moker et al. (2018). The following model physical parameterizations used are: Yonsei University (YSU) PBL scheme (Hong et al. 2006), WSM6 explicit microphysics (Hong and Lim 2006), RRTMG longwave radiation (Iacono et al. 2008), Goddard shortwave radiation (Chou and Suarez 1994), and Unified Noah Land surface scheme (Tewari et al. 2004). These are listed in Table 1 along with details on the nested domain configuration.

We use the GFS (0.25° horizontal grid spacing) model for initial conditions (ICs) and 6-h lateral boundary conditions (LBCs). Moker et al. (2018) found that the GFS had a lower precipitation and moisture bias than the NAM model (32-km horizontal grid spacing). Like Moker et al. (2018), the Rapid Refresh (RAP) model is used to initialize the volumetric soil moisture and temperature because these data have finer spatial resolution than the GFS model.

c) Assimilation Scheme

1) INITIAL ENSEMBLE PERTURBATIONS AND SPIN-UP

After the WRF Preprocessing System (WPS) is used to generate the initial WRF fields and LBCs from the 0000 UTC run of the GFS model, a 20-member ensemble is created from the mean state in the 30-km domain (d01). With the WRF Data Assimilation (WRF-DA; Barker et al. 2012) software, we use the CV3 option to generate 20 sets of random errors from the default global covariance file via the NMC method (Parrish and Derber 1992) with the default values for variance, horizontal scale, and vertical scale (see Table 3). Then, these errors are added back to the mean state (d01) to serve as the ICs for the ensemble members. Then, the LBCs are also updated using a method described by Torn et al. (2006). The ensemble "spins up" for 6 h until 0600 UTC to allow for the

117

perturbations in d01 to propagate to the convective-allowing 2.5-km domain (d03) (row 3 of Fig. 2). The ensemble at 0600 UTC serves as the "forecast" or "prior" for the data assimilation algorithm described in the next subsection.

2) EAKF DATA ASSIMILATION ALGORITHM

We use the ensemble adjustment Kalman Filter (EAKF; Anderson 2001, 2003) within the Data Assimilation Research Testbed (DART) software (Anderson et al. 2009) to assimilate the GPS-PWV observations into the WRF forecasts. Using the configuration parameters in Table 3, we employ a "local least-squares" framework with equations adapted from Anderson (2003) and is described in the following steps and eqs. 1-6:

1) Start with a forecast (or "prior") ensemble (denoted by an f superscript). At the first assimilation time (0600 UTC), use the result of the 6-h spin-up.

2) Obtain an ensemble sample of the observation by using the PWV observation operator H (defined in the next subsection [Subsection 2c3]) on the WRF state variables x (Table 6) for each member i (eq. 1).

$$PWV_i^f = H(x_i^f) \qquad (1)$$

3) Retrieve a GPS-PWV observation PWV^o and its error σ^o (Table 3).

4) Adaptively inflate the prior ensemble spread σ^f by multiplying it by a value that evolves in time and varies in space according to the method described in Anderson (2009). For the first DA cycle, use the default initial values in Table 3.

5) Calculate the analysis ensemble mean (or "posterior") of the model-equivalent PWV ($\overline{PWV^a}$) (eq. 2).

118

$$\overline{PWV^a} = \left(\frac{1}{\frac{1}{(\sigma^f)^2} + \frac{1}{(\sigma^o)^2}} \right) \left(\frac{\overline{PWV^f}}{(\sigma^f)^2} + \frac{PWV^o}{(\sigma^o)^2} \right) \tag{2}$$

6) For each ensemble member i, calculate the analysis of the model-equivalent PWV (PWV_i^a) (eq. 3) and observation increment ΔPWV_i (eq. 4).

$$PWV_i^a = \sqrt{\frac{(\sigma^o)^2}{(\sigma^f)^2 + (\sigma^o)^2}} \left(PWV_i^f - \overline{PWV^f} \right) + \overline{PWV^a} \tag{3}$$

$$\Delta PWV_i = PWV_i^a - PWV_i^f \tag{4}$$

7) For each ensemble member i, obtain an analysis value for each grid point for each of the WRF variables (Table 6). The WRF variable $QVAPOR$ (water vapor mixing ratio) is used as an example variable to be adjusted and shown in eqs. 5-6. For a given grid point, obtain analysis $QVAPOR_i^a$ by first multiplying the observation increment ΔPWV_i by the forecast ensemble sensitivity of $QVAPOR_i^f$ to PWV_i^f (linear regression) to get the increment $\Delta QVAPOR_i$ (eq. 5). Then, multiply the ensemble sensitivity by a 5th-order piecewise covariance localization function (Gaspari and Cohn 1999) (explained in Subsection 2c4). Finally, add the forecast ensemble sample $QVAPOR_i^f$ to the increment $\Delta QVAPOR_i$ to obtain the analysis $QVAPOR_i^a$ (eq. 6).

$$\Delta QVAPOR_i = \frac{cov\left(QVAPOR_i^f, PWV_i^f \right)}{var\left(PWV_i^f \right)} \Delta PWV_i \tag{5}$$

$$QVAPOR_i^a = QVAPOR_i^f + \Delta QVAPOR_i \tag{6}$$

8) Continue until all 5-min GPS-PWV observations are processed (+/- 30 minutes of observation time) for each site and WRF variables (Table 6) are updated to the

analysis. Do this for each domain. Then, advance the ensemble members one hour to the next observation time using WRF-ARW.

Hourly DA cycles are defined as: 1) 1-h ensemble forecasts that are then ensemble-averaged and 2) an ensemble update of the mean and each member with a new analysis from the assimilated observations. We run 6.5 DA cycles from 0600 UTC to 1200 UTC before initializing the forecast; the first cycle at 0600 UTC is a considered a half-cycle because the ensemble is updated immediately after the spin-up.

3) GPS-PWV OBSERVATION OPERATOR

The model-equivalent GPS-PWV observation is calculated in the obs_def_tpw_mod.f90 Fortran module in DART. The following steps closely approximate the code contained within this module and can be thought of as the RHS of eq. 1.

1) For the four grid points surrounding the observation location, construct vectors of pressure and water vapor mixing ratio from surface values and mass points above.

2) Bilinearly interpolate these vectors horizontally to the observation location to obtain one vector each of pressure and water vapor mixing ratio.

3) Using eq. 7, summate the products of the mean water vapor mixing ratio (q) and the pressure (p) difference for each level (k) across all vertical levels. Then, negate and divide by the gravitational acceleration constant to get PWV in mm.

$$PWV = -\frac{1}{9.81} \sum_{k=1}^{27} \left(\frac{q_k + q_{k+1}}{2}\right)(p_{k+1} - p_k) \tag{7}$$

The first layer is integrated between the surface ($k = 1$) and the mass point at the first model level. All other layers are integrated between the mass points at each model level.

The top half of the top model level is not included in the integration; this has virtually no impact on the calculation because moisture is negligible above 100 hPa. This PWV calculation is the integration within a column rather than the observed cone implying that the atmosphere away from the site may less adequately represented the further one moves away from the surface.

4) HORIZONTAL AND VERTICAL COVARIANCE LOCALIZATION

Covariance localization in three dimensions must be applied to the adjustments of state variables to limit spurious correlations. We use the Gaspari-Cohn function (Gaspari and Cohn 1999) that is a 5th-order piecewise polynomial function similar to a Gaussian function with a given cutoff in radians that indicates the half-width – in this case the distance from the observation where the function is halfway to zero (not to be confused with a function value of 0.5). At the observation location, the function value is 1; at $2 \times$ cutoff, the function value is 0. The cutoff is multiplied by the mean radius of earth (6380 km) to obtain the horizontal-equivalent distance. The cutoff is multiplied by a set normalization height to obtain the vertical-equivalent distance. In our study, we set the normalization height to 50 km.

With a Gaspari-Cohn localization cutoff of 0.07 radians (~450-km horizontal-equivalent distance), the mean 0.5 function value contour across all GPS-PWV locations sites forms an area of ~150,000 km^2 that matches the magnitude of the spatial scale of an MCS in general (Maddox 1980) and is centered near the actual MCS initiation location on that day (Fig. 3b, panel 3). In observed soundings during the 2013 NAM season, the 3.5-km vertical-equivalent distance fits well with the highest correlations of PWV to water vapor mixing ratio just above the surface at CHIH (~0.7 to 0.9) and MAZT (~0.6 to

0.8) (Fig. 4). Being limited to the Gaspari-Cohn localization function precludes having a secondary maximum that would be needed to capture the secondary region high correlation of 0.65 at ~7.5 km MSL in MAZT, but we are confident that not addressing the second maximum does not have a noticeable negative impact due to lower water vapor amounts at that level.

d) Experimental design

Our study consists of three main hindcast experiments, which are listed in the schematic in Fig. 2 and the top three rows of Table 4, using the WRF-ARW configuration listen in Table 2. The first two experiments do not involve data assimilation: *cold_start* experiment (Fig. 1, row 1) and *no_assim* experiment (Fig. 1, row 2). The *cold_start* experiment is identical to the WRF-GFS experiments of Moker et al. (2018) in which the IC and 6-h LBCs are taken from the 0.25° 1200 UTC GFS run and then WRF-ARW is run for 24 h as a deterministic forecast. The *no_assim* experiment uses ICs and LBCs from the perturbations of the 0000 UTC GFS run using WRF-DA that creates 20 ensemble members as described in Section 2c1. Then, these ensemble members are advanced 12 h via WRF-ARW until the 1200 UTC forecast initialization time when a 24-h deterministic forecast is initialized with the ensemble mean. The *assim* experiment (Fig. 2, row 3) contains the assimilation of GPS-PWV in hourly cycles for the 6 hours leading up to the forecast initialization time of 1200 UTC. We adjust all variables including winds as listed in Table 6. The *assim* experiment uses the same ensemble as *no_assim* from 0000 to 0600 UTC, but then hourly DA cycles are run with WRF-DA and the EAKF within DART from 0600 to 1200 UTC. This DA algorithm is described in detail in Section 2c2 with a configuration overview in Table 3. Finally, a 24-h deterministic

forecast is initialized at 1200 UTC from the ensemble mean of the final analysis.

e) Hindcast verification

The cloud shields of simulated MCSs are verified against the infrared window channel (~11 μm) of the gridded Satellite (GridSat-B1) latitude equal-angle 0.07° dataset (Knapp et al. 2011). To identify an MCS, we use the definition from Maddox (1980):

1) A cloud shield colder than -32 °C that has an area of at least 100,000 km^2.

2) An interior cloud shield colder than -52 °C that has an area of at least 50,000 km^2.

3) The cloud shield has an eccentricity of at least 0.7 at the time of maximum extent.

4) The size conditions in 1 and 2 must persist for at least 6 hours.

The cloud shield of the MCS is observed from the GridSat-B1 brightness temperature is compared with the model-equivalent brightness temperature that is converted from the outgoing longwave radiation variable (OLR) WRF variable via the Stefan-Boltzmann law. Only condition 1 is used for the WRF hindcasts because modeled cloud shields with temperatures < -52 °C were rarely found in our experiments. After defining the MCS using the 3-h GridSat-B1, the percentage of the area of the modeled cloud shield < -32 °C that is matched with that of the observed is the metric for determining the hindcast performance of the simulated MCS. The WRF grid is regridded to the GridSat-B1 grid using the "bilinear" option of the EMSF_regrid NCL function.

The gridded rainfall dataset that we use for visual verification is the Tropical Rainfall Measuring Mission (TRMM) Multisatellite Precipitation Analysis (TMPA) 3B42 version 7 dataset (Huffman et al. 2007), hereafter referred to as just TRMM. Its 3-h temporal resolution captures the diurnal cycle of convective precipitation. Moker et al. (2018) concluded that TRMM was preferable to other satellite precipitation datasets.

Additionally, they found that TRMM gave the lowest root mean square difference compared to their hindcast simulations for the 2013 monsoon season.

To verify the initialized temperature, humidity, and wind profiles for our experiments, we use 1200 UTC sounding data from radiosondes launched from Mazatlán, Sinaloa (MAZT) and Chihuahua, Chihuahua (CHIH). The instrumental errors are calculated from the Vaisala Radiosonde RS92-SGP data sheet (available online at https://www.vaisala.com/sites/default/files/documents/RS92SGP-Datasheet-B210358EN-F-LOW.pdf, accessed 28 April 2019). CHIH is the only upper air site collocated with a GPS-PWV site, but it is in a relatively dry region ~200 km east of the SMO crest where afternoon convection initiates. MAZT is located along the GOC in southern Sinaloa, which is near the southern edge of the convection-allowing domain (d03), ~150 km south of the closest GPS-PWV station (BGTO), and > 300 km from the location of the MCS that we are investigating. This region was also being affected by the outer bands of Tropical Storm Erick. For future investigations involving similar MCSs in this region, we recommend choosing a day when the 1200 UTC sounding from Empalme, Sonora is available; it is the best RAOB site location for model verification of the atmospheric profile in that it is located along the GOC ~100 km from both KINO and ONVS and near the region of high moisture and convective activity.

3. Assimilation experiments

Because the NAM is a novel region for GPS-PWV assimilation, we take an engineering approach to the data assimilation algorithm to determine the most physically important options for the minimization of errors of the initial model fields (e.g., moisture) and creating the environment necessary for a well-simulated MCS several hours later. We

explore this with a variety of sensitivity and impact experiments, including the radius of influence of the observation (Subsection a), adjusting thermodynamic and dynamic variables (Subsection b), and other uncertainties (Subsection c) like assimilating individual observations versus one mean observation ("super-obbing"), changing the number of ensemble members, and running an ensemble forecast, for example. A list of these sensitivity experiments is in Table 4.

a) Spatial localization

We investigate the sensitivity of covariance localization radius using the cutoff values of 0.03, 0.05, 0.07, and 0.10 radians of the Gaspari-Cohn function (Gaspari and Cohn 1999) to adjust the WRF state variables (Table 6). Table 5 shows the horizontal- and vertical-equivalent distances for each cutoff. Fig. 3a shows the function graphically for each GPS-PWV site. Fig. 3b shows the mean horizontal component of that function across all sites for each cutoff. We used 0.07 radians for the *assim* experiment with the rationale explained in Section 2c4.

b) Adjusted variable localization

We investigate the sensitivity of the variables we adjust in eqs. 5-6 of the assimilation scheme (Table 6) for each cutoff that is described in the previous subsection. All variables (thermodynamic and dynamic) are adjusted in the *all_vars* experiments whereas only thermodynamic variables (no horizontal wind or vertical motion) are adjusted in the *no_winds* experiments.

c) Other impacts

We also explore other impacts. The *super_ob* experiment looks at the impact of "super-obbing" (e.g., Lorenc 1981; Purser et al. 2000), which is a "data thinning"

procedure that combines multiple observations into one datum that has been used mainly for satellite observations to reduce computing time. In our case, during the DA cycles for each site, we assimilate the mean of the 5-min observations +/- 30 min analysis time rather than the 5-min observations individually like in the *assim* experiment.

The *40_mem* experiment looks at the impact of using 40 ensemble members instead of the 20 ensemble members during the spin-up and DA cycles in the *assim* experiment. The *ens_fx* experiment looks at the impact of an ensemble forecast instead of the deterministic forecast in the *assim* experiment. The *warm_start* experiment is a deterministic forecast that begins at 0000 UTC and runs for 36 h that includes a 12-h spin-up and 24-h forecast. The forecast is compared to the *cold_start* experiment in order to investigate the impact of a spin-up.

4. Results

a) Overall performance

1) PWV

From the perspective of the GPS-PWV, we evaluate the assimilation of PWV in the 6 hourly DA cycles and its effect on the first 7 hours of the forecast. At the time of the first DA cycle 6 hours before forecast initialization, PWV RMSE across all stations decreases from 3 mm to 0.5 mm after the posterior of the final cycle (Fig. 4a, top panel). The 0.5 mm in the *assim* experiment is compared to just above 2 mm for the *no_assim* experiment and just under 2 mm for the *cold_start* experiment. The average difference in RMSE between prior and posterior is ~1 mm. The mean bias starts at +1 mm at 0600 UTC for the prior but then becomes negative by the 0700 UTC posterior (Fig. 4a, bottom panel). From 0700-0800 UTC, the model advance decreases the bias for the only time

during the cycles and reaches a minimum of -1 mm for the prior at 0800 UTC. The posterior at this time brings the bias closer to zero. With the model advance from 0800-0900 UTC and the 0900 UTC analysis, the bias stays nearly constant. After this time, the mean bias increases to +1 mm once again after advancing to 1000 UTC followed by two more cycles of increasing the bias ~0.5 mm with model advance and decreasing ~1 mm in the posterior. By 1200 UTC, the posterior mean bias is nearly zero.

One can see that the PWV RMSE is minimum and mean bias is near zero for the *assim* experiment at forecast initialization (Fig. 4a) that matches well with the low biases across all stations (Fig. 4b, bottom panel). After the forecast starts, the *assim* experiment RMSE rises rapidly from 0.5 mm to 2 mm in 1 hour. This can be attributed to biases of +4.0 mm bias at PSCO, -2.0 mm at ONVS, +2.1 mm at MULT, and +2.4 mm at BGTO. The other 5 stations have biases less than +/- 1 mm. Between 1 and 2 hours, the *assim* experiment RMSE approaches the *no_assim* experiment RMSE. By 2 hours into the forecast, the RMSE of the *assim* experiment is just under 3 mm while the *no_assim* experiment is just above 2 mm. Contributing to the higher RMSE is the relatively higher moist bias at PSCO and MOCH. At PSCO, the bias is +4.5 mm in the *no_assim* experiment while it is +5.9 mm in the *assim* experiment. At MOCH, the bias is -2.1 mm in the *no_assim* experiment while it is -3.6 mm in the *assim* experiment. Between 2 and 3 hours into the forecast, the *assim* experiment passes the the *cold_start* experiment RMSE. At 3 hours into the forecast, there is a high moist bias at 2 sites that are mainly driving the high RMSE in the *assim* experiment: PSCO (+7.7 mm) and BGTO (+4.5 mm). From 3-6 hours into the forecast, the RMSE decreases as the mean biases decrease. After 6 hours into the forecast, the mean biases become negative and the RMSEs begin to rise.

The adjustments of PWV from the *assim* experiment advects northward from the model initialization at 12z through 19z (Fig. 5). For instance, the drying out in the northernmost site of PSCO (-4.1 mm) has the moisture return for the next 2 h from added moisture just to the south of the site over the GOC. By 14z, PSCO has +1.4 mm.

2) CLOUD COVER AND RAINFALL

According to Maddox (1980), an MCS must have a cloud shield that is at least 100,000 km^2 with a temperature of < -32 °C and an inner colder area with an area of at least 50,000 km^2 with a temperature of < -52 °C, as explained in Section 2e. These criteria must be met for at least 6 h and the region at maximum coverage must have an eccentricity of 0.7.

The MCS lasted 9 h, from 3-12z Jul 9, as shown by the progression of the cloud top temperature every 3 h from the GridSat-B1 infrared brightness temperature (Fig. 6a, row 1). Fig. 6b shows the associated 24-h accumulated rainfall. A small area of convection begins along the SMO near MULT in the western slope of the northern SMO by 21z Jul 8. By 00z Jul 9, the convection had built south and west across the ONVS-MULT-BASC transect sites in a developing MCS with coldest brightness temperatures -70 to -65 °C (~150 hPa, ~14 km, ~ 46 kft). However, the MCS criteria had not been met yet. At 03z Jul 9, cloud tops become the coldest ~100 km north of ONVS with values of -80 to -75 °C (~115 hPa, ~16 km, ~ 52 kft) and this is when the convection meets the MCS Maddox (1980) criteria. The cloud shield expands north and west with the anvils over the GOC coast at KINO and south. By 06z Jul 09, the anvil expands and west reaching halfway across the GOC between KINO and PSCO. Active convection continues with colder cloud tops around -75 to -70 °C. By 09z Jul 09, the coldest areas

warm slightly as the cloud shield maximizes in coverage. By 12z Jul 09, the clouds warm

significantly with a decrease in convection; only two areas are near -70 °C but most of

the area is in the -60 to -40 °C range and MCS criteria are still met. At 15z Jul 09 (not

shown), the cloud tops had warmed to where the MCS criteria were no longer met.

The WRF-ARW hindcast experiments rarely displayed cloud top temperatures

colder than -52 °C and rarely did the region colder than < -32 °C form an area of at least

100,000 km^2. Therefore, we match up the modeled areas < -32 °C that match < -32 °C

observed area meeting the first criteria of Maddox (1980). At 21z Jul 8, the *cold_start*

experiment produced convection that started earlier and was more widespread than the

observed and different from *no_assim* and *assim* experiments (Fig. 6a, row 2). It also did

not resemble the shape of an MCS. By 00z Jul 9, the convection in *cold_start* experiment

was strongest as indicated by the coldest cloud top temperatures. By 03z Jul 9, the

convection becomes most widespread north of ONVS-MULT-BASC, but convection had

developed just to the south of there that was not present in the observed, with 42.3% of

the MCS area matching. From 06-09z, 34.3% and 19.4% matching, respectively,

convection builds south and east between ONVS-MULT-BASC and MOCH-BGTO

while convection ends north of ONVS-MULT-BASC. Looking at the 24-hour TRMM

rainfall, this resulted in a lack of rainfall north of ONVS-MULT-BASC and an

abundance of rainfall south of there (Fig. 6b, panel 2).

The *assim* and *no_assim* experiments (Fig. 6a, rows 3-4) were most similar to

each other and to the observations, especially just to the north of the ONVS-MULT-

BASC region, with the time of strongest convection at 03z Jul 8. The *assim* experiment

performed better with the area matching the MCS going from 43.4% to 56.3%.

Convection in these experiments began at 21z Jul 8 around MULT on the western slope

of the northern SMO. From 00-03z, the convection developed into an MCS where the one

in the *assim* experiment was closer to observations and further south than *no_assim*

experiment. By 06z, the intensities were similar and greater than the already-weakened

one in the *cold_start* experiment. Also, at that time, the MCS in *no_assim* experiment

was drifting even farther north and west from the *assim* experiment and observations with

the centroid being ~150 km due east of PSCO. The MCS in the *assim* experiment was

most aligned with observations (50.3% match) with the centroid centered 100-150 km

southeast of the *no_assim* experiment (41.6% match). At 09-12z, the *assim* experiment

(33.4% and 8.8% match, respectively) had more cold clouds associated with the MCS

than *no_assim* experiment (20.5% and 5.6% match, respectively), but the observed MCS

was still stronger at this time than any of the experiments. Although there was a deficit in

rainfall directly over the ONVS-MULT-BASC region for the MCS compared to

observations, the total rainfall aligned best in the *assim* experiment with the rainfall

maximizing too far north in the *no_assim* experiment (Fig. 6b).

In all experiments, the proportion of cloud temperatures colder than -60 °C is less

widespread than the observed. Also, in the steeper terrain of the western slope of the

central SMO, south of ONVS-MULT-BASC, the convection is more widespread

(although less organized) leading to more rainfall. This southern area of convection also

occurs separately from the MCS whereas that separation is less apparent in the

observations.

3) INITIALIZATION FROM SOUNDINGS

At 12z, the CHIH and MAZT sounding verification of water vapor mixing ratio

(QVAPOR), temperature (T), and horizontal winds (U and V) is shown in Fig. 7a and 7b, respectively. While CHIH is collocated with a GPS-PWV site, MAZT is located near the southern edge of the convective-allowing domain (d03) about 150 km from the closest site (BGTO). As a result, it is no surprise that the adjustments are greater over CHIH. At CHIH, QVAPOR was most improved (Fig. 7a, row 1) in the *assim* experiment from the *no_assim* experiment. From the surface to 700 hPa, a bias of +1 to +2.5 g kg^{-1} was decreased to -0.5 to +0.5 g kg^{-1}, falling within the range of the instrument error. Compared to the *cold_start* experiment, there was a slight improvement of bias. In the mid-levels of the atmosphere (~ 650-500 hPa), QVAPOR was dry in the *cold_start* and *no_assim* experiments but was overcorrected to a positive bias. Below 650 hPa, the cold bias of -1 to 2.5 °C in both the *cold_start* and *no_assim* experiments was minimized (Fig. 7a, row 2). However, above that level, the DA introduced a cold bias of -1 to -2 °C while the *cold_start* and *no_assim* experiments was in between -1 and +1 °C. There was negligible change to QVAPOR and T profiles (Fig. 7b, rows 1-2) at MAZT.

