Statistical Analyses of Differences in Wind between ARW and NMM WRF Cores

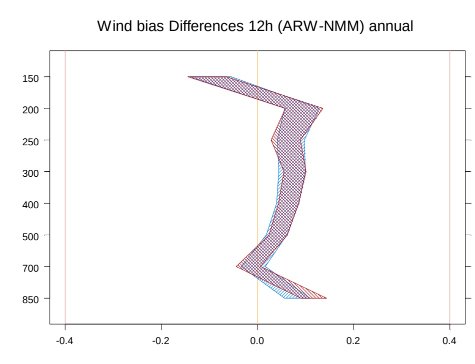
August 3rd, 2006

Elizabeth Weatherhead
Gregory Noonan

Weather forecasts were produced simultaneously using two different cores (ARW and NMM). An original report (Statistical Analyses of Differences in Relative Humidity, Temperature, Wind and Precipitation between ARW and NMM, July 27, 2006) gives an overview of the differences observed in temperature, relative humidity and wind using sonde data. This study takes a closer look at bias and RMSE of winds using both aircraft reports and sondes. The primary question addressed is whether the differences observed in upper level winds—roughly 200mb—is dependent on which source of wind data is used for evaluation. This report focuses on wind bias and RMSE using sondes and aircraft. The analyses are completed for CONUS, CONUS East and CONUS West, and all calculations were performed using all data as well as seasonal data. Both Phase-1 physics and Phase-2 physics are included in this report.

This analysis relies heavily on pairwise comparison tests. The comparison fundamentally looks at differences in pairs of numbers. In its most basic form, the pairwise test assumes that these differences are normally distributed, independent and of equal accuracy. This is augmented in this report to look more carefully at how the number of observations available for verification may influence the results. Of particular concern is whether the results from sondes are biased by the fact that on very windy days, some sondes can blow out of communication range and data from the highest elevations may not be available for analysis.

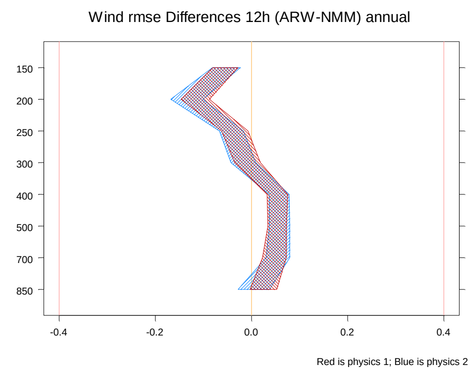
Wind Bias



Wind bias results from sondes are in general consistent from 850mb through 200mb although NMM shows a smaller magnitude of bias for these levels. At 150mb, ARW shows a smaller magnitude of bias. These results are somewhat seasonally dependent, particularly for the behavior above 300mb. Aircraft data generally show that NMM has a smaller magnitude of bias, at all levels, although the difference in bias is smaller. The major difference is in the highest altitude. Aircraft results were analyzed up to 200-250mb. There are no aircraft results which correspond to the highest sonde level (150mb) where sondes show that ARW has a smaller magnitude of bias. In Autumn and Winter, measurements from ARW - NMM show different results for 200mb, with the sondes showing that ARW has a much lower magnitude of bias than NMM while the aircraft show a small advantage to NMM. These differences in bias are based exclusively on differences in forecasts for the locations at which observations are available and not on the value of the measurements themselves. More information is on this subject is available in Appendix 3 and Appendix 4.

Wind RMSE

For both the 12h and 24h forecasts, NMM exhibits a smaller RMSE for the lowest part of the atmosphere (between 850mb and 300mb). Above 300mb, ARW has a smaller RMSE for all regions and all physics packages. This difference is most clear for the 12h forecast and becomes less clear with the 24h forecast, although both forecasts show similar patterns. These patterns are repeated seasonally, although at 850mb in Autumn and Winter, ARW has a smaller RMSE. The results are similar from both sondes and aircraft, although the aircraft do not offer results for levels as high as 150mb. The seasonal information is the same for both aircraft and sondes, although the error bars from the aircraft data are smaller, presumably because the results are based on more observations. For 24 hour forecasts, the RMSE values are smaller for ARW than NMM for almost all levels from the aircraft data, while this is less true for the sonde data. This difference is particularly notable in Autumn and Winter. More information is on this subject is available in Appendix 5 and Appendix 6.



Model Initialization.

The pairwise approach was also used to examine whether the ARW and NMM models were initialized similarly. Based on an examination of sonde winds it was clear that the initialization of the models was statistically different. However, the RMSE differences between ARW and NMM at initialization are extremely small when the aircraft data are used. Using sonde data, many of the differences while small, were highly statistically significant. Using aircraft data, we observe even smaller differences with most being statistically significant. More information on the model initialization of winds is available in Appendix 7.

Dependence of Results.

An important question, when sonde data and aircraft data give different results, is why the results are different. In order to address this question we look at inherent biases in each of the datasets. For sonde data, it has been proposed that lost sondes at high altitudes could affect the results. We examine this in a stepwise approach. First it is verified that fewer measurements are available at 150mb than at 200mb. On average there are a few less sondes for the higher altitudes, but this difference is very small. Second, it is shown that at high altitudes (150mb and 200mb), high average zonal winds are associated with lower numbers of sonde measurements; this is not true lower in the atmosphere where the number of sonde measurements is roughly independent of average winds. Finally, we note that the estimates of RMSE and the differences in RMSE between ARW and NMM are roughly independent of the number of measurements collected. Thus, while the dropoff rate with sondes is real, and likely due to the ambient conditions, it does not appear to have a large influence on the conclusion regarding RMSE and bias. More information is available on this in Appendix 8.

Appendices

References

Von Storch, H. and F. Zwiers, “Statistical Analysis in Climate Research”, Cambridge University Press, 1999.

