UCLA
UCLA Electronic Theses and Dissertations

Title
The Evaluation of a Southern California Physics Ensemble

Permalink
https://escholarship.org/uc/item/8ct5c0bv

Author
Wilson, Travis Harold

Publication Date
2012

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

University of California

Los Angeles

The Evaluation of a Southern California Physics Ensemble.

A thesis submitted in partial satisfaction

of the requirements for the degree

Master of Science in Atmospheric and Oceanic Sciences

by

Travis Harold Wilson

2012

Abstract of the Thesis

The Evaluation of a Southern California Physics Ensemble.

by

Travis Harold Wilson

Master of Science in Atmospheric and Oceanic Sciences

University of California, Los Angeles, 2012

Professor Robert G. Fovell, Chair

Using the Weather Research and Forecasting (WRF) model, an intra-model physics ensemble was created for July 2005 in hopes of finding a physics configuration that could most accurately recreate the typical summertime conditions of Southern California. The ensemble consisted of twenty-four members that mixed different Land Surface Models (LSM) with various Planetary Boundary Layer (PBL) parameterizations currently offered in WRF. It was found that the Pleim-Xiu LSM coupled with either the asymmetric convective model version 2 or the Yonsei University PBL scheme created the most skillful ensemble member.

In addition to the physics ensemble, short integrations along with different initializations proved to be beneficial. Initializing the model off of the North American Regional Reanalysis (NARR) instead of the North American Mesoscale (NAM) model was, in this case, extremely useful. The most accurate simulation of July 2005 was found when taking integration methods, initializations, and model physics into consideration.

The thesis of Travis Harold Wilson is approved.

Yongkang Xue

Kuo-Nan Liou

Robert G. Fovell, Committee Chair

University of California, Los Angeles

2012

Table of Contents

List of Figures

List of Tables

CHAPTER 1

Introduction

In November 2000 a regional Numerical Weather Prediction (NWP) model that was capable of handling relatively small horizontal length scales (1-10 kilometers (km)) was released to the public. This new model, named the Weather Research and Forecasting Model (WRF), was a collaborative effort by university scientists, the National Center for Atmospheric Research (NCAR), the National Centers for Environmental Prediction (NCEP), the Air Force Weather Agency (AFWA) and the Forecast Systems Laboratory (FSL) [18]. When first proposed, WRF would be designed with three different dynamical cores under which you could choose an array of physics parameterizations to optimize the model for your area [5]. Though as the WRF project progressed, only two dynamical cores were retained: the Nonhydrostatic Mesoscale Model (NMM) and the Advanced Research WRF (ARW). In this paper we will focus on the WRF-ARW which is a fully compressible, nonhydrostatic model using terrain-following coordinates.

As of 2012 the WRF-ARW, which will be referred to as WRF, contains eleven different Planetary Boundary Layer (PBL) parameterizations along with four Land Surface Models (LSM). As previously mentioned, all of these parameterizations can be mixed and matched for optimization in the area of interest. This is particularly useful because it allows for intra-model comparisons along with sensi-

tivity studies. Gilliam and Pleim's intra-model WRF comparison has found that the Pleim-Xiu LSM (coupled with the Asymmetric Convective Model (ACM2) PBL) was superior to the Noah LSM (coupled with the Yonsei University (YSU) PBL) for the Eastern United States summertime 2 meter (m) temperature, mixing ratio, and 10m wind [11]. In contrast, wintertime simulations proved the Noah LSM better in forecasting both the 2m temperature and mixing ratio. A different study done by Argüeso et al. — which consisted of a 10 year ensemble for Southern Spain — found that WRF compared best with surface observations when using the Noah LSM coupled with either the ACM2 or YSU PBL[1]. All combinations of physics parameterizations mentioned in these studies are available for use in the current version of WRF.

Both the YSU and ACM2 are non-local boundary layer schemes, which means they can mix momentum, temperature, and moisture over the depth of the PBL rather than just mixing with adjacent model levels. These schemes are generally considered less sophisticated since they do not have any prognostic variables. In contrast the Mellor-Yamada-Janjic (MYJ) which was one of WRFs original PBL options, uses Turbulent Kinetic Energy (TKE) as a prognostic variable and has local mixing that is a function of the prognostic TKE [17]. Despite being less sophisticated, it has been found that non-local PBL schemes such as the YSU and ACM2 more accurately reproduce meteorological features in the turbulent flow of dry convective boundary layers [10].

Not all intra-model ensembles have proven a specific boundary layer or land surface model as superior. Ruiz and Saulo found that summertime simulations of South America did not favor a unique model design for the surface parameters they tested[16]. In other words, a certain physics combination that did well

forecasting temperatures did not perform favorably for other variables. This paper is an anomalous case as most intra-model ensembles discover certain physics combinations that are most suitable for the area of interest. This is because parameterizations designed and optimized for a certain climate will most likely not be suitable for other locations. Essentially, one can fine tune the model — in this case the WRF model — for each specific climate using the different physics parameterizations.

Our area of interest is the densely populated region of Southern California. If indeed a unique physics configuration is found for this area it will not only assist ongoing improvements to physics parameterizations, but a relatively large population could benefit from improved forecasts.

1.1 Model Design and Setup

A twenty-four member physics ensemble was created for Southern California during July 2005. Using the Advanced Research core of the WRF model, an evaluation of eight different boundary layer schemes along with three land surface models was performed. The ultimate goal of such a large physics ensemble is to determine what physics parameterization combination is most suitable for this region under typical summertime conditions. To the author's knowledge, no one has ever created this large of an intra-model WRF ensemble for Southern California.

The relative skill of each ensemble member was validated using both surface based weather stations and observations obtained from aircraft. While aircraft data used to verify ensemble results is highly variable both spatially and tempo-

rally, it offers the most abundant and sometime the only direct observations in the lower atmosphere. Surface meteorological parameters verified were the 1.5m temperature, dew point, and 10m wind speed. The dew point, which is a measure of atmospheric water vapor, was not measured on aircraft during the time of interest. Thus, vertical profiles of the numerical simulation were only verified with temperature and wind.

The ensemble used 51 vertical levels in a triply-nested design. The outer domain, which will be referred to as domain 1, has a horizontal resolution of 36km (Fig. 1.1). The inner nests, domains 2 and 3, are centered along the Southern California coast and have resolutions of 12 and 4km, respectively. We chose to position domain 3 towards the coast in order to resolve the correct location of the large summertime temperature gradients that so commonly exist between the land and sea. In addition to this, we wanted the computational power focused on the area with the highest population.

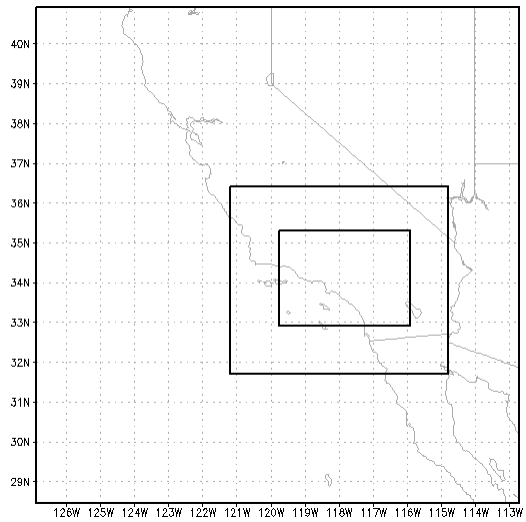

Figure 1.1: Domains 1,2, and 3 shown.

Simulations were initialized off of the ETA model on July 3rd 2005. The ETA

model was an operational NWP model for the Untied Stated that was renamed the North American Mesoscale (NAM) model in January 2005. Since the name changed happened before July 2005 this paper will be referring to the ETA as the NAM. During June 2006 the NAM model was replaced by the WRF-NMM and is similar to the Advanced Research WRF used in this physics ensemble; the major difference lies in the dynamical core of each model. Another large distinction lies in the physics parameterization offered; the WRF-ARW has a wide array of parameterizations while the WRF-NMM is much more limited. Though we did not test the WRF-NMM, Gallus and Bresch suggest that grater sensitivities to different physics parameterization will occur in the WRF-NMM than in the WRF-ARW used in this paper[9]. This does not mean that the WRF-NMM is closer to reality but rather has a larger spread among possible outcomes when using different physics.

After the initialization the ensemble was integrated for 72 hours, stopped, then reinitialized 24 hours earlier and integrated for a subsequent 72 hours. The purpose of creating a large number of short simulations with 24 hours of overlap was to mitigate the affects of 'drift' and 'spin-up' in the WRF model, more of which will be discussed later. Simulations ended on the 30th of July resulting in a total of thirteen different 72 hour runs for each ensemble member. Considering the eight different boundary layer schemes tested with each of the three land surface models, a total of 312 simulations were created. Five of these simulations consistently failed for unknown reasons and are not reported in this paper; however, the effect of this missing data is minute given our large sample size.

The three land surface models chosen to be tested in July 2005 were the Noah, Pleim-Xiu (PX), and the 5-layer Thermal Diffusion (TD) scheme. Each of these

land surface models use a different number of soil layers to predict soil moisture and/or soil temperature, though the number of soil layers do not necessarily correlate to the LSMs skill. The TD scheme, which was imported from the Fifth-Generation NCAR/Penn State Mesoscale Model (MM5), has five soil layers that predict soil temperatures but does not explicitly solve soil moisture or snow cover. The moisture availability in this land surface model is based on the land-use category only. On the other hand, the Noah land surface model predicts soil moisture in four different layers in addition to soil temperature, snow cover, and canopy moisture. Lastly, the PX is similar to the Noah LSM in terms of prognostic variables but has only two soil layers [6].

Arguably the most important parameters calculated by the land surface models are the sensible and latent heat fluxes. The sensible heat flux is an exchange in heat via conduction and convection. Simply put, a warm ground surface could exchange heat with the atmosphere resulting in a cooler ground surface and warmer atmosphere. A simplified version of the sensible heat flux can be presented as

$$H = -\rho c_p C_H(\theta_1 - \theta_s) \tag{1.1}$$

$$C_H = \frac{k u_*}{ln(\frac{z_1}{z_{0T}}) - \psi_H} \tag{1.2}$$

$$u_* = \frac{k U_1}{ln(\frac{z_1}{z_{0M}}) - \psi_M} \tag{1.3}$$

where H is the sensible heat flux, ρ is density, c_p is the specific heat at constant pressure, and θ_1, θ_s are the potential temperatures at sigma level one (the first

model layer above the surface) and the surface, respectively [17]. One can see that the sensible heat flux increases as the difference in potential temperature from the surface and the lowest model level increase. In eq. 1.1, θ_s is the only value directly calculated in the LSM.

The exchange coefficient, C_H (eq. 1.2), is calculated in the surface layer where k is the von Kármán constant, z_1 is the height of sigma level one, z_{0T} is surface roughness length for heat (which increases for uneven surfaces), and ψ_H is the stability function for heat which varies between surface layers schemes. The surface layer in a numerical model is the link between the PBL and LSM parameterizations; it calculates the momentum flux along with the described exchange coefficients for heat and moisture. In the WRF model, the surface layer parameterization changes based on the selected PBL.

Finally, the exchange coefficient is dependent on the friction velocity, u_* (eq. 1.3), in which U_1 is the wind speed at the lowest sigma level and the subscript M in the stability function and roughness length represents momentum. While these three equations can be interpreted a number of ways, one can see that an increase in wind speed would increase the friction velocity. In turn, this would increase the exchange coefficient along with the sensible heat flux. Essentially, the faster the wind, the more heat we can exchange with the surrounding atmosphere.

The latent heat flux, which is the heat absorbed or released during a phase change for water, is also calculated by the LSM. In its simplest form the latent heat flux is defined by

$$LH = L_v(E_{dir} + E_c + E_t) \tag{1.4}$$

where L_v is the latent heat of vaporization, E_t is the transpiration, E_c is canopy reevaporation, and E_{dir} is the direct evaporation from bare soil [17]. Imagine water evaporating from soil by the absorption of solar radiation: as the water evaporates the latent heat absorbed would be transferred into the air, however, it would not change the temperature of the atmosphere. The energy is in a sense, hidden. If there was no water to evaporate, the absorption of solar radiation would simply allow the ground to warm which would in turn warm the atmosphere. From this example we can see that the partitioning of the sensible and latent heat fluxes, which are dependent on the land surface model chosen, are crucial for low-level temperatures.

It should be noted that the Noah and PX LSMs used in the ensemble require additional fields for initialization. These fields, which include soil temperature, moisture, and snow water equivalent, are provided in the NAM dataset but are not from observations. Rather, the variables come from offline models which incorporate observations (rainfall, temperature, humidity, radiation, and wind) to estimate soil fields [6]. So when using a regional model such as WRF, the "downscaling" that occurs means there will be inconsistencies from the original grid in elevation, vegetation, and soil type. To accommodate the downscaling one should run an offline model with the same grid to 'spin-up' the soil parameters in the LSM. In this study we did not run an offline model to alleviate the spin-up effects of the LSM; instead, we decided to remove the first 24 hours that were previously mentioned as an overlap period.

In order to determine if the 24 hour spin-up period will suffice for the surface parameters of interest, the Mean Absolute Errors (MAE) of temperature, dew point, and wind speed averaged over all simulations by forecast hour (including

the spin-up period) are shown in Fig. 1.2. The MAE is defined by

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|F_i - O_i| \tag{1.5}$$

where O_i represents an observation from a station and F_i is the average forecast of 9 grid points around the station. Since all model integrations started at midnight Coordinated Universal Time (UTC), which is also referred to as Zulu (Z) time, note that forecast hours 0, 24, 48, and 72 represent 5pm Pacific Daylight Time (PDT). From the figure there is very little spin-up for the 10 meter wind; this is shown by the small and rapid adjustment of the error during the initial few hours of the simulation. Both the forecasted temperature and dew point take roughly 12 hours after the start of the simulation to minimize errors. This is in part due to the spin-up but also because of the diurnal cycle of MAEs which can be seen throughout the 72 hour forecast period. This demonstrates that a 24 hour spin-up period for our land surface model is sufficient for this summer month. While this figure shows the first 24 hours of all simulations, it will not be considered in any statistics reported in this study.