For both CHIH and MAZT, different from the QVAPOR and T, the *cold_start* experiment had the lowest bias for the winds while the *assim* and *no_assim* experiments had a greater bias and were also close to each other. At CHIH, the horizontal winds have highest bias especially closest to the ground (below 700 hPa). The U wind bias of +3 to +4 m s^{-1} between 950 and 700 hPa actually changed the direction from easterly to westerly compared to the *no_assim* and *cold_start* experiments (Fig. 6a, rows 3-4). The -10 to -5 m s^{-1} V wind bias below 750 hPa gave near zero winds compared to southerly winds in the *no_assim* and *cold_start* experiments effectively stopping the southerly flow (gulf surge) at the northern end of the GOC (not shown). Looking at MAZT, there were

no noticeable changes to winds from the DA, but there was generally a greater bias throughout the profile in both U and V (Fig. 7b, rows 3-4).

b) Model increments

Across the DA cycles, the PWV increments and increments of variables at 850 hPa and 750 hPa show a pattern of increments near the MCS formation in northwest Mexico (Fig. 8a). Increments at low levels (e.g., ~ 950 hPa) are more chaotic owing to the non-linear advective influences of the GOC and the model moisture bias.

1) MOISTURE

In general, the magnitude of PWV increments is 0.5 to 1 mm (Fig. 8a, left panel). The highest magnitude of PWV increments are along the GOC, both positive and negative. There is an average positive PWV increment in the northern SMO and west through the GOC. However, there chaotic patterns of negative PWV increments north of ONVS-MULT that are mostly contributed by the QVAPOR increment (Fig. 8a, column 1) at 850 hPa at which level overall has the greatest magnitude of increments near 1 g kg^{-1} in this region. South and southwest of ONVS-MULT to MOCH over the GOC, the increment is mainly negative. Although there is a mixture of positive and negative QVAPOR increment contributions at 950 hPa, there is a positive QVAPOR increment at 850 hPa, especially in the middle of the GOC. South of MOCH, positive PWV increments due to generally positive QVAPOR increments at 950 hPa and 850 hPa and are associated with the outer bands of Tropical Cyclone Erick. East of the SMO crest (CUAH-CHIH), the negative PWV increment is reflected by the 750 hPa positive QVAPOR increment.

2) TEMPERATURE

At 950 hPa, the temperature increments over the GOC generally demonstrate a noisy pattern of both negative and positive values (Fig. 8a, column 2). The highest of ~ +1 K in the shape of an arc just to the east of PSCO associated with an outflow boundary from an MCS in the northern part of the domain that can also be seen at 850 hPa. At 950, 850, and 750 hPa, there were generally positive temperature increments from KINO to the ONVS region and to the north towards PSCO. South of ONVS-MULT-BASC, there is a general negative temperature increment. East of the SMO is a positive increment.

3) U WINDS

The U increments (Fig. 7a, column 3) at 950 hPa were greatest (+/- 1 m s^{-1}) in the western half of the GOC with alternating positive and negative streaks perpendicular to the Baja peninsula that are associated with sea breezes from the Pacific Ocean blowing east through the Baja terrain. Also, there was a positive and negative arc-shaped couplet of U increment just east of PSCO that matches up with an outflow bounding mentioned in the previous subsection. Otherwise, from an imaginary line halfway between KINO and MOCH, the U increment was negative to the north of that line and negative to the south. At 850 hPa, this rule generally holds. From just north of ONVS-MULT to the US border, although there is an underlining negative U increment, there are blotches of positive U increments due to MCS convection in that region. In the eastern slope, the U increments are positive. At 750 hPa, there is large area of positive U increment between ONVS-MULT-BASC and MOCH-BGTO, across the highest elevations of the SMO, and east of the SMO crest. There is a negative U increment west of the SMO crest and north of ONVS.

4) V WINDS

The V increments (Fig. 8a, column 4) at 950 hPa were of similar magnitude to the U increments. Also, south of KINO in the western half of the GOC there were alternating positive and negative streaks perpendicular to the Baja peninsula similar to the U increments associated with the Pacific Ocean sea breeze through the terrain. In the northern part of the GOC, there was a general negative increment near 1 m s^{-1}, Other than this, the V increments at 950 and 850 hPa west of the SMO crest were similar to the U increments. Northeast of the SMO had a negative increment. At 750 hPa, the underlining V increment is negative along the SMO transect sites and to the north with blotches of positive increments associated with convection. South of an imaginary line halfway between ONVS and MOCH, north and west of BGTO, and west of the SMO, there was an area with small positive V increment.

5) RESULTING CHANGE TO FORECAST INITIALIZATION

At forecast initialization time (12z), Fig. 8b shows the difference in the meteorological fields from the DA (*assim – no_assim*). As a general rule, these maps are less noisy than the increment maps (Fig. 8a). There is increased PWV from KINO to just to the north of ONVS-MULT in that is much larger than what was shown in the increments (Fig 8b, left panel). This is contributed from increase of QVAPOR at lowest levels of 950 and 850 hPa (Fig. 8b, column 1). South of KINO along the GOC, there is decreased PWV contributed mostly by 750 hPa QVAPOR near KINO to 950 hPa QVAPOR near the GOC mouth. South of BASC, there is a tripole band of increase-decrease-increase of PWV parallel to the SMO that is separated by elevation: increase from the coast to ~ 500 m, decrease from ~ 500-1500 m, and increase above ~ 1500 m. The patterns of the differences in temperature are essentially opposite to the QVAPOR

134

differences (Fig. 8b, column 1), i.e., if there is an increase in QVAPOR, there is decrease in T, and vice versa.

Looking at the winds, west of PSCO the change in V has the highest decrease at 950 hPa that was also seen at 850 hPa that matched up well with the increments in association with the gulf surge. Across the KINO and ONVS region, there was a decrease of U with a mix of positive and negative differences of V. The highest elevations of the SMO had an increase in U and a decrease in V that matched up with the pattern in the increments.

c) Sensitivity tests and other impacts

1) IMPACT OF SPATIAL AND ADJUSTED VARIABLE LOCALIZATION

In this subsection, we explore the sensitivity of the DA and simulated MCS to covariance localization radius and adjusted WRF variables. In the *all_vars* experiments, there is a general steady decrease in PWV RMSE through the cycles (Fig. 9a, column 1), while in the *no_winds* experiments, the RMSE is minimized by DA cycle 3 before increasing quickly after model advance from cycle 3-4 (Fig. 9b, column 1). In the *all_vars* experiments, there is a mean moist bias in the priors throughout all cycles (~ +1 mm) except for cycle 3 where it is a dry bias (~ -1 mm) (Fig. 8a, column 2). In contrast, in the *no_winds* experiments, there is mean moist bias in the priors throughout all cycles except for cycles 2-4 in the 0.10 radian cutoff (Fig. 9b, column 2). The moist bias generally decreases with larger cutoffs. Cycles 5-7 have higher prior biases in *no_winds* experiments (~ +1 to +1.5 mm) than in the *all_vars* experiments (~ +1 mm). The posterior for the final cycle has a lower RMSE for the *all_vars* experiments with a range ~ 0.25-0.75 mm than for the *no_winds* experiments with a range of ~ 0.50-1 mm. In the

135

all_vars experiments, the RMSE increases to 1-2 mm by fh01. In the 0.07 and 0.10

radian cutoffs, the RMSE rises to ~ 3 mm for at least fh02-fh04 before decreasing to near

2mm by fh06. In the 0.03 and 0.05 radian cutoffs, the PWV RMSE rises to ~2 by fh02-

fh04 before stabilizing at 2 mm. Except for 0.10, the RMSE in the *no_winds* experiments

increases to 2-3 mm by fh01, then decreases and stabilizes around 2 mm until fh06. In

both the *all_vars* and *no_winds* experiments, the RMSE increased quickly to near 3 mm

by fh07.

In the *all_vars* experiments, larger cutoff results in a better development of the

MCS from the brightness temperature that can be seen with larger spatial scale and colder

cloud tops from 03-09z (Fig. 10a, columns 3-5). At 6z, the MCS is matched with

observation 31.1% for 0.03, 41.1% for 0.05, 50.3% for 0.07, and 49.7% for 0.10. This

results in a greater rainfall falling in the MCS region (within the 15-mm isopleth) with

0.07 and 0.10 radian cutoffs (Fig. 10c, bottom row). In *no_winds* experiments, 0.03, 0.05,

and 0.07 radian cutoffs show an MCS progression from 03-09z (Fig. 9b, rows 1-3,

columns 3-5) that are similar to each other where the localization does not have much

impact. At 6z, the MCS is matched with observation 36.2% for 0.03, 40.8% for 0.05,

42.0% for 0.07, but decreases to 24.8% for 0.10. Looking at the 0.07 rad cutoff at 03z,

the time that the MCS is strongest, the cloud tops are colder and are uniform in the

all_vars experiments (Fig. 9a, row 3, column 3) whereas in the *no_winds* experiments the

coldest cloud tops indicate more separate updrafts in a ring-like formation (Fig. 10c, row

3, column 3). Unlike in the *all_vars* experiments, the highest rainfall is located the

northern region and shows a linear pattern in the *no_winds* experiments for the 0.03, 0.05,

and 0.07 radian cutoffs (Fig. 10d, rows 1-3). With the 0.10 radian cutoff, any semblance

of an MCS disappears with a smaller region of cold cloud tops well to the north at 03z that dissipates by 06z (Fig. 10b, rows 4, columns 3-4) resulting in an obvious absence of rainfall in the observed MCS region (Fig. 10d, row 4).

The change in MCAPE (calculated by the wrf_cape_2d NCL function that uses the parcel with the highest θ^e) from a change in cutoff and adjusted variables for the DA cycles is shown from initialization (fh00) to 19z (fh07), around the time convection begins in the SMO, for the *all_vars* (Fig. 11a) and *no_winds* experiments (Fig. 11b). For the *all_vars* experiments, although unclear in fh00 (Fig. 11a, column 1), with increased localization radius, there is increased MCAPE in the area of the MCS along and north of an imaginary line from KINO-ONVS at fh07 (Fig. 11a, column 8). For the *no_winds* experiments, there is no obvious MCAPE trend with increasing localization radius (Fig. 11b). The change in CAPE at f00, with the exception of the 0.03 rad cutoff, are of higher magnitude than the *no_winds* experiments, including a large area of MCAPE decrease of more than 1500 J/kg in the mid-GOC and a large increase of CAPE over the MCS region more expansive and of higher magnitude than the *all_vars* experiments. Despite the greater positive MCAPE change over the MCS region, the negative region over the GOC is advected north replacing much of the positive area by fh07. The *all_vars* experiments cutoffs of 0.05 and 0.07 have the highest MCAPE change remaining in the MCS region at this time.

Looking at the 12z CHIH sounding for the *all_vars* experiments, the water vapor mixing ratio bias decreases with increased localization radius for with a range of less than 1 g kg^{-1} across all experiments throughout the column (Fig. 12a, row 1). For the *no_winds* experiments, there was no trend in biases with increased cutoff and the range is between

1 and 1.5 g kg^{-1} (Fig. 12c, row 1). The temperature has no trend with increased cutoff for both the *all_vars* (Fig. 12a, row 2) and *no_winds* experiments (Fig. 12c, row 2). The range is less than 1°C for the *all_vars* experiments and 1-2 °C for *no_winds* experiments. For the verification of winds, the *all_vars* and *no_winds* experiments acted opposite to each with increased cutoffs. Increased cutoffs lowered U biases in the profile with the *no_winds* experiments, but only U winds biases for the *all_vars* experiments were decreased above 700 hPa and actually increased below 700 hPa. For V winds, increased cutoff lowered biases below 750 hPa for the *all_vars* experiments, but actually raised biases for the*no_winds* experiments.

2) OTHER IMPACTS

The impact of an ensemble forecast (*ens_fx* experiment) is explored versus a deterministic forecast (*assim* experiment). (Note that the *ens_fx* experiment uses a 0.05 rad cutoff while the *assim* experiment uses a 0.07 radian cutoff.) In the *ens_fx* experiment, the PWV RMSE and mean bias amplitude is lower. By fh02, the RMSE increases 1.5 mm from 0.5 mm at initialization to where it stabilizes at 2 mm through at least fh07 while the mean bias decreases slowly from +5 mm (moist bias) to -1 mm (dry bias) (Fig. 13a, row 2), resembling a smoothed version of the *assim* experiment hindcast (Fig. 13b, row 1). The coldest brightness temperature associated with the MCS at the strongest time of 03z Jul 9 is tapered from -80 °C in the *assim* experiment (Fig. 14a, row 1, column 3) to -60 °C in *ens_fx* (Fig. 14a, row 3, column 3) demonstrating that the ensemble mean removes the extremes in the updraft locations of widespread convection whose locations within the MCS are not collocated throughout the convection. At 6z, the MCS is only matched 36.3%. The coldest brightness temperature in the *ens_fx*

experiment is a small circle of -80 °C at 21z Jul 8 just south of MULT (Fig. 14a, row 3, column 1) near the beginning of the diurnal convection before the MCS forms. This circle indicates a common initialization point across the ensemble whereas the brightest areas in the actual MCS are less uniform.

In the *super_ob* experiment, we investigate the impact of assimilating a mean hourly observation as compared to individual 5-min within the +/- 30 minutes of analysis time. Compared to all other experiments including the *assim* experiment, the RMSE PWV increments across all stations are smaller than the *super_ob* experiment (Fig. 13a, row 3). At forecast initialization, the posterior is 1 mm which is higher than the *assim* experiment of 0.5 mm. The PWV biases within the increments are also smaller at ~0.25 mm and they do not become negative as in the *assim* experiment. Further, the mean RMSE increases to 2 mm at f01 and remains stable through f06. The shape of the MCS during its development at 03z Jul 9 has two main areas of convection instead of one larger one (Fig. 13a, row 2). This leads to less rainfall distributed in the middle of the highest rainfall in the observed MCS region (Fig. 14b, top-right panel). At 6z, the MCS was 41.7% matched, about 8% lower than in the *assim* experiment. The changes in MCAPE are of lower magnitude at f00 with an overall decrease in MCAPE by fh07 near the time of convective initiation (Fig. 15, row 2).

We look at doubling the number of ensemble members from 20 to 40 in the *40_mem* experiment and its impact on the DA and forecast. Surprisingly, this resulted in the poorest forecast. From the first 4 cycles, both the prior and posterior PWV RMSE decrease quickly to 0.5 mm before stabilizing throughout the final cycles (Fig. 13a, row 4). The increment is less than 1 mm until cycle 4 where at which time the increment

becomes ~1.5 mm. The mean PWV bias increment becomes zero by cycles 3 and 4 when the model advance increases the mean bias 1 mm for the last cycles. At fh00, the posterior PWV is just above 0.5, slightly higher than the *assim* experiment. From fh00 to fh02, the PWV RMSE increases to 2.5 mm where it stabilizes. At 6z, the MCS is very low at 23.9% matched. The poorer MCS simulation is shown in the lack of a continuous area of cold cloud tops throughout the observed MCS time (Fig. 14a, row 3) thereby minimizing rainfall (Fig. 14c, row 3). The MCAPE difference is negative along the SMO coast up to 100 km inland at fh00 (Fig. 15, row 3). This negative difference is advected eastward into the MCS region by fh07.

Finally, we investigate the impact of 12-hour spinup with the *warm_start* experiments versus no spinup with the *cold_start* experiment. At fh00, the PWV RMSE for the *cold_start* experiment is ~2 mm while the *warm_start* experiment is 2.5 mm (Fig. 13b). The higher RMSE in the *warm_start* experiment is contributed by a higher bias (~1.5 mm) compared to the *cold_start* experiment (1 mm). Biases trend to become negative throughout the first 7 hours of the forecast. Despite the higher bias and RMSE, the *warm_start* experiment produces a better forecast of MCS in terms of the progression of the clouds via the brightness temperature (Fig. 14b) where at 6z the MCS is 37.2% matched up from the 34.3% match in the *cold_start* experiment. A conclusion of an improved MCS forecast in the *warm_start* experiment is not as straightforward by looking at the total rainfall, but superfluous precipitation has decreased in the western slope of the southern SMO (Fig. 14d)

5. Discussion

The assimilation of GPS-PWV into the convection-allowing WRF forecasts is

140

successful in lowering the PWV RMSE. There is also limited positive impacts to the simulation of the MCS in our study. This does not appear to be an initial value problem but an error within the model.

a) Decreased PWV RMSE at the site locations

For all experiments, during the 6-hour DA partial cycling period leading up to forecast initialization, analysis PWV RMSE across all 9 stations decreases with each cycle (Figs. 9a and 13a), indicating that the assimilation worked. In general, the PWV RMSE increments are smallest at the beginning and increased past the first 2-4 cycles. At the beginning of the cycling period, the RMSE is ~ 3 mm, but decreases to a range of 0.25-1 mm after the final analysis. The mean PWV bias, however, did not display this trend of decreasing throughout the cycling period (Figs. 9b and 13b). In the first DA cycle, the prior mean PWV bias was +1 mm. After the first cycle brings the PWV to ~ 0 mm, the model advance in subsequent cycles in general increases the bias 0.5-1 mm. An anomaly to this behavior is the model advance between cycles 2-3 where the model advance decreases the PWV bias resulting in a dry bias in the *all_vars* experiments whereas the model advance does not change the PWV bias (near 0 mm) in the *no_winds* experiments. And important consideration is that the RMSE cannot be evaluated outside of the 9 site locations.

b) Response of MCS

Without the use of an ensemble or DA, there was improvement of MCS shape and progression just by adding a 12-h spin-up. This improvement was greater from the *cold_start* experiment (no spin-up) to the *warm_start* experiment (12-h spin-up) (Fig. 14b) than from the *no_assim* experiment (same as the *warm_start* experiment, but with

an ensemble spin-up) to the *assim* experiment (Fig. 5a, rows 3-4). Despite adding a moist bias and wind degradation in *warm_start* experiment, a continuous MCS is simulated especially at 03-06z Jul 9. In the spin-up, flow and circulations have time to become more realistic at the convective-allowing domain from the 0.25° GFS ICs and LBCs allowing for the formation of an MCS. (The ensemble during spin-up also improves the MCS in terms of intensity and coverage.) No experiment was able to simulate the expanse and duration of the observed MCS.

Despite a reduction of PWV RMSE, we did not find a robust relationship between localized improved PWV and better skill of rainfall that begins 7+ hours after initialization. This may appear at first to go against Kursinski et al. (2008), but there were 3 important differences in that study: a) QVAPOR was multiplied by 5% difference, b) no other variables were adjusted, c) no localization (adjustments were grid-wide). Assimilating the observations from the nine GPS-PWV sites does not necessarily improve the location and strength of the MCS from the *no_assim* experiment; improvement depends on the turning parameters such as covariance localization radius and adjusted variable localization. In the *all_vars* experiments, 0.03 and 0.05 radian cutoffs bring the location of the MCS further south to along with observation but degrade the shape of the MCS whereas 0.07 and 0.10 rad cutoffs improve the shape of MCS along with its rainfall. In the *no_winds* experiments, the highest rainfall is confined to the northern edge of the observed MCS in 0.03, 0.05, and 0.07 radian cutoffs with a slight improvement from the *no_assim* experiments. In the 0.10 radian cutoff, the semblance of an MCS is missing.

c) Representativeness of GPS-PWV observations

The conversion of GPS to PWV with GIPSY software results in a cone that represents an atmosphere with a 15-20 km radius centered at the site location. During the 6-hour DA cycle period, there was convection from an MCS from the previous evening nearby the ONVS-MULT-BASC and Tropical Storm Erick whose northern edge was in close proximity to MOCH-BGTO. According to Seko et al. (2011), convective rainfall changes the PWV in a localized area making the representation radius much smaller. Convection combined with complex terrain may not allow for realistic spatial sensitivities to other state variables. PWV east of the SMO crest correlated with PWV west of the SMO where the air masses are very different although one can affect the other. A new covariance function can be introduced that stops at terrain. Further, there is no unique solution for QVAPOR adjustments when assimilating PWV as it is a vertically-integrated quantity. One could adjust QVAPOR layers differently to achieve identical PWV.

d) Sources of error

   1) IMPERFECT MODEL

Throughout all experiments, there is a short memory of initial WRF moisture adjustments indicating an imperfect model giving way to model error. In other words, the adjustments in QVAPOR at forecast initialization do not maintain for more than a couple of hours after the simulation begins. For instance, in the *assim* experiment, from 0.5 mm RMSE at f00, the RMSE increases to or surpasses the 2 mm in the *no_assim* experiment between fh01 and fh02. RMSE stabilizes at ~2-3 mm after fh01. Mean moist PWV biases of ~1 mm decrease with time after forecast initialization and becoming negative by fh09.

An imperfect model can also be attributed to performance of the *40_mem* experiment where one would expect the 40-member ensemble to be more representative of the errors with double the degrees of freedom. Even though the model reaches minimum posterior PWV RMSE after 3 cycles, the increment is still large (> 1 mm) between cycles. Out of the all our experiments, this one simulates the MCS the poorest.

2) INSTABILITY FROM ASSIMILATING OBSERVATIONS INDIVIDUALLY

Inconsistent analysis is observed between the *super_ob* experiment with smaller increments and the *assim* experiment with relatively larger increments. In the *assim* experiment, 12 5-min observations are assimilated +/- 30 minutes from the hourly observation time at each site. In the *super_ob* experiment, the mean of the observations in the 1-h window is assimilated. Nerger (2015) notes that combination of serial processing of observations combined with localization can cause destabilization of the analysis in the EnKF. They also show that different orders of the observations result in different analyses as an indication of the instability.

3) LACK OF STATIONS AND OBSERVATION VARIABLES

Only 9 stations are used to assimilate PWV across northwest Mexico in complex terrain. In order for the assimilation to have an impact on more than just a bull eye around each site, larger cutoffs are needed. Covariances across terrain, even with the use of localization, may not be realistic especially with the model error in mind. Variables that are far away from the site may have spurious correlations despite the use of localization. Assimilating other variables than PWV would benefit the forecast. Wind data assimilated from soundings, for example, would correct the bias that develops during the 12-hour period (6-hour spinup + 6 hourly DA cycles). Winds do not usually improve if they are

adjusted with only the covariances relative to the GPS-PWV observations in our study.

 4) INCORRECT OBSERVATION OPERATOR

The observation operator uses the 4 closest grid points to calculate PWV. This equates to an area that is 5 km x 5 km. However, the GPS-PWV observation is a cone representative of an area that has a radius of 20-30 km across a proportion of the lower atmosphere where the majority of the moisture resides. An example observation operator that would capture the GPS-PWV footprint would be to use a 7 x 7 grid box representing a 17.5 km x 17.5 km area. Additionally, instead of the site location treated a point where all grid boxes are bilinearly interpolated, each box would contribute 1/49 to the PWV at the site.

6. Summary and Recommendations

a) Summary

We performed the assimilation of GPS-PWV in a series of hindcast simulations on 8 July 2013, a day that Moker et al. (2018) had identified as "weakly forced", using the "linear least squares" implementation (Anderson 2003) of the EAKF (Anderson 2001) via DART software (Anderson et al. 2009). The control experiments consisited of the *cold_start*, *no_assim*, and *assim* experiments (Fig. 2). For the *assim* experiment, a Gaspari and Cohn (1999) location cutoff of 0.07 radians was chosen because the mean function across all locations matched up with the magnitude of an MCS (Fig. 3b, column 3). The vertical-equivalent distance also included the highest correlations of PWV to vertical moisture distribution in observed soundings during the NAM season (Fig. 3c). All WRF state variables were adjusted in the EAKF (Table 6).

The simulation of the MCS improves with the spin-up (12 h) because the meso-γ-

scale flow is being established in the convection-resolving domain during that time. We use the criteria defined in Maddox (1980) to determine the observed MCS and then as a metric determine the proportion of the simulated MCS matched the location of the observed MCS. At 6z, the *cold_start* experiment showed a 34.3% MCS match. Then, with the influence of spin-up, the *warm_start* experiment (deterministic spin-up) showed 37.2% match and the *warm_start* experiment (ensemble spin-up) showed a 41.6% match, Then, the GPS-PWV DA improves the MCS location bringing it south closer to the observation as well as increasing intensity to where at 6z there is a 50.2% match in the *assim* experiment.