Rowlings, J., S Pantula and D. Dickey, “Applied Regression Analysis,” Springer-Verlag, 1998.

Neter, J., W. Wasserman and M. Kutner, “Applied Linear Statistical Models,” Irwin, 1990.

Wilks, Daniel S., “Satistical Methods in the Atmospheric Sciences,” Academic Press, 1995.

Box, G., G. Jenkins, and G. Reinsel, “Time Series Analysis: forecasting and control,” Prentice-Hall, 1994.

Weatherhead, E., et al., “Statistical considerations in the detection of trends.” JGR, 103, 17,149-17161, 1998.

Efron, B. and R. Tibshirani, “An Introduction to the Bootstrap,” Chapman and Hall/CRC, New York, 1994.

Appendix 1: Background Information: WRF-Rapid Refresh Core Test

Model Summary:

The WRF (Weather Research and Forecasting Model) is a next generation mesoscale numerical weather prediction system developed cooperatively by NCAR, NCEP, GSD and others to serve operational forecasting needs. The WRF features two dynamical cores/solvers and a 3-D variational data assimilation system setting up the initial conditions before the model run. Currently in use are:

- The WRF w/ ARW (Advanced Research WRF) core/solver – This model was developed primarily at NCAR, but is operated via NCEP's computers.

- WRF w/ NMM (Non-Hydrostatic Mesoscale Model) core/solver was developed by NCEP, but currently supported to the community by the WRF DTC (Developmental Test Bed Center); The WRF DTC is a facility where the NWP research and operational communities interact to accelerate testing and evaluation of new models and techniques for research applications and operational implementation, without interfering with current operations.

NOAA/NCEP's EMC (Environmental Modeling Center) also maintains an array of models with the RUC being one these. The RUC (Rapid Update Cycle) model in use today was developed by NOAA's GSD, formerly FSL. It is comprised of a numerical forecast model and an analysis/assimilation system needed to initialize the model. It serves users needing frequently updated short-range weather forecasts, such as in aviation and the severe weather forecasting community. It is maintained and run through NCEP.

By 2008, full implementation of a replacement to the current RUC is anticipated. The WRF-RR (Rapid Refresh) model is the next generation version of the 1-hour cycle system in use by the RUC today. The Rapid Refresh model will use a version of NCAR's WRF model along with a GSI (Gridpoint Statistical Analysis) developed at NCEP's EMC. GSD's RUC Development Group and the DTC are both currently involved with testing towards the selection of a model core and a physics package.

The comparison/verification below between the two cores/solvers and two physics packages (phases) is at the heart of our report.

ARW core/solver (NCAR)		**NMM core/solver (NCEP)**	
Phase 1 physics / NMM package	Phase 2 physics / RUC package	Phase 1 physics / NMM package	Phase 2 physics / RUC package

Appendix 2: Data Inventory and File Structure

Two types of files were made available to us: verification database files (VSDB files) and QPF files. We looked at a small amount of the information available from these files. Specifically, we examined the temperature, relative humidity and vector wind fields for only three regions: Conus and the sub-regions of Conus West and Conus East. We focused on upper air comparisons for these parameters, as well as surface precipitation from QPF files.

VSDB (Verification Database) File inventory:

	ARW		NMM	
	CONUS 13		CONUS 13	
	Phase-1 physics	Phase-2 physics	Phase-1 physics	Phase-2 physics
Summer 2005	2005071500&12 thru 2005081500&12 63/64 possible days	2005071500&12 thru 2005081500&12 56/64 possible days	2005071500&12 thru 2005081500&12 63/64 possible days	2005071500&12 thru 2005081500&12 63/64 possible days
Autumn 2005	2005110100&12 thru 2005113000&12 59/60 possible days	2005110100&12 thru 2005113000&12 57/60 possible days	2005110100&12 thru 2005113000&12 59/60 possible days	2005110100&12 thru 2005113000&12 59/60 possible days
Winter 2006	2006011600&12 thru 2006021500&12 60/64 possible days	2006011600&12 thru 2006021500&12 59/64 possible days	2006011600&12 thru 2006021500&12 59/64 possible days	2006011600&12 thru 2006021500&12 59/64 possible days
Spring 2006	2006032500&12 thru 2006042500&12 59/64 possible days	2006032500&12 thru 2006042500&12 58/64 possible days	2006032500&12 thru 2006042500&12 59/64 possible days	2006032500&12 thru 2006042500&12 58/64 possible days

A total of 241 out of 252 total possible files exist for ARW phase-1. For ARW phase-2, 233 out of 252 exist. For NMM phase-1, 240 out of 252 files exist. For NMM phase-2, 239 out of 252 files exist. In summary, 953 out of a possible 1008 total files exist

Upon opening any dated directory, two files exist for all 0Z run times. They are as follows:

1) **wrfrr_ARW_fc_2005071500_218_3h_QPF.vsdb** 126 kb

wrf rapid refresh_arw_forecast_7-15-05 at 0Z_218_3 hour_quantitative precipitation forecast.vsdb

2) **wrfrr_sfcupa_2005071500.vsdb** 334 kb

wrf rapid refresh_surface and upper air_7-15-05 at 0Z.

Upon opening a 0z run time directory, three files exist. They are as follows:

1) **wrfrr_ARW_fc_2005071512_218_3h_QPF.vsdb** 126 kb
wrf rapid refresh_arw_forecast_7-15-05 at 12Z_218_3 hour_quantitative precipitation forecast.vsdb

2) **wrfrr_ARW_fc_2005071512_218_24h_QPF.vsdb** 16 kb
wrf rapid refresh_arw_forecast_7-15-05 at 12Z_218_24 hour_quantitative precipitation forecast.vsdb

3) **wrfrr_sfcupa_2005071512.vsdb** 329 kb
wrf rapid refresh_surface and upper air_7-15-05 at 12Z.