While the LSM and surface layer calculate the surface heat, momentum, and moisture fluxes, the PBL takes on an entirely different task. The PBLs role in the numerical model is to accomplish the sub-grid scale mixing of heat, momentum, and moisture that would otherwise not happen due to resolution constraints. How the mixing (more commonly known as eddy diffusion) is performed is very much dependent on the PBL used. The schemes tested in this paper included the Yonsei University (YSU) and the Mellor-Yamada-Janjic (MYJ) scheme, the latter of which is currently used in the operational NAM. The Quasi-Normal

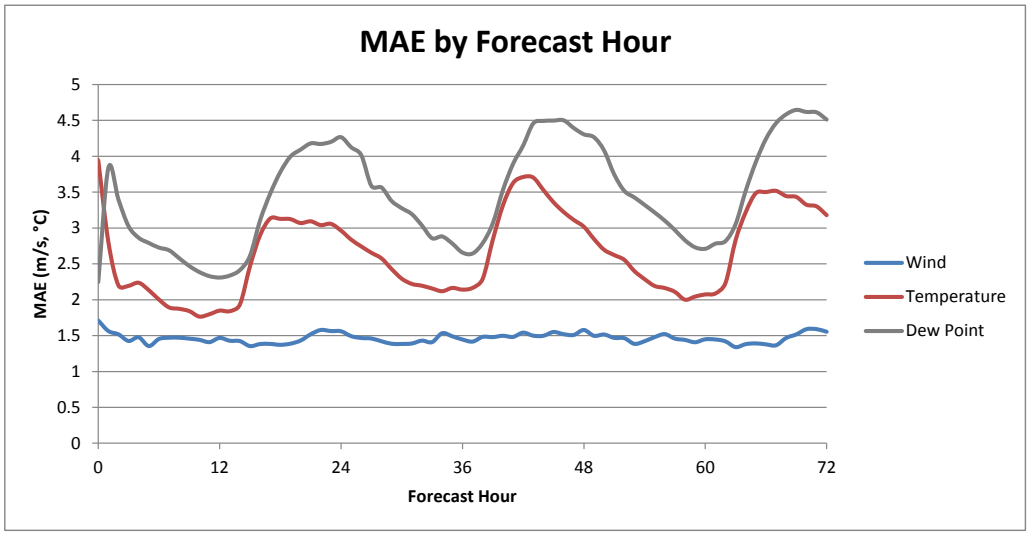

Figure 1.2: MAEs averaged over all ensemble members by forecast hour. The 10 meter wind along with the 1.5 meter temperature and dew point error are shown. Observations used in verification are provided by Automated Surface Observing System (ASOS) stations in domain 3.

Scale Elimination (QNSE), Mellor-Yamada Nakanishi and Niio (MYNN2) level 2.5, the Asymmetric Convective Model version 2 (ACM2), and the Bougeault-Lacarrere (BouLac) PBL scheme, all of which were added in 2008-09 were also examined. Lastly, we evaluated the newest PBL schemes that were added to WRF in 2011. These include the Total Energy - Mass Flux (TEMF) and the University of Washington (UW) — which uses physics from the Community Earth System Model (CESM) — PBL scheme. Along with our three land surface models, these eight different planetary boundary layer schemes create the twenty-four member ensemble.

1.2 July 2005

In the state of California, July 2005 was warmer than normal with temperatures +3.2 degrees Fahrenheit (°F) above average [7]. This was due to a strong ridge of high pressure aloft centered near the four corners region that regularly expanded into California. The desert locations of Southern California finished off the month +3.5°F above normal with Lancaster leading the group with a +4.6°F anomaly. In contrast, a dominate surface high pressure created a rather persistent coastal marine layer that allowed Downtown Los Angeles to finish the month off -1.2°F below normal with most coastal location reporting similar anomalies. As with typical summertime months, relatively little precipitation fell in Southern California during July. The only significant precipitation came from mountain thunderstorms across Los Angeles and Ventura counties where flash flooding was report on the 23rd and 24th [4].

A look at observed surface temperatures and dew points in domain 3 (Fig. 1.3) shows two pronounced warm periods on the 13th and 22nd of July. This figure was created by combining observations at stations across Southern California that will be introduced in the next chapter. Comparing this to the temperatures observed at the Los Angeles International Airport (KLAX) we see only one distinct heat wave in the latter half of the month. This is due to the previously described marine layer which kept coastal locations in a relatively stagnant weather pattern for the first 20 days of the month. Winds in domain 3 were no different; they were characterized by the typical sea breeze that would develop and strengthen throughout the day due to the large temperature gradients between the land and ocean. For this July month, averaged winds in domain 3 peaked around 4pm

local time which is consistent with the sea breeze.

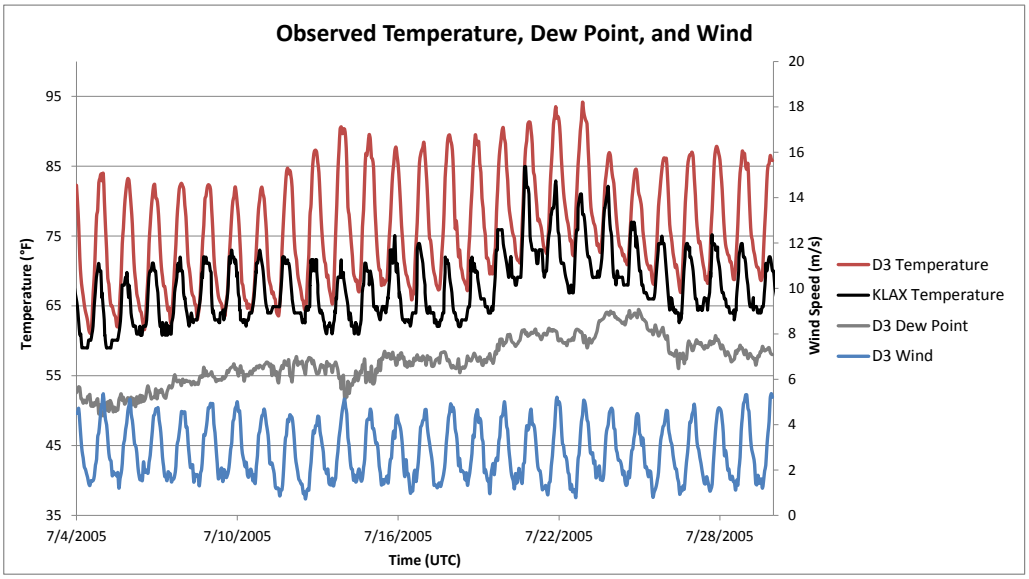

Figure 1.3: Observed temperatures and dew points in domain 3 during July 2005. Observations represent an average of all ASOS stations in domain 3.

CHAPTER 2

Surface Statistics

Station data used for the validation of surface model fields were made available through the Meteorological Assimilation Data Ingest System (MADIS) database. From this, we used data from the Automated Surface Observing System (ASOS) and the Automated Weather Observing System (AWOS) stations; their exact locations can be seen in Fig. 2.1. Since the data is primarily composed of ASOS stations we will be referring to both stations types as ASOS. It should be noted that a majority of these stations are located within the Los Angeles basin with very few in the higher elevations. For this reason, the results will mostly be representative of the lower, more densely populated elevations.

2.1 Ensemble Members

The average MAE of wind speed, temperature, and dew point for each ensemble member with respect to the ASOS stations can be seen in Fig. 2.2. Despite the unit inconsistencies, all three parameters are graphed on the same axis for comparison. Of the 24 physics combinations shown, the Pleim-Xiu land surface model coupled with the YSU PBL scheme had the most accurate representation of surface parameters. In fact, the YSU PBL was always the front runner in each of the three land surface models when wind speed, temperature, and dew

point were taken into consideration. Since the YSU PBL can accurately recreate conditions despite the land surface model used, it hints at the fact that it could be creating the most accurate representation of the lower atmosphere; however, this will be discussed more in depth in later chapters. In contrast to this, the recently added TEMF PBL always verified as the worst PBL with each land surface model.

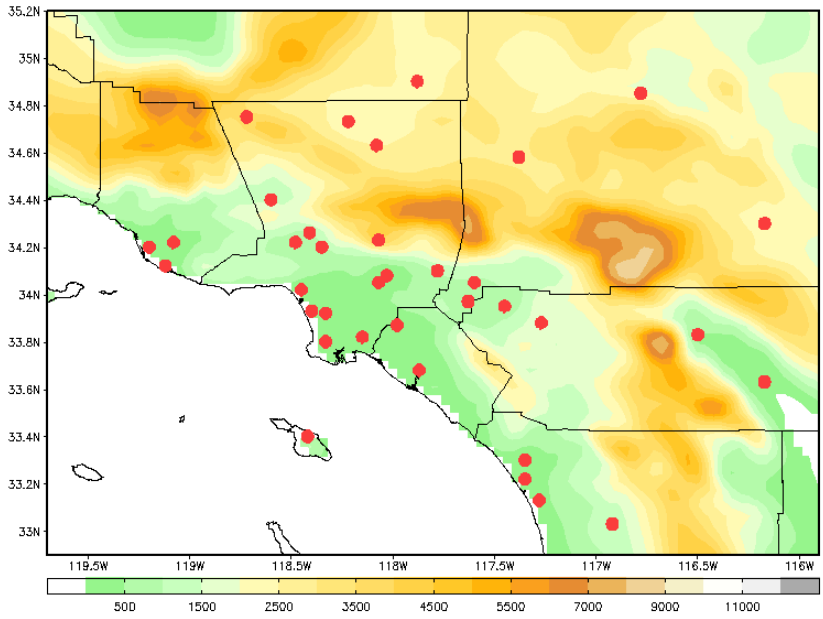

Figure 2.1: Red dots indicate the locations of the ASOS and AWOS stations used in the analysis. The shaded field represents actual resolved modeled terrain in feet above mean sea level.

Despite the least amount of soil layers, the PX LSM holds the top six spots (out of the eight PX ensemble members tested). But as stated earlier, July 2005 was a dry month for the lower elevations; this meant there was no additional soil moisture added to the LSM though precipitation. As one could hypothesize, different results — especially for the dew point — might be obtained from the LSMs during a wetter month.

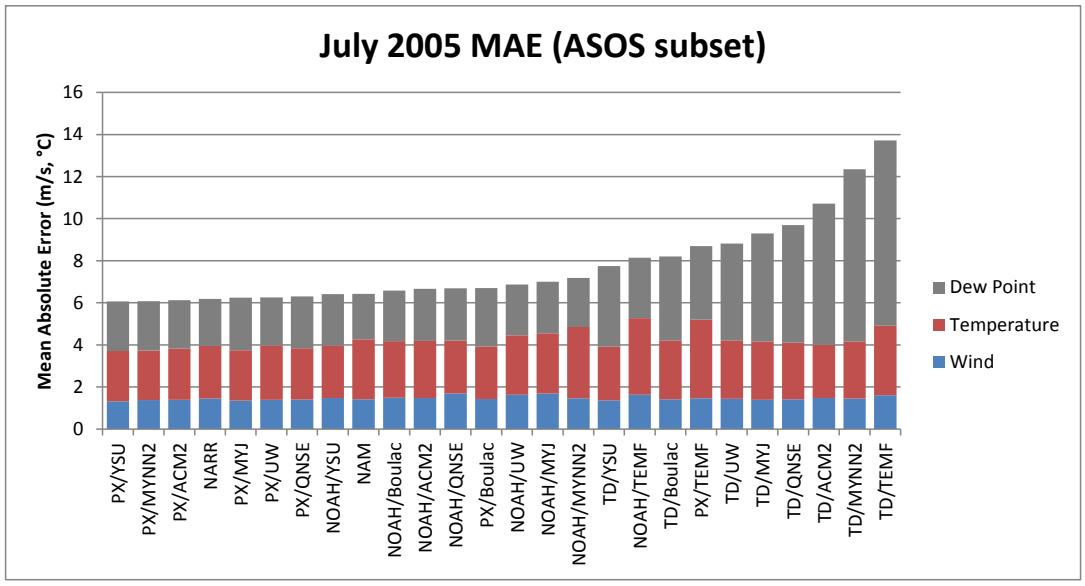

Figure 2.2: Mean Absolute Errors for wind, temperature, and dew point averaged over all ASOS stations. Ensemble members were verified hourly while the NAM and NARR are only verified every 3 hours.

Performing in the middle of the group, and not far behind the PX LSM, was the Noah LSM. Despite being a more popular choice amongst model users, the Noah LSM is clearly not as accurate. The TD LSM had the worst performance overall. If one looks solely at temperature and wind speed the TD LSM is quite competitive; the majority of the errors originate from the dew point. This is indicative of the fact that this five layer soil model only predicts temperature and not soil moisture. This is the main difference that separates the TD from the PX and Noah LSMs, and is illustrated well in these results.

Verification of the NAM analysis along with the North American Regional Reanalysis (NARR) was done for comparison in Fig. 2.2. The NARR dataset is a 3 hour regional reanalysis with a horizontal resolution of 32km. Using information from models and observations (radiosondes, dropsondes, aircraft, surface weather stations (neglecting 2m temperature), etc.) to reconstruct the atmosphere, the

NARR is also seen as a suitable dataset for initialization of a regional model. Both the NARR and NAM datasets verify particularly well even when considering their course resolutions of 32 and 40km, respectively [14].

Fig. 2.3 shows the mean dew point by forecast hour which further demonstrates the LSMs differences. As previously discussed, the TD LSM is not predicting soil moisture but is rather using values from a climatological table; this creates erroneously high dew points during the daytime. The large diurnal cycle seen in the TD LSM is a result of high afternoon dew points that are untimely reduced by the minimum temperature. The afternoon dew points seen in the PX and Noah LSMs are much lower so overnight minimums rarely reach the dew point, this creates only small diurnal variation.

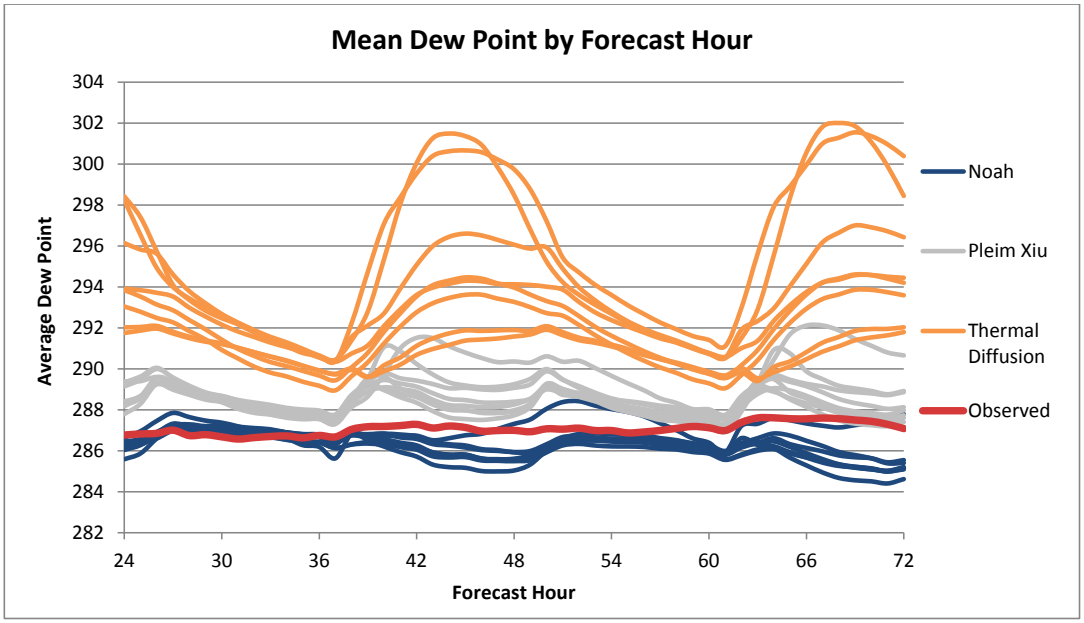

Figure 2.3: Dew points averaged over all runs for each ensemble member between the forecast hours of 24 and 72. Each LSM is shown individually with the 8 PBLs tested.