Within the hourly DA cycles leading up to forecast initialization, PWV RMSE time series across the 9 sites minimizes with time indicating that the DA is functioning properly. At CHIH, the water vapor mixing ratio improved to within instrument error below 700 hPa. The T profile also improved but the wind profile shows mixed results. At MAZT, there is negligible changes to QVAPOR as it is 150 km away from the closest GPS site (BGTO). Both upper air sites show an increased wind bias from the spin-up. In addition, at the time of convective initiation (18z), there is increased MUCAPE in the region where MCS reaches peak intensity a few hours later (near KINO-ONVS).

In addition to determining the role of spin-up in the *warm_start* experiment, we also run a series of other sensitivity experiments involving spatial localization cutoff distances, localization of the adjusted variables (thermodynamic variables only versus all variables), ensemble size, "super-obbing", and deterministic versus ensemble forecast. In the *all_vars* experiments, the simulation of the MCS improves with increased cutoff length with 0.07 and 0.10 radian being the best. In the *no_winds* experiments, there is not

146

much change between 0.03-0.07 radian cutoffs with the rainfall was focused towards the northern edge of the observed MCS rainfall. By 0.10 rad, the simulated MCS was degraded. The *super_ob* experiment provided the smallest increments during the DA cycles and simulated an MCS close to the *assim* experiment. The *40_mem* experiment minimized PWV error quickest during the DA cycles but produced the poorest MCS simulation out of all experiments that is attributed to model error.

b) Recommendations

A lack of vertical information is a source of uncertainty when assimilating an integrated quantity such as GPS-PWV. Expanding on NAME over a decade ago, we recommend a field campaign in this region to assess statistical connections of quantities of moisture and winds in three dimensions. One way to address this is with the addition of lidars that can continuously observe atmospheric variables a few km from the ground.

Addressing model configuration and model error is independent of DA. For the configuration, add more vertical layers to the model as in other convection-allowing ensemble DA forecasting studies. For example, Schwartz et al. (2014) used 40 vertical levels. For model error, investigate the role of model dynamics and parameterizations such as cloud microphysics. One way to address this is to create a 20-member ensemble with random microphysics schemes.

Acknowledgements

References

Adams, D. K., and A. C. Comrie, 1997: The North American Monsoon. *Bull. Am. Meteorol. Soc.*, **78**, 2197–2213, doi:10.1175/1520-0477(1997)078<2197:TNAM>2.0.CO;2.

——, R.M. Fernandes, and J.M. Maia, 2011: GNSS Precipitable Water Vapor from an Amazonian Rain Forest Flux Tower. *J. Atmos. Oceanic Technol.*, **28**, 1192–1198, https://doi.org/10.1175/JTECH-D-11-00082.1

——, C. Minjarez, Y. Serra, A. Quintanar, L. Alatorre, A. Granados, E. Vazquez, and J. Braun, 2014: Mexican GPS tracks convection from North American Monsoon. *Eos (Washington. DC).*, **95**, 61–62, doi:10.1002/2014EO070001.

Ancell, B. C., C. F. Mass, and G. J. Hakim, 2011: Evaluation of Surface Analyses and Forecasts with a Multiscale Ensemble Kalman Filter in Regions of Complex Terrain. *Mon. Weather Rev.*, **139**, 2008–2024, doi:10.1175/2010mwr3612.1.

Anderson, J., T. Hoar, K. Raeder, H. Liu, N. Collins, R. Torn, and A. Arellano, 2009: The data assimilation research testbed a community facility. *Bull. Am. Meteorol. Soc.*, **90**, 1283–1296, doi:10.1175/2009BAMS2618.1.

——, 2001: An Ensemble Adjustment Kalman Filter for Data Assimilation. *Mon. Weather Rev.*, **129**, 2884–2903, doi:10.1175/1520-0493(2001)129<2884:AEAKFF>2.0.CO;2. http://journals.ametsoc.org/doi/abs/10.1175/1520-0493%282001%29129%3C2884%3AAEAKFF%3E2.0.CO%3B2.

——, 2003: A Local Least Squares Framework for Ensemble Filtering. *Mon. Weather*

Rev., **131**, 634–642, doi:10.1175/1520-0493(2003)131<0634:ALLSFF>2.0.CO;2. http://journals.ametsoc.org/doi/abs/10.1175/1520-0493%282003%29131%3C0634%3AALLSFF%3E2.0.CO%3B2.

——, 2009: Spatially and temporally varying adaptive covariance inflation for ensemble filters. *Tellus, Ser. A Dyn. Meteorol. Oceanogr.*, **61 A**, 72–83, doi:10.1111/j.1600-0870.2008.00361.x.

Barker, D., and Coauthors, 2012: The weather research and forecasting model's community variational/ensemble data assimilation system: WRFDA. *Bull. Am. Meteorol. Soc.*, **93**, 831–843, doi:10.1175/BAMS-D-11-00167.1.

Bevis, M., S. Businger, T. A. Herring, C. Rocken, R. A. Anthes, and R. H. Ware, 1992: GPS meteorology: Remote sensing of atmospheric water vapor using the global positioning system. *J. Geophys. Res.*, **97**, 15787, doi:10.1029/92JD01517. http://doi.wiley.com/10.1029/92JD01517.

Bieda, S. W., C. L. Castro, S. L. Mullen, A. C. Comrie, and E. Pytlak, 2009: The relationship of transient upper-level troughs to variability of the North American monsoon system. *J. Clim.*, **22**, 4213–4227, doi:10.1175/2009JCLI2487.1.

Castro, C. L., H. I. Chang, F. Dominguez, C. Carrillo, J. K. Schemm, and H. M. H. Juang, 2012: Can a regional climate model improve the ability to forecast the North American monsoon? *J. Clim.*, **25**, 8212–8237, doi:10.1175/JCLI-D-11-00441.1.

Chou, M.-D., and M. J. Suarez, 1999: A solar radiation parameterization for atmospheric studies. Tech. Memo. NASA/TM1999-104606, Vol. 15, 38 pp., http://gmao.gsfc.nasa.gov/pubs/docs/Chou136.pdf.

Douglas, M. W., R. A. Maddox, K. Howard, and S. Reyes, 1993: The Mexican monsoon.

J. Clim., **6**, 1665–1677, doi:10.1175/1520-0442(1993)006<1665:TMM>2.0.CO;2.

Duan, J., and Coauthors, 1996: GPS meteorology: Direct estimation of the absolute value of precipitable water. *J. Appl. Meteor.*, **35**, 830–838.

Gao, X., J. Li, and S. Sorooshian, 2007: Modeling intraseasonal features of 2004 North American monsoon precipitation. *J. Clim.*, **20**, 1882–1896, doi:10.1175/JCLI4100.1.

Gaspari, G., and S. E. Cohn, 1999: Construction of correlation functions in two and three dimensions. *Q. J. R. Meteorol. Soc.*, **125**, 723–757, doi:10.1256/smsqj.55416.

Gochis, D. J., A. Jimenez, C. J. Watts, J. Garatuza-Payan, and W. J. Shuttleworth, 2004: Analysis of 2002 and 2003 Warm-Season Precipitation from the North American Monsoon Experiment Event Rain Gauge Network. *Mon. Weather Rev.*, **132**, 2938–2953, doi:10.1175/mwr2838.1.

——, C. J. Watts, J. Garatuza-Payan, and J. Cesar-Rodriguez, 2007: Spatial and temporal patterns of precipitation intensity as observed by the NAME Event Rain gauge Network from 2002 to 2004. *J. Clim.*, **20**, 1734–1750, doi:10.1175/JCLI4092.1.

Guo, Y.-R., Y.-H. Kuo, J. Dudhia, D. Parsons, and C. Rocken, 2000: Four-Dimensional Variational Data Assimilation of Heterogeneous Mesoscale Observations for a Strong Convective Case. *Mon. Weather Rev.*, **128**, 619–643, doi:10.1175/1520-0493(2000)128<0619:FDVDAO>2.0.CO;2. https://doi.org/10.1175/1520-0493(2000)128%3C0619:FDVDAO%3E2.0.CO.

Gutman, S. I., S. R. Sahm, S. G. Benjamin, B. E. Schwartz, K. L. Holub, J. Q. Stewart, and T. L. Smith, 2004: Rapid Retrieval and Assimilation of Ground Based GPS Precipitable Water Observations at the NOAA Forecast Systems Laboratory: Impact on Weather Forecasts. *J. Meteorol. Soc. Japan*, **82**, 351–360,

doi:10.2151/jmsj.2004.351.

http://joi.jlc.jst.go.jp/JST.JSTAGE/jmsj/2004.351?from=CrossRef.

Ha, S.-Y., Y.-H. Kuo, Y.-R. Guo, and G.-H. Lim, 2003: Variational Assimilation of
Slant-Path Wet Delay Measurements from a Hypothetical Ground-Based {GPS}
Network. Part {I}: Comparison with Precipitable Water Assimilation. *Mon. Weather
Rev.*, **131**, 2635–2655, doi:10.1175/1520-
0493(2003)131<2635:VAOSWD>2.0.CO;2.
http://journals.ametsoc.org/doi/abs/10.1175/1520-
0493(2003)131%3C2635:VAOSWD%3E2.0.CO;2.

Higgins, W., and D. Gochis, 2007: Synthesis of Results from the North American
Monsoon Experiment (NAME) Process Study. *J. Clim.*, **20**, 1601–1607,
doi:10.1175/JCLI4081.1.

——, and Coauthors, 2006: The NAME 2004 field campaign and modeling strategy.
Bull. Am. Meteorol. Soc., **87**, 79–94, doi:10.1175/BAMS-87-1-79.

Hong, S.-Y., Y. Noh, and J. Dudhia, 2006: A New Vertical Diffusion Package with an
Explicit Treatment of Entrainment Processes. *Mon. Weather Rev.*, **134**, 2318–2341,
doi:10.1175/mwr3199.1. https://doi.org/10.1175/MWR3199.1.

——, and J. Lim, 2006: The WRF single-moment 6-class microphysics scheme (WSM6).
J. Korean Meteor. Soc, **42**, 129–151, doi:10.5194/acp-2016-162.
http://www2.mmm.ucar.edu/wrf/users/phys_refs/MICRO_PHYS/WSM6.pdf.

Huffman, G. J., and Coauthors, 2007: The TRMM Multisatellite Precipitation Analysis
(TMPA): Quasi-Global, Multiyear, Combined-Sensor Precipitation Estimates at
Fine Scales. *J. Hydrometeorol.*, **8**, 38–55, doi:10.1175/JHM560.1.

https://doi.org/10.1175/JHM560.1.

Hunt, B. R., E. J. Kostelich, and I. Szunyogh, 2007: Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter. *Phys. D Nonlinear Phenom.*, **230**, 112–126, doi:https://doi.org/10.1016/j.physd.2006.11.008. http://www.sciencedirect.com/science/article/pii/S0167278906004647.

Iacono, M. J., J. S. Delamere, E. J. Mlawer, M. W. Shephard, S. A. Clough, and W. D. Collins, 2008: Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models. *J. Geophys. Res. Atmos.*, **113**, doi:10.1029/2008JD009944. https://doi.org/10.1029/2008JD009944.

Janowiak, J. E., V. J. Dagostaro, V. E. Kousky, and R. J. Joyce, 2007: An examination of precipitation in observations and model forecasts during NAME with emphasis on the diurnal cycle. *J. Clim.*, **20**, 1680–1692, doi:10.1175/JCLI4084.1.

Johnson, R. H., P. E. Ciesielski, B. D. McNoldy, P. J. Rogers, and R. K. Taft, 2007: Multiscales variability of the flow during the North American Monsoon experiment. *J. Clim.*, **20**, 1628–1648, doi:10.1175/JCLI4087.1.

Kain, J. S., 2004: The Kain–Fritsch Convective Parameterization: An Update. *J. Appl. Meteorol.*, **43**, 170–181, doi:10.1175/1520-0450(2004)043<0170:TKCPAU>2.0.CO;2. https://doi.org/10.1175/1520-0450(2004)043%3C0170:TKCPAU%3E2.0.CO.

Knapp, K. R., and Coauthors, 2011: Globally Gridded Satellite Observations for Climate Studies. *Bull. Am. Meteorol. Soc.*, **92**, 893–907, doi:10.1175/2011BAMS3039.1. https://doi.org/10.1175/2011BAMS3039.1.

Kuo, Y.-H., Y.-R. Guo, and E. R. Westwater, 1993: Assimilation of Precipitable Water

Measurements into a Mesoscale Numerical Model. *Mon. Weather Rev.*, **121**, 1215–1238. https://doi.org/10.1175/1520-0493(1993)121%3C1215:AOPWMI%3E2.0.CO;2.

——, X. Zou, and Y.-R. Guo, 1996: Variational Assimilation of Precipitable Water Using a Nonhydrostatic Mesoscale Adjoint Model. Part I: Moisture Retrieval and Sensitivity Experiments. *Mon. Weather Rev.*, **123**, 122–147, doi:10.1175/1520-0493(1996)124<0122:VAOPWU>2.0.CO;2.

Kursinski, E. R., D. K. Adams, and M. Leuthold, 2008: GPS observations of precipitable water and implications for the predictability of precipitation during the North American monsoon. *CLIVAR Exchanges*, No. 45, International CLIVAR Project Office, Southampton, United Kingdom, 14, 19–21.

Lang, T. J., D. A. Ahijevych, S. W. Nesbitt, R. E. Carbone, S. A. Rutledge, and R. Cifelli, 2007: Radar-observed characteristics of precipitating systems during NAME 2004. *J. Clim.*, **20**, 1713–1733, doi:10.1175/JCLI4082.1.

Li, J., S. Sorooshian, W. Higgins, X. Gao, B. Imam, and K. Hsu, 2008: Influence of spatial resolution on diurnal variability during the north American monsoon. *J. Clim.*, **21**, 3967–3988, doi:10.1175/2008JCLI2022.1.

Liu, Y., and Coauthors, 2012: Advancing data assimilation in operational hydrologic forecasting: Progresses, challenges, and emerging opportunities. *Hydrol. Earth Syst. Sci.*, **16**, 3863–3887, doi:10.5194/hess-16-3863-2012.

Lorenc, A. C., 1981: A global three-dimensional multivariate statistical interpolation scheme. *Mon. Wea. Rev.*, **109**, 701–721.

Maddox, R. A., 1980: Mesoscale convective complexes. *Bull. Amer. Meteor. Soc.*, **61**,

1374–1387, doi:https://doi.org/10.1175/1520-0477(1980)061<1374:MCC>2.0.CO;2.

Meng, Z., and F. Zhang, 2007: Tests of an Ensemble Kalman Filter for Mesoscale and Regional-Scale Data Assimilation. Part II: Imperfect Model Experiments. *Mon. Weather Rev.*, **135**, 1403–1423, doi:10.1175/MWR3352.1. http://journals.ametsoc.org/doi/abs/10.1175/MWR3352.1.

——, and ——, 2008: Tests of an Ensemble Kalman Filter for Mesoscale and Regional-Scale Data Assimilation. Part III: Comparison with 3DVAR in a Real-Data Case Study. *Mon. Weather Rev.*, **136**, 522–540, doi:10.1175/2007MWR2106.1. http://journals.ametsoc.org/doi/abs/10.1175/2007MWR2106.1.

Minjarez-Sosa, C. M., C. L. Castro, K. L. Cummins, E. P. Krider, and J. Waissmann, 2012: Toward Development of Improved QPE in Complex Terrain Using Cloud-to-Ground Lightning Data: A Case Study for the 2005 Monsoon in Southern Arizona. *J. Hydrometeorol.*, **13**, 1855–1873, doi:10.1175/jhm-d-11-0129.1.

Mo, K. C., E. Rogers, W. Ebisuzaki, R. W. Higgins, J. Woollen, and M. L. Carrera, 2007: Influence of the North American monsoon experiment (NAME) 2004 enhanced soundings of NCEP operational analysis. *J. Clim.*, **20**, 1821–1842, doi:10.1175/JCLI4083.1.

Moker, J. M., C. L. Castro, A. F. Arellano, Y. L. Serra, and D. K. Adams, 2018: Convective-permitting hindcast simulations during the North American Monsoon GPS Transect Experiment 2013: Establishing baseline model performance without data assimilation. *J. Appl. Meteorol. Climatol.*, **57**, 1683–1710, doi:10.1175/JAMC-D-17-0136.1.

Moncrieff, M. W., and M. J. Miller, 1976: The dynamics and simulation of tropical cumulonimbus and squall lines. *Q. J. R. Meteorol. Soc.*, **102**, 373–394, doi:10.1002/qj.49710243208.

Nerger, L., 2015: On Serial Observation Processing in Localized Ensemble Kalman Filters. *Mon. Weather Rev.*, **143**, 1554–1567, doi:10.1175/mwr-d-14-00182.1.

Nesbitt, S. W., D. J. Gochis, and T. J. Lang, 2008: The Diurnal Cycle of Clouds and Precipitation along the Sierra Madre Occidental Observed during NAME-2004: Implications for Warm Season Precipitation Estimation in Complex Terrain. *J. Hydrometeorol.*, **9**, 728–743, doi:10.1175/2008JHM939.1. http://journals.ametsoc.org/doi/abs/10.1175/2008JHM939.1.

Newman, A., and R. H. Johnson, 2012: Mechanisms for Precipitation Enhancement in a North American Monsoon Upper-Tropospheric Trough. *J. Atmos. Sci.*, **69**, 1775–1792, doi:10.1175/JAS-D-11-0223.1. http://journals.ametsoc.org/doi/abs/10.1175/JAS-D-11-0223.1.

Oigawa, M., T. Tsuda, H. Seko, Y. Shoji, and E. Realini, 2018: Data assimilation experiment of precipitable water vapor observed by a hyper-dense GNSS receiver network using a nested NHM-LETKF system. *Earth, Planets Sp.*, **70**, doi:10.1186/s40623-018-0851-3. https://doi.org/10.1186/s40623-018-0851-3.

Parrish, D. F., and J. C. Derber, 1992: The National Meteorological Center's Spectral Statistical-Interpolation Analysis System. *Mon. Weather Rev.*, **120**, 1747–1763, doi:10.1175/1520-0493(1992)120<1747:TNMCSS>2.0.CO;2. https://doi.org/10.1175/1520-0493(1992)120%3C1747:TNMCSS%3E2.0.CO.

Pearson, K. J., G. M. S. Lister, C. E. Birch, R. P. Allan, R. J. Hogan, and S. J.

Woolnough, 2014: Modelling the diurnal cycle of tropical convection across the "grey zone." *Q. J. R. Meteorol. Soc.*, **140**, 491–499, doi:10.1002/qj.2145.

Poli, P., and Coauthors, 2007: Forecast impact studies of zenith total delay data from European near real-time GPS stations in Météo France 4DVAR. *J. Geophys. Res. Atmos.*, **112**, doi:10.1029/2006JD007430. https://doi.org/10.1029/2006JD007430.

Purser, R. J., D. F. Parrish, and M. Masutani, 2000: Meteorological observational data compression; An alternative to conventional "super-obbing". NCEP Office Note 430, 12 pp. [Available online at

http://www.emc.ncep.noaa.gov/mmb/papers/purser/on430.pdf].

Pytlak, E., M. Goering, and A. Bennett, 2005: Upper tropospheric troughs and their interaction with the North American monsoon. *19th Conf. on Hydrology*, San Diego, CA, Amer. Meteor. Soc., P2.3, https://ams.confex.com/ams/pdfpapers/85393.pdf.

Ray, A.J., G.M. Garfin, M. Wilder, M. Vásquez-León, M. Lenart, and A.C. Comrie, 2007: Applications of Monsoon Research: Opportunities to Inform Decision Making and Reduce Regional Vulnerability. *J. Climate*, **20**, 1608–1627, https://doi.org/10.1175/JCLI4098.1

Rocken, C., R. H. Ware, T. V. Hove, F. Solheim, C. Alber, J. Johnson, M. Bevis, and S. Businger, 1993: Sensing atmospheric water vapor with the Global Positioning System. *Geophys. Res. Lett.*, **20**, 2631–2634

SAITO, K., J. ISHIDA, K. ARANAMI, T. HARA, T. SEGAWA, M. NARITA, and Y. HONDA, 2007: Nonhydrostatic Atmospheric Models and Operational Development at JMA. *気象集誌. 第2輯*, **85B**, 271–304, doi:10.2151/jmsj.85B.271.

Schwartz, C. S., G. S. Romine, K. R. Smith, and M. L. Weisman, 2014: Characterizing

and Optimizing Precipitation Forecasts from a Convection-Permitting Ensemble
Initialized by a Mesoscale Ensemble Kalman Filter. *Weather Forecast.*, **29**, 1295–
1318, doi:10.1534/genetics.108.098772.
http://journals.ametsoc.org/doi/abs/10.1175/WAF-D-13-00145.1.

Seastrand, S., Y. Serra, C. Castro, and E. Ritchie, 2015: The dominant synoptic-scale
modes of North American monsoon precipitation. *Int. J. Climatol.*, **35**, 2019–2032,
doi:10.1002/joc.4104. http://doi.wiley.com/10.1002/joc.4104.

Seko, H., T. Miyoshi, Y. Shoji, and K. Saito, 2011: Data assimilation experiments of
precipitable water vapour using the LETKF system: Intense rainfall event over
Japan 28 July 2008. *Tellus, Ser. A Dyn. Meteorol. Oceanogr.*, **63**, 402–414,
doi:10.1111/j.1600-0870.2010.00508.x.

——, T. Tsuyuki, K. Saito, and T. Miyoshi, 2013: Development of a Two-way Nested
LETKF System for Cloud-resolving Model BT - Data Assimilation for
Atmospheric, Oceanic and Hydrologic Applications (Vol. II). S.K. Park and L. Xu,
Eds., Springer Berlin Heidelberg, Berlin, Heidelberg, 489–507
https://doi.org/10.1007/978-3-642-35088-7_20.

Serra, Y. L., and Coauthors, 2016: The north American monsoon GPS transect
experiment 2013. *Bull. Am. Meteorol. Soc.*, **97**, 2103–2115, doi:10.1175/BAMS-D-
14-00250.1.

Skamarock, W. C., and Coauthors, 2008: A Description of the Advanced Research WRF
Version 3. *NCAR Tech. NOTE*, **NCAR/TN**–**47**,
doi:10.1080/07377366.2001.10400427.

Smith, T. L., S. G. Benjamin, S. I. Gutman, and S. Sahm, 2007: Short-Range Forecast

Impact from Assimilation of GPS-IPW Observations into the Rapid Update Cycle. *Mon. Weather Rev.*, **135**, 2914–2930, doi:10.1175/MWR3436.1. http://journals.ametsoc.org/doi/abs/10.1175/MWR3436.1.

Suarez, A., H. D. Reeves, D. Wheatley, and M. Coniglio, 2012: Comparison of Ensemble Kalman Filter–Based Forecasts to Traditional Ensemble and Deterministic Forecasts for a Case Study of Banded Snow. *Weather Forecast.*, **27**, 85–105, doi:10.1175/waf-d-11-00030.1.

Tewari, M., and Coauthors, 2004: Implementation and verification of the unified NOAH land surface model in the WRF model. 20th Conf. on Weather Analysis and Forecasting/16th Conf. on Numerical Weather Prediction, Seattle, WA, Amer. Meteor. Soc., 14.2A, https://ams.confex.com/ams/pdfpapers/69061.pdf.

Torn, R. D., 2010: Performance of a Mesoscale Ensemble Kalman Filter (EnKF) during the NOAA High-Resolution Hurricane Test. *Mon. Weather Rev.*, **138**, 4375–4392, doi:10.1175/2010mwr3361.1.

——, G. J. Hakim, and C. Snyder, 2006: Boundary Conditions for Limited-Area Ensemble Kalman Filters. *Mon. Weather Rev.*, **134**, 2490–2502, doi:10.1175/MWR3187.1.

Whitaker, J. S., and T. M. Hamill, 2002: Ensemble Data Assimilation without Perturbed Observations. *Mon. Weather Rev.*, **130**, 1913–1924, doi:10.1175/1520-0493(2002)130<1913:EDAWPO>2.0.CO;2. http://journals.ametsoc.org/doi/abs/10.1175/1520-0493%282002%29130%3C1913%3AEDAWPO%3E2.0.CO%3B2.

Wolfe, D. E., and S. I. Gutman, 2000: Developing an operational, surface-based, GPS,

water vapor observing system for NOAA: Network design and results. *J. Atmos. Oceanic Technol.*, **17**, 426–440

Yan, X., V. Ducrocq, G. Jaubert, P. Brousseau, P. Poli, C. Champollion, C. Flamant, and K. Boniface, 2009: The benefit of GPS zenith delay assimilation to high-resolution quantitative precipitation forecasts: a case-study from COPS IOP 9. *Q. J. R. Meteorol. Soc.*, **135**, 1788–1800, doi:10.1002/qj.508. https://doi.org/10.1002/qj.508.