Reading the File Structure:

This is a QPF file: Quantitative Precipitation Forecast 3h
C:\VSDBFiles\ARW\conus13\ph1\2005071500\verifprd\wrfrr_ARW_fc_2005071500_218_3h_QPF.vsdb

1	2	3	4	5	6	7	8	9					
V01	**WRFRR-ARW-PH1R3**	**03**	**2005071503**	**MC_PCP**	**G218/RFC**	**FHO>.01**	**APCP/03**	**SFC**	**=**	**73212**	**0.23196**	**0.08682**	**0.17783**

A typical file is separated into two parts; the header field is separated from the data field by the separator field designated with an equal sign (=).

Looking at the header field, the 9 components above are described using the following format:

1) Verification database version: V01 = version 1.
2) Forecast model being verified: WRFRR – ARW core – Phase-1 physics.
3) Forecast hour: 03, 06, 09, 12, 15, 18, 21, 24. So, what is the model forecasting 3, 6, 9…hours from initial time of 0Z in this case. Every 3 hours
4) Verifying date: 20050715**03**, 20050715**06**, 20050715**09**, 20050715**12**, 20050715**15**, 20050715**18**, 20050715**21**, 20050715**24**. Last two numbers match forecast hour.
5) Verifying data source or analysis: All MC_PCP i.e. Mike C's precipitation analysis.
6) Verifying grid or region: All G218 followed by 15 different regions: this pattern of grid regions repeats itself for each forecast hour and each statistic type (field 7).
7) Statistic type: FHO >.01, .02, .05, .10, .15, .25, .35, .50, 1.0: Active statistics.
8) Parameter identifier: Always APCP/03 = 3 hour accumulated precipitation. At 03 forecast hour, how much precipitation fell between 0Z and 3Z. At 06, how much fell between 3Z and 6Z etc.
9) Level identifier: SFC = surface. Model is predicting precipitation at the surface.

The data fields which follow are described using the following format:

Column #1) The number of values used – grid points or observations.

Column #2) F = Forecasted fraction above/below threshold. The forecasted amount of precipitation generated by the model over a region for a three hour period. This amount must be over some threshold amount - .01 in this case. The line above is forecasting .23196 inches of precipitation <u>somewhere</u> in the RFC region between 0Z and 3Z; we don't have an output as to where in the region the model is predicting.

Column #3) H = Correct fraction above/below threshold. How much area within the region actually received precipitation that the model forecast for that region.

Column #4) O = Observed fraction above/below threshold. How much precipitation fell total in the region, not just in some area of the region.

This is a SFCUPA file: Surface and Upper Air
C:\VSDB Files\ARW\conus13\ph1\2005071500\verifprd\ wrfrr_sfcupa_2005071500.vsdb

V01 WRFRR-ARW-PH1R3 00 2005071500 ADPUPA G164 SL1L2 Z P1000 = 72. 0.888472140E+02 0.925555500E+02 0.100121380E+05 0.991984670E+04 0.102059990E+05

Looking at the header field, the 9 components above are described using the following format:

1) <u>Verification database version:</u> V01 = version 1.
2) <u>Forecast model being verified:</u> WRFRR – ARW core – Phase-1 physics.
3) <u>Forecast hour:</u> 00, 03, 06, 09, 12, 15, 18, 21, 24. So, what are the model's initialization conditions at 00. Followed by 3, 6, 9…hours from initial time of 0Z in this case.
4) <u>Verifying date:</u> 20050715**00**, 20050715**03**, 20050715**06**, 20050715**09,** 20050715**12**, 20050715**15**, 20050715**18**, 20050715**21,** 2005071500. Last two numbers match forecast hour.
5) <u>Verifying data source or analysis:</u>
    - 1/5 = ADPUPA = Conventional upper-air
    - 1/5 = ANYAIR = Any upper-air data source
    - 1/5 = PROFLR = Profiler
    - 1/5 = VADWND = VAD WSR88D wind profiles
    - 1/5 = ONLYSF = Surface data verified against 2/10-m forecast data
6) <u>Verifying grid or region:</u>
   For the above data sources 4/5 use G164, G165, G166.
    - G164 = entire CONUS
    - G165 = ½ CONUS to west
    - G166= ½ CONUS to east
7) <u>Statistic type:</u> SL1L2 = active type; used with Z, T, RH, and SLP (in ONLYSF).
   VL1L2 = active type; used with all VWND fields.
8) <u>Parameter identifier:</u>
    - Z = geopotential height (only in ADPUPA fields)
    - T = temperature (sensible) (in ADPUPA, ANYAIR, ONLYSF fields)
    - RH = relative humidity (in ADPUPA and ONLYSF fields)
    - VWND = vector wind (in ADPUPA, ANYAIR, PROFLR, VADWND, ONLYSF)
    - SLP = sea level pressure (only in ONLYSF fields)
9) <u>Pressure level:</u>

- **ADPUPA** (Z, T, VWND) = 1000, 850, 700, 500, 400, 300, 250, 200, 150, 100 (10 levels)
  (RH) = 1000, 850, 700, 500, 400, 300 (6 levels)
- **ANYAIR** (T, VWIND) = 1000-850, 850-700, 700-550, 550-400, 400-300, 300-250, 250-200 (7 levels)

- **PROFLR** (VMND) = 1000-850, 850-700, 700-550, 550-400, 400-300, 300-250, 250-200 (7 levels)
- **VADWND** (VMND) = 1000-850, 850-700, 700-550, 550-400, 400-300, 250-200 (6 levels)
- **ONLYSF** (SLP, T, RH, VWND,) Only at the surface!!

The data fields which follow are described using the following format:

Column #1) The number of values used – grid points or observations.

Column #2) Mean [f] = mean of the forecast.
The model spits out a forecast temperature for each grid point. The actual grid where observations are made (the cites e.g.) is moved onto the model grid. Temperatures are then assigned to the cities via interpolation. A mean is then derived from these interpolated values. The first column in the data fields corresponds to this number.