It can also be seen that the Noah LSM has little bias at hour 24 but then slowly trends towards a cool bias throughout the simulation length. On the other

hand, the Pleim-Xiu LSM has a warm dew point bias that trends to neutral. It is unclear whether this slow drift towards cooler dew points would continue if we extended the simulation beyond 72 hours. It should be noted that average dew points show large variations between LSMs whereas average temperatures and winds do not (not shown).

2.2 The Top Performers

From the simulations, the Yonsei University PBL had the most accurate recreation of surface conditions for each LSM when taking temperature, dew point, and wind speed into consideration. Because of this finding, this section will focus less on the ensemble as a whole and more on the top performers from each LSM. For the month of July 2005 the PX/YSU physics combination was overall the best, having a 1.32m/s wind speed MAE, a 2.39°C MAE for temperature, and a 2.35°C MAE for dew point. The best Noah ensemble member, the Noah/YSU, had a 1.48m/s wind speed MAE, a 2.49°C MAE for temperature, and a 2.44°C MAE for dew point. Lastly, the TD/YSU member had a 1.37m/s wind speed MAE, a 2.54°C MAE for temperature, and a 3.84°C MAE for dew point.

Fig. 2.4 shows the wind bias for the PX/YSU physics combination using only ASOS data. Here and throughout the rest of the paper the bias can be defined by

$$Bias = \frac{1}{n}\sum_{i=1}^{n}(F_i - O_i) \tag{2.1}$$

where O_i represents an observation at a station and F_i is the average forecast

of 9 grid points around the station. From the figure, cooler colors represent winds that were underforecasted for the month of July while warmer colors depict overforecasted winds. Apart from the few exceptions, we can see that the winds were slightly underforecasted. The high wind bias seen at both San Clemente and Catalina Island is probably due to the sub-grid scale turbulence that could not be resolved at the 4km resolution. Catalina Island is only slightly visible in the 4km model resolved terrain while San Clemente does not appear at all.

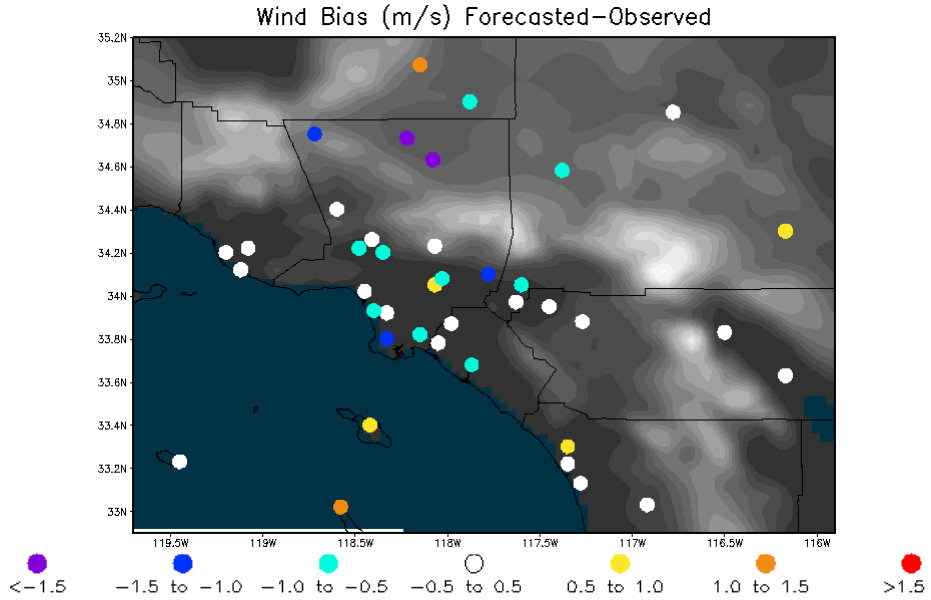

Figure 2.4: Wind bias computed for domain 3 using only the ASOS stations for the PX/YSU physics combination. Each dot represents the monthly averaged bias for July 2005.

When one considers using only Remote Automated Weather Stations (RAWS) in Southern California, something interesting happens; the underforecasted winds now become overforecasted (Fig. 2.5). An explanation for this could be that RAWS stations are generally not as well sited as the ASOS stations which are usually located at airports. Obstruction upwind from a poorly sited RAWS station could be decreasing the observed wind creating a high forecasted wind bias.

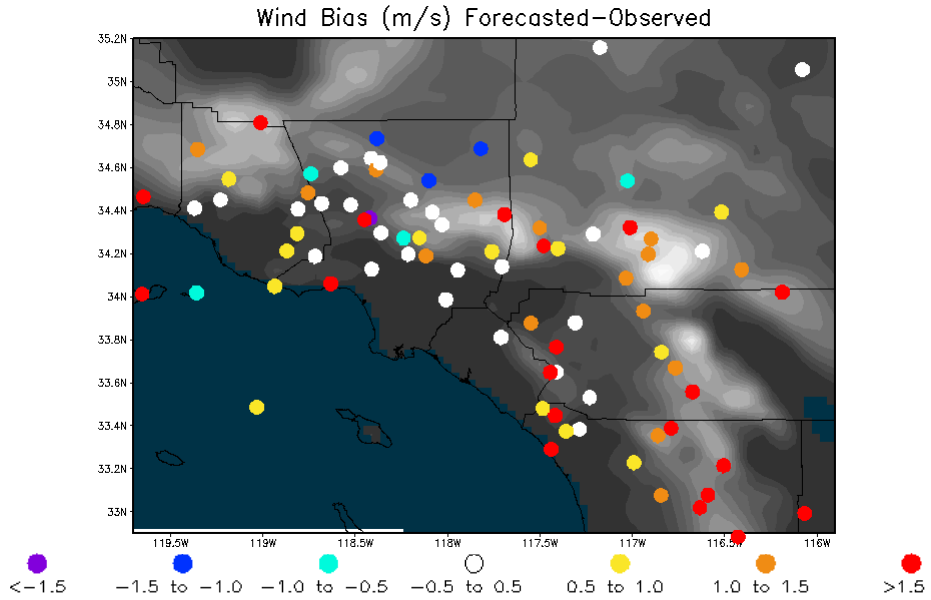

Figure 2.5: Wind bias computed for domain 3 using only RAWS stations for the PX/YSU physics combination. Each dot represents the monthly averaged bias for July 2005.

However, we acknowledge the fact that the RAWS stations are generally located in more remote locations where unresolved subgrid-scale terrain is not accounted for [12]. This unresolved terrain can create a sink for momentum that is not parameterized in this version of WRF. On the other hand, some RAWS stations are located in unresolved valleys and gaps that could channel the flow and increase the wind speeds; this would essentially decrease the high wind bias. As of now, we do not have enough ASOS stations in rough terrain to test these hypotheses, however, there are a few RAWS stations located in relatively smooth terrain (along the south coast in Fig. 2.5). Of these stations, one can see that a large number of them do have a high wind bias whereas the closely located ASOS stations have little to no bias.

In addition to this, the RAWS temperature sensor can be mounted anywhere between 1.2-2.4m above the ground while the ASOS stations permit less variance

in sensor height (1.5-2m). During the day and night time hours, temperatures can vary quite dramatically over the first few meters above the ground. Since there is such a large spread in sensor height at RAWS stations, data are often erroneous and sometimes useless. In the rest of this chapter, we will continue to look at RAWS stations but only when temperatures are averaged amongst a large number of stations.

The temperature bias for the PX/YSU physics combination is dependent on the region (Fig. 2.6) rather than the station type. Warm colors represent an overprediction of temperature in July 2005 while cool colors are an underprediction. Both RAWS and ASOS stations were combined in this figure because it was not clear whether there was any specific bias with respect to station type when looking at average temperature. Most coastal locations along with the Inland Empire show WRF overforecasting temperatures. A cool bias begins at the San Gabriel and San Bernardino Mountains that further amplifies as one moves into the Mojave and Colorado Deserts. While some of the bias can be correlated to differences in modeled and resolved terrain height, it is unlikely to cause significant errors and would only influence areas of rough terrain. As of now, the magnitude of this large regional bias remains unclear. It should be noted that the regional bias is independent of the land surface model used as very similar patterns are present in both the NOAH and TD land surface models (not shown).

Lastly, a look at the dew point bias for the PX/YSU physics combination reveals a warm bias (Fig. 2.7) throughout all regions. Just as in the previous figure, ASOS and RAWS stations are combined because it was not apparent if there was any specific bias with respect to station type. The majority of places have a warm dew point bias which is consistent with Fig. 2.3. Unlike surface

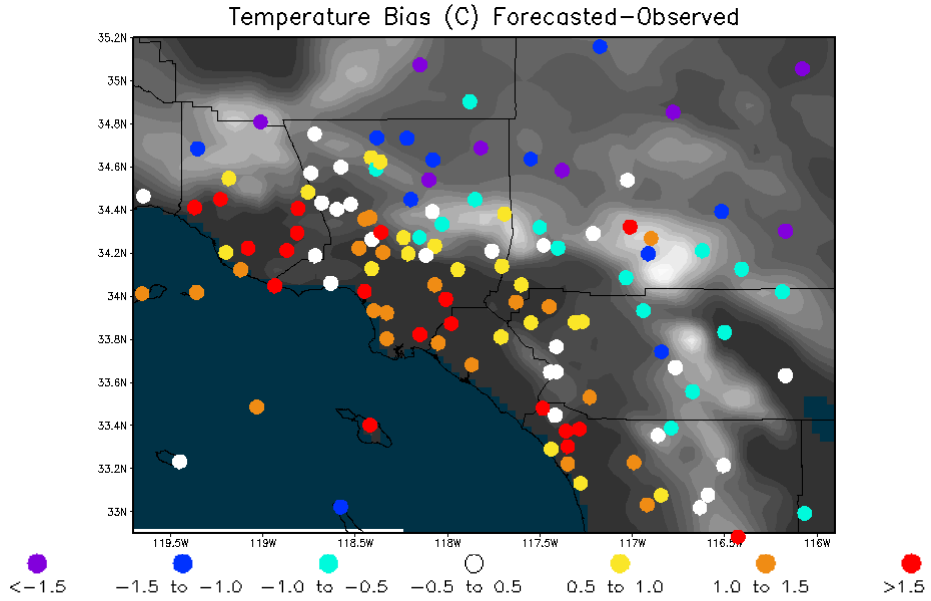

Figure 2.6: Temperature bias computed for domain 3 using both RAWS and ASOS stations for the PX/YSU physics combination. Each dot represents the monthly averaged bias for July 2005.

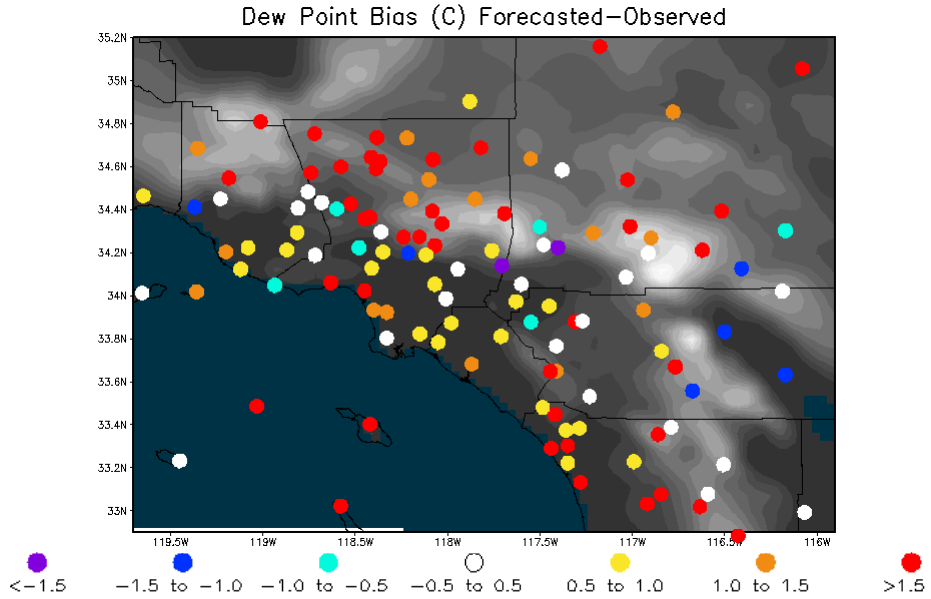

Figure 2.7: Dew point bias computed for domain 3 using both RAWS and ASOS stations for the PX/YSU physics combination. Each dot represents the monthly averaged bias for July 2005.

temperature, our dew point bias is dependent on the land surface model. The TD scheme has an extremely warm dew point bias at almost every location while the Noah scheme has a cool dew point bias, neither of which are shown.

While a 2.47°C average MAE in temperature was found for the top PBLs of each LSM (PX/YSU, Noah/YSU, TD/YSU), July 2005 had periods of very large errors. Fig. 2.8 shows the domain observed and forecasted temperatures averaged over these three ensemble members. From the beginning of the month until the 10th, MAEs along with the forecast bias were low. This period was dominated by the typical summertime onshore flow which can be seen by the KDAG-KLAX (Daggett to Los Angeles International Airport) Mean Sea Level Pressure (MSLP) gradient. Following this, the onshore flow weakens and MAEs climbed to a staggering 6.68°C on the 17th. A similar pattern to this erroneous mid-month temperature bias is found at the end of July as well, which also coincides with a weak onshore flow. Forecasting temperatures in weak pressure gradients is often difficult at coastal locations due to the fact that the cool sea breeze could be replaced by a much warmer and drier off-shore breeze. As we seen from Fig. 2.6, much of this temperature bias originates at coastal locations.