Zavaleta, F., and R. Vargas, 2012: WB/Mexico: Modernization of National Meteorological Service for improved climate change adaptation. Accessed 2 April 2019, World Bank. [Available online at www.worldbank.org/en/news/press-release/2012/05/17/Mexico-modernization-national-meteorological-service-for-improved-climate-change-adaptation.]

Tables

Table 1: Information on the datasets used.

Data	Resolution	Reference(s)	Source	Use
GPS-PWV	5 min	Adams et al. (2014) Serra et al. (2016)	ResearchWorks UW archive http://hdl.handle.net/1773/37267	Assimilation and IC / Forecast Verification
NCEP Global Forecast System (GFS) model fields	0.25°, 6 h	NCAR	https://rda.ucar.edu/datasets/ds084.1	IC / LBC for WRF-ARW simulations
TRMM/TMPA 3B42 TRMM Rainfall Estimate Data V7	0.25°, 3 h	Huffman et al. (2007)	https://disc.gsfc.nasa.gov/datasets/TRMM_3B42_V7/summary	Forecast Verification (rainfall)
Geostationary IR Channel Brightness Temperature - GridSat B1	0.07°, 3 h	Knapp et al. (2011)	https://www.ncdc.noaa.gov/gridsat/gridsat-index.php?name=data	Forecast Verification (MCS cloud shield)
Radiosonde data (CHIH and MAZT)	12 h	Earth System Research Laboratory (ESRL)	https://ruc.noaa.gov/raobs/	IC Verification (moisture, temperature, and winds)

Table 2: Configuration of the WRF-ARW for the simulations.

Parameter	Domain d01	Domain d02	Domain d03
Horizontal grid	159×99, $\Delta x = 30$ km	270×231, $\Delta x = 10$ km	460×548, $\Delta x = 2.5$ km
Vertical grid	27 levels, ptop = 50 hpa	Same as d01	Same as d01
Cumulus scheme	Kain-Fritch	Same as d01	None
PBL scheme	YSU	Same as d01	Same as d01
Explicit microphysics	WSM6	Same as d01	Same as d01
Radiation (longwave)	RRTMG	Same as d01	Same as d01
Radiation (shortwave)	Goddard	Same as d01	Same as d01
Land surface scheme	Unified Noah	Same as d01	Same as d01

Table 3: Configuration of DART parameters.

Parameter	Value
Filter type	EAKF (Ensemble Adjustment Kalman Filter)
CV3 variance scale factor	0.25
CV3 horizontal length scale factor	1.00
CV3 vertical length scale factor	1.50
Covariance localization	Gaspari-Cohn, 0.07 rad cutoff Horizontal distance: 446 km Vertical distance: 3.5 km
Adaptive inflation (prior only)	1.0, 0.6 (initial mean, spread)
Adaptive inflation flavor	Spatially-varying time-evolving
Adaptive localization threshold	disabled
Ensemble members	20
Observation type	GPS-PWV
Observation error	0.075 cm

Table 4: Description of assimilation experiments

|	Name	Description	Values/Notes
Control Experiments	*cold_start*	See Fig. 2, top panel	N/A
	no_assim	See Fig. 2, middle panel	N/A
	assim	See Fig. 2, bottom panel	Deterministic fx, 0.07 rad cutoff, 20 members, *all_vars*, no superob
Sensitivity Experiments	Spatial localization	Assess impact of assimilating observations using differing Gaspari-Cohn covariance cutoffs	Cutoffs of 0.03, 0.05, 0.07, and 0.10 rad (equivalent distances listed in Table 5)
	Adjusted variable localization	Assess impact of adjusting all variables (*all_vars*) vs. thermodynamic variables only (*no_winds*)	See Table 6 for adjusted variables for *all_vars* and *no_winds*.
	super_ob	Assess impact of assimilating mean hourly observations ("super ob")	Use 1-hr mean of 5-min PWV observations for DA cycles
	40_mem	Assess impact of 40 ensemble members for spinup/DA cycles	Same schem
	ens_fx	Assess impact of ensemble forecast	Use ensemble members of the posterior at 1200 UTC in *assim*
	warm_start	Assess impact of forecast spin-up from a warm start (12-hour spin-up) vs. cold start (0-h spin-up)	No assimilation. Deterministic runs.

Table 5: Gaspari-Cohn covariance function cutoffs and equivalent lengths.

Cutoff (radians)	Equivalent Distance (km)	
	Horizontal	Vertical
0.03	191 km	1.5 km
0.05	319 km	2.5 km
0.07	446 km	3.5 km
0.10	637 km	5.0 km

Table 6: Adjusted WRF variables in DART for *all_vars* and *no_winds* experiments.

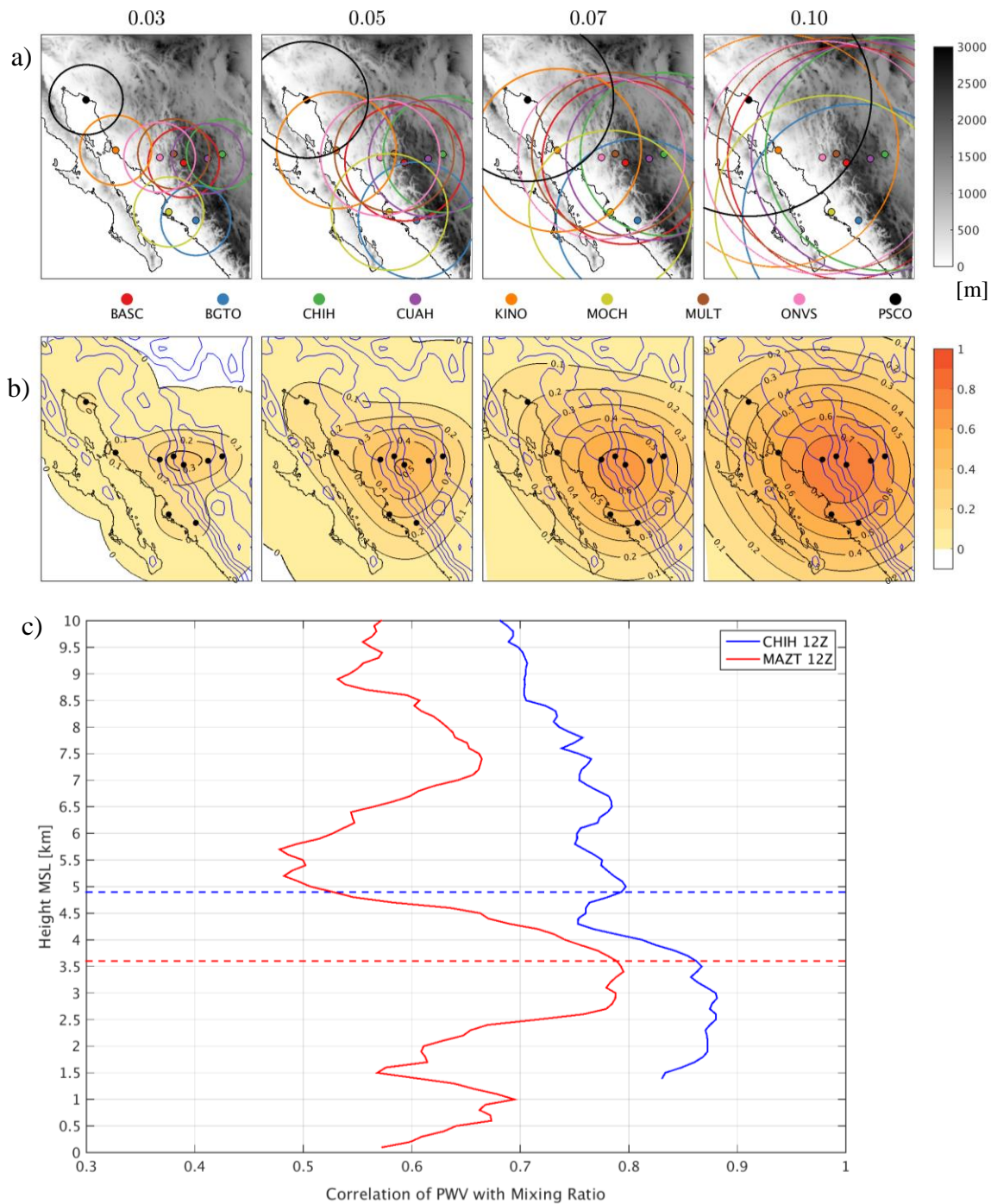| Variable description	Variable name	Dimensions	Variable type	*all_vars*	*no_winds*
Perturbation geopotential	PH	3-D	state	✓	✓
Perturbation potential temperature	T	3-D	state	✓	✓
Perturbation dry air mass in column	MU	2-D	state	✓	✓
Water vapor mixing ratio	QVAPOR	3-D	state	✓	✓
Cloud water mixing ratio	QCLOUD	3-D	state	✓	✓
Rain water mixing ratio	QRAIN	3-D	state	✓	✓
Ice mixing ratio	QICE	3-D	state	✓	✓
Snow mixing ratio	QSNOW	3-D	state	✓	✓
2-meter temperature	T2	2-D	diagnosed	✓	✓
x-wind component	U	3-D	state	✓	
y-wind component	V	3-D	state	✓	
z-wind component	W	3-D	state	✓	
2-meter potential temperature	TH2	2-D	diagnosed	✓	✓
2-meter specific humidity	Q2	2-D	diagnosed	✓	✓
Surface pressure	PSFC	2-D	diagnosed	✓	✓
10-meter U-wind component	U10	2-D	diagnosed	✓	
10-meter V-wind component	V10	2-D	diagnosed	✓	

Figures
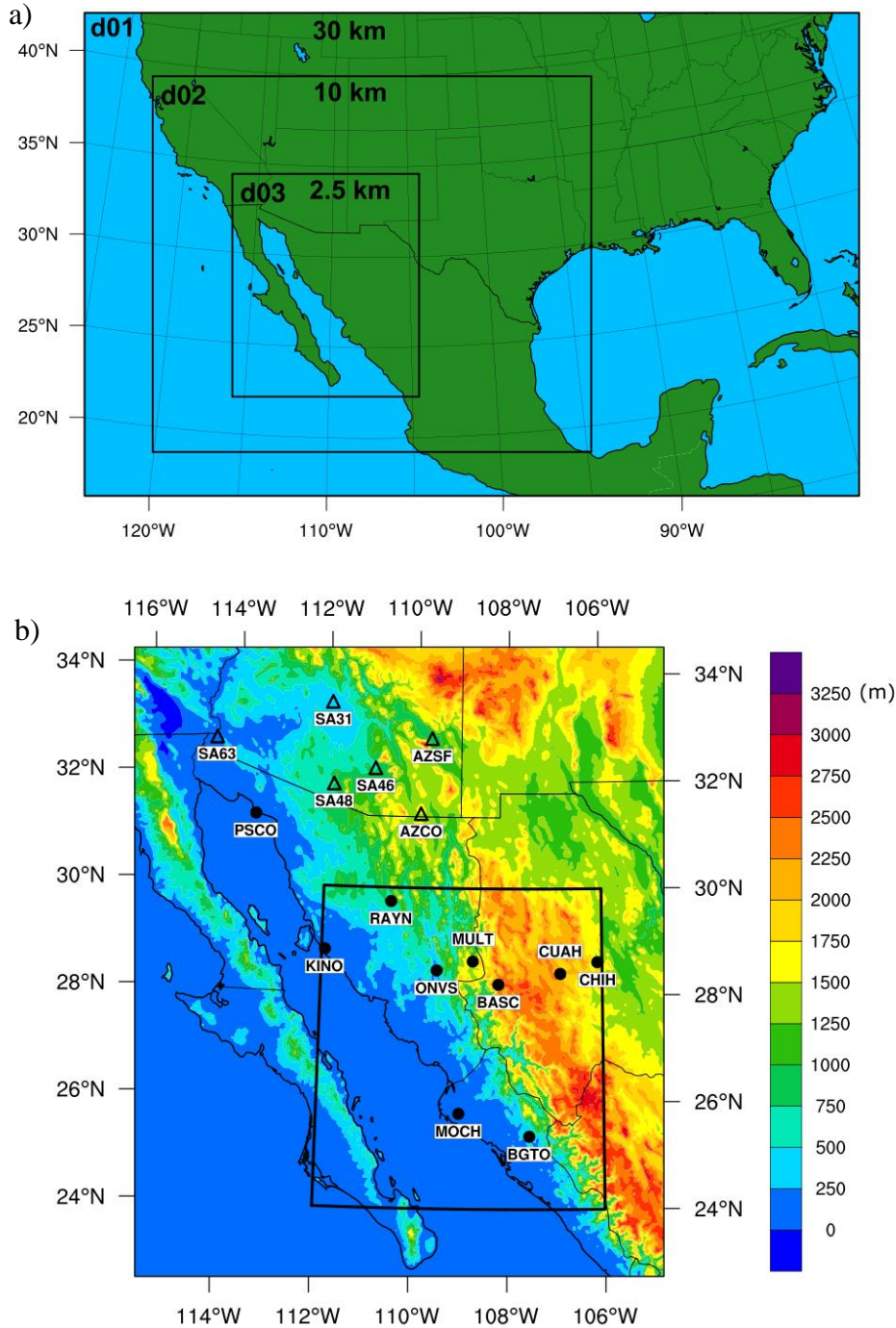
Figure 1: a) Identical to the setup of Moker et al. (2018) (Fig. 1 of that reference), the WRF-ARW nested domain (d01, d02, and d03) boundaries in thick black lines. Details of the domain configuration are listed in Table 1.

b) The convective-allowing domain (d03) with the locations of the Transect 2013 GPS-PWS sites in filled black circles. A thick black line outlines the NAM core region (24°-30° N and 112°-106° W) with the SMO transect (KINO-CHIH) situated in the northern part. Terrain is shaded every 250 m.

Figure 2: Schematic of the main experiments. Ensemble adjustment Kalman filter (EAKF) analysis are indicted in red circles and 1-h model advances are indicated in blue arrows.

Figure 3: a) Horizontal-equivalent distances for each GPS site location of the Gaspari-Cohn covariance localization function for the cutoffs of 0.03 (panel 1), 0.05 (panel 2), 0.07 (panel 3), and 0.10 radians (panel 4). Terrain shaded in gray [m]. b) Mean horizontal component of the function across all 9 GPS sites for each cutoff (panel) c) Correlation of PWV to mixing ratio from 1200 UTC soundings at Chihuahua (CHIH; blue line) and Mazatlán (MAZT; red line) during the 2013 NAM. Dashed lines indicate the vertical-equivalent distances for CHIH and MAZT for the cutoff of 0.07 radians (3.5 km). All horizontal- and vertical-equivalent distances for each cutoff are listed in Table 5.

169
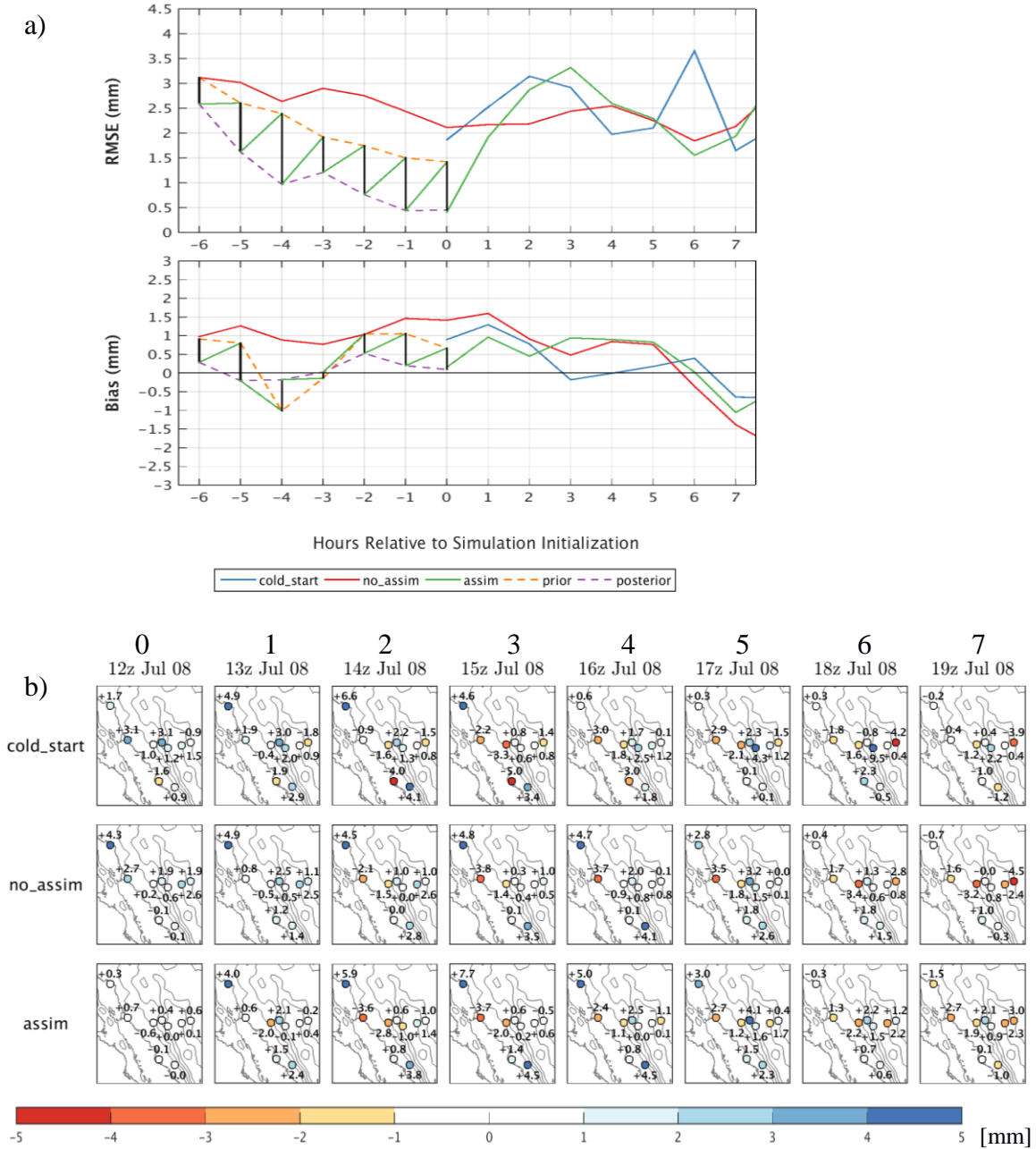
Figure 4: a) PWV RMSE (top panel) and mean bias (model – observation; bottom panel) (mm) across the 9 GPS-PWV stations in Transect 2013 for the hourly DA cycles (06-12z Jul 8) and for the first 7 hours of the forecast (12-19z Jul 8) for *cold_start* (solid blue line), *no_assim* (solid red line), and *assim* (solid green line), as well as prior (dashed orange line) and posterior (dashed purple line) during the DA cycles.
b) Hourly PWV biases (mm) at the GPS-PWV stations for the first 7 hours of the forecast (12-19z Jul 8) for *cold_start* (row 1), *no_assim* (row 2), and *assim* (row 3). Red indicates a dry model bias and blue indicates a wet model bias.
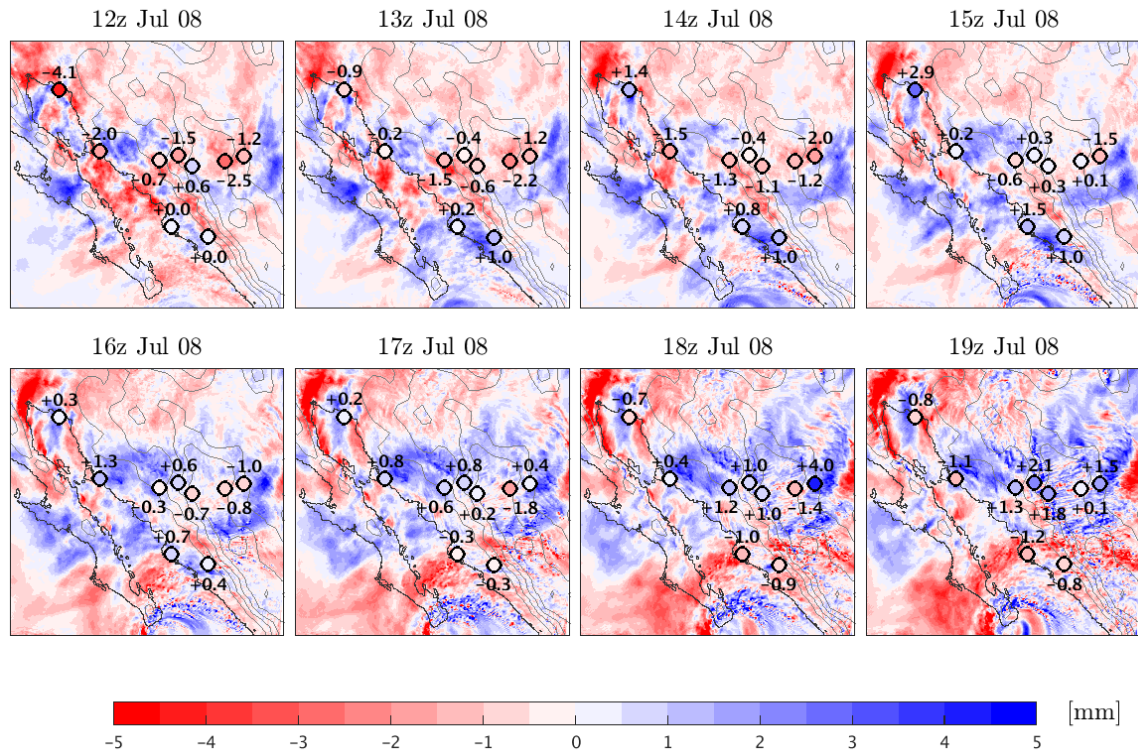
Figure 5: Hourly PWV biases (mm) (*assim – no_assim*) from 12-19z Jul 8 across the convection-allowing domain (d03) and the 9 GPS-PWV sites in Transect 2013. Blue indicates higher PWV in the *assim* experiment. Terrain contoured every 500 m in gray.
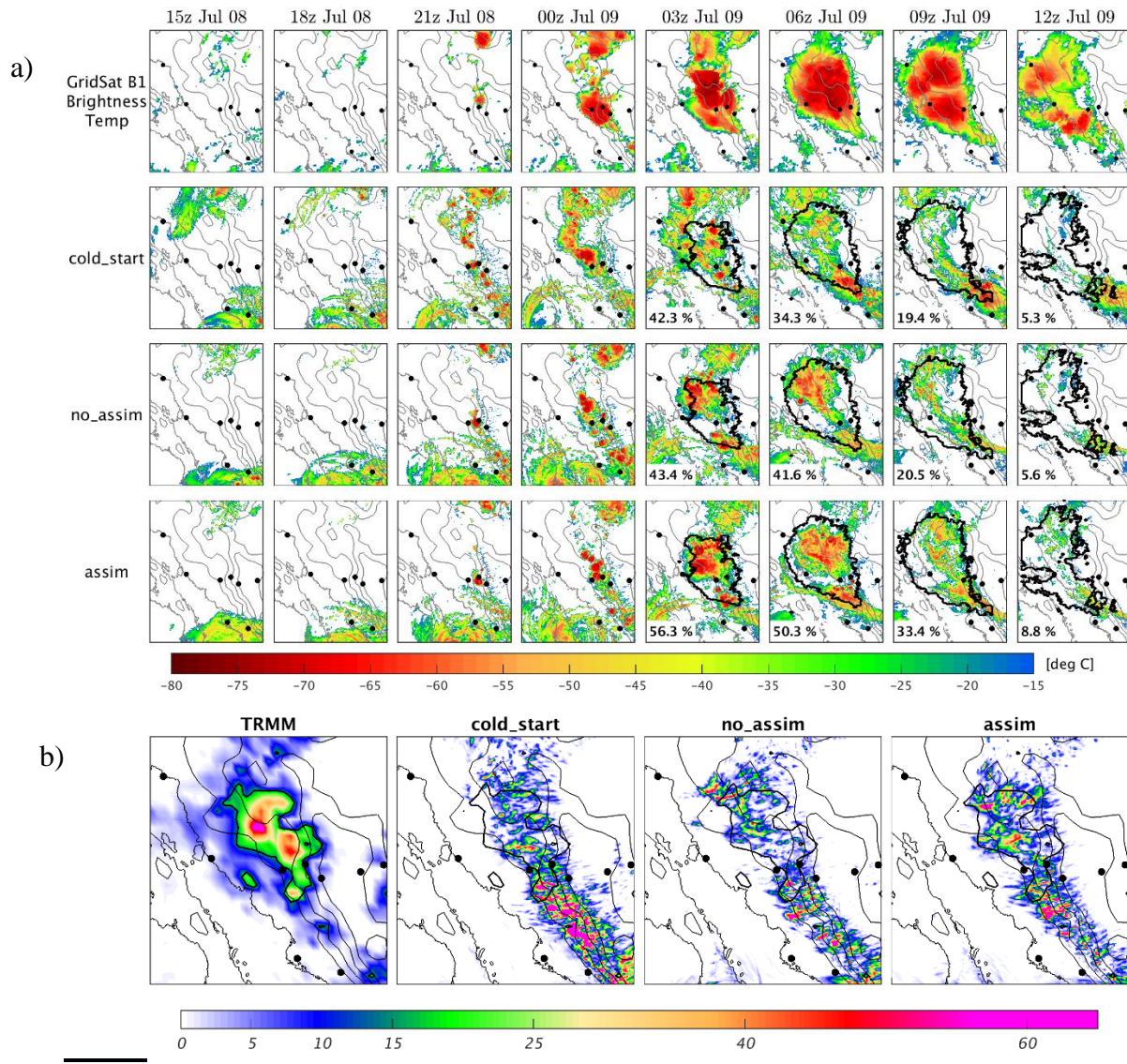
Figure 6: a) Infrared brightness temperature (°C) from GridSat B1 (row 1) *cold_start* (row 2), *no_assim* (row 3), and *assim* (row 4) every 3 hours from 15z Jul 8 to 12z Jul 9. The WRF OLR (outgoing longwave radiation) variable is converted to brightness temperature using the Stefan-Boltzmann Law. One of the criteria for an MCS is a cloud shield colder than -32 °C that has an area of at least 100,000 km^2 (Maddox 1980). When the observation meets this threshold, the -32 °C observed isotherm is outlined in the experiments. The percentage of the area that falls within the observed isotherm that is itself colder than -32 °C is displayed.