Column #3) Mean [o] = mean of the observed.
Those cities that fall under the model grid take an observation at the time of the model prediction. A mean of these observations is taken to produce one number.

Column #4) Mean [f*o] = mean of the forecast value x observed value.

Column #5) Mean [f*f] = mean of the forecast value x forecast value.

Column #6) Mean [o*o] mean of the observed value x observed value.

Appendix 3: Summary of Wind Bias using Sonde Data

Wind results are compared here for both the ARW and NMM core with 2 sets of physics, Phase-1 and Phase-2. Three regions: CONUS, CONUS-West, and CONUS-East along with two forecast hours (12h and 24h) are analyzed separately for comparison.

The wind bias is generally consistent from 850mb through 200mb with NMM showing a smaller magnitude of bias for these levels. At 150mb, the ARW model shows a smaller magnitude of bias. These results are somewhat seasonally dependent, particular for the behavior above 300mb. For Winter and Autumn, the ARW model shows a much smaller magnitude of bias than the NMM model.

This section consists of two subsections. The first subsection addresses the 12 hour forecasts while the second subsection addresses 24 hour forecasts. Initialization values are addressed in Appendix 7. Each subsection begins with a series of summary plots followed by tables of values. The tables are organized by region: CONUS, CONUS West, and CONUS East. For each region, The first table, labeled “annual,” gives results using all of the data; this table is followed by four tables giving the results when using only data for each of the four seasons.

We calculate a mean derived from daily (0Z and 12Z) bias calculations, we calculate both the Standard Error (sigma) and a Pair-Wise Comparison (Z-val) on our results. In columns 5 and 6, we show the results of a comparison made on bias, sigma, and Z-val for both cores with the same physics packages.

The following tables summarize results of bias calculations performed on the wind values generated for the WRF comparison. A description of each column is as follows:

Column 1: The first cell of the first column describes which season is examined: Annual means all available data are used (7-15-05 thru 4-25-06); Winter indicates data from 1-16-06 thru 2-15-06; Spring indicates data from 3-25-06 thru 4-25-06; Summer indicates data from 7-15-05 thru 8-15-05; and Autumn indicates data from 11-1-05 thru 11-30-05. The remainder of Column 1 helps describe the data and are grouped to describe the results for each pressure level. Eight pressure levels are examined: 150mb, 200mb, 250mb, 300mb, 400mb; 500mb, 700mb, 850mb. For each pressure level there are three rows of information in the tables. The first row gives the mean estimate; the second row gives the standard error of the estimate and the third row gives the Z-value which is the mean divided by the standard error. More information on the statistical techniques used are available in the main document and appendices 13 and 14.

Column 2: ARW.ph1 The mean of the ARW bias using Phase-1 physics. 00Z and 12Z forecasts were used. Only forecasts for which both ARW and NMM bias values for Phase-1 physics were available were included. Below the mean estimate is the standard error on the estimate; below the standard error is the associated Z-value obtained by dividing the mean by the standard deviation. This is calculated and recorded for all eight pressure levels.

Column 3: NMM.ph1 The mean of the NMM bias using Phase-1 physics. 00Z and 12Z forecasts were used. Only forecasts for which both ARW and NMM bias values for Phase-1 physics were available were included. Below the mean estimate is the standard error on the estimate; below the standard error is the associated Z-value obtained by dividing the mean by the standard deviation. This is calculated and recorded for all eight pressure levels.

Column 4: ARW.ph2 The mean of the ARW bias using Phase-2 physics. 00Z and 12Z forecasts were used. Only forecasts for which both ARW and NMM bias values for Phase-2 physics were available were included. Below the mean estimate is the standard error on the estimate; below the standard error is the associated Z-value obtained by dividing the mean by the standard deviation. This is calculated and recorded for all eight pressure levels.