As a means of addressing this mid-month spike in MAEs, an analysis of the temperature bias and error was done for RAWS and ASOS stations at elevations less than 1000ft. Apart from a few exceptions, most ensemble members show relatively the same skill in forecasting temperature (not shown). From the top 16 ensemble members, MAEs ranged from 2.62°C to 3.03°C with biases from +0.81°C to +2.22°C. The TD/QNSE combination offered the smallest bias for temperatures while maintaining a low mean absolute error (2.69°C) for stations less than 1000ft. It is worth noting that the TD/QNSE scheme is overall one of

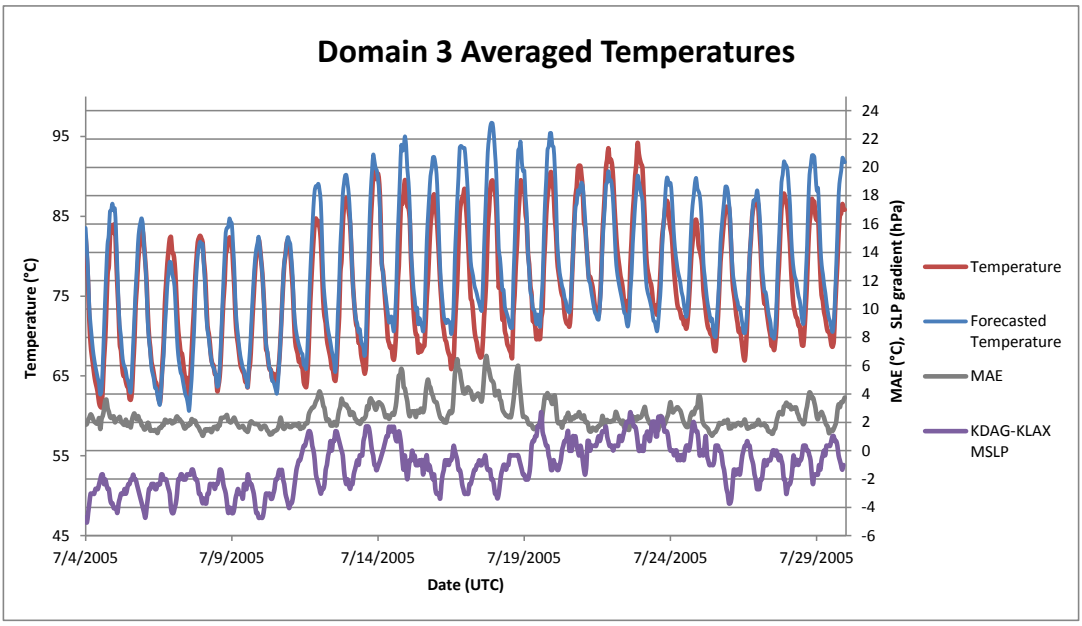

Figure 2.8: Forecasted and observed temperatures in domain 3 average over our top PBLs from each LSM (PX/YSU, Noah/YSU, TD/YSU) using ASOS stations. The KDAG-KLAX MSLP gradient was also derived from the ASOS observation.

the worst physics combinations (Fig. 2.2) because of its large dew point errors. The PX/YSU, Noah/YSU, and TD/YSU combinations were all within a 0.25°C MAE range of the TD/QNSE; however, they average a warm bias of 1.86°C which is more than double the value from the TD/QNSE.

The temperature bias by station for the TD/QNSE ensemble member is shown in Fig. 2.9a from the hours of 20UTC the 17th until 00UTC the 18th (the day of our highest MAEs). Fig. 2.9b-d shows our top performers from each land surface model. While the TD/QNSE scheme has the smallest bias for low elevation stations, it was at the expense of making inland locations even colder. The other members (PX/YSU, Noah/YSU, and TD/YSU), all of which use the YSU PBL, have little bias in the mountains and inland deserts on this day but have an exceptionally warm bias towards the coast. Since the problem does persist

across multiple ensemble members with different LSM/PBL combinations, one can hypothesize that a fix lies somewhere other than model physics; however, this will be discussed in depth later.

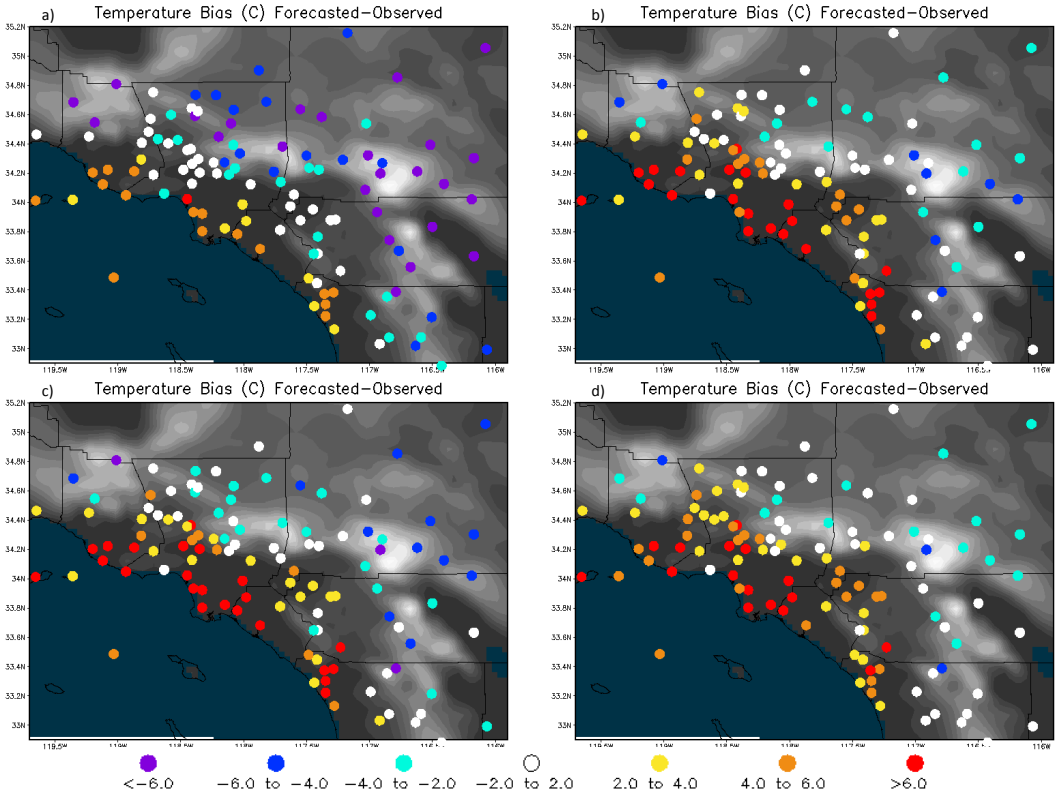

Figure 2.9: The forecasted temperature bias averaged from 20UTC July 17th to 00UTC the 18th. The TD/QNSE, PX/YSU, TD/YSU, and Noah/YSU ensemble members are shown in figure a,b,c, and d, respectively.

The TD/QNSE is not necessarily a cure-all, although, it does provide a more realistic temperature forecast for the densely populated regions of Southern California (Fig. 2.10). Overall, the afternoon bias is quite small despite the few days around the 22nd where it largely underestimated temperatures. This is compared to the PX/YSU which grossly overestimates temperatures for the majority of the month. Notice that neither combination can accurately forecast the overnight

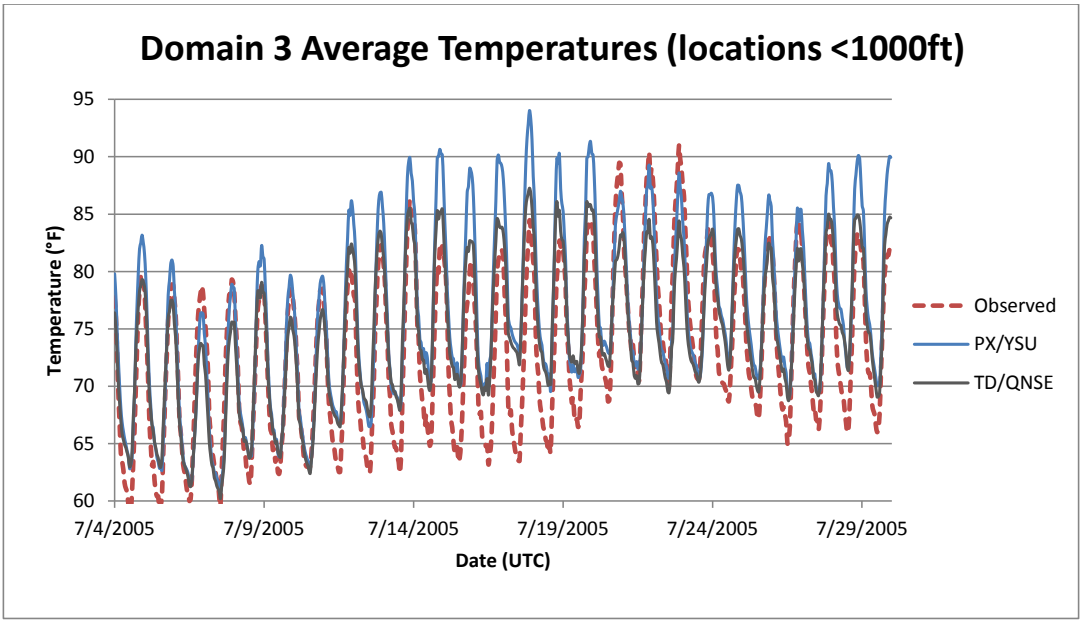

Figure 2.10: Observed and forecasted temperatures average over stations less than 1000 feet using both RAWS and ASOS data.

temperatures, especially towards the middle of the month.

Up to this point we have been using RAWS stations to discuss the temperature and dew point bias for the ensemble members. Despite this, we cannot depend on them to give us an accurate measure of the MAEs due to reasons earlier discussed. From this point forward we will no longer be using RAWS stations in any diagram or statistic.

CHAPTER 3

The Vertical Profile

The vertical profile of temperature and wind is quite influential on the surface parameters tested in the previous chapter, although, the converse is also true. This is because the temperature aloft controls how warm the 2m temperature can get due to the limit of the dry adiabatic lapse rate. Of course, the temperatures aloft are dependent on those at the surface. In addition to this, how temperature changes with height directly affects atmospheric stability which will not only affect rainfall, but can control how turbulent the atmosphere will be. The vertical wind profile is also very important and is similar in the sense that is also controls atmospheric turbulence. From this, one can see that the conditions aloft are closely linked to those at the surface.

In numerical models, the structure of the vertical profile is controlled largely by the planetary boundary layer and in some cases, even more so by the land surface model. Since the land surface model controls the surface heat flux, it strongly influences how warm the lower atmosphere will get. Temperature biases found in the lower atmosphere may be a direct result from the land surface model or from other model deficiencies. Biases could also arise from the surface layer and/or boundary layer, a result of too little or too much mixing. The vertical profile of wind is only slightly different because the momentum fluxes

are controlled by the surface layer and not the LSM. Since the conditions aloft are directly linked to those of the surface, one could hypothesize that a certain physics combination that correctly captures surface parameters probably does a respectable job at forecasting those aloft.

3.1 ACARS Data

While the last chapter focused solely on the surface verification of the ensemble, we will now turn the attention to the accuracy of the vertical profile by using data from the Aircraft Meteorological Data Relay (AMDAR). The United States contribution to AMDAR is called the Meteorological Data Collection and Reporting System (MDCRS) which is most often referred to as Aircraft Communications Addressing and Reporting System (ACARS). Because of these reasons, we will refer to all aircraft data as ACARS. The data is comprised of seven commercial airlines (American, Delta, Federal Express, Northwest, Southwest, United, and United Parcel Service) that take automated weather reports throughout their flights [8]. As one could imagine, these weather observations taken from aircraft are biased with respect to the time of day and geographic location; however, for Southern California, it is the best option to validate the vertical structure of the model due to the abundance of observations.

The geographic location used to validate the vertical profile of the model is outlined by the white polygon in Fig. 3.1. The polygon, which will be referred to as the coastal subset from this point forward, is limited to a geographic area where the model terrain does not exceed 500ft (roughly 990hPa). This is done because the standard pressure coordinate output is not easily converted into

constant height coordinates. By limiting the polygon to areas less than 500ft, a given pressure level would be approximately the same height above the surface throughout the polygon. At the bottom of our vertical profile we will again be using ASOS data for validation, this will include roughly 15 surface stations that are continuously operating (refer to Fig. 2.1 for actual locations).

Approximately 49,000 ACARS observations comprised of wind and temperature will be used to validate the month; roughly 50% of this data is below 900hPa. The spatial distribution of these observations can been seen on Fig. 3.1. Since wind and temperature observations are almost always taken simultaneously, one dot on the figure consists of both a temperature and wind observation. Upon further inspection one can see three prominent areas where aircraft take-off and land; these areas can been seen by the warmer colors. The northernmost location is the Los Angeles International Airport, followed by the Long Beach and John Wayne Airports. Due to the large number of observations around these airports, the statistics from the vertical profile will be representative of those areas.

The number of temperature observations displayed by the hour of day for both ACARS and ASOS data is shown in Fig. 3.2. It can be seen that the 950-925hPa layer has the most observations while the highest level, 750-700hPa, has the fewest. This can be explained by looking at the observations in the previous figure; as a aircraft takes off from the airport it quickly moves out of our domain limiting the observations at the higher levels. Fig. 3.2 also depicts the temporal variation of the ACARS data showing relatively few observations taken between 9-11 UTC. Conversely, the automated ASOS stations have little variation throughout day and, of course, are fixed in space. As mentioned earlier, wind observations are almost identical to those of temperature; for this reason,

one can also interpret Fig. 3.2 as the wind observations by the hour of day.