b) Beginning at 12z July 8, 24-h rainfall accumulation (mm) from TRMM TMPA (panel 1), *cold_start* (panel 2), *no_assim* (panel 3), and *assim* (panel 4). Terrain contoured every 500 m. Locations of the Transect 2013 GPS-PWV sites in black circles.
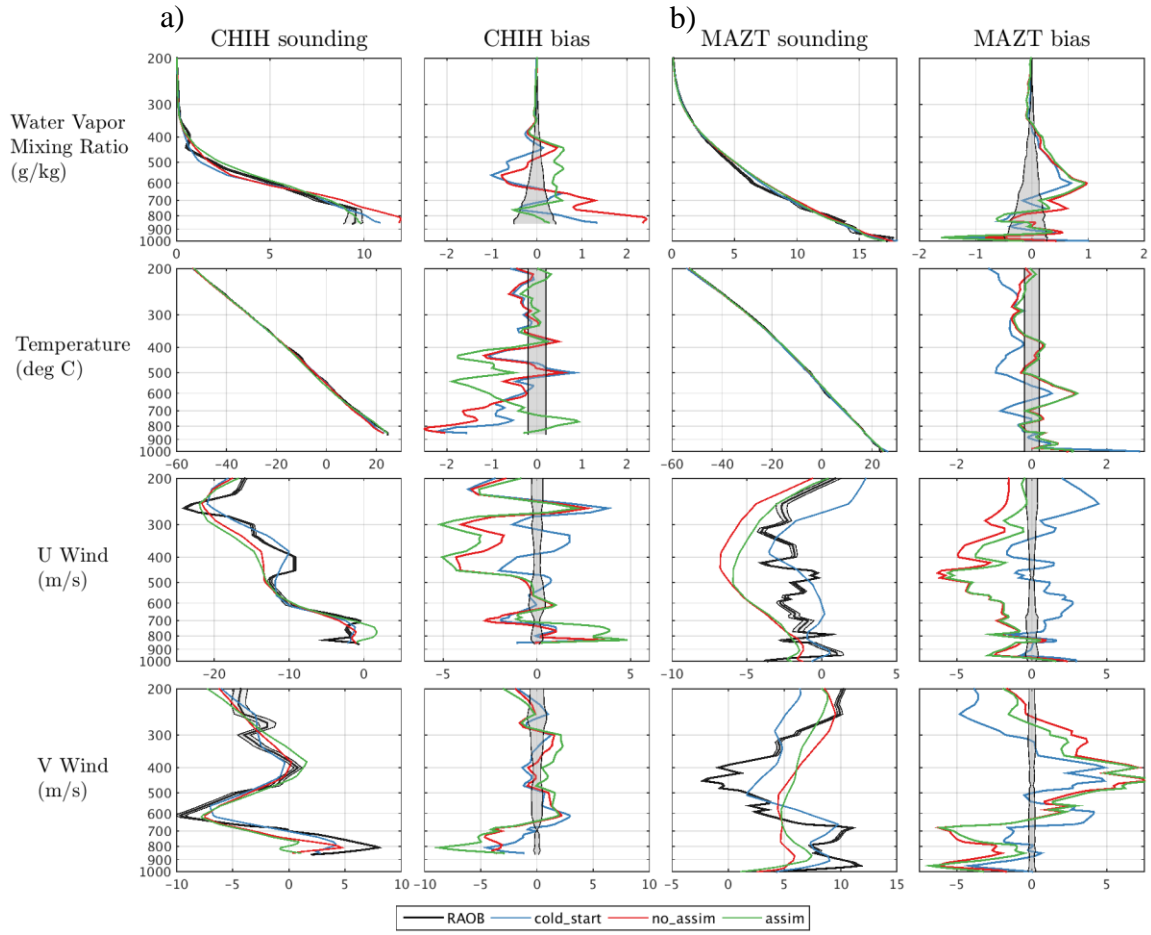
Figure 7: 12z Jul 8 sounding verification for a) CHIH and b) MAZT (sounding in column 1; bias in column 2) for water vapor mixing ratio (g/kg; row 1), temperature (°C row 2), U wind (m/s; row 3), and V wind (m/s; row 4). The error of the instruments on board the Vaisala RS92 radiosonde is shaded in gray.
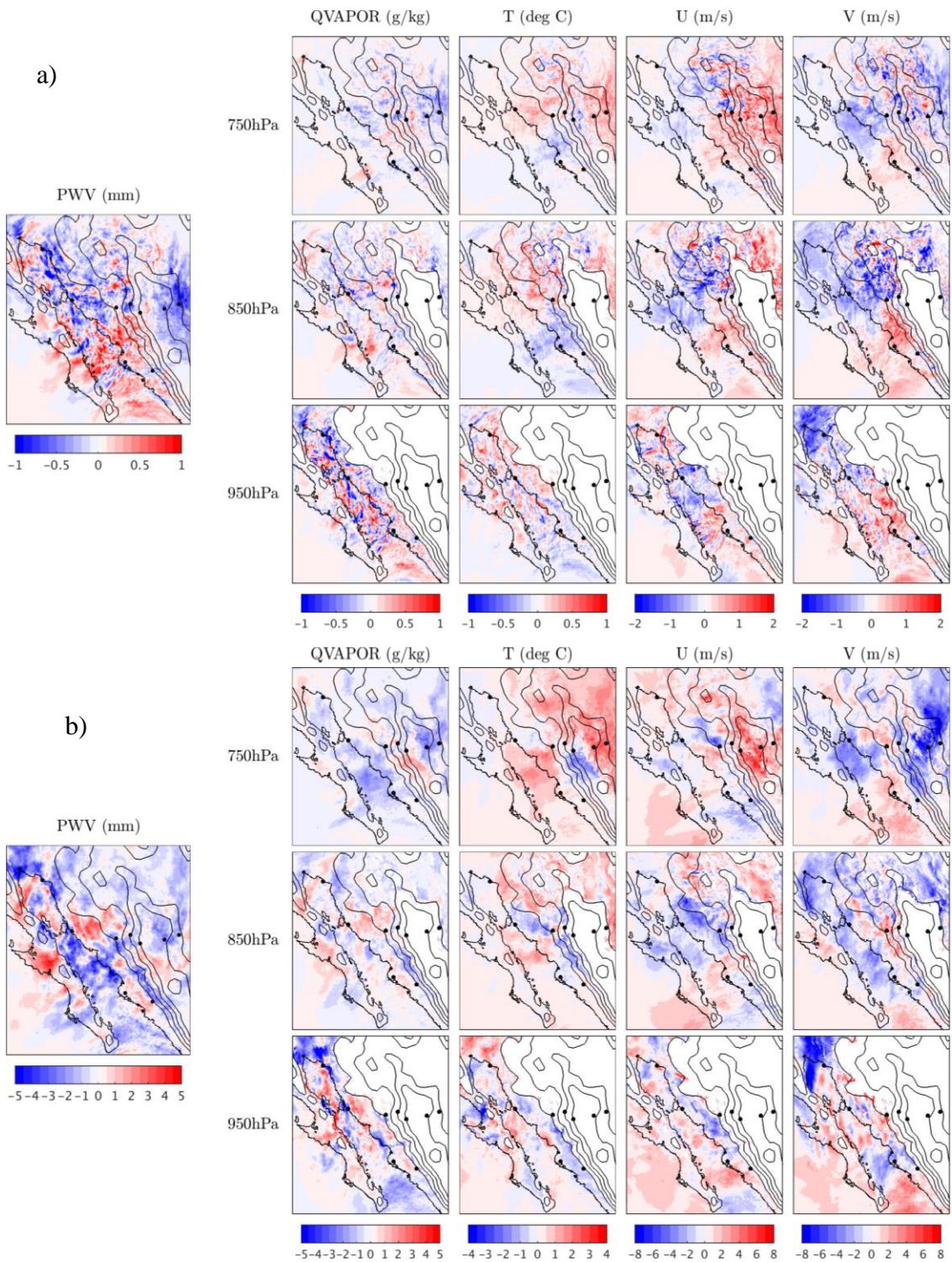
Figure 8: a) Mean increments (posterior – prior) across all hourly assimilation cycles (06-12z Jul 8) and b) 12z ensemble mean initial field bias (*assim – no_assim*) for PWV (mm) (separate left panel) and water vapor mixing ratio (g/kg) (column 1), T (°C) (column 2), U (m/s) (column 3), and V (m/s) (column 4) at the 750 hPa (row 1), 850 hPa (row 2), and 950 hPa (row 3) levels.
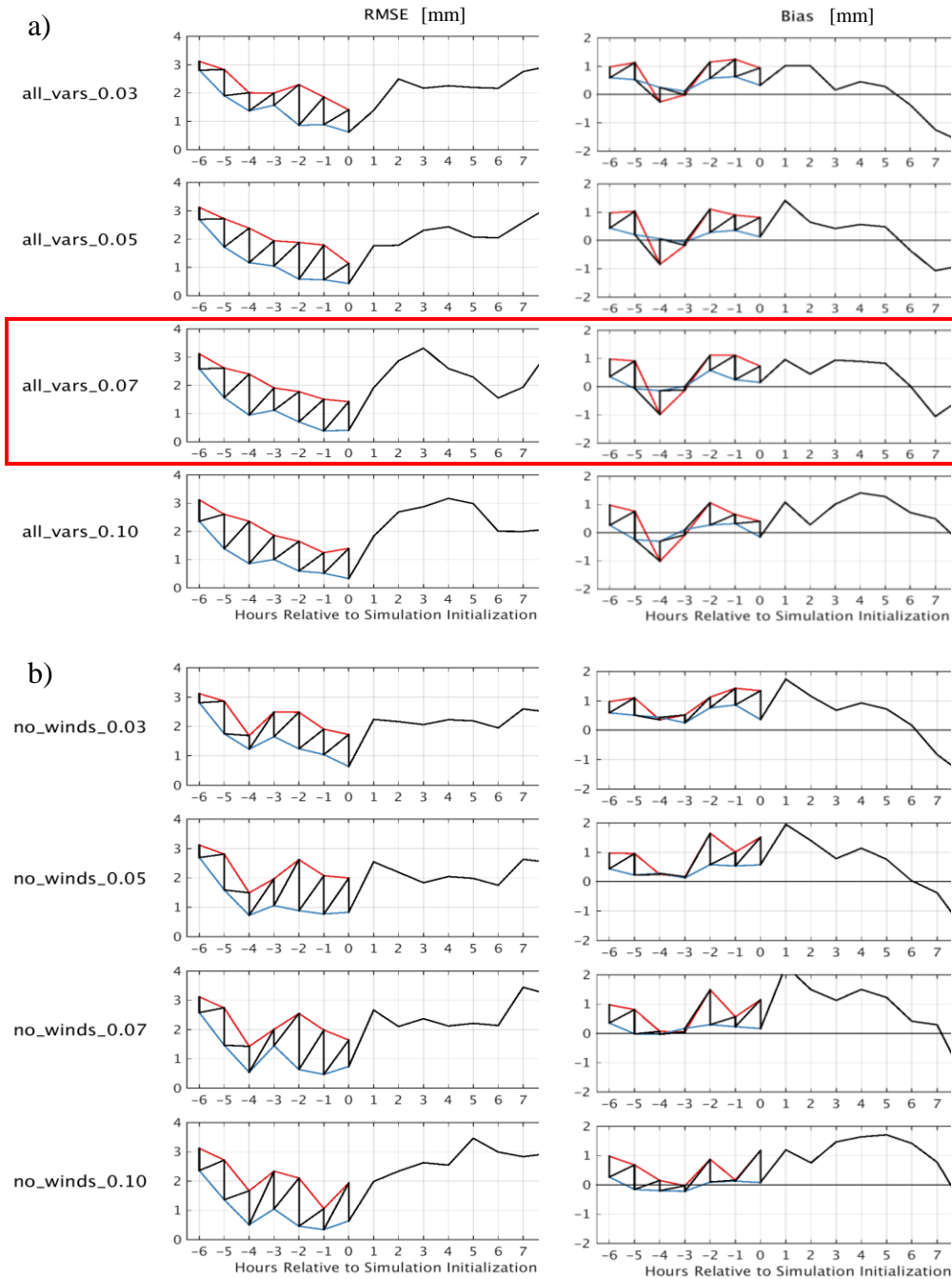
Figure 9: PWV (mm) RMSE (left column) and mean bias (model – observation; right column) across the 9 GPS stations for the prior (red line) and posterior (blue line) for the hourly DA cycles (6-12z Jul 8) and for the first 7 hours of the forecast (black line) (12-19z Jul 8) in a) *all_vars* and b) *no_winds* experiments. Covariance function cutoffs of 0.03 (row 1), 0.05 (row 2), 0.07 (row 3), and 0.10 radians (row 4) are shown. *Assim* experiment is outlined in a red box.
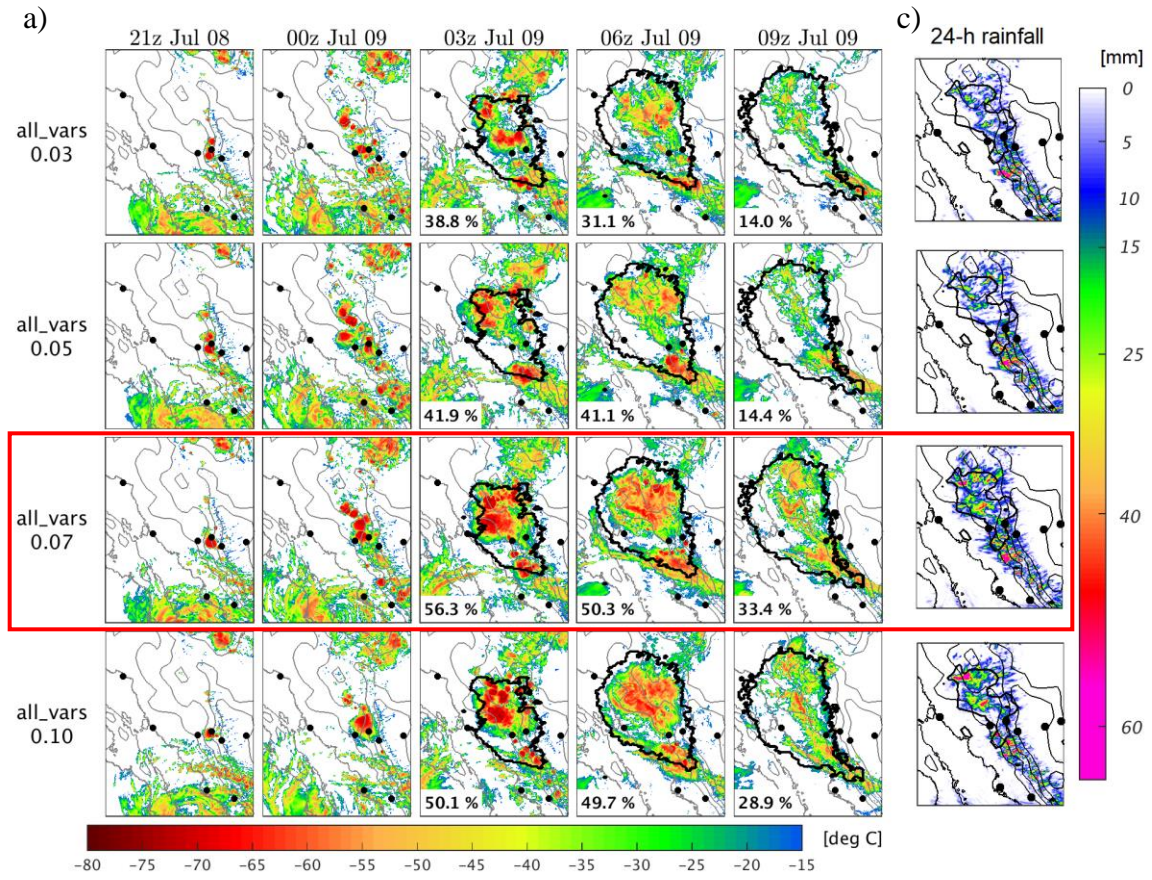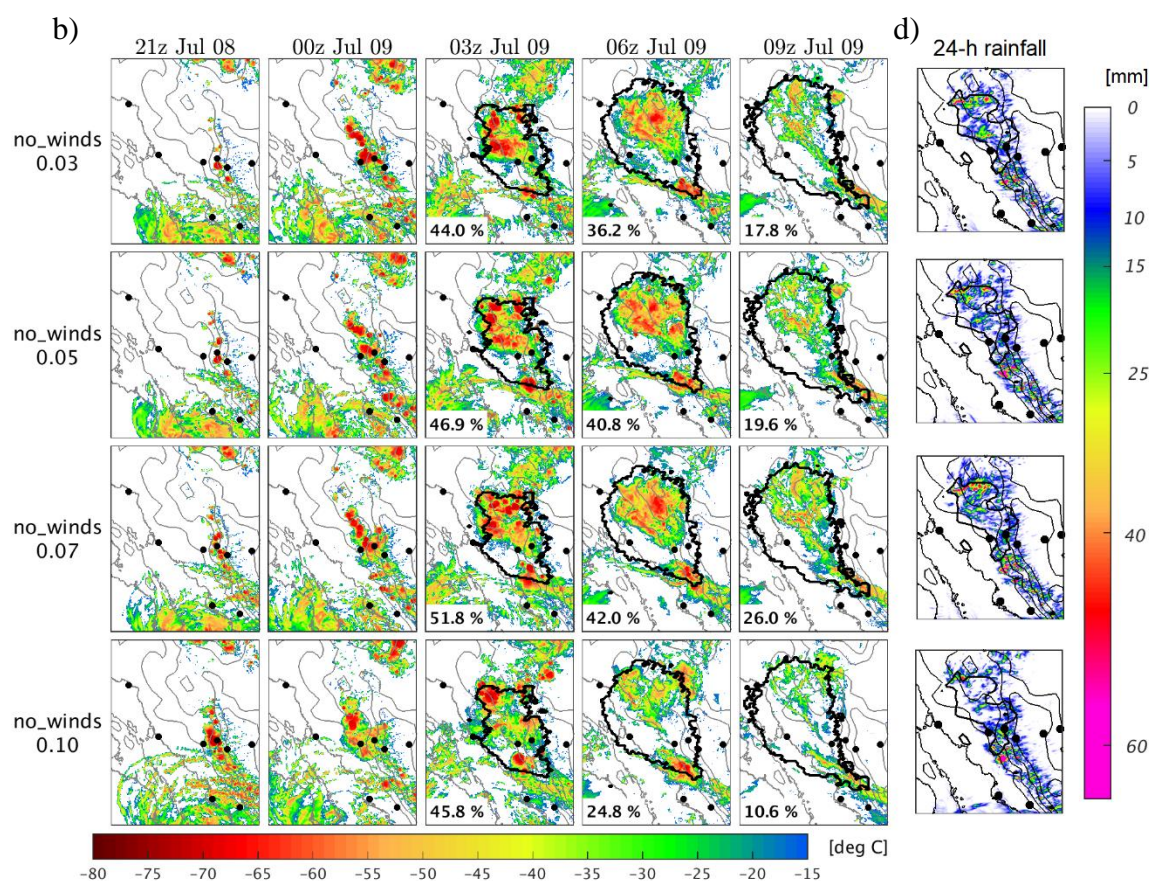
Figure 10: Infrared brightness temperature (°C) every three hours from 21z Jul 8 to 9z Jul 9 for covariance function cutoffs of 0.03 (row 1), 0.05 (row 2), 0.07 (row 3), and 0.10 radians (row 4) for a) *all_vars* and b) *no_winds* experiments. More information listed in the Fig. 6a caption and om Table 5. Refer to Fig. 6 caption for MCS verification technique.
Beginning at 12Z Jul 8, 24-hour rainfall accumulation (mm) for c) *all_vars* and d) *no_winds* experiments. *Assim* is outlined in a red box.
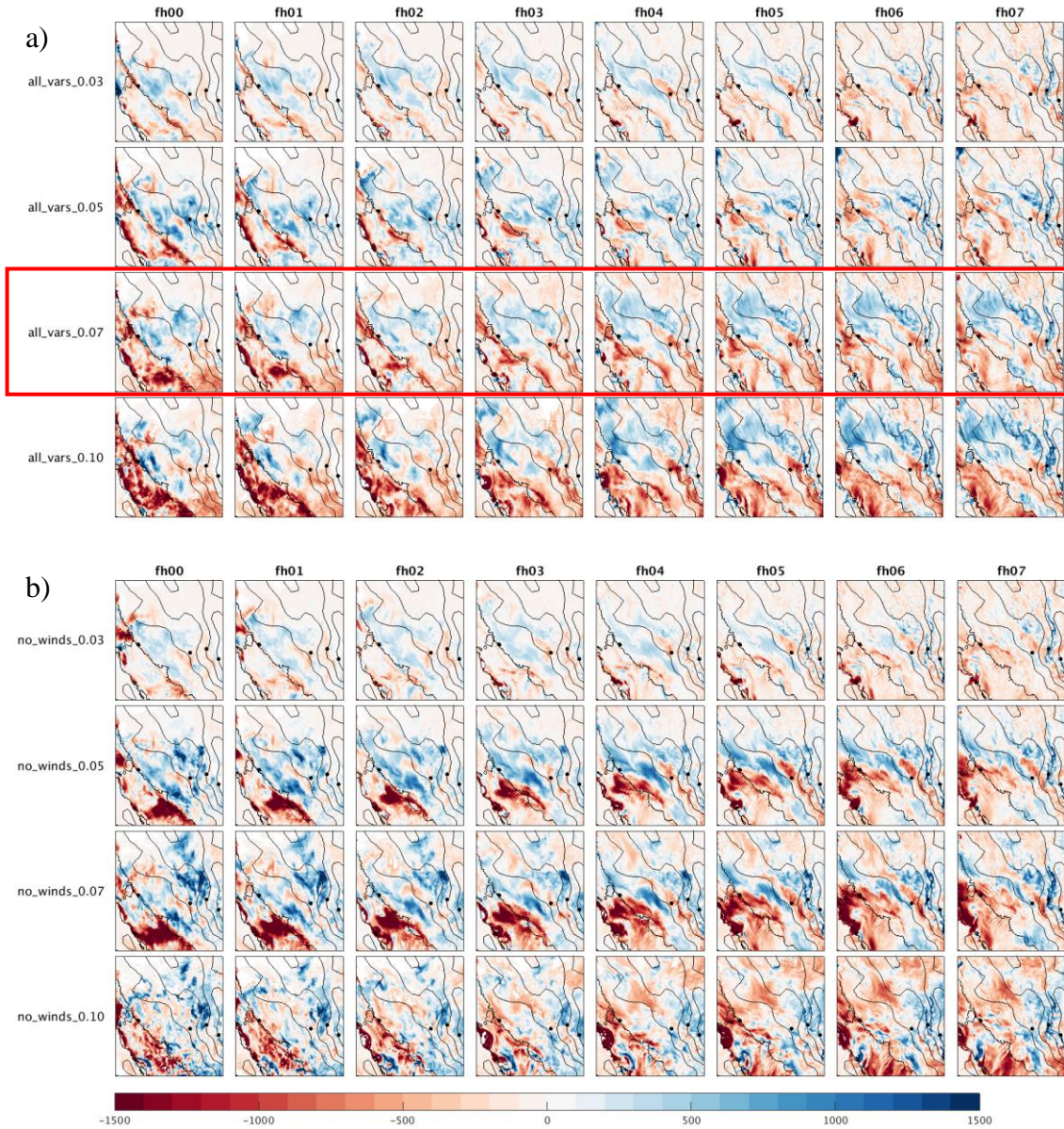
Figure 11: Using the parcel with highest theta-e, CAPE (J/kg) bias (model – observation; right column) for the first 7 hours of the forecast (12-19z Jul 8) in a) *all_vars* and b) *no_winds* experiments. Covariance function cutoffs of 0.03 (row 1), 0.05 (row 2), 0.07 (row 3), and 0.10 radians (row 4) are shown. *Assim* is outlined in a red box.
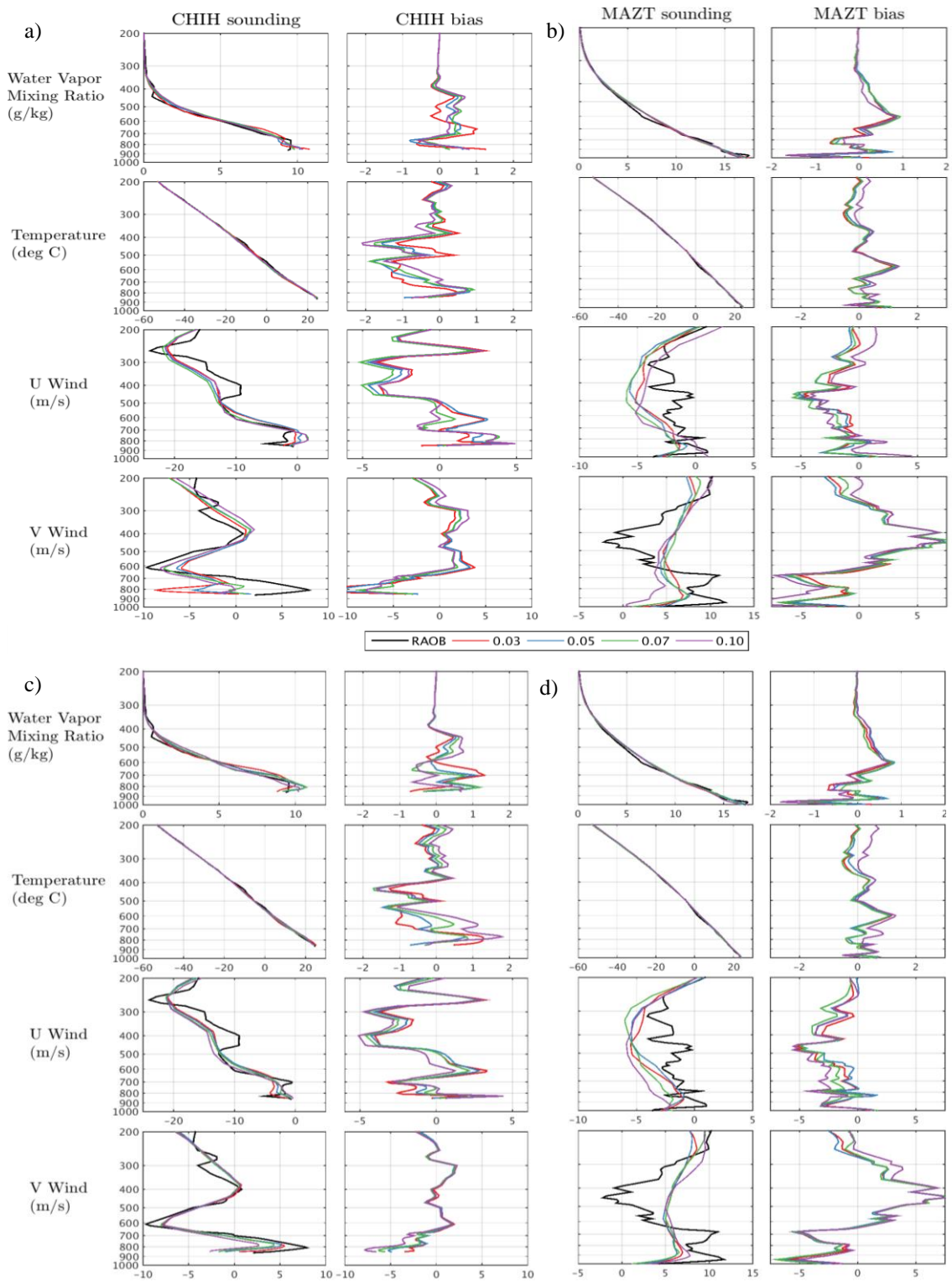
Figure 12: 12Z Jul 8 sounding verification for a) CHIH *all_vars*, b) MAZT *all_vars*, c) CHIH *no_winds*, and d) MAZT *no_winds* (sounding in column 1; bias in column 2) for water vapor mixing ratio (g/kg; row 1), temperature (°C; row 2), U wind (m/s; row 3), and V wind (m/s; row 4) for 0.03 (red), 0.05 (blue), 0.07 (green), and 0.10 rad (magenta) covariance function cutoffs.
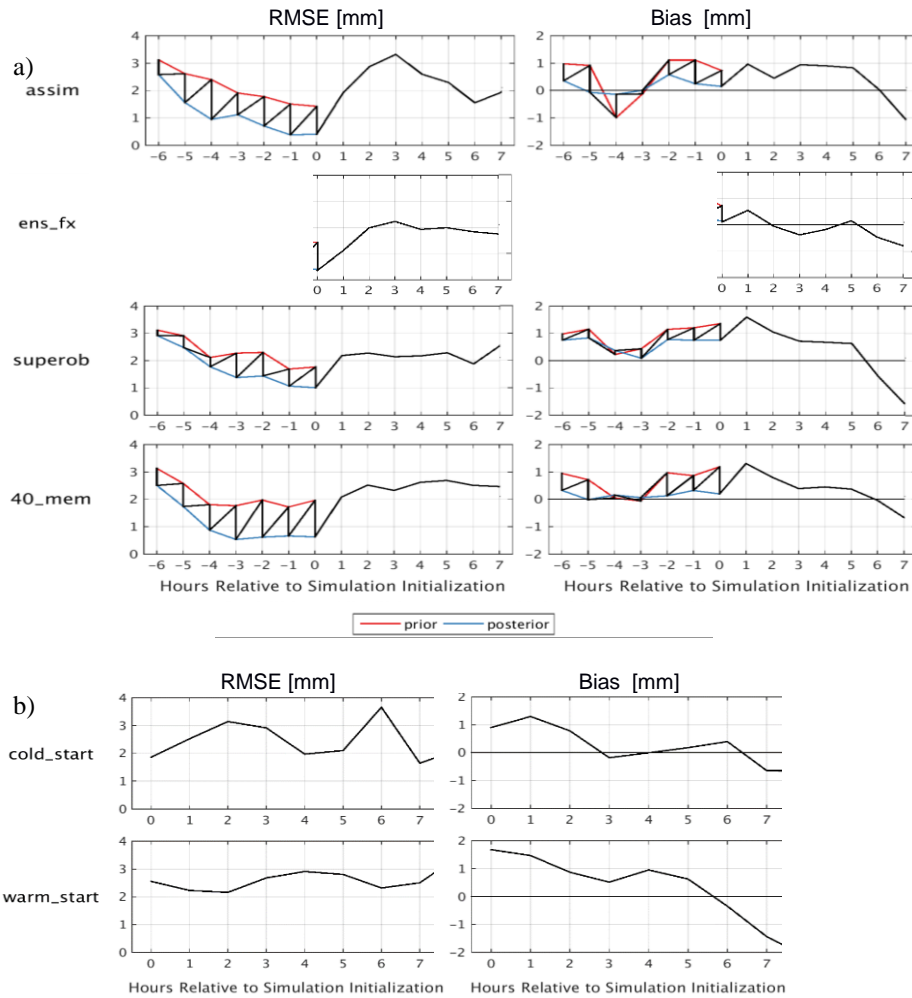
Figure 13: a) PWV RMSE (left column) and mean bias (model – observation; right column) across the 9 GPS stations for the prior (red) and posterior (blue) for the hourly DA cycles (6-12z Jul 8) and for the first 7 hours of the forecast (black) (12-19z Jul 8) in *assim* (row 1), *ens_fx* (row 2), *superob* (row 3), and *40_mem* (row 4).
b) PWV RMSE (left column) and mean bias (right column) across the 9 GPS stations for the first 7 hours of the forecast in *cold_start* (row 1) and *warm_start* (row 2).
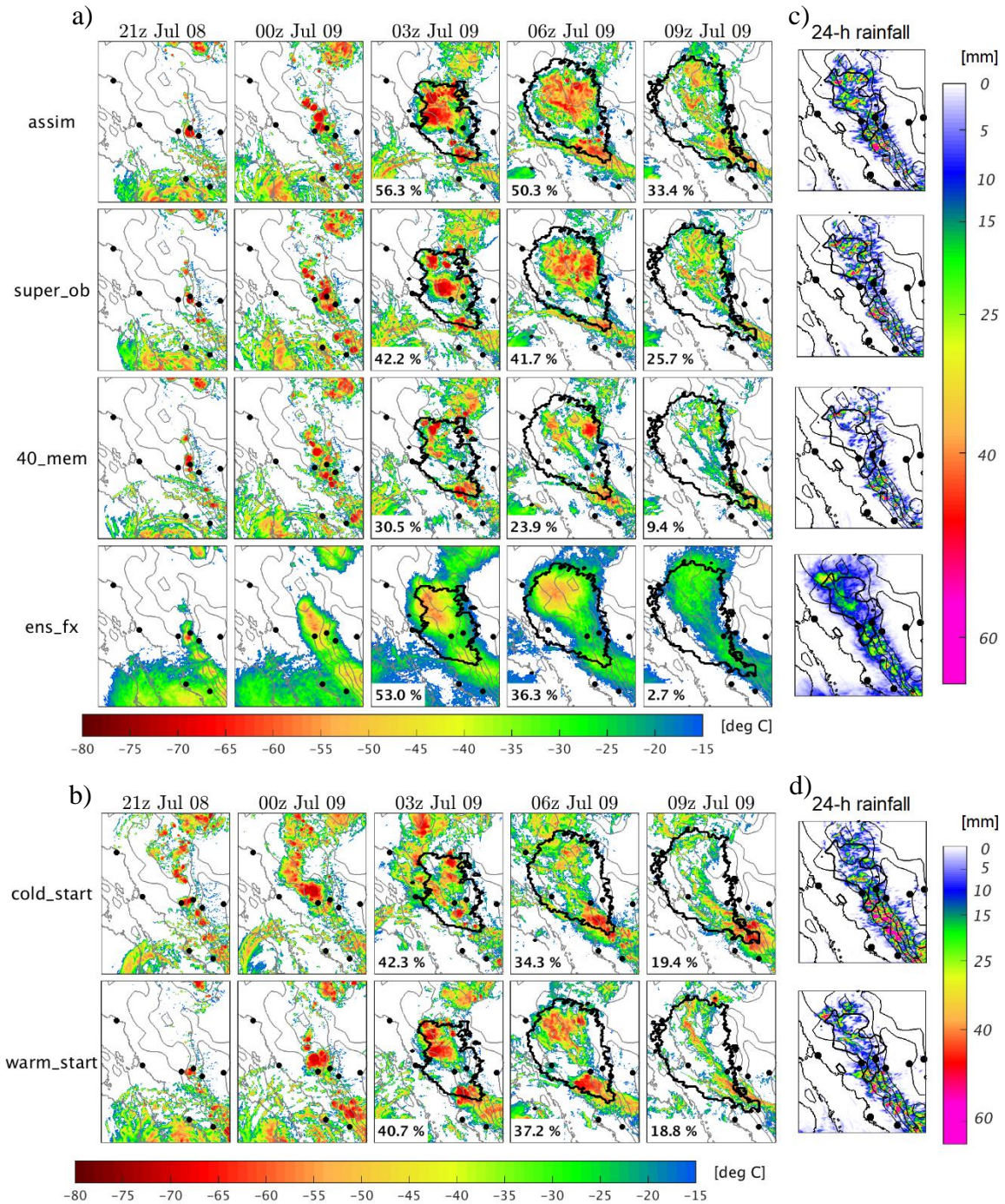
Figure 14: Infrared brightness temperature (°C) every three hours from 21z Jul 8 to 09z Jul 9 for a) *assim* (row 1), *superob* (row 2), *40_mem* (row 3), and *ens_fx* (row 4) and b) *cold_start* (row 1) and *warm_start* (row 2). More information listed in the Fig. 5a caption. Refer to Fig. 6 caption for MCS verification technique.
Beginning at 12z Jul 8, 24-hour rainfall accumulation (mm) for c) *assim*, *super_ob*, *40_mem*, and *ensemble_fx* and d) *cold_start* and *warm_start*.