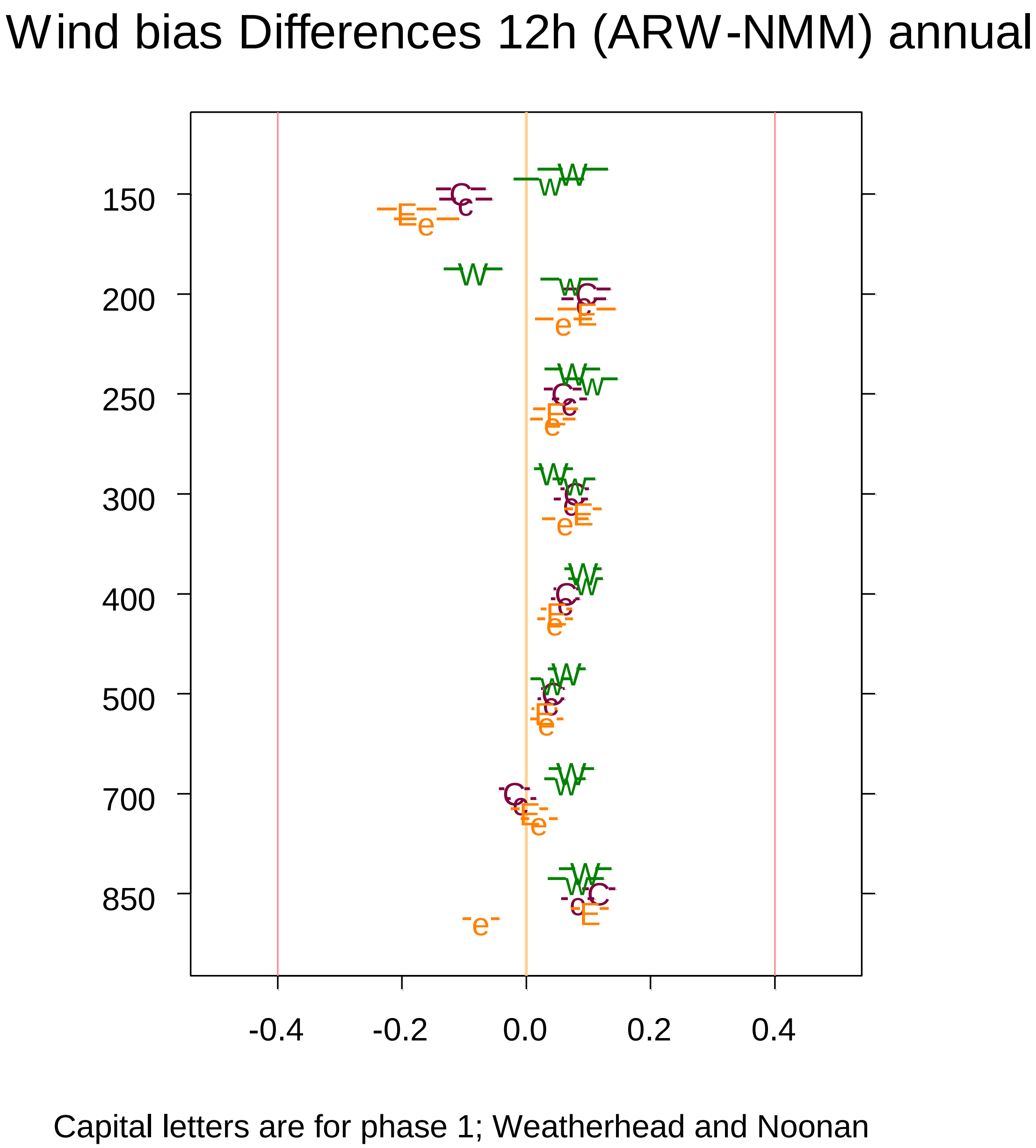
Column 5: NMM.ph2 The mean of the NMM bias using Phase-2 physics. 00Z and 12Z forecasts were used. Only forecasts for which both ARW and NMM bias values for Phase-2 physics were available were included. Below the mean estimate is the standard error on the estimate; below the standard error is the associated Z-value obtained by dividing the mean by the standard deviation. This is calculated and recorded for all eight pressure levels.

Column 6: Ph1.diff The mean of the pairwise differences of ARW bias and NMM bias using Phase-1 physics. Below this mean is the calculated standard error in the mean; below the standard error is the associated Z-value obtained by dividing the mean by the standard deviation.

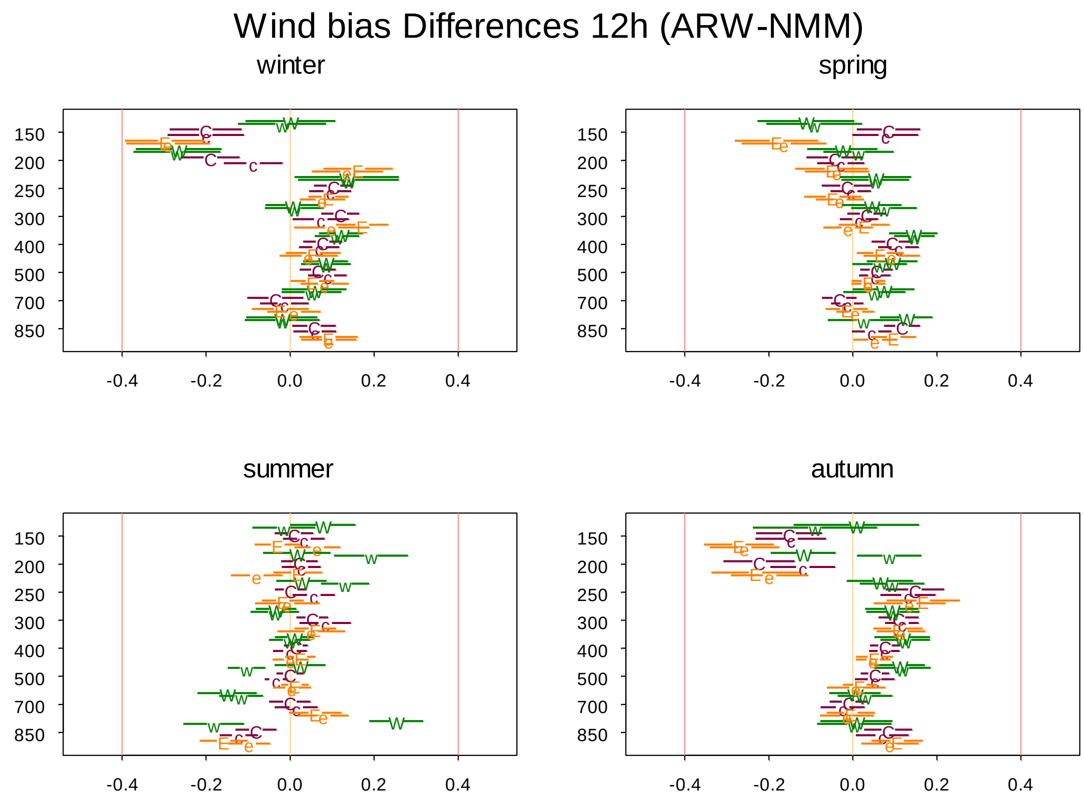
Column 7: Ph2.diff The mean of the pairwise differences of ARW bias and NMM bias using Phase-2 physics. Below this mean is the calculated standard error in the mean; below the standard error is the associated Z-value obtained by dividing the mean by the standard deviation.

Column 8: # of Obs The total number of observations used in deriving the results in the prior columns. The first number is the number of forecasts used for the calculations. The second and third numbers are the average number of observations used to derive the individual forecast values for Phase-1 and Phase-2 respectively.