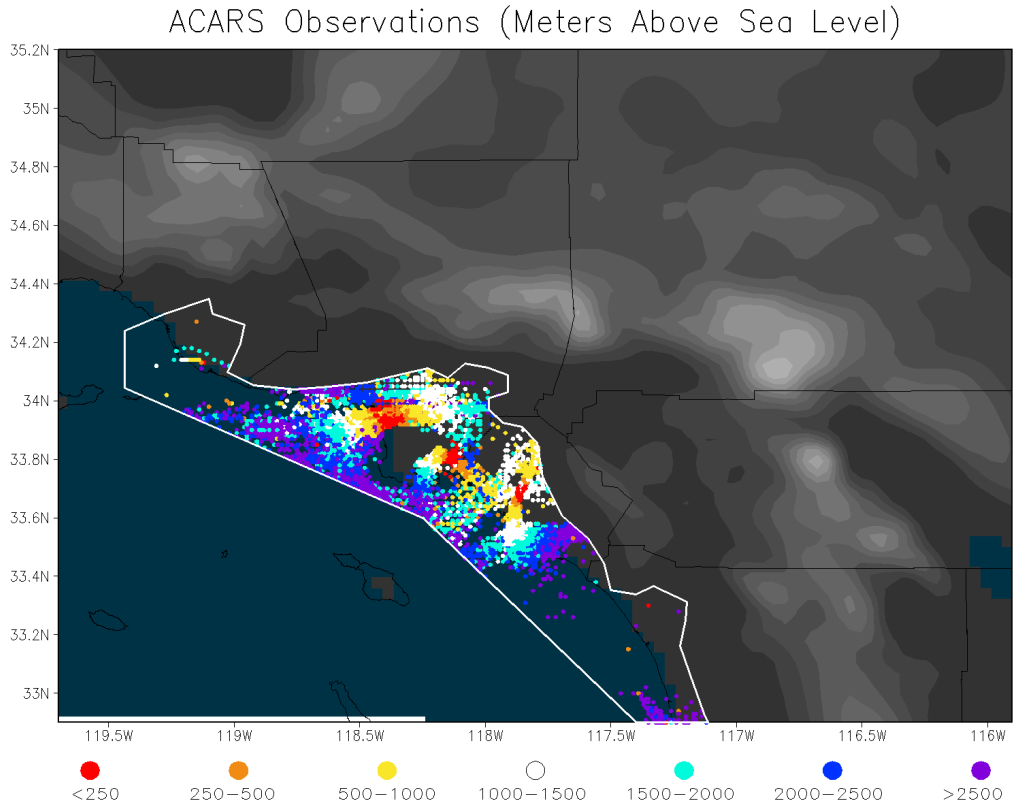

Figure 3.1: The area used for the ACARS and ASOS data is outlined by the white polygon and will be referred to as the coastal subset. Each dot represents an ACARS observation at a certain height above sea level (m) and is limited to 3150m (roughly 700hPa) above sea level. The ASOS data which is not shown, includes roughly 16 wind and 14 temperature observation stations, respectively.

While the ACARS data are numerous, the accuracy of the observations has been debated in journals. In 1995, Schwartz and Benjamin compared aircraft wind and temperature profiles with radiosondes and found a standard deviation temperature difference of 0.59°C [20]. Their findings also revealed that vector wind speeds did not do as well, reporting a Root-Mean-Squared (RMS) error of 4.0 m/s. However, four years later the same authors (joined by Cole) revised their estimates of wind vector RMS error to 1.8 m/s and temperature error to 0.5

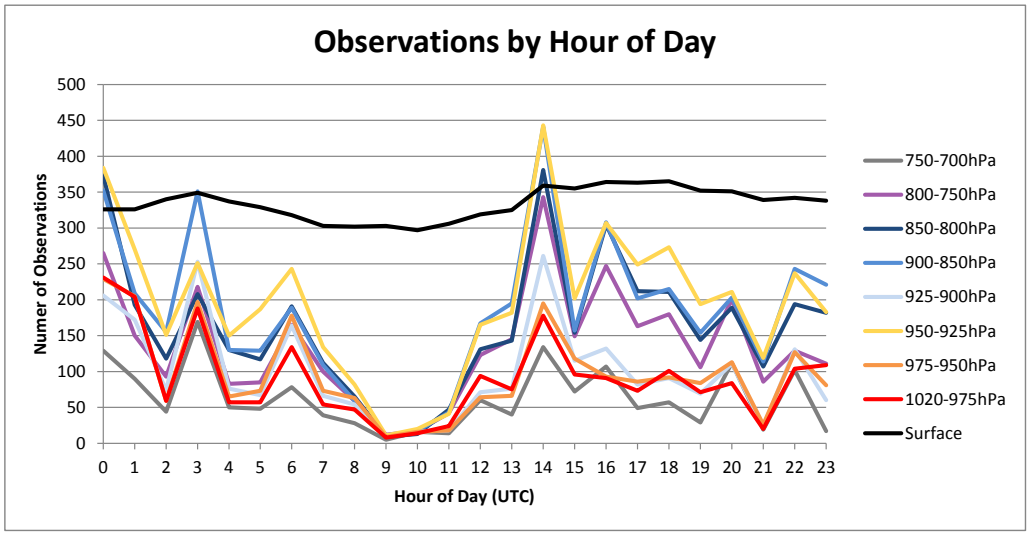

Figure 3.2: Temperature observations from the ACARS and ASOS data by the hour of day. The ASOS data only pertains to the surface observations only.

degree Celsius [3]. Since then, no new articles have emerged about the validity of the ACARS data.

3.2 Ensemble Members

In this section, we will be inspecting the vertical profile from all ensemble members. It should be noted that statistic were averaged by the hour of day to eliminate the influence from the high number of daytime observations. This means that nighttime statistics are composed of fewer observations but are equally weighted with respect to the daytime statistics.

Using the coastal subset, the MAEs from July 2005 can be seen in Fig. 3.3 for various pressure levels. For comparison purposes, the NAM and NARR dataset were included in the figure and rank in the top two positions. While these datasets are somewhat coarse compared to the 4km ensemble members, they are extremely

accurate. When looking at the NARR and NAM datasets together, they seem to perform equally well; however, on closer inspection one can see that the NAM has large errors below 975hPa and relatively small errors above 950hPa. This is because large surface temperature gradients that exist on the coast during the summertime cannot be properly resolved with the coarse resolution of the NAM. The NARRs horizontal resolution is only slight better (32km) yet has a much better handle on this large surface temperature gradient. It is worth noting that the NARR and NAM were verified every three hours whereas the verification of ensemble members was every hour.

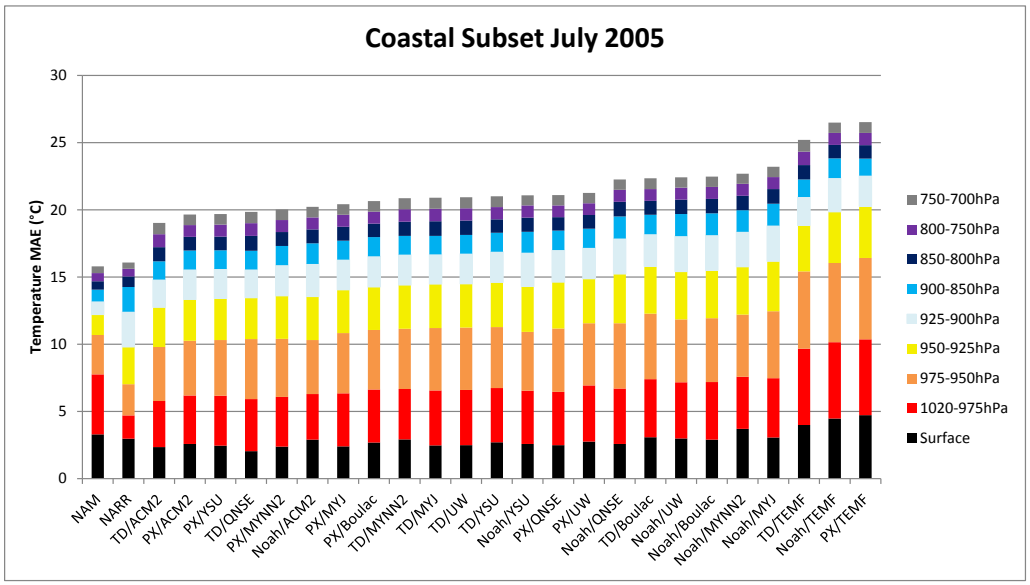

Figure 3.3: MAEs for temperatures from the ACARS and ASOS data by ensemble member. The ASOS data is confined to the surface only. Ensemble members were verified hourly while the NAM and NARR were only verified every 3 hours.

The ensemble members that are shown in Fig. 3.3 did not perform as well as their initialization dataset (NAM). This can be expected because as one ascends in the atmosphere, the large temperature gradients that exist at the surface will vanish. So while a 40km dataset might not be able to accurately recreate the large

temperature gradients at the surface, it is more than suitable for conditions aloft. Overall, the TD/ACM2 has the most accurate representation for atmospheric temperatures, however, we have seen from previous results that the TD LSM has trouble handling surface moisture. While we cannot verify atmospheric moisture, one could hypothesize that improper prediction of surface moisture could have a rather large impact aloft.

The PX/ACM2 and PX/YSU members rank two and three for all ensemble members. From previous results we have found that these members were the most accurate for surface conditions when considering wind, dew point, and temperature (Fig. 2.2). In contrast, the TEMF boundary layer has the largest errors independent of the LSM chosen; and again, from our previous results (Fig. 2.2) the TEMF PBL is the worst for each respective LSM. Surprisingly, one of the worst physics combinations one could choose for this month is actually one of the most popular, the Noah/MYJ. This physics combination is used in multiple NCAR simulations including the spring real-time convection forecast, the daily real-time runs, and the the Antarctic Mesoscale Prediction System as noted in the WRF Users Guide [2].

In Fig. 3.4, a graph similar to that of temperature is shown for wind speed. This time the ensemble spread is quite small and there are two probable causes. Firstly, the ACARS data has a 1.8m/s RMS when considering each component of the wind vector, this would essentially create less variation between members. Secondly, the pressure gradients in the summer months are relatively small, producing weak winds and thus, small MAEs. As seen from Fig. 3.3, the overall MAE variance between members is mostly controlled by levels below 950hPa, above this, different LSMs and PBLs have little variation in MAEs. This makes

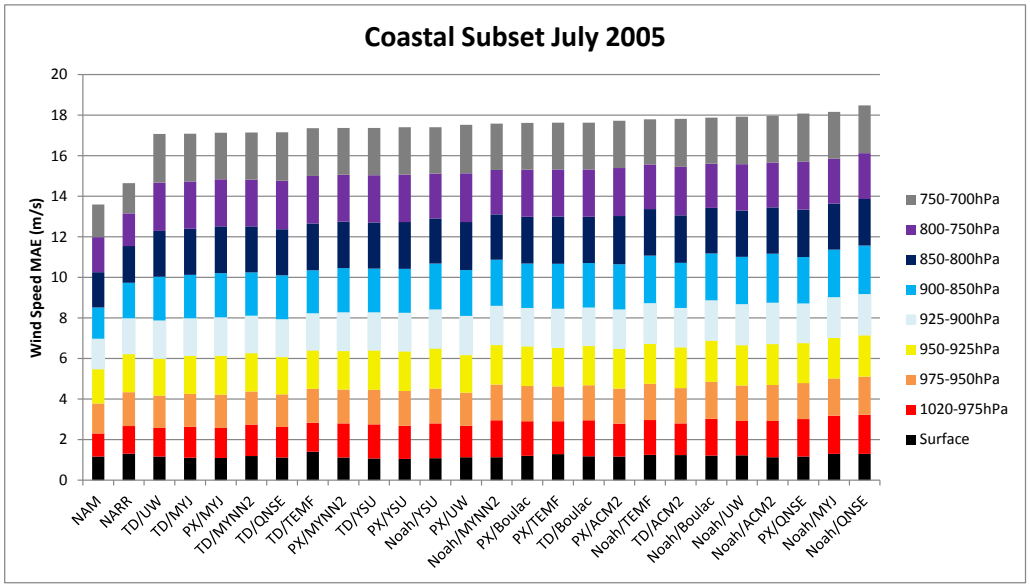

Figure 3.4: MAEs for wind speed from the ACARS and ASOS data by ensemble member. The ASOS data is confined to the surface only. Ensemble members were verified hourly while the NAM and NARR were only verified every 3 hours.

sense because one would expect all members to convergence on a solution aloft since we only vary the PBL and LSM.

3.3 The Pleim-Xiu/Yonsei University

As a means of addressing errors and biases in the model we will now begin to focus our attention on a single ensemble member. The PX/YSU was found to be the best member for surface variables and the third best for the vertical profile of temperature. Since this member has proven itself accurate in several areas it will be chosen for further investigation.

The MAEs by forecast hour for the PX/YSU can be seen in Fig. 3.5. The forecast hours between 0-24 are considered a spin-up period and are not used in the statistics, however, they are shown in order to understand how the errors

grow with time. At the surface one can see that the model initializes with erroneous temperatures but these are mitigated as the model integrates forward. From Fig. 3.6, one can see that the bias is almost identical to the MAE at the surface, meaning that a warm bias at all coastal locations accounts for the errors. As shown here, the warm bias is never fixed by the PX/YSU throughout the simulation; this is congruent with all ensemble members (not shown). When one does not consider the spin-up period the PX/YSU is ranked 6th in surface temperature bias (+1.95°C). This is comparable to the TD/ACM2 which was the last bias member (+1.35°C).

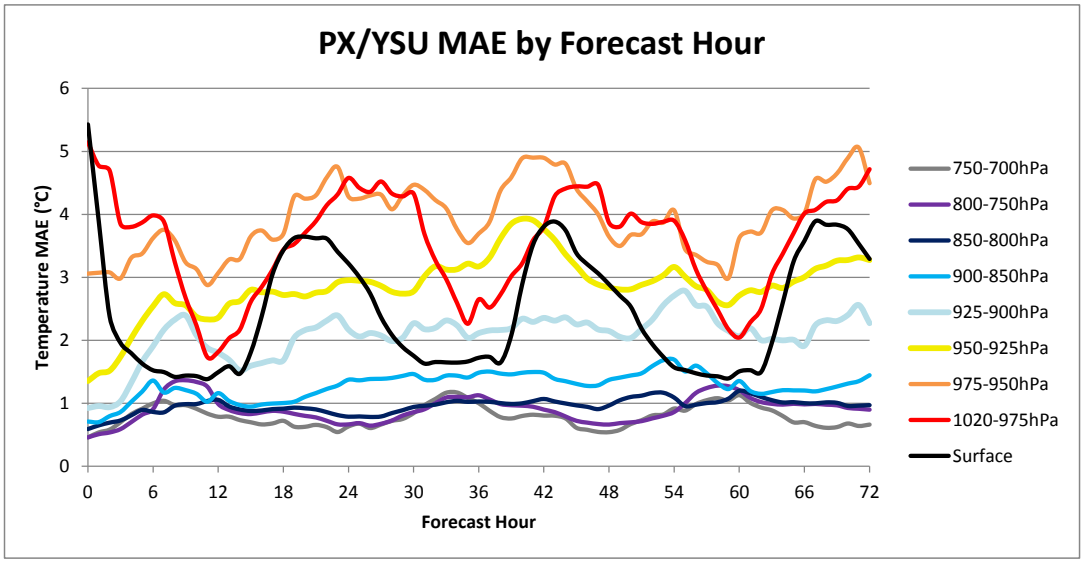

Figure 3.5: MAEs by forecast hour is shown for the PX/YSU. The ACARS data (1020hPa and up) is averaged ±2 hours around the forecast hour whereas the ASOS data (surface) is not.