181
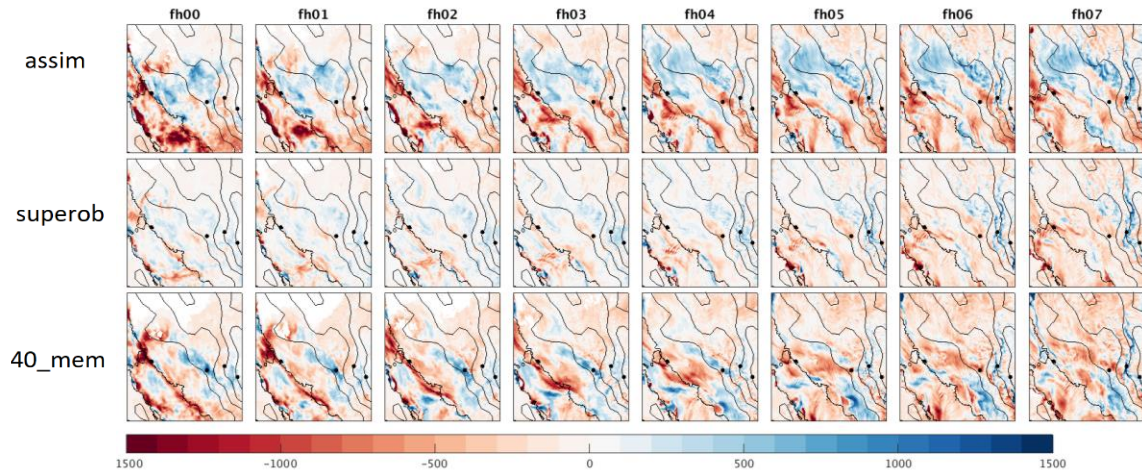
Figure 15: Same as Fig. 11, but for *assim* (row 1), *super_ob* (row 2), and *40_mem* (row 3).

APPENDIX C: DATA ASSIMILATION BACKGROUND WITH AN APPLICATION TO ENSEMBLE KALMAN FILTERING AND GPS-PWV

James M. Moker, Jr.

Department of Hydrology and Atmospheric Sciences, University of Arizona, Tucson, Arizona

1. Overview of Data Assimilation

The objective of this document is to describe how GPS precipitable water vapor (PWV) is assimilated into the WRF model along with background literature review of data assimilation. DA is a discipline that uses new information (e.g., observations) to update the estimate of the state of a system to reduce the error of the state. In the context of atmospheric science, DA attempts to create a consistent picture of the atmosphere in space and time using information and statistics from irregularly spaced observations (Kalnay 2002). On a model grid, DA adjusts unobserved atmospheric data with statistical connections with observed variables and the background structure of the state on that grid. A main problem with DA is the accuracy of estimates of the errors of a non-linear system with linear approximations. This section is organized by starting with DA being framed through a Bayesian lens in Subsection a. Then, DA methods can be divided into sequential and variational where they are discussed in Subsections b and c, respectively.

This document is organized as follows: a background of data assimilation (DA) in Section 1, a description of the Kalman Filter in Section 2, an explanation of the Ensemble Kalman Filter (EnKF) with an application to convection-resolving modeling in Section 3, and an exploration of past studies of GPS-PWV assimilation with a description of my WRF/DART setup in Section 4.

a. DA from a Bayesian perspective

The state of the atmosphere may be approximated by a gaussian probability density function (PDF). Conveniently, a multivariate gaussian PDF can represented by simply the mean and the covariance matrix. Bayes' Rule states that the state sample can

be updated ("posterior" in Bayesian parlance) from the pre-assimilation state ("prior") and the observations ("data likelihood") (Wikle and Berliner 2007; Lorenc 1986).

Let y resemble the observation distribution, x resemble the state distribution, and p(·) notation indicate probability. The probability of observations given all previous information is $p(y|x)$ and is called the "data likelihood" PDF. The probability of the state before the current observations is $p(x)$ and is called the "prior" PDF. The probability of the state given all observations including the current one is $p(x|y)$ and is called the "posterior" PDF. Bayes' Rule states that the "posterior" PDF is proportional to the product of the "prior" PDF and the "data likelihood" PDF (eq. 1).

$$p(x|y) \propto p(x) \, p(y|x) \tag{1}$$

The "posterior" PDF is gaussian since it is proportional to the product of two gaussians. All observations (current and previous) are considered part of the same distribution; therefore, each observation can be assimilated individually since they do not change the statistics of the "posterior" PDF.

b. Sequential methods

One may use sequential methods to assimilate observations into a model. The model fields are updated whenever observations are available (called an "analysis"). Between observation times, the model is advanced forward in time until the next observation(s) are available. At the analysis time, the goal is to find an optimal weight ("least squares") that includes statistical information about the observation and background state errors to estimate to the state that has minimum variance and is unbiased. This can be done in one dimension (a scalar) and is demonstrated in this section. After expanding this framework spatially to a grid, we can evolve it in time in the

Kalman Filter (Kalman 1960) and extended Kalman Filter (EKF). These methods are discussed in Section 2. For large models, we must represent the background state with an ensemble in the ensemble Kalman Filter (EnKF; Evensen 1994) discussed in Section 3.

Consider PWV where its unknowable true value is indicated by PWV_t. Given a PWV measurement, PWV_o, and a background PWV measurement, PWV_b, we can find the best estimate, PWV_a, based on a linear relationship between the measurements and statistical assumptions about the measurements (adapted from Kalnay 2002). Eq. 2 describes this scenario where A and B are linear coefficients for PWV_o and PWV_b, respectively.

$$PWV_a = A \cdot PWV_o + B \cdot PWV_b \qquad (2)$$

The sum of the coefficients is unity as a result of the unbiased assumption of PWV_a (eq. 3).

$$A + B = 1 \qquad (3)$$

We define the errors as the deviation of the measurement or analysis from the true value PWV_t in eq. 4-6:

$$\varepsilon_a \equiv PWV_a - PWV_t \qquad (4)$$

$$\varepsilon_o \equiv PWV_o - PWV_t \qquad (5)$$

$$\varepsilon_b \equiv PWV_b - PWV_t \qquad (6)$$

The expectation operator, $E[\cdot]$, is equivalent to taking the arithmetic mean in the equations that follow. The measurements are assumed to be unbiased and therefore the analysis is unbiased. In other words, the expected values of their errors are zero (eq. 7).

$$E[\varepsilon_o] = E[\varepsilon_b] = E[\varepsilon_a] = 0 \qquad (7)$$

Additionally, we assume that the errors between the two PWV measurements are independent and therefore uncorrelated (eq. 8).

$$E[\varepsilon_o \varepsilon_b] = 0 \tag{8}$$

Error variances are defined as the expected value of the squared errors (eq. 9).

$$\sigma_o^2 \equiv E[\varepsilon_o^2], \qquad \sigma_b^2 \equiv E[\varepsilon_b^2], \qquad \sigma_a^2 \equiv E[\varepsilon_a^2] \tag{9}$$

Substitute B in terms of A from eq. 2 into eq. 1 and solve for PWV_a to obtain the analysis equation (eq. 9).

$$PWV_a = PWV_b + A(PWV_o - PWV_b) \tag{10}$$

Substitute PWV_o and PWV_b in terms of their errors and PWV_t from eq. 4-6 and solve for analysis error ε_a to obtain the error analysis equation (eq. 11).

$$\varepsilon_a = \varepsilon_b + A(\varepsilon_o - \varepsilon_b) \tag{11}$$

Using the assumption that the measurement errors are uncorrelated (eq. 7) and the definition of error variances (eq. 8), the analysis error variance equation (eq. 12) is obtained in terms of A and the error variances by squaring and then taking the expected value of eq. 11.

$$\sigma_a^2 = (1 - A)^2 \sigma_b^2 + A^2 \sigma_o^2 \tag{12}$$

The optimal weight A (eq. 13) for the analysis (eq. 10) is obtained by minimizing the analysis error variance σ_a^2 (eq. 12). This is done by taking the derivative of eq. 12 with respect to A and setting it to zero. Note that according to the 'Minimum Variance Method', minimizing the variance is equivalent to minimizing the mean square error since ε_a is unbiased.

$$A = \frac{\sigma_b^2}{\sigma_b^2 + \sigma_o^2} \tag{13}$$

The optimal weight A is comprised of the ratio of variance of background measurement to the sum of variances of the background and observation and has a range of 0 to 1. If the observation error variance is zero ($\sigma_o^2 = 0$), then the optimal weight is unity ($A = 1$) and the analysis equals the observation ($PWV_a = PWV_o$). Similarly, if the background error variance is zero ($\sigma_b^2 = 0$), then the optimal weight is zero ($A = 0$) and the analysis equals the background ($PWV_a = PWV_b$). If the background and observation variances are equal (but not zero), then $A = 0.5$ and the analysis is simply the arithmetic mean of the background and observation ($PWV_a = 0.5(PWV_b + PWV_o)$).

The least squares estimate of the analysis error variance is determined by substituting the optimal weight A into eq. 12 and solving for σ_a^2 (eq. 14).

$$\sigma_a^2 = \frac{\sigma_b^2 \sigma_o^2}{(\sigma_b^2 + \sigma_o^2)} = A\sigma_o^2 = (1 - A)\sigma_b^2 = \left(\frac{1}{\sigma_b^2} + \frac{1}{\sigma_o^2}\right)^{-1} \tag{14}$$

The minimized analysis error variance is also equal to the inverse of the sum of the background error and observation error precisions (inverses of variances). Substituting A into eq. 10, we obtain the analysis in terms of the observation, observation variance, background, and background variance (eq. 15). This linear combination of the background and observation measurements and their accuracies results in the best linear unbiased estimate that has minimum variance.

$$PWV_a = PWV_b + \left(\frac{\sigma_b^2}{\sigma_b^2 + \sigma_o^2}\right)(PWV_o - PWV_b) \tag{15}$$

After all observations are processes to get the new analysis (eq. 15), then it and its errors (eq. 14) are advanced in time until observations are next available. This can be done using a Kalman Filter (described in Section 2).

c. Variational methods

An alternative to the sequential approach is the variational or continuous approach. This involves minimizing a cost function, $J(x)$, that is comprised of the squares of the differences between the analysis and each measurement (eq. 16). We will use a scalar case once again.

$$J(x) = \frac{1}{2}\left\{\frac{(PWV_b - PWV)^2}{\sigma_b^2} + \frac{(PWV_o - PWV)^2}{\sigma_o^2}\right\} \tag{16}$$

To minimize $J(x)$, we take its derivative with respect to PWV, set it to zero, and solve for PWV (eq. 17). This will result in PWV that is equal to PWV_a in the analysis equation obtained via the optimal weight least-squares approach (eq. 15).

$$PWV = \left(\frac{\sigma_o^2}{\sigma_b^2 + \sigma_o^2}\right)PWV_b + \left(\frac{\sigma_b^2}{\sigma_b^2 + \sigma_o^2}\right)PWV_o \tag{17}$$

A prevalent variational method is 3D-VAR (Lorenc 1986), where 3D-VAR stands for "three-dimensional variational" and the three dimensions are that of space, and allows for all observations at a particular time to be assimilated at once. In eq. 16, the scalars are replaced with vectors and the cost function $J(x)$ becomes a three-dimensional bowl. The gradient of $J(x)$ must be minimized instead of minimizing the derivative as in the scalar case. Finding the gradient of $J(x)$ is not a trivial task and must be done incrementally.