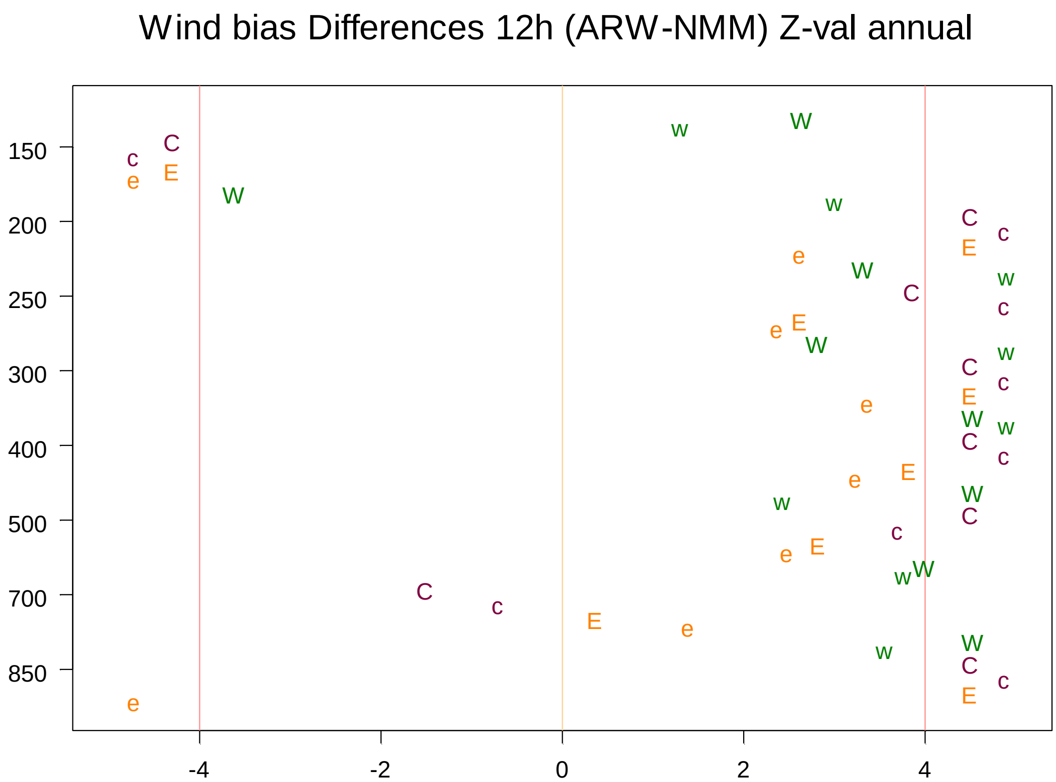
Bias for Sonde Vector Wind – Forecast Hour 12



Caption: The differences in the absolute values of wind bias are presented for three regions: Conus (C for physics 1 and c for physics 2), Conus West (W for physics 1 and w for physics 2)and Conus East (E for physics 1 and e for physics 2). Model bias for each level may be either positive or negative. The absolute value of the bias means that the results can be interpreted more directly. If the magnitude of the bias of ARW is larger, the results will be on the right side of the plot. If the magnitude of the bias of NMM is larger, the results will be on the left side of the plot. The accompanying error bars represent a 95% confidence limit on the difference of the wind bias. Results are presented for eight pressure levels, with the three regions slightly offset vertically for clarity.
(For comparison: see 12h aircraft bias on page 24)



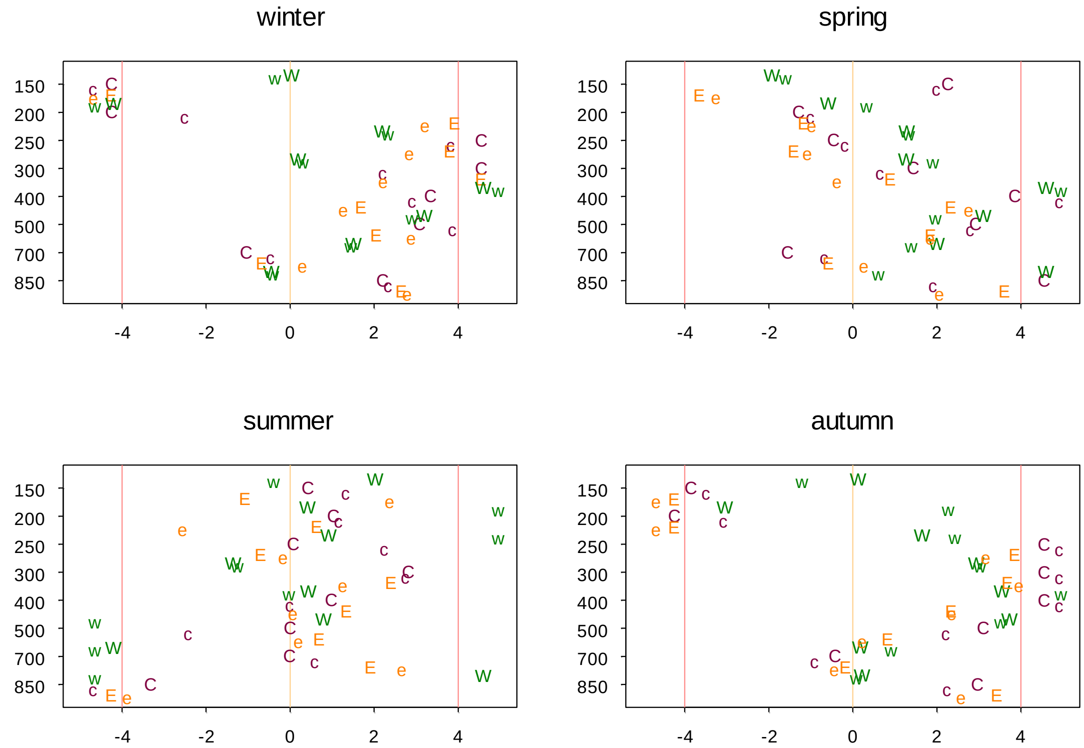
Caption: The differences in the absolute values of wind bias are presented for three regions: Conus (C for physics 1 and c for physics 2), Conus West (W for physics 1 and w for physics 2)and Conus East (E for physics 1 and e for physics 2). Model bias for each level may be either positive or negative. The absolute value of the bias means that the results can be interpreted more directly. If the magnitude of the bias of ARW is larger, the results will be on the right side of the plot. If the magnitude of the bias of NMM is larger, the results will be on the left side of the plot. The accompanying error bars represent a 95% confidence limit on the difference of the wind bias. Results are presented for eight pressure levels, with the three regions slightly offset vertically for clarity.