Levels slightly higher in the atmosphere (950-900hPa) show MAEs that are quite low at initialization but grow by roughly a factor of two by the time we reach the end of our 24 hour spin-up period. It should be noted that among all the vertical levels, the 950-900hPa range experiences the highest error growth

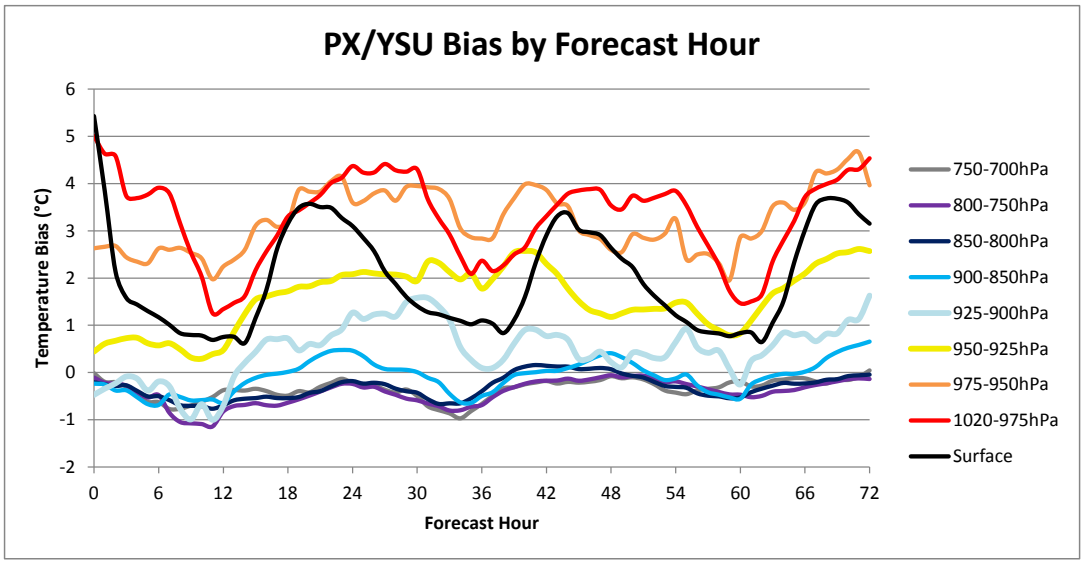

Figure 3.6: Bias by forecast hour is shown for the PX/YSU. The ACARS data (1020hPa and up) is averaged ±2 hours around the forecast hour whereas the ASOS data (surface) is not.

rate between hours 0-24. This is in part due to the warm bias at the surface mixing upwards which is depicted in Fig. 3.6.

At 900hPa we reach the top of the average summertime inversion (shown later). Above this is the free atmosphere where our simulations have only small error growths with time. Since these levels are separated from the surface by a strongly stratified atmosphere, the erroneous surface temperatures have little impact on temperatures aloft. This can also be verified by looking at biases in Fig. 3.6; the warm bias at the surface has difficulty penetrating above 900hPa.

CHAPTER 4

Initialization and Integration Methods

Among the 24 physics ensemble members tested, the PX/YSU and PX/ACM2 have proven themselves as superior physics parameterizations for July 2005. However, the temperatures simulated were initialized with a warm bias at the surface which could not be mitigated despite a large number of physics options tested. In this chapter we will continue to explore ways to improve our ensemble without modifying model physics.

4.1 Initialization

Looking at the NAM and NARR datasets in Fig. 4.1 we see substantial differences in afternoon temperatures. When using this sort of comparison, one can now get a feel for how grossly overestimated temperatures are for the NAM dataset and the PX/YSU. The warm bias for the NAM peaks on the 15th of July at +9.05°C (21Z). In contrast, the NARR handles afternoon temperatures quite well, the only problem coming from the latter half of the month (21st-23rd) where it fails to capture the coastal heat wave. Overnight temperatures are handled quite well in both datasets with the NARR being slightly cooler than the NAM.

From earlier statistics (Fig. 3.3) one can see that the NAM is a better dataset

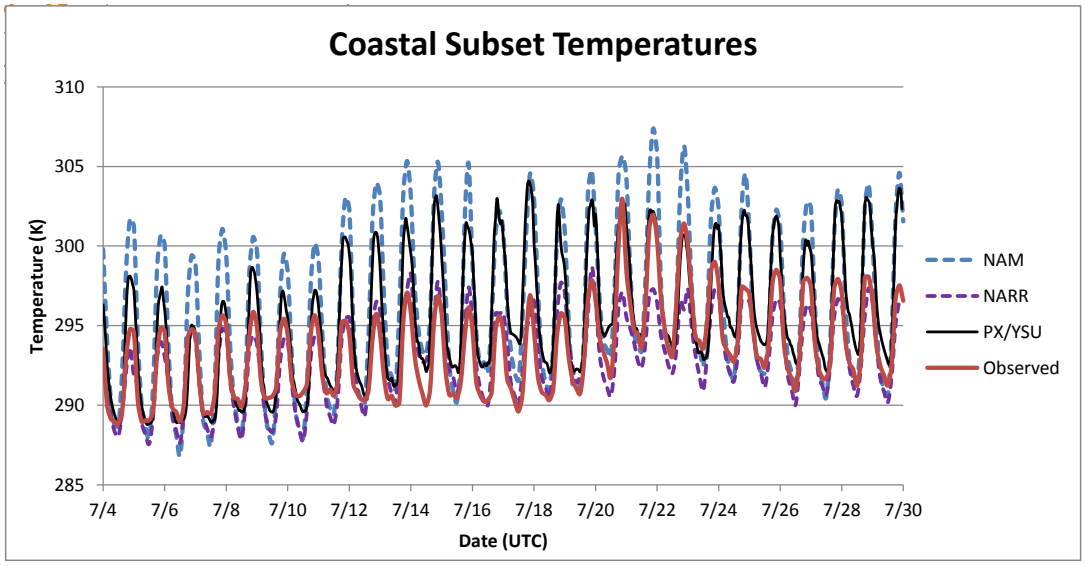

Figure 4.1: Monthly temperatures for the coastal subset using ASOS data.

when considering the vertically integrated temperature MAEs. While this statistic includes how the NAM performs throughout all hours of the day, Fig. 4.2 shows the 00Z and 12Z vertical temperature profiles along with the MAEs. The already apparent warm bias during initialization (00Z) does not exist in the NARR, however, at the surface, it does have comparable MAE errors (4.65°C) to that of the NAM(5.38°C). We see that as one ascends in the atmosphere the NAM fails to capture the base of the inversion — only developing an isothermal layer — whereas the NARR does an acceptable job. Slightly higher around 900hPa, both datasets fail to recreate the strength of the inversion resulting in a cold bias.

When compared with the 00Z NAM, the 12Z NAM initialization has much smaller errors in the boundary layer. It was hypothesized and tested that initializing an ensemble member off of the 12Z NAM could perhaps improve the simulation due to the relatively small vertically integrated MAEs. This however was not the case. To get an idea for vertically integrated MAEs one should refer

back to Fig. 3.3 where errors from the PX/YSU summed to 19.67°C. When the PX/YSU was initialized off of the NAM at 12Z, vertically integrated errors in the coastal subset only decreased by 0.11°C to 19.56°C. Even though the warm biases below 950 actually did better, decreasing from 2.90°C to 2.44°C (24-hour average over the coastal subset), initializing at different times does not make a substantial difference.

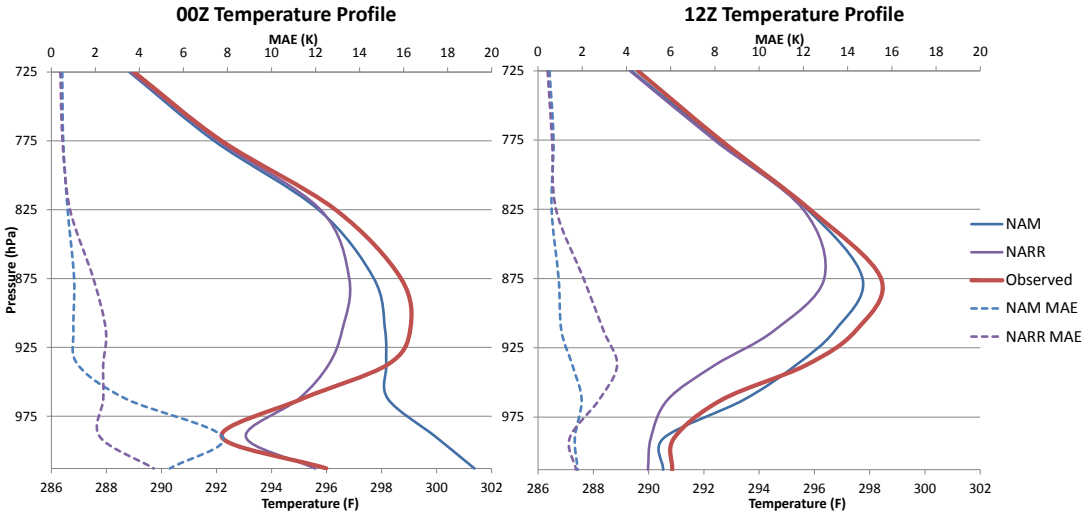

Figure 4.2: Vertical temperature profiles shown for the NAM and NARR datasets averaged over the coastal subset. Datasets are compared to ASOS (bottom data point only) and ACARS data.

An explanation for this relatively small change in vertically integrated MAEs despite a much better temperature initialization could be as follows. The parent model used to help create the NAM dataset will behave similarly to that of the regional model WRF. Even though the dynamical core and all physics parameterizations between the two models are different, they are in a sense similar because they are designed to handle all atmospheric phenomenons as realistically as possible. This means that modeling the same atmospheric event with two entirely different models will give you a similar answer most of the time. This explains

that when initializing off the less erroneous 12Z NAM we expect results to look similar to what the NAM would produce 12 hours later at 00Z. Also, one must not overlook that fact that our lateral boundaries are always forced by the parent dataset, thus further influencing the regional model.

If the regional model often behaves similarly to that of the parent model, one would expect different results when initializing the ensemble off of the NARR dataset despite the fact that it is a reanalysis product. In order to test this hypothesis we created a new ensemble member that was initialized off of the NARR dataset. Just as with the past simulations, the lateral boundary had inputs every three hours — this time from the NARR — while integrating the model for only three days at a time. Because of the previous findings, this new ensemble member uses the PX land surface model along with the YSU planetary boundary layer and will be referred to as PX/YSU NARR.

PX/YSU NARR dramatically improves the surface conditions while slightly increasing the free atmosphere errors. Fig. 4.3 shows the vertically integrated MAEs of the PX/YSU NARR being equal to the NARR and NAM datasets. One big difference that should not be overlooked is the pressure levels where the errors accumulate. At the lower levels (higher pressures), the PX/YSU NARR has small errors relative it to its counterpart, the PX/YSU. As one ascends in the atmosphere the errors actually grow slightly larger for the PX/YSU NARR compared to the PX/YSU. This is due to the fact that the PX/YSU parent model (NAM) has smaller errors aloft when compared to the PX/YSU NARR parent model (NARR). Therefore, the PX/YSU should have slightly smaller errors above the inversion layer because the initialization was more accurate.

The inversion height from the coastal subset is further exemplified in Fig. 4.4.

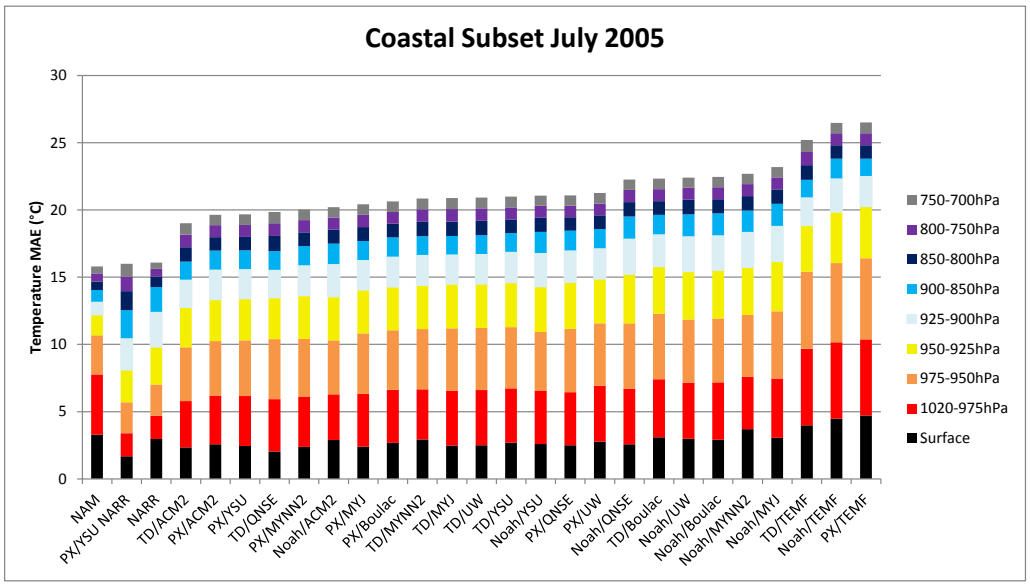

Figure 4.3: A replica of Fig. 3.3 with the newly added ensemble member.

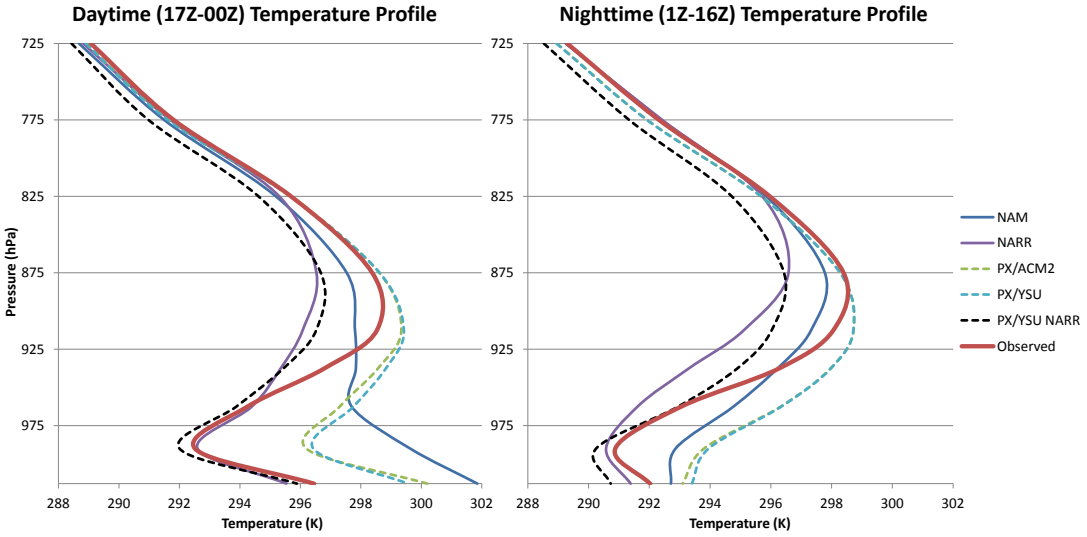

Figure 4.4: The vertical temperature profile in the coastal subset for both day and night. Result are compared to observations from ASOS (bottom data point only) and ACARS data.