Other variational methods include but are not limited to 4D-VAR (Le Dimet and Talagrand 1986; Lewis and Derber 1985) and the Physical-space Statistical Analysis System (PSAS; Cohn et al. 1998). 4D-VAR is an extension of 3D-VAR in that accounts for the time dimension and includes all observations in space and time in the cost function. Because it takes into consideration time, it requires the use of a model requiring more computational resources and code for a tangent linear model (TLM) and an adjoint

model (AM). The PSAS solves 3D-VAR or 4D-VAR in observation space instead of model space using OI equations. Additionally, there is an incremental version of 4D-VAR (Courtier et al. 1994) that solves for the analysis increment instead of the analysis itself resulting in a computationally cheaper computation.

2. Kalman Filter

A Kalman Filter (KF) is a sequential method of data assimilation that evolves the state in time via a forward model (Kalman 1960). An analysis is computed that includes an optimal weight (or gain in multiple dimensions) whenever observations become available that follow the logic of eq.15. In NWP, the state is not just one location; it is vector that lies on a model grid. The conversion from scalar to multivariate is described in subsection a, the KF equations are described in subsection b, and the extended KF, which allows for slightly non-linear systems, is described in subsection c.

a. Extension of variables from scalar to multivariate

We are using the KF to propagate a state of the atmosphere and that is represented on a grid along with its error characteristics. We are also using multiple observations. Therefore, since are dealing with multi-dimensionality, we must extend our variables from scalars to vectors as shown in Table 1. The errors of the variables are also vectors. For brevity, we use the same symbol as the scalar case, $\boldsymbol{\varepsilon}$, to represent the errors, but they will become column vectors. A covariance matrix is generated by multiplying an error vector by its transpose and then taking the expected value. In the KF, $\boldsymbol{P}^f$ ($m \times m$) is used to indicate the forecast (background) covariance matrix (eq. 18).

$$\boldsymbol{P}^f = E[\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^T] \tag{18}$$

In practice, P^f is estimated from a first guess of a model background. It represents the relationship of the model grid points to one another and has information depicting how the grid will change based on the assimilation of an observation.

R $(n \times n)$ is the observation error covariance matrix (eq. 19). If the same instrument is used, then observation errors are uncorrelated and usually identical and R is defined by the identity matrix multiplied by the observation error variance. Therefore, only the variances along the diagonal exist.

$$R = I_n \sigma_o^2 \tag{19}$$

Because observations are not usually located exactly on the model grid, an observation operator, H, is required. It maps the observation to model space and therefore has dimensions $n \times m$. The observation operator performs interpolation in space from the model grid to the observation location as well as computation for variables that are not explicitly represented by the state vector (such as PWV that is discussed in Section 4).

b. Kalman Filter equations

The KF has two phases. The "update" phase is the analysis performed (usually on a grid) using a least-squares optimal matrix (gain) and results in a new state and covariance matrix. The "predict" phase then propagates the new state and covariance matrix in time using a model.

To get the KF analysis equation (eq. 20) in the update phase, we substitute the vector definitions from Table 1 into the same format as the least-squares scalar analysis equation (eq. 10). We also generalize the PWV to state vector x.

$$x^a = x^f + K(y - Hx^f) \tag{20}$$

The gain matrix $\boldsymbol{K}$ replaces scalar weight A and is multiplied by the innovation vector

$\boldsymbol{y} - \boldsymbol{H}\boldsymbol{x}^f$ to get the analysis increment vector $\boldsymbol{x}^a - \boldsymbol{x}^f$. Since the innovation vector is in

observation space, the observation operator $\boldsymbol{H}$ must be multiplied by the forecast state

vector $\boldsymbol{x}^f$ to get the model-equivalent observation so it can be subtracted from the

observation vector $\boldsymbol{y}$. The gain matrix $\boldsymbol{K}$ is defined as the ratio between the error in

model-observation space over the total error in observation space (eq. 21) and is similar

to the scalar least-squares optimal weight (eq. 13).

$$\boldsymbol{K} = \boldsymbol{P}^f \boldsymbol{H}^T \left(\boldsymbol{H}\boldsymbol{P}^f \boldsymbol{H}^T + \boldsymbol{R} \right)^{-1} \tag{21}$$

Finally, the analysis error covariance matrix, $\boldsymbol{P}^a$, is updated (eq. 21) similar to the

scalar analysis error variance equation (eq. 13). The analysis error covariance matrix $\boldsymbol{P}^a$

is determined by reducing the background error covariance matrix $\boldsymbol{P}^b$ by a factor of

$(\boldsymbol{I} - \boldsymbol{K}\boldsymbol{H})$.

$$\boldsymbol{P}^a = (\boldsymbol{I} - \boldsymbol{K}\boldsymbol{H})\boldsymbol{P}^f \tag{226}$$

In eq. 20-22, all variables are valid at time t so their time indices are removed for brevity.

In the predict phase of the KF (eqs. 22-24), the model state and error covariance matrices

are advanced in time from t to $t + 1$. The forecast state vector $\boldsymbol{x}^f$ of dimension $m \times 1$ is

advanced in time by a linear model matrix $\boldsymbol{M}$ of dimension $m \times m$ (eq. 22).

$$\boldsymbol{x}^f_{t+1} = \boldsymbol{M}_{t \to t+1}(\boldsymbol{x}^f_t) \tag{23}$$

The model error covariance matrix $\boldsymbol{Q}$ is the model error that accumulated with

time from t to $t + 1$. It is constructed by the expected value of the model error vector $\boldsymbol{\eta}$

multiplied by its transpose (eq. 24). This error is difficult to estimate, but the KF will

become unstable if $\boldsymbol{Q}$ is too large.

$$Q = E[\boldsymbol{\eta\eta}^T] \tag{24}$$

The predicted error covariance matrix $\boldsymbol{P}^f_{t+1}$ is propagated in time (eq. 25) by the linear

model operator matrix $\boldsymbol{M}$ and its transpose and then added to model error covariance

matrix Q.

$$\boldsymbol{P}^f_{t+1} = \boldsymbol{M}\boldsymbol{P}^a\boldsymbol{M}^T + \boldsymbol{Q} \tag{25}$$

This predict phase continues until new observations are available to assimilate. At that

time, we return to the update phase with eqs. 20-23 and set $\boldsymbol{P}^f_t$ to $\boldsymbol{P}^f_{t+1}$ and $\boldsymbol{x}^f_t$ to $\boldsymbol{x}^f_{t+1}$.

c. Extended Kalman Filter

In the KF, linearity is assumed in the observation operator $\boldsymbol{H}$ and forward model

operator $\boldsymbol{M}$. However, sometimes these operators are non-linear and are represented as $\boldsymbol{\mathcal{H}}$

and $\boldsymbol{\mathcal{M}}$, respectively. In the Extended Kalman Filter (EKF), the non-linear versions

replace the linear versions of these matrices. Eq. 20 is replaced by eq. 26 and eq. 23 is

replaced by eq. 27.

$$\boldsymbol{x}^a = \boldsymbol{x}^f + \boldsymbol{K}(\boldsymbol{y} - \boldsymbol{\mathcal{H}}\boldsymbol{x}^f) \tag{26}$$

$$\boldsymbol{x}^f_{t+1} = \boldsymbol{\mathcal{M}}_{t\rightarrow t+1}(\boldsymbol{x}^f_t) \tag{27}$$

The other KF equations remain the same. The $\boldsymbol{\mathcal{H}}$ and $\boldsymbol{\mathcal{M}}$ matrices can be

linearized by taking the Jacobian at the current time to allow for the computations of

equations 21, 22, and 25. $\boldsymbol{\mathcal{H}}$ is linearized at time t (eq. 28) and $\boldsymbol{\mathcal{M}}$ is linearized at $t \rightarrow$

$t + 1$ (eq. 29).

$$\boldsymbol{H} = \left.\frac{\partial\boldsymbol{\mathcal{H}}}{\partial\boldsymbol{x}}\right|_t \tag{28}$$

$$\boldsymbol{M} = \left.\frac{\partial\boldsymbol{\mathcal{M}}}{\partial\boldsymbol{x}}\right|_{t\rightarrow t+1} \tag{29}$$

The non-trivial tangent linear model (TLM) $\boldsymbol{M}$ and its adjoint (AM) $\boldsymbol{M}^T$ must be coded to be run alongside the non-linear model $\boldsymbol{\mathcal{M}}$ in the EKF.

The EKF can be use on only slightly non-linear systems because the linearization of highly non-linear systems will cause high instability in the filter. In NWP, both eq. 20 and 24 would not be computationally feasible with current technology because of the high dimensionality of $\boldsymbol{P} \sim 10^{14}$. There would also be difficulty computing the matrix inversion of $(\boldsymbol{H}\boldsymbol{P}^f\boldsymbol{H}^T + \boldsymbol{R})^{-1}$ in the Kalman gain $\boldsymbol{K}$ (eq. 21) and the propagation in time of the error covariance matrix with the TLM and AM (eq. 23). A remedy for this is to approximate the state using an ensemble and not explicitly store $\boldsymbol{P}$. This is called an Ensemble Kalman Filter (EnKF) and is described in the next section.

3. Ensemble Kalman Filter (EnKF) and NWP convective applications

In convective-allowing NWP, the EnKF has been shown to lower error in initialization resulting in promising results in an operational setting (e.g. Schwartz et al. 2015). The EnKF uses an ensemble to represent the state vector PDF so the KF equations can be computed. A common way to implement the EnKF is via the Data Assimilation Research Testbed (DART; Anderson et al. 2009), which is a community facility that allows for the application of EnKF methodologies, especially the Ensemble Adjustment Kalman Filter (EAKF; Anderson 2001). The general formulation of EnKF is discussed in subsection a, DART is described in subsection b, and past studies involving EnKF/DART in convective-allowing models are discussed in subsection c.

a. Description of EnKF

The EnKF uses Monte Carlo methods to sample the PDF of the atmosphere to approximate error covariances via ensembles. A gaussian state PDF is assumed. The error

covariance matrices are not explicitly constructed so the analysis error covariance matrix is never stored. The EnKF advances the state with a nonlinear model. No AM or TLM code is needed. In the context of NWP, unlike the KF or EKF, the implementation and computation of the EnKF is feasible because of the reduced dimensionality and not needing to propagate the covariance matrix of the state explicitly. The update step is computationally cheap relative to 4D-VAR. There is a flow-dependence to the error covariances as they propagate in time that resemble the dynamic structure of the atmosphere.

The EnKF is a square root filter (SRF) in that the matrix square root of the error covariance matrix is propagated instead of the entire matrix. It was first proposed by Evensen (1994) for use in geostatistical applications, but it was eventually realized that the observations needed to be treated like random variables. Houtekamer and Mitchell (1998) developed an EnKF that addressed the treatment of observations this way. They defined the state as $\overline{x^f}$ where the overbar indicates the ensemble mean (eq. 30).

$$\overline{x^f} = \frac{1}{N_{ens}} \sum_{i=1}^{N_{ens}} x_i^f \qquad (30)$$

The forecast error covariance matrix P^f is approximated by a sample mean of the sum of the differences between each ensemble member and the ensemble mean multiplied by its transpose (eq. 31).

$$P^f = x^f(x^f)^T = \frac{1}{N_{ens}-1} \sum_{i=1}^{N_{ens}} \left(x_i^f - \overline{x^f}\right)\left(x_i^f - \overline{x^f}\right)^T \qquad (31)$$

There are two general types of EnKF, "stochastic" and "deterministic", whose differences deal with the method in which the analysis error covariance matrices are

represented. Both types are considered SRFs (Tippett et al. 2003). The stochastic algorithm perturbs the observations to represent the observation error (Houtekamer and Mitchell 1998; Burgers et al. 1998), but this method introduces an additional source of sampling error that reduces the analysis error covariance accuracy and increases the probability of underestimating the analysis error covariance matrix (Whitaker and Hamill 2002). The deterministic algorithm, in contrast, adjusts the analysis ensemble itself instead of perturbing the observations (Whitaker and Hamill 2002). There are several types of deterministic EnKFs. Examples are the Ensemble Transform Kalman Filter (ETKF; Bishop et al. 2001), the Local Ensemble Transform Kalman Filter (LETKF; Hunt et al. 2007), and the EAKF (Anderson 2001).

Because of the sampling error from the state being represented by an ensemble, there are two main problems that arise in EnKFs that may render the filter not responsive to observations (filter divergence). First, since the covariance matrices are approximated and therefore not full rank, spurious correlations will occur as one moves away from the observation site. To mitigate this, error covariances must be localized in some fashion (Anderson and Anderson 1999). Second, the model error could be higher than the observation error. To mitigate this, the covariance is inflated. Methods for addressing these modifications are discussed in more detail in the next subsection.

b. Data Assimilation Research Testbed

DART (Anderson et al. 2009) is a community DA facility that allows for the application of ensemble algorithms in operational forecasting. Many options are available including the EAKF (Anderson 2001), EnKF (EnSRF in Whitaker and Hamill 2002),

196

rank histogram filter (Anderson 2010), and particle filter (Poterjoy 2015), etc. The recommended algorithm is EAKF so that will be explored in depth.

EAKF creates an updated ensemble from the prior ensemble members and observations and approximates the KF. However, for the EAKF to accurately compute the KF, it would need a linear forecast model, linear observation operator, gaussian observation errors, and an ensemble larger than the state vector. We are using a non-linear forecast model and an ensemble much smaller than the state vector, so errors are introduced. The EAKF algorithm takes advantage of Bayes' rule introduced in Section 1a.

For an observed variable, the EAKF shifts the prior ensemble so it becomes the same mean as the posterior. Then, the ensemble linearly contracts around the posterior mean resulting in a standard deviation equal to that of the posterior. Fig. 1 shows this process using 5 ensemble members. Once we arrive at the posterior values of the observed variables, the increments of each component of the prior state vector are computed from the observation increments via linear regression (Fig. 2). Anderson (2003) applies a "local least-squares" framework to the EAKF where each ensemble member at each model grid space is updated with scalars instead of vectors for each observation. Once an ensemble is established (discussed in Section 4), the steps to assimilate an observation is as follows (shown graphically in Fig. 3):

1) Start with an ensemble of the state of the atmosphere. This is the "prior" state.

2) Get an ensemble sample of the observation using the observational operator on the state variable for each member i: $y_i^f = h(x_i^f)$. (Fig. 3, step 2).

3) Get the observation from the instrument y^o and the observation error distribution $(\sigma^o)^2$ (Fig. 3, step 3).

4) Inflate the prior variances $(\sigma^f)^2$ to maintain model spread because of the limited sample size and model error (Hamill et al. 2001; Anderson 2001; Whitaker and Hamill 2002). This can be done multiplicative (Anderson 2001) or additive (Hamill and Whitaker 2005), but the use of a relaxation can limit excessive ensemble spread in data-sparse regions caused by the previous two techniques (Zhang et al. 2004). A newer technique is to use adaptive inflation which evolves in time and varies in space (Anderson 2009).

5) Calculate the analysis ensemble mean of the observation $\overline{y^a}$ (eq. 32) and then for each ensemble member i compute the analysis observation y_i^a (eq. 33) and observation increment Δy_i (eq. 34) (Fig. 3, step 4).

$$\overline{y^a} = \left(\frac{1}{\frac{1}{(\sigma^f)^2} + \frac{1}{(\sigma^o)^2}} \right) \left(\frac{\overline{y^f}}{(\sigma^f)^2} + \frac{y^o}{(\sigma^o)^2} \right) \tag{32}$$

$$y_i^a = \sqrt{\frac{(\sigma^o)^2}{(\sigma^f)^2 + (\sigma^o)^2}} \left(y_i^f - \overline{y^f} \right) + \overline{y^a} \tag{33}$$

$$\Delta y_i = y_i^a - y_i^f \tag{34}$$

6) Linearly regress the observation increments Δy_i onto state variable increments Δx_i for each ensemble member (eq. 35) and add to the prior state ensemble sample x_i^f to get the analysis value for the state variable x_i^a (eq. 36) (Fig. 3, step 5).

$$\Delta x_i = \frac{cov(x_i^f, y_i^f)}{var(y_i^f)} \Delta y_i \tag{35}$$

$$x_i^a = x_i^f + \Delta x_i \tag{36}$$

7) Continue until all observations are processed and state variables are updated to the analysis. Then, advance the ensemble members to the next observation time via the non-linear forward model (Fig. 3, step 6).

To limit filter divergence from spurious correlations far away from the observation (Houtekamer and Mitchell 2001; Hamill et al. 2001) which could render the EAKF useless, covariance localization must be multiplied by the RHS of eq. 35. A popular localization technique is using a 5th order polynomial and setting a cutoff radius both horizontally and vertically (Gaspari and Cohn 1999).

c. EnKF in convective-allowing NWP

In Zhang et al. (2006), the performance of an EnKF in conjunction with the fifth-generation Pennsylvania State University–National Center for Atmospheric Research Mesoscale Model (MM5) was analyzed. They assimilated synthetic sounding and surface observations with a perfect-model assumption and found that an EnKF used with 40 members kept the analysis close to the "truth". They also concluded that larger ensemble sizes, smaller cutoff radii (localization half-widths), and the implementation of a variance relaxation method (Zhang et al. 2004) resulted in larger ensemble spread to possibly remedy filter divergence. In an extension of this work, Meng and Zhang (2008) modeled a mesoscale convective vortex (MCV) on 10-12 June 2003 using the Weather Research and Forecast Model (WRF; Skamarock et al. 2008). They showed that the EnKF outperformed the 3D-VAR in assimilating one or more meteorological variables.

NCAR's Experimental Real-Time Convection-Allowing Ensemble Prediction System (Schwartz et al. 2015) was a prototype of the use of EnKF in an operational limited-area convective-permitting ensemble forecasting system. This setup used DART in conjunction with the WRF in a continuously-cycling system that was used to initialize the ensembles. They assimilated rawinsonde temperature, wind, moisture, and altimeter, aircraft temperature and wind, satellite winds, surface temperature, wind, altimeter, and moisture from ship, buoy, METAR, surface synoptic observations, and Oklahoma Mesonet sites. Although NCAR's system used a continuous-cycling framework, the use of partial cycling systems may lead to better forecasts by eliminating the buildup of biases. The downside is that they fail robustly to assess the model biases (Romine et al. 2013). Also, the external model forcing is only responsible for the lateral boundary conditions (LBCs) in continuously-cycling systems whereas there may be dynamically-inconsistent issues from the external model ICs in partially-cycling systems.

4. Assimilating GPS PWV via EnKF

a. PWV from GPS observations

Ground-based stations can measure precipitable water vapor (PWV) from Global Positioning System (GPS) satellites that have been in place since 1992 (Bevis et al. 1992). The stations measure Zenith Total Delay (ZTD) based on the delay of L-band radio signals in the troposphere between the transmitters and station. The ZTD is dependent on the surface pressure (hydrostatic dry delay) and the total column PWV (wet delay) above the station. The ZTD can be converted to PWV with temperature and pressure data from the station using a method described in Bevis et al. (1992). Poli et al. (2007) suggests that the ZTD/PWV for each station is representative of a 17-km radius

with most of the contribution in the lowest 3 km. The error in PWV is 1-2 mm for small values. The long wavelength from the GPS transmitter makes PWV an all-weather observable meteorological variable.

b. Value of PWV in NWP

Over the central US, Gutman et al. (2004) and a companion study Smith et al. (2007) showed that lower troposphere moisture fields in the RUC model improved upon the assimilation of PWV. This was done via an optimal interpolation method where the shape of the moisture profile is assumed to be correct and the values are adjusted by a certain percentage as demonstrated in Kuo et al. (1993) and Smith et al. (2000). Smith et al. (2007) showed greatest impact to 3-6 hour forecasts with a noticeable impact on the 9-12 hour forecasts.

In a European study, Poli et al. (2007) examine the impact of the assimilation of ZTD into at 4D-VAR system implemented with Meteo-France global forecast system. In addition to constraining the 1-4 day synoptic patterns for all seasons over France, this system also improves rainfall forecasts 12-36 hours out. In another European study as part of the Convective and Orographically-induced Precipitation Study (COPS) campaign, Yan et al. (2009) assimilate ZTD into a 3D-VAR/AROME Meteo-France system with 2.5-km horizontal grid spacing (convective-permitting). This study used a dense GPS network in the complex terrain of northeast France and southwest Germany that were a part of COPS. Improved forecast skill of rainfall were attributed to the assimilation of the ZTD where the lower tropospheric moisture profiles were improved similar to the Central US studies of Gutman et al. (2004) and Smith et al. (2007).

Seko et al. (2011) were one of the first to assimilate PWV using an ensemble DA system. They used a Local Ensemble Transform Kalman Filter (LETKF) and analyzed fields of the mesoscale 4D-VAR system of the Japan Meteorological Agency (JMA) to improve rainfall in a case study. The horizontal resolution was 1.6 km with 20 ensemble members. The assimilation of PWV improved the lower tropospheric moisture and rainfall forecasts in the heavy bands. They hypothesize that the assimilation of more variables, having a higher resolution, and increasing the ensemble members will allow a better simulation of the rain bands.

c. Experimental Design

1) PERTURBATIONS AND SPINUP

In my data assimilation algorithm, I use DART with the EAKF filter and WRF-ARW version 3.41 to advance the model. I initialize the model at 0000 UTC with the same configuration as Moker et al. (2018). WPS and real.exe produce the wrfbdy and wrfinputs from ICs and LBCs using the 0.25-degree GFS and 30-km NAM in three nested domains: 30 km, 10 km, 2.5 km. Soil temperature and moisture information is obtained from RAP. Then, I perturb the 30-km domain to form 20 ensemble members using WRF-DA (Barker et al. 2012). Here, the "cv3" option is chosen that uses the NMC method for background error covariance where the statistics are based on the difference between the 12h and 24h GFS forecasts on a global scale (Parrish and Derber 1992). The model "spins up" from 0000 UTC to 0600 UTC in order for the perturbations in the 30-km domain to propagate to the innermost 2.5-km domain and there is flow-dependent background error statistics.

2) ASSIMILATION ALGORITHM

At 0600 UTC, we begin hourly assimilation cycles using the EAKF scalar update (Anderson 2003) that was generalized in Section 3b. During Transect 2013, 5-min PWV data from 9 stations across the SMO and lowest elevations of northwest Mexico. At each hour and for each station, the 5-min observations falling within +/- 30 min of the observation time are assimilated sequentially in the following 5 steps. When all hourly observations are assimilated, the models advance for 1 hour with WRF-ARW using the new wrfinputs. These hourly cycles continue until 1200 UTC where there is a final analysis. We assimilate a PWV observation in the following steps:

1) Compute an ensemble sample of the PWV observation at the GPS site $PWV_i^f = H(x_i^f)$ using an operational operator H. First, construct a column of specific humidity q and pressure p at 4 model grid points surrounding the site. Then, bilinearly interpolate these columns to form one column above the observation site. Finally, multiply the average q in each model layer by the change in p, sum across all levels, then multiple by the negative reciprocal of the average force of gravity g to obtain the PWV in mm (eq. 37).

$$PWV = -\frac{1}{g} \sum_{i=1}^{nlev-1} \left(\frac{q_i + q_{i+1}}{2}\right)(p_{i+1} - p_i) \tag{37}$$

2) Use adaptive inflation on the prior variance with a method that evolves in time and varies in space (Anderson 2009).

3) Get the PWV observation from the instrument using an observation error of 0.75 mm. Note: this is half of that in the literature because the ensembles were being adjusted too greatly resulting in some members crashing during the predict stage.

4) For each ensemble member i, compute the PWV observation increment ΔPWV_i (Δy_i in eq. 34) from the analysis PWV observation PWV_i^a (y_i^a in eq. 33) after first computing the analysis ensemble mean of the PWV observation $\overline{PWV^a}$ ($\overline{y^a}$ in eq. 32).