Capital letters are for phase 1; Weatherhead and Noonan Mon Jul 24 17:57:33 MST 2006

Caption: The number of standard errors represents how strong the statistical significance is between ARW bias and NMM bias. The values are presented for three regions: Conus (C for physics 1 and c for physics 2), Conus West (W for physics 1 and w for physics 2)and Conus East (E for physics 1 and e for physics 2). When the significance is more than four standard errors, either positively or negatively, the symbol for that region is listed in the end region. If there is little significant difference between the two models, the symbol for that region will appear near zero. If the statistical analysis indicates that it is highly likely that the bias of ARW is larger than NMM and the symbol will appear more to the right. Similarly, if the statistical analysis indicates that it is highly likely that the bias of NMM is larger than ARM, the symbol will appear on the left. The actual difference in bias may be very small. This plot only indicates how confident we are that the difference is significant in a statistical sense.

Wind bias Differences 12h (ARW-NMM) Z-val



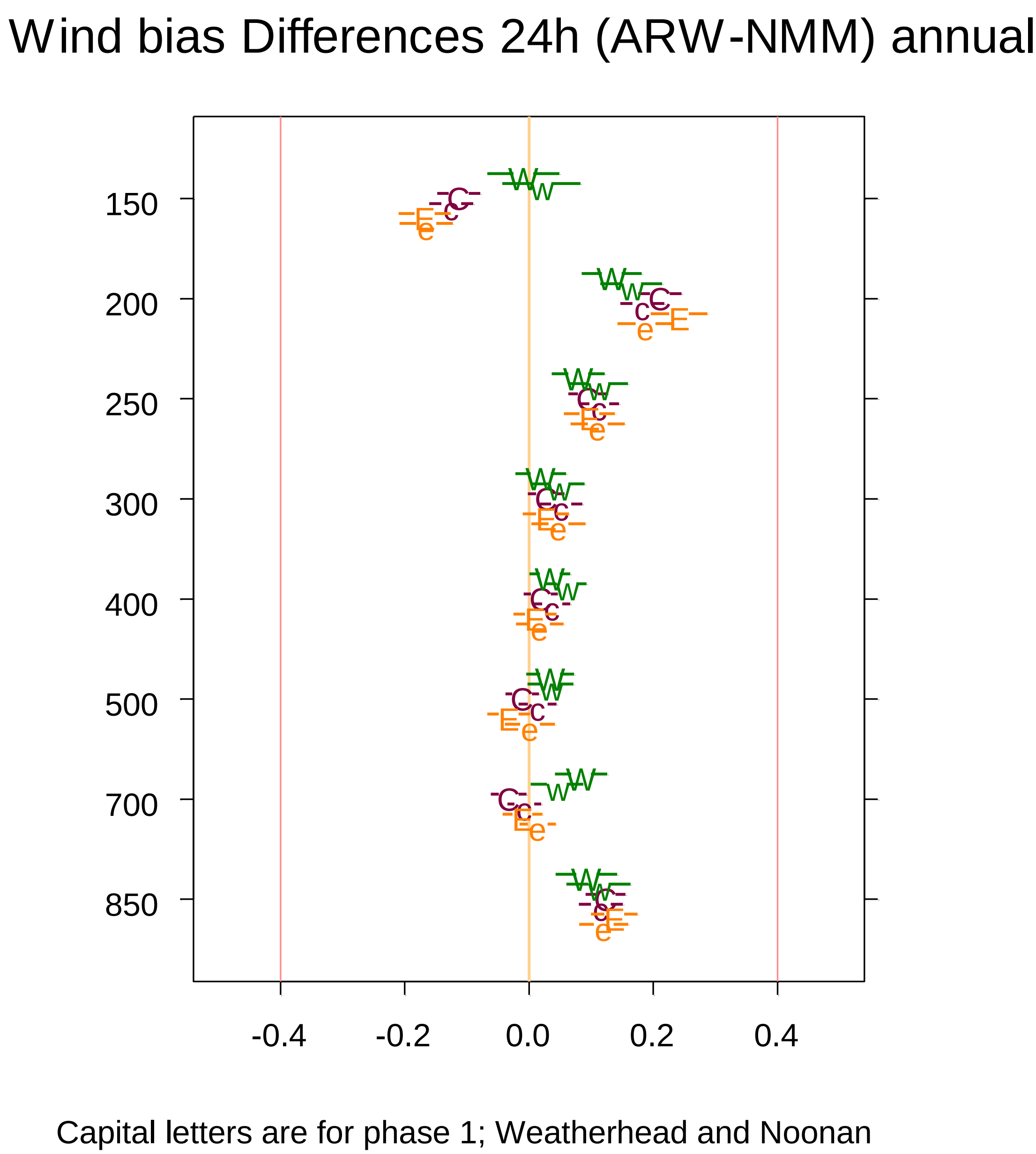
Caption: The number of standard errors represents how strong the statistical significance is between ARW bias and NMM bias. The values are presented for three regions: Conus (C for physics 1 and c for physics 2), Conus West (W for physics 1 and w for physics 2)and Conus East (E for physics 1 and e for physics 2). When the significance is more than four standard errors, either positively or negatively, the symbol for that region is listed in the end region. If there is little significant difference between the two models, the symbol for that region will appear near zero. If the statistical analysis indicates that it is highly likely that the bias of ARW is larger than NMM and the symbol will appear more to the right. Similarly, if the statistical analysis indicates that it is highly likely that the bias of NMM is larger than ARM, the symbol will appear on the left. The actual difference in bias may be very small. This plot only indicates how confident we are that the difference is significant in a statistical sense.