The figure, which uses ASOS data at the surface and ACARS aloft, compares the vertical profile from the two datasets along with a few ensemble members to observations. Looking first at the daytime profile, which is from 17Z-00Z (10am-5pm PDT), one can see that the NAM's surface warm bias extends up to 925hPa where it is succumbed by the summertime inversion. The NARR dataset in relatively unbiased below 950hPa; beyond this, it fails to capture the strength of the inversion and develops a cold bias by the time it reaches the top of the inversion. These results are very similar to Fig. 4.2. Looking at the three ensemble members shown, it is interesting to note that they all have very similar profile shapes despite being initialized off of different datasets. One can see that the members initialized off the warmer dataset seem to be shifted towards warmer temperature. The PX/ACM2 was included to show how little variation there is between the different boundary layers whereas the same physics (PX/YSU vs PX/YSU NARR) with different initializations produces dramatic differences. Though not shown, the PX LSM coupled with a MYJ PBL produces results very similar to the PX/YSU and PX/ACM2. As mentioned earlier, the profile shapes can sometimes be more influenced by the land surface model than by different boundary layer schemes, while it is not completely shown in the figure, this was the case for this particular time period.

The nighttime profile has similar characteristics to that of the daytime with the most notable difference coming from a stronger inversion. Once again, the ensemble members show very similar profile shapes, all of which fail to capture the strength of the inversion. Since none of the initialized datasets can properly reproduce the inversion, it is unclear whether this could be fixed with a better initialization. Despite the differences seen, all members and datasets are in agree-

Ensemble Member	Temp.(°C)	Wind Speed(m/s)	Dew Point(°C)
PX/YSU Bias	+0.84	-0.26	+1.19
PX/YSU NARR Bias	-1.36	-0.24	+0.09
PX/YSU MAE	2.39	1.32	2.35
PX/YSU NARR MAE	2.02	1.21	1.84

Table 4.1: Surface statistic for domains 3 using ASOS stations only. The PX/YSU results shown here are identical to Fig. 2.2

ment with respect to the inversion height (not strength) for both the day and night.

From Fig. 4.3 and Fig. 4.4 we have seen that initializing the model off of the NARR dataset substantially improves results. However, we have not looked at how the results from the PX/YSU NARR affects the entire domain. Table 4.1 demonstrates this by showing the errors and biases from the PX/YSU (also seen in Fig. 2.2) and PX/YSU NARR. The PX/YSU NARR has a MAE that is smaller than its counterpart and all ensemble members for all three surface parameters tested. The surface temperatures from the PX/YSU NARR suffers from a rather large cold bias, however, its MAEs are still lower than every ensemble member. The wind speeds only benefited slightly from the new initialization while the dew point had a rather large improvement — decreasing the MAE by 0.51°C.

It should be understood that these results show the entire domain which includes the coastal subset. So while the errors decrease in the coastal subset it does not demonstrate how the NARR initialization affects other regions. In order to understand how the errors change with respect to each geographic region we bring our attention to Fig. 4.5. This 2 panel figure shows the MAE for specific geographic locations for both the PX/YSU (at left) and the PX/YSU NARR (at right). The smaller dots represent low MAEs while the larger ones show places

of higher MAEs.

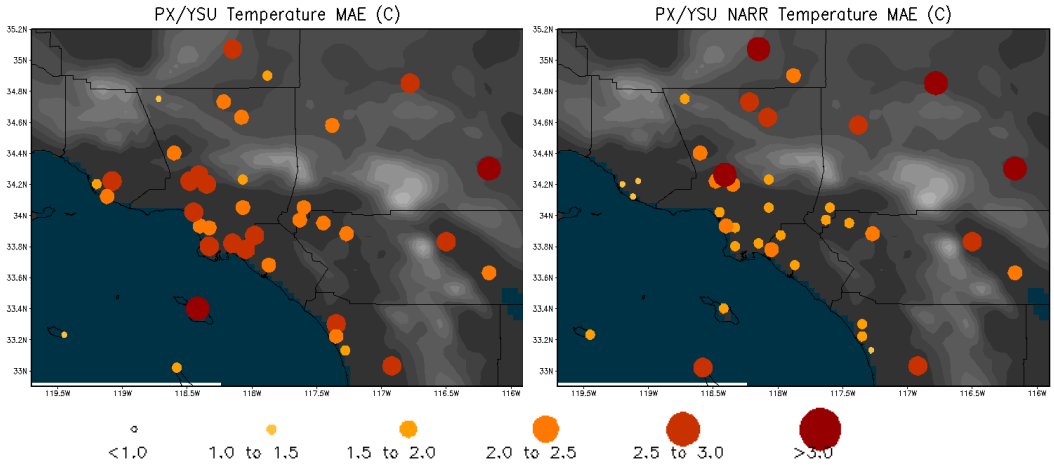

Figure 4.5: The temperature MAE for specific geographic locations for both the PX/YSU (left) and PX/YSU NARR(right). Each dot represents the monthly averaged MAE for a specific ASOS station.

When looking solely at locations in the Mojave Desert, we see slightly higher MAEs at almost all locations for the PX/YSU NARR. In order to understand why, Fig. 4.6 shows the temperature bias for the same geographic locations during both the day (17Z-00Z) and night time (01Z-16Z) hours. During the overnight hours the PX/YSU (Fig. 4.6b) has a similar cold bias in Mojave Desert as compared to the PX/YSU NARR (Fig. 4.6d); although, some locations are slightly colder for the PX/YSU NARR. During the daytime hours, the PX/YSU NARR (Fig. 4.6c) has a stronger cold bias than the PX/YSU (Fig. 4.6a) for almost all desert locations. This accounts for the majority of the MAE increase seen over the Mojave Desert. While it is not shown here, the 00Z surface temperatures from the NARR and NAM datasets experienced very little bias and were very similar between individual desert locations. Due to insufficient aircraft data in this area we could not verify temperatures aloft. Though a plausible explanation for the PX/YSU and PX/YSU NARR differences could be due to variations in

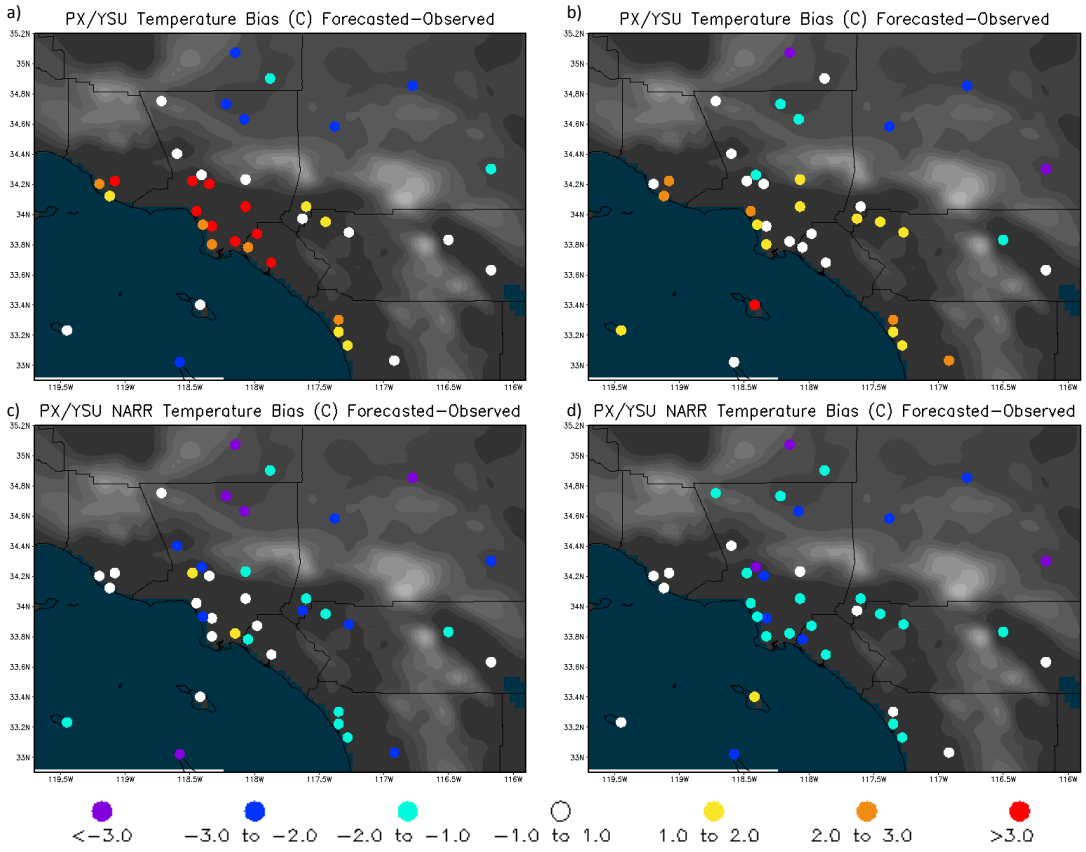

Figure 4.6: The temperature bias for specific geographic locations for both the PX/YSU (a-b) and PX/YSU NARR(c-d). The daytime (17Z-00Z) bias is shown in a,c while the nighttime (01Z-16Z) bias is b,d.

temperatures aloft at initialization.

In contrast, all coastal locations, the LA basin, and the Inland Empire benefit substantially from the NARR initialization. As discussed earlier, the afternoon warm bias seen at the coastal locations becomes non-existent when initializing off of the NARR dataset. The overnight hours at coastal locations do develop a cold bias for the PX/YSU NARR while the PX/YSU has a slight warm bias. Nevertheless, the few occurrences where the PX/YSU NARR actually does worse are outweighed by its overall superior performance.

Up to this point we have seen major differences between the PX/YSU and PX/YSU NARR, but have not uncovered the underlying reason as to why. We have seen that initializing the PX/YSU with a warm bias produces an overall warm bias and initializing the PX/YSU NARR with little bias produces a member with minimal bias. It is deceptive to think that initializing with or without bias will mean the difference between a bias or unbiased simulation. One can recall that we initialized the PX/YSU off the unbiased 12Z NAM and still produced a warm bias that was only slightly better (cooler) than the 00Z initialization. This suggests that there is something else driving this warm bias which is not connected to the biased temperature initialization over land.

It was hypothesized that the differences seen between NARR and NAM initialized simulations were a result of low levels clouds, or lack thereof. Though not shown, the NARR has low level (975hPa) relative humidity values that are roughly 15% higher than that of the NAM for both 12Z and 00Z initialization. This tongue of higher relative humidity covers Downtown Los Angeles, all coastal areas, and all Pacific waters in domains 3. On average, the 00Z initialized low level relative humidity over the coastal Pacific waters were around 65% with the

NAM and 80% with the NARR. The 12Z initialization offers the same difference but slightly higher relative humidities of 70% and 85% for the NAM and NARR respectively.

The differences in relative humidity are large enough to greatly influence the amount of low-level clouds and thus, the amount of solar radiation received. If the ground is receiving more solar radiation due to insufficient cloud coverage, the temperature will ultimately increase. This will in turn lower the relative humidity which overall, decreases the chances of forming clouds. This is a positive feedback which could very well be sensitive to the initialized low-level relative humidity.

The amount of shortwave radiation received at the surface of the earth is an excellent indicator for the areal cloud coverage. Ignoring the effects of the seasons, the solar radiation received at the surface of the earth is dependent on the reflection/absorption of shortwave radiation by atmospheric gases, the reflection due to the earth's surface, and that due to clouds. While atmospheric gases will reflect/absorb some solar radiation, the percentage it does reflect will be relatively fixed in time and space. The radiation reflected by the earth surface will vary in space but these values are fixed between each simulation. This means a difference in the surface shortwave radiation flux between simulations is a result of clouds.

Ideally one would like to look at the surface shortwave radiation flux in the initialized dataset but this is not readily available. Instead, we can look at the average surface short wave flux for July 2005 between the PX/YSU and PX/YSU NARR to understand how the different initializations affect cloud coverage. This is shown in Fig. 4.7 where the PX/YSU (at left) is compared to the PX/YSU NARR (at right). From the figure we see relatively little differences in the short-

wave flux over the inland areas of Southern California. This can be expected because hardly any clouds are observed during the summer months at inland locations.

Over the Pacific and coastal waters one can see a rather large difference (roughly 50 W/m^2) in the average surface shortwave radiation flux. This is due to the increased cloud coverage from the NARR initialization. In contrast, over the area of ASOS stations used to verify the coastal subset one can see that there is only a modest decrease in shortwave flux of approximately 10 W/m^2. This means that the decrease in temperatures over land seen in the PX/YSU NARR can mostly be attributed to an enhanced marine layer over the ocean which creates a cooler afternoon sea breeze.

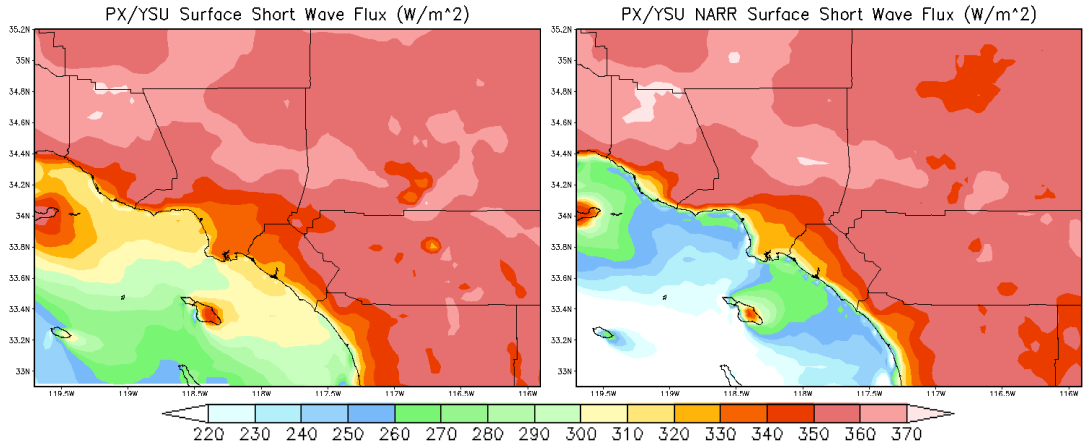

Figure 4.7: Surface shortwave radiation flux (W/m^2) average over July 2005. The PX/YSU is shown (at left) against the PX/YSU NARR (at right).