5) For each ensemble member i, linearly regress the PWV observation increments ΔPWV_i onto state variable increments Δx_i using eq. 38.

$$\Delta x_i = \beta \frac{cov\left(x_i^f, PWV_i^f\right)}{var\left(PWV_i^f\right)} \Delta PWV_i \tag{38}$$

This equation is eq. 34 mutiplied by β, the localization value (Gaspari and Cohn 1999) between 0 and 1 that decreases as one moves both horizontally and vertically away from the observation site. With a cutoff value of 0.07, the horizontal-equivalent distance is ~450 km and the vertical-equivalent distance of 3.5 km. We regress onto all state variables (thermodynamic and dynamic) and on all three domains. Finally, add the state variable increment Δx_i to the prior state ensemble sample x_i^f to get the analysis value for the state variable x_i^a (eq. 36).

3) DETERMINISTIC FORECASTS

After the final analysis occurs at 1200 UTC, the ensemble mean is then computed to initialize the WRF deterministic run which is then run for 24 hours as in Moker et al. (2018).

References

Anderson, J. L., T. Hoar, K. Raeder, H. Liu, N. Collins, R. Torn, and A. Arellano, 2009: The data assimilation research testbed a community facility. *Bull. Am. Meteorol. Soc.*, **90**, 1283–1296, doi:10.1175/2009BAMS2618.1.

——, 2001: An Ensemble Adjustment Kalman Filter for Data Assimilation. *Mon. Weather Rev.*, **129**, 2884–2903, doi:10.1175/1520-0493(2001)129<2884:AEAKFF>2.0.CO;2. http://journals.ametsoc.org/doi/abs/10.1175/1520-0493%282001%29129%3C2884%3AAEAKFF%3E2.0.CO%3B2.

——, 2003: A Local Least Squares Framework for Ensemble Filtering. *Mon. Weather Rev.*, **131**, 634–642, doi:10.1175/1520-0493(2003)131<0634:ALLSFF>2.0.CO;2. http://journals.ametsoc.org/doi/abs/10.1175/1520-0493%282003%29131%3C0634%3AALLSFF%3E2.0.CO%3B2.

——, 2009: Spatially and temporally varying adaptive covariance inflation for ensemble filters. *Tellus, Ser. A Dyn. Meteorol. Oceanogr.*, **61 A**, 72–83, doi:10.1111/j.1600-0870.2008.00361.x.

——, 2010: A Non-Gaussian Ensemble Filter Update for Data Assimilation. *Mon. Weather Rev.*, **138**, 4186–4198, doi:10.1175/2010MWR3253.1. https://doi.org/10.1175/2010MWR3253.1.

——, and S. L. Anderson, 1999: A Monte Carlo Implementation of the Nonlinear Filtering Problem to Produce Ensemble Assimilations and Forecasts. *Mon. Weather Rev.*, **127**, 2741–2758, doi:10.1175/1520-0493(1999)127<2741:AMCIOT>2.0.CO;2. https://doi.org/10.1175/1520-

0493(1999)127%3C2741:AMCIOT%3E2.0.CO.

Barker, D., and Coauthors, 2012: The weather research and forecasting model's community variational/ensemble data assimilation system: WRFDA. *Bull. Am. Meteorol. Soc.*, **93**, 831–843, doi:10.1175/BAMS-D-11-00167.1.

Bevis, M., S. Businger, T. A. Herring, C. Rocken, R. A. Anthes, and R. H. Ware, 1992: GPS meteorology: Remote sensing of atmospheric water vapor using the global positioning system. *J. Geophys. Res.*, **97**, 15787, doi:10.1029/92JD01517. http://doi.wiley.com/10.1029/92JD01517.

Bishop, C. H., B. J. Etherton, and S. J. Majumdar, 2001: Adaptive Sampling with the Ensemble Transform Kalman Filter. Part I: Theoretical Aspects. *Mon. Weather Rev.*, **129**, 420–436, doi:10.1175/1520-0493(2001)129<0420:ASWTET>2.0.CO;2. https://doi.org/10.1175/1520-0493(2001)129%3C0420:ASWTET%3E2.0.CO.

Burgers, G., P. Jan van Leeuwen, and G. Evensen, 1998: Analysis Scheme in the Ensemble Kalman Filter. *Mon. Weather Rev.*, **126**, 1719–1724, doi:10.1175/1520-0493(1998)126<1719:ASITEK>2.0.CO;2. https://doi.org/10.1175/1520-0493(1998)126%3C1719:ASITEK%3E2.0.CO.

Cohn, S. E., A. da Silva, J. Guo, M. Sienkiewicz, and D. Lamich, 1998: Assessing the Effects of Data Selection with the DAO Physical-Space Statistical Analysis System. *Mon. Weather Rev.*, **126**, 2913–2926, doi:10.1175/1520-0493(1998)126<2913:ATEODS>2.0.CO;2. https://doi.org/10.1175/1520-0493(1998)126%3C2913:ATEODS%3E2.0.CO.

Courtier, P., J.-N. Thépaut, and A. Hollingsworth, 1994: A strategy for operational

implementation of 4D-Var, using an incremental approach. *Q. J. R. Meteorol. Soc.*, **120**, 1367–1387, doi:10.1002/qj.49712051912. https://doi.org/10.1002/qj.49712051912.

Le Dimet, F.-X., and O. Talagrand, 1986: Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects. *Tellus A*, **38A**, 97–110, doi:10.1111/j.1600-0870.1986.tb00459.x. https://doi.org/10.1111/j.1600-0870.1986.tb00459.x.

Evensen, G., 1994: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.*, **99**, 10143–10162, doi:10.1016/0379-0738(94)01632-F.

Gaspari, G., and S. E. Cohn, 1999: Construction of correlation functions in two and three dimensions. *Q. J. R. Meteorol. Soc.*, **125**, 723–757, doi:10.1256/smsqj.55416.

Gutman, S. I., S. R. Sahm, S. G. Benjamin, B. E. Schwartz, K. L. Holub, J. Q. Stewart, and T. L. Smith, 2004: Rapid Retrieval and Assimilation of Ground Based GPS Precipitable Water Observations at the NOAA Forecast Systems Laboratory: Impact on Weather Forecasts. *J. Meteorol. Soc. Japan*, **82**, 351–360, doi:10.2151/jmsj.2004.351. http://joi.jlc.jst.go.jp/JST.JSTAGE/jmsj/2004.351?from=CrossRef.

Hamill, T. M., and J. S. Whitaker, 2005: Accounting for the Error due to Unresolved Scales in Ensemble Data Assimilation: A Comparison of Different Approaches. *Mon. Weather Rev.*, **133**, 3132–3147, doi:10.1175/MWR3020.1. https://doi.org/10.1175/MWR3020.1.

——, J. S. Whitaker, and C. Snyder, 2001: Distance-Dependent Filtering of Background

Error Covariance Estimates in an Ensemble Kalman Filter. *Mon. Weather Rev.*, **129**, 2276–2790, doi:10.1080/00365519009091567.

Houtekamer, P. L., and H. L. Mitchell, 1998: Data Assimilation Using an Ensemble Kalman Filter Technique. *Mon. Weather Rev.*, **126**, 796–811, doi:10.1175/1520-0493(1998)126<0796:DAUAEK>2.0.CO;2. http://journals.ametsoc.org/doi/abs/10.1175/1520-0493%281998%29126%3C0796%3ADAUAEK%3E2.0.CO%3B2.

——, and ——, 2001: A Sequential Ensemble Kalman Filter for Atmospheric Data Assimilation. *Mon. Weather Rev.*, **129**, 123–137, doi:10.1175/1520-0493(2001)129<0123:ASEKFF>2.0.CO;2. http://journals.ametsoc.org/doi/abs/10.1175/1520-0493%282001%29129%3C0123%3AASEKFF%3E2.0.CO%3B2.

Hunt, B. R., E. J. Kostelich, and I. Szunyogh, 2007: Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter. *Phys. D Nonlinear Phenom.*, **230**, 112–126, doi:https://doi.org/10.1016/j.physd.2006.11.008. http://www.sciencedirect.com/science/article/pii/S0167278906004647.

Kalman, R. E., 1960: A New Approach to Linear Filtering and Prediction Problems. *J. Basic Eng.*, **82**, 35–45. http://dx.doi.org/10.1115/1.3662552.

Kalnay, E., 2002: *Atmospheric Modeling, Data Assimilation and Predictability*. Cambridge University Press, Cambridge, https://www.cambridge.org/core/books/atmospheric-modeling-data-assimilation-and-predictability/C5FD207439132836E85027754CE9BC1A.

Kuo, Y.-H., Y.-R. Guo, and E. R. Westwater, 1993: Assimilation of Precipitable Water

    Measurements into a Mesoscale Numerical Model. *Mon. Weather Rev.*, **121**,

    1215–1238. https://doi.org/10.1175/1520-

    0493(1993)121%3C1215:AOPWMI%3E2.0.CO;2.

Lewis, J. M., and J. C. Derber, 1985: The use of adjoint equations to solve a variational

    adjustment problem with advective constraints. *Tellus A*, **37A**, 309–322,

    doi:10.1111/j.1600-0870.1985.tb00430.x. https://doi.org/10.1111/j.1600-

    0870.1985.tb00430.x.

Lorenc, A. C., 1986: Analysis methods for numerical weather prediction. *Q. J. R.*

    Meteorol. Soc., **112**, 1177–1194, doi:10.1002/qj.49711247414.

    https://doi.org/10.1002/qj.49711247414.

Meng, Z., and F. Zhang, 2008: Tests of an Ensemble Kalman Filter for Mesoscale and

    Regional-Scale Data Assimilation. Part III: Comparison with 3DVAR in a Real-

    Data Case Study. *Mon. Weather Rev.*, **136**, 522–540,

    doi:10.1175/2007MWR2106.1.

    http://journals.ametsoc.org/doi/abs/10.1175/2007MWR2106.1.

Parrish, D. F., and J. C. Derber, 1992: The National Meteorological Center's Spectral

    Statistical-Interpolation Analysis System. *Mon. Weather Rev.*, **120**, 1747–1763,

    doi:10.1175/1520-0493(1992)120<1747:TNMCSS>2.0.CO;2.

    https://doi.org/10.1175/1520-0493(1992)120%3C1747:TNMCSS%3E2.0.CO.

Poli, P., and Coauthors, 2007: Forecast impact studies of zenith total delay data from

    European near real-time GPS stations in Météo France 4DVAR. *J. Geophys. Res.*

    Atmos., **112**, doi:10.1029/2006JD007430. https://doi.org/10.1029/2006JD007430.

Poterjoy, J., 2015: A Localized Particle Filter for High-Dimensional Nonlinear Systems.
Mon. Weather Rev., **144**, 59–76, doi:10.1175/MWR-D-15-0163.1.
https://doi.org/10.1175/MWR-D-15-0163.1.

Romine, G. S., C. S. Schwartz, C. Snyder, J. L. Anderson, and M. L. Weisman, 2013:
Model Bias in a Continuously Cycled Assimilation System and Its Influence on
Convection-Permitting Forecasts. *Mon. Weather Rev.*, **141**, 1263–1284,
doi:10.1175/MWR-D-12-00112.1.
http://journals.ametsoc.org/doi/abs/10.1175/MWR-D-12-00112.1.

Schwartz, C. S., G. S. Romine, R. A. Sobash, K. R. Fossell, and M. L. Weisman, 2015:
NCAR's Experimental Real-Time Convection-Allowing Ensemble Prediction
System. *Weather Forecast.*, **30**, 1645–1654, doi:10.1175/WAF-D-15-0103.1.
http://journals.ametsoc.org/doi/10.1175/WAF-D-15-0103.1.

Seko, H., T. Miyoshi, Y. Shoji, and K. Saito, 2011: Data assimilation experiments of
precipitable water vapour using the LETKF system: Intense rainfall event over
Japan 28 July 2008. *Tellus, Ser. A Dyn. Meteorol. Oceanogr.*, **63**, 402–414,
doi:10.1111/j.1600-0870.2010.00508.x.

Skamarock, W. C., and Coauthors, 2008: A Description of the Advanced Research WRF
Version 3. *NCAR Tech. NOTE*, **NCAR/TN**–**47**,
doi:10.1080/07377366.2001.10400427.

Smith, T. L., S. G. Benjamin, B. E. Schwartz, and S. I. Gutman, 2000: Using GPS-IPW
in a 4-D data assimilation system. *Earth, Planets Sp.*, **52**, 921–926,
doi:10.1186/BF03352306.

——, ——, S. I. Gutman, and S. Sahm, 2007: Short-Range Forecast Impact from

Assimilation of GPS-IPW Observations into the Rapid Update Cycle. *Mon. Weather Rev.*, **135**, 2914–2930, doi:10.1175/MWR3436.1. http://journals.ametsoc.org/doi/abs/10.1175/MWR3436.1.

Tippett, M. K., J. L. Anderson, C. H. Bishop, T. M. Hamill, and J. S. Whitaker, 2003: Ensemble Square Root Filters. *Mon. Weather Rev.*, **131**, 1485–1490.

Whitaker, J. S., and T. M. Hamill, 2002: Ensemble Data Assimilation without Perturbed Observations. *Mon. Weather Rev.*, **130**, 1913–1924, doi:10.1175/1520-0493(2002)130<1913:EDAWPO>2.0.CO;2. https://doi.org/10.1175/1520-0493(2002)130%3C1913:EDAWPO%3E2.0.CO.

Wikle, C. K., and L. M. Berliner, 2007: A Bayesian tutorial for data assimilation. *Phys. D Nonlinear Phenom.*, **230**, 1–16, doi:10.1016/j.physd.2006.09.017.

Yan, X., V. Ducrocq, G. Jaubert, P. Brousseau, P. Poli, C. Champollion, C. Flamant, and K. Boniface, 2009: The benefit of GPS zenith delay assimilation to high-resolution quantitative precipitation forecasts: a case-study from COPS IOP 9. *Q. J. R. Meteorol. Soc.*, **135**, 1788–1800, doi:10.1002/qj.508. https://doi.org/10.1002/qj.508.

Zhang, F., C. Snyder, and J. Sun, 2004: Impacts of Initial Estimate and Observation Availability on Convective-Scale Data Assimilation with an Ensemble Kalman Filter. *Mon. Weather Rev.*, **132**, 1238–1253, doi:10.1175/1520-0493(2004)132<1238:IOIEAO>2.0.CO;2. https://doi.org/10.1175/1520-0493(2004)132%3C1238:IOIEAO%3E2.0.CO.

——, Z. Meng, and A. Aksoy, 2006: Tests of an Ensemble Kalman Filter for Mesoscale and Regional-Scale Data Assimilation. Part I: Perfect Model Experiments. *Mon.*

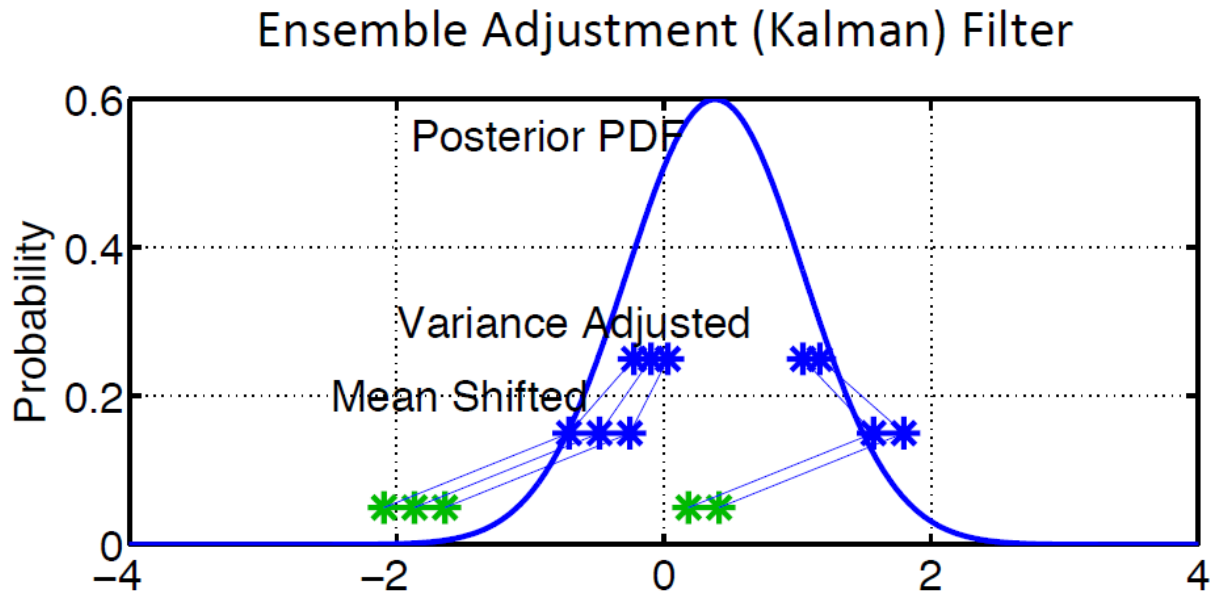*Weather Rev.*, **134**, 722–736, doi:10.1175/MWR3101.1.

https://doi.org/10.1175/MWR3101.1.

Tables

Table 1: Conversion from scalar to vector quantities

	Scalar	Vector	Dimension
Background	x_b	$\boldsymbol{x}^b \equiv \begin{bmatrix} x_0^b \\ \vdots \\ x_n^b \end{bmatrix}$	$m \times 1$
Analysis	x_a	$\boldsymbol{x}^a \equiv \begin{bmatrix} x_0^a \\ \vdots \\ x_n^a \end{bmatrix}$	$m \times 1$
Observation	x_o	$\boldsymbol{y} \equiv \begin{bmatrix} y_0 \\ \vdots \\ y_n \end{bmatrix}$	$n \times 1$

Figures

Ensemble Adjustment (Kalman) Filter

Figure 1: Using a 1-dimensional EAKF, 5 ensemble members that are also observable are adjusted from their prior values (green *) to their posterior values (blue *) based on observation(s) and error characteristics (not shown). The posterior ensemble members are a sample of resulting posterior PDF (blue line) after they are bias-corrected (mean shifted) and error-reduced (variance adjusted). Source: DART Tutorial.
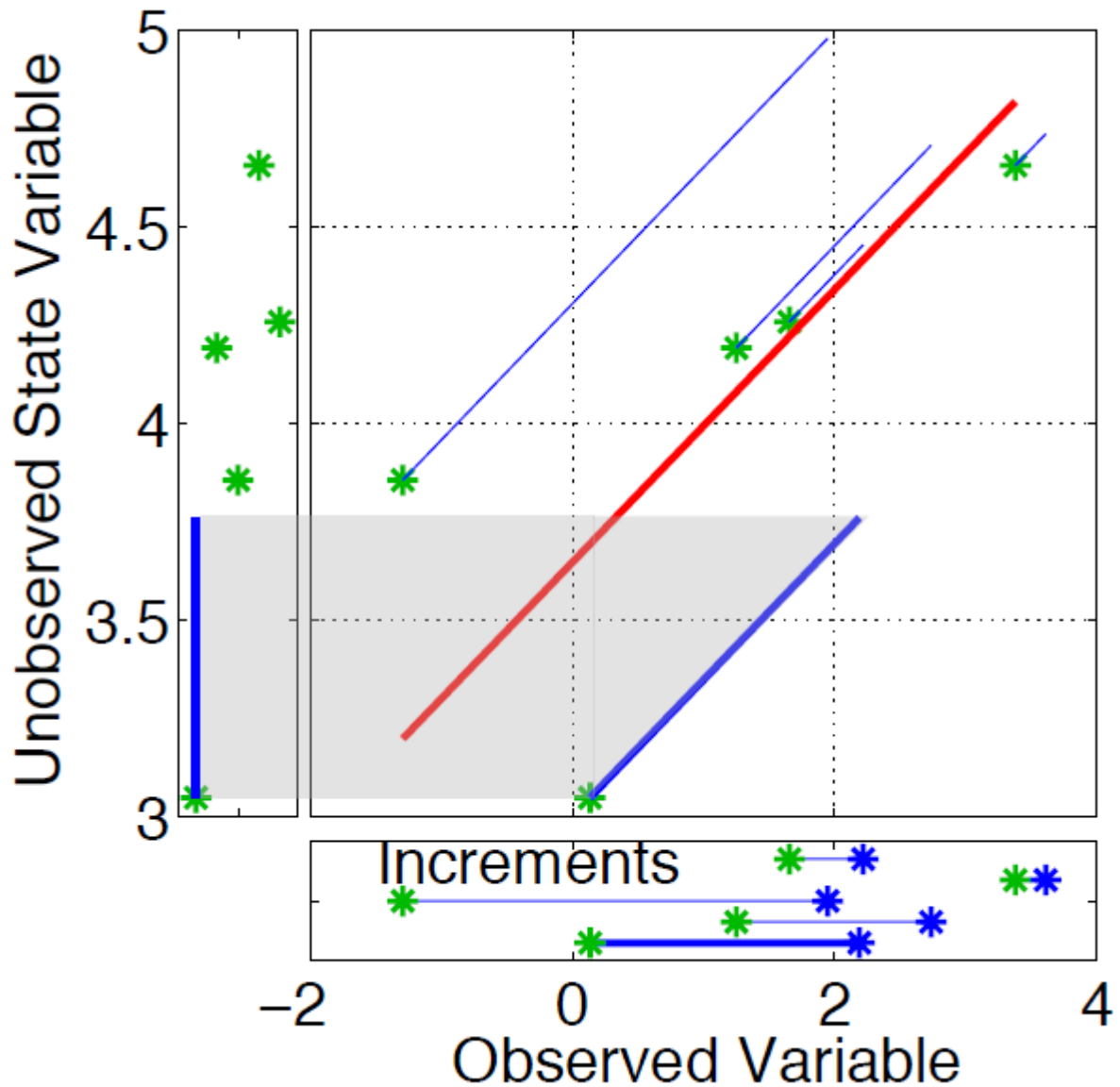
Figure 2: Unobserved variables (on the ordinate) are adjusted from the increments (blue lines) from observed state variables (on the abscissa) using a scalar version of EAKF and 5 ensemble members. The increments are the change from the prior (green *) to the posterior (blue *). In other words, the observed variable increments are linearly regressed onto the unobserved state variable increments in the joint space. The red line indicates the regression coefficient. Source: DART Tutorial.
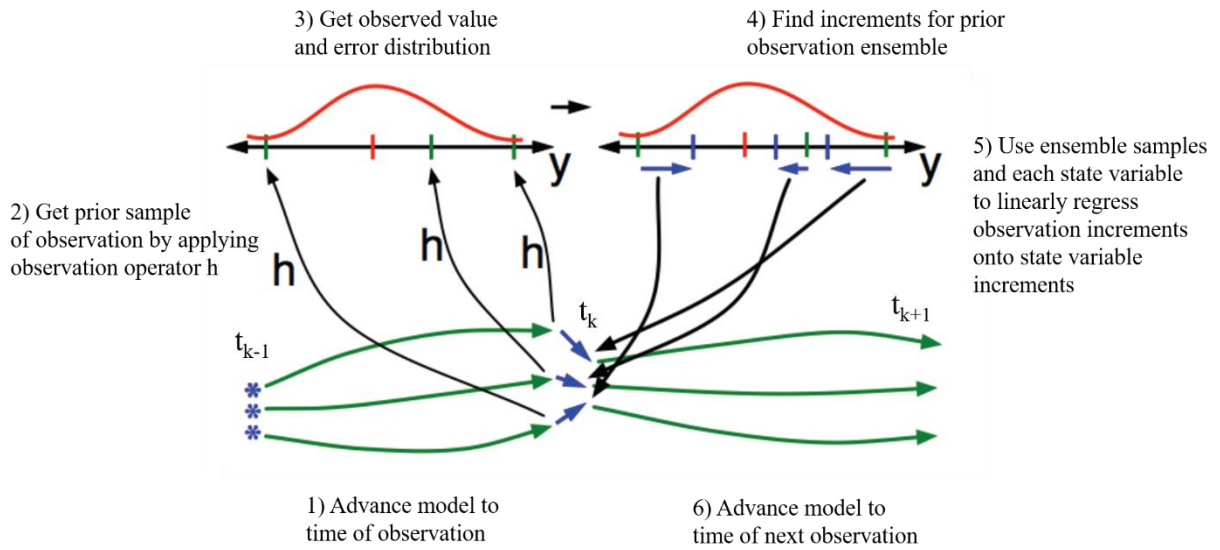
Figure 3: The steps involved in assimilating an observation via the EAKF algorithm using 3 ensemble members. Green indicates prior, blue indicates posterior, and red indicates observation. Source graphic from DART Tutorial.