Another important difference between the two datasets can be seen in the Sea Surface Temperatures (SST) . In most coastal waters, the NARR is roughly 2°C cooler than the NAM and is an alarming 8°C cooler in the Santa Monica Bay (not shown). We mention this because the odd spatial pattern of sea surface temperatures seen in the NAM — which is probably due to inconsistencies between

the parent and regional model grids — is not a problem in the NARR. It should be noted that the WRF model was not configured to predict SSTs, but rather to keep the initialized SSTs constant throughout the entire run. So it was thought that colder SSTs found in the NARR dataset would contribute to cooler daytime temperatures.

In order to test the influence of ocean temperatures on air temperatures we replaced only the SSTs in the NAM dataset with those of the NARR. Other than this, the initialized dataset reflected that of the NAM. Rerunning only one of the PX/YSU simulations (July 17th-20th), we observed overnight temperatures that were roughly 3°C cooler for Downtown Los Angeles and 5°C cooler for the Los Angeles International Airport when compared to the original PX/YSU. In contrast, we saw little to no decrease in afternoon temperatures for inland places, such as Downtown Los Angeles, but a 1-3°C decrease for coastal locations. This did not agree with the PX/YSU NARR which had afternoon temperatures that were 3.5°C cooler for Downtown Los Angeles. So it seems that initializing with colder SSTs does not help the afternoon warm bias unless you are on the immediate coast, it does however create colder minimum temperatures.

4.2 Integration Methods

All work up to this point has focused on changing the physics parameterization or initialized datasets to improve simulations for July 2005. This last section will discuss the use of different integration methods and how they impact results.

Rather than ending our simulation every 3 days we looked at longer integrations (up to 30 days) in hopes that the temperature bias would be removed

with time. We found that the bias for the PX/YSU decreased from +1.95°C to +1.05°C when we integrated the model for a month (not shown). However, vertically integrated MAEs increased slightly, though not substantially, in the month-long simulation from 19.6°C to 20.1°C. Because of this one may assume that it is appropriate to integrate the WRF model for a month at a time; however, this is not the case.

This same experiment was done with the PX/YSU NARR and yielded different results. This time, vertically integrated MAEs increased by 0.92°C but the main difference was that errors began accumulating rapidly towards the end of the month. This is demonstrated in Fig. 4.8 where the PX/YSU NARR month simulation begins to drift away from its counterpart, the PX/YSU NARR. This figure, which depicts surface temperature, illustrates how errors can accumulate with time in a numerical model creating an entirely different result. It is interesting to note that the NARR initialization produces a drift whereas the NAM initialization does not. It is possible that the NAM initialization has temperatures that are more in tune with the model's climate. This means that long integrations would not show the model drifting towards its climatic bias since it was initialized with it. Based on the information seen here, integrating the model for long periods of time is never recommended.

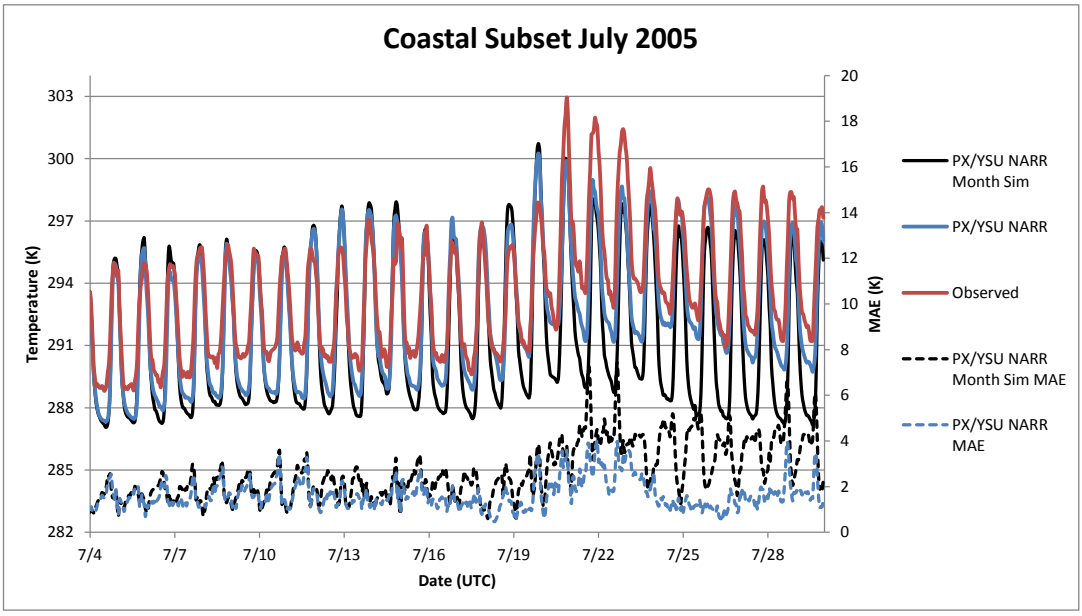

Figure 4.8: Observed and forecasted temperatures in the coastal subset using ASOS data. The PX/YSU NARR month simulation is seen to 'drift' away from its parent model due to the long model integration.

CHAPTER 5

Conclusions

July 2005 was an above average month in terms of temperature, but by no means was it abnormal. Successfully modeling this typical summertime weather can prove to be challenging due to the large temperature gradients that so commonly exist between the land and ocean. We discovered that modeling this large temperature gradient meant that low level relative humidities must not be overlooked. In particular, the shallow marine layer that exists in this summer month can be easily missed by certain datasets. While we have only tested July 2005 and found the NARR to be superior, later years may reveal different results due to changes in the NAM model.

We have also discovered that certain model physics are superior to others in Southern California. When talking solely about surface parameters, the PX LSM almost always had the most realistic recreation of the month. In fact, out of the 8 PX ensemble members tested, 6 of them were in the top 8 spots when considering surface temperature, dew point and wind. The TD LSM was generally the worst due to the fact that it does not explicitly solve for soil moisture. Thus, dew point forecasts were not as accurate as those of the Noah or PX LSMs. Most of the Noah simulations were not far behind those of the PX and were also seen as quite skillful.

In addition to the land surface model, the boundary layer schemes also displayed systematic ranking. The non-local YSU PBL always performed the best for the surface variables while the TEMF was always the worst and is not recommended for use. Surprisingly, the more sophisticated and widely used MYJ PBL always ranked near the end of our 8 PBLs tested.

The boundary layer parameterizations were less systematic than the LSMs and were generally in no particular order in terms of skill. For the temperature profile, the ACM2 proved to be the most accurate PBL; the YSU performed second best. Just as with the surface, the TEMF was significantly worse than all other PBL options and again, should not be used in this area at this resolution. Since dew point was not measured in the vertical profile we could not verify the ensemble's moisture above 2 meters. However, one could hypothesize that accurate moisture fields at the surface would be indicative of a PBL that is properly handing the vertical moisture diffusion. For wind, we did have an abundant amount of data although MAEs produced variances so small that we found them not to be of significance.

From the 24 member ensemble, 2 physics combinations proved themselves as quite skillful amongst the various verification tests. These being the PX/YSU and PX/ACM2. Based off their performance of dew point, temperature, and winds at the surface along with the vertical temperature profile, they produce the most realistic results for the month. Not surprisingly, both of these ensemble members use a very similar non-local PBL scheme. The major difference between the two is that the ACM2 has non-local mixing upwards — which matches the YSU — but has only local transport downward.

Along the coast, the ensemble was initialized with a positive temperature bias

that originated from 00Z NAM initialization. We found that the 00Z positive temperature bias was not an issue for the 12Z initialization. However, starting the simulations at 12Z still created an afternoon warm bias along the coast and was an ineffective means of improving the models performance. Even integrating the model for a month at a time proved to be unsuccessful at removing the warm temperature bias. Eventually, it was discovered that initializing the model with a different dataset would be extremely beneficial at removing the bias.

The main differences between the NAM and NARR datasets were the low level relative humidities that contributed to the formation of clouds along the coastal waters. The simulation initialized with the NARR developed more clouds than that of the NAM. This in turn decreased the amount of solar radiation received and completely removed the warm bias found along the coast. Another important difference between the two datasets was the alarming difference in sea surface temperature. The NARR was on average 2°C cooler than the NAM and up to 8°C cooler in the Santa Monica Bay. Replacing the NAM dataset SSTs with those of the NARR did in fact create cooler daytime temperatures but only along the immediate coast. The biggest impact were the overnight minimum temperatures which were roughly 3°C to 5°C cooler from downtown Los Angeles to the immediate coast. Due to the impact on coastal temperatures, it is suggested to initialize the model with an accurate SST product.

While the data used to initialize the model is extremely important for coastal areas, we cannot recommend one dataset over the other since the NAM has likely changed throughout the years. In addition to this, short model integrations are highly suggested as it will limit the amount of error that accumulates from day-to-day. While we have tested a wide array of physics options, we can confidently

recommend using the PX/YSU or PX/ACM2 physics combination for summertime simulations in Southern California.

References

[1] Argüeso, D., J. M. Hidalgo-Muñoz, S. R. Gámiz-Fortis, M. J. Esteban-Parra, J. Dudhia, Y. Castro-Díez, 'Evaluation of WRF Parameterizations for Climate Studies over Southern Spain Using a Multistep Regionalization', *Journal of Climate*, Volume 24, Issue 21 (November 2011) pp. 5633-5651, accessed 10 March 2012 from AMS Journals Online.

[2] 2012, *ARW Version 3 Modeling System User's Guide*, University Corporation for Atmospheric Research, accessed 5 May 2012, <http://www.mmm.ucar.edu/wrf/users/docs/user_guide_V3.3/ARWUsersGuideV3.pdf>

[3] Benjamin, S. G., B. E. Schwartz and R. E. Cole 'Accuracy of ACARS Wind and Temperature Observations Determined by Collocation': *Weather and Forecasting*, Volume 14, Issue 6 (December 1999) pp. 1032-1038, accessed 6 March 2012 from AMS Journals Online.

[4] Bruno, Kaplan, Gomberg 2006, *Public Information Statement: Southern California Weather in 2005*, NOAA's National Weather Service Western Regional Headquarters, accessed 25 March 2012, <http://www.wrh.noaa.gov/lox/archive/pns_2005summary.pdf>

[5] Cheng, William Y. Y. and W. J. Steenburgh, 'Evaluation of Surface Sensible Weather Forecasts by the WRF and the Eta Models over the Western United States', *Weather and Forecasting* Volume 20, Issue 5 (October 2005) pp. 812-821, accessed 9 March 2012 from AMS Journals Online.

[6] Dudhia, J., *WRF Physics Options*, University Corporation for Atmospheric Research, accessed 8 March 2012, <http://www.mmm.ucar.edu/wrf/users/tutorial/200807/WRF_Physics_Dudhia.pdf>

[7] Edwards, L. and J. Null 2005, *California Climate Watch*, California Climate Data Archive, accessed 25 March 2012, <http://www.calclim.dri.edu/climatewatch/CalClim200507-2.pdf>

[8] *ESRL/GSD Aircraft Data (AMDAR) Information*, NOAA/ESRL/GSD, accessed 6 March 2012, <http://amdar.noaa.gov/FAQ.html>

[9] Gallus, W. A. Jr. and J. F. Bresch, 'Comparison of Impacts of WRF Dynamic Core, Physics Package, and Initial Conditions on Warm Season Rainfall Forecasts', *Monthly Weather Review*, Volume 134, Issue 9 (September 2006) pp. 2632-2641, accessed 10 March 2012 from AMS Journals Online.

[10] Gibbs, J. A., E. Fedorovich, A. M. J. van Eijk, 'Evaluating Weather Research and Forecasting (WRF) Model Predictions of Turbulent Flow Parameters in a Dry Convective Boundary Layer', *Journal of Applied Meteorology and Climatology*, Volume 50, Issue 12 (December 2011) pp. 2429-2444, accessed 9 March 2012 from AMS Journals Online.

[11] Gilliam, R. C. and J. E Pleim, 'Performance Assessment of New Land Surface and Planetary Boundary Layer Physics in the WRF-ARW', *Journal of Applied Meteorology and Climatology*, Volume 49, Issue 4 (April 2010) pp. 760-774, accessed 9 March 2012 from AMS Journals Online.

[12] Jiménez P. A. and J. Dudhia, 'Improving the Representation of Resolved and Unresolved Topographic Effects on Surface Wind in the WRF Model', *Journal of Applied Meteorology and Climatology*, Volume 51, Issue 2 (February 2012) pp. 300-316, accessed 8 March 2012 from AMS Journals Online.

[13] *Log Of Operational ETA Model Changes*, National Weather Service's National Centers For Environmental Prediction's Environmental Modeling Center, accessed 8 March 2012, <http://www.emc.ncep.noaa.gov/mmb/research/eta.log.html>

[14] Mesinger, F., D. Geoff et al. 2005, *North American Regional Reanalysis*, National Weather Service's National Centers For Environmental Prediction's Environmental Modeling Center, accessed 6 March 2012, <http://www.emc.ncep.noaa.gov/mmb/rreanl/narr_bams.pdf>

[15] Mesinger, 2004, *The Eta Model: Design, History, Performance, What Lessons have we Learned?* National Oceanic Atmospheric Administration, accessed 8 March 2012, <http://docs.lib.noaa.gov/rescue/JNWP/50th_Symp_2004_CD.PDF/Extended%20Abstracts/Session%203%20%20Evolution%20of%20Operational%20Numerical%20Weather%20Prediction/033.pdf>

[16] Ruiz, J. J., C. Saulo and J. Nogués-Paegle, 'WRF Model Sensitivity to Choice of Parameterization over South America: Validation against Surface Variables', *Monthly Weather Review*, Volume 138, Issue 8 (August 2010) pp. 3342-3355, accessed 10 March 2012 from AMS Journals Online.

[17] Shin, H. H., S.-Y. Hong and J. Dudhia, 'Impacts of the Lowest Model Level Height on the Performance of Planetary Boundary Layer Parameterizations', *Monthly Weather Review*, Volume 140, Issue 2 (February 2012) pp. 664-682, accessed 10 March 2012 from AMS Journals Online.

[18] Skamarock, W. C., J. B. Klemp, and J. Dudhia, 2001, *Prototypes for the WRF (Weather Research and Forecasting) model.* Preprints, Ninth Conf. on Mesoscale Processes, Fort Lauderdale, FL, Amer. Meteor. Soc., CD-ROM, J1.5., accessed 9 March 2012.

[19] Staudenmaier M. Jr., 1996, *The Initialization Procedure In The Meso Eta Model*, NOAA's National Weather Service Western Regional Headquarters, accessed 8 March 2012, <http://www.wrh.noaa.gov/wrh/96TAs/TA9630/ta96-30.html>

[20] Schwartz, B. and S. G. Benjamin, 'A Comparison of Temperature and Wind Measurements from ACARS-Equipped Aircraft and Rawinsondes', *Weather and Forecasting*, Volume 10, Issue 3 (September 1995) pp. 528-544, accessed 25 March 2012 from AMS Journals Online.