Bias for Sonde Vector Wind / Conus / Forecast Hour 12

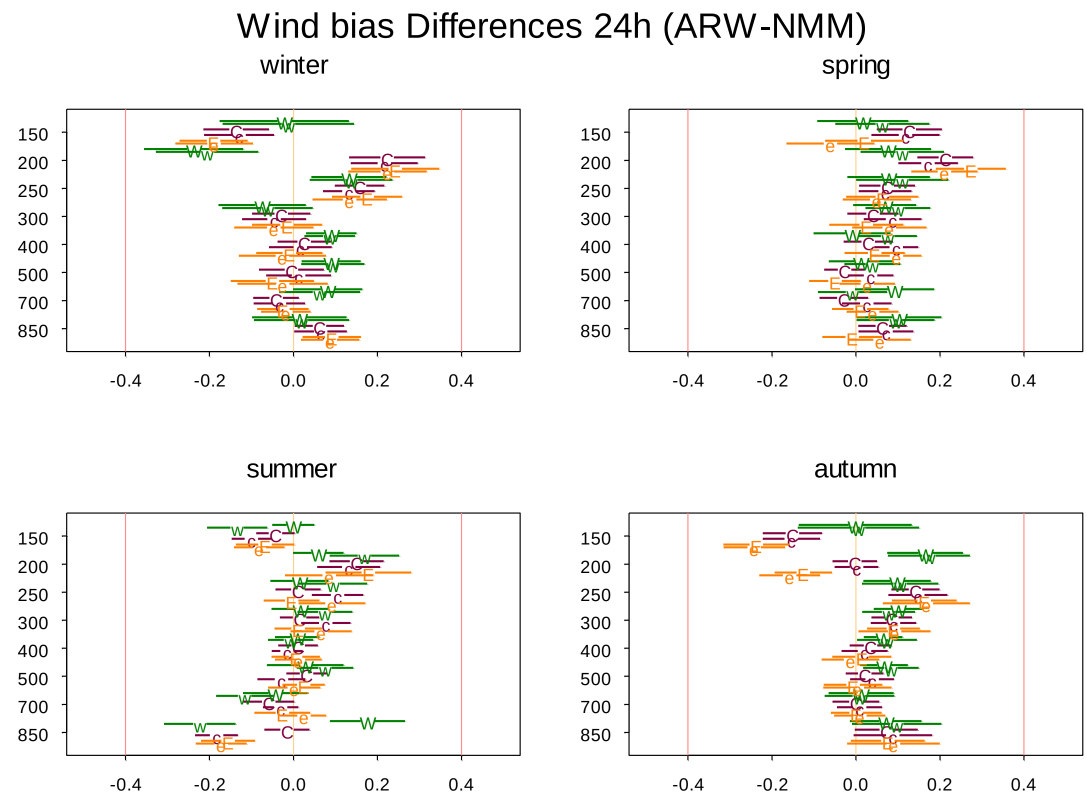
ANNUAL

	ARW.ph1	NMM.ph1	ARW.ph2	NMM.ph2	ph1.diff	ph2.diff	# of Obs
P150	0.22571	0.33075	0.20528	0.30265	-0.10504	-0.09737	67.97377
sigma	0.08350	0.07815	0.08415	0.07782	0.02002	0.02128	232.00000
p-val	2.70321	4.23217	2.43961	3.88902	-5.24823	-4.57548	232.00000
P200	-0.13707	-0.04004	-0.25614	-0.16375	0.09703	0.09239	69.61385
sigma	0.06414	0.06789	0.06552	0.06971	0.01947	0.01797	232.00000
p-val	-2.13704	-0.58984	-3.90927	-2.34891	4.98379	5.14148	232.00000
P250	-0.38248	-0.32379	-0.49747	-0.42816	0.05870	0.06930	71.76810
sigma	0.05986	0.06104	0.06307	0.06607	0.01524	0.01411	232.00000
p-val	-6.38964	-5.30485	-7.88799	-6.48066	3.85173	4.91005	232.00000
P300	-0.41026	-0.33222	-0.43206	-0.36019	0.07804	0.07186	72.14900
sigma	0.04969	0.04862	0.05085	0.05084	0.01151	0.01381	232.00000
p-val	-8.25656	-6.83312	-8.49675	-7.08431	6.78279	5.20536	232.00000
P400	-0.41423	-0.34967	-0.38102	-0.31818	0.06456	0.06284	73.86569
sigma	0.04211	0.03925	0.04182	0.03894	0.01047	0.01159	232.00000
p-val	-9.83779	-8.90851	-9.10998	-8.17031	6.16670	5.42222	232.00000
P500	-0.40718	-0.36420	-0.38749	-0.34786	0.04297	0.03963	72.83945
sigma	0.04492	0.04137	0.04133	0.03770	0.00951	0.01073	232.00000
p-val	-9.06380	-8.80402	-9.37452	-9.22589	4.51637	3.69226	232.00000
P700	-0.19616	-0.21522	-0.21950	-0.22845	-0.01906	-0.00895	73.26128
sigma	0.03382	0.03312	0.03427	0.03386	0.01256	0.01255	232.00000
p-val	-5.79960	-6.49896	-6.40499	-6.74654	-1.51755	-0.71343	232.00000
P800	0.19711	0.08023	0.08993	-0.00707	0.11688	0.08286	67.86464
sigma	0.03531	0.03810	0.03789	0.04174	0.01331	0.01325	232.00000
p-val	5.58204	2.10586	2.37380	-0.16935	8.78205	6.25225	232.00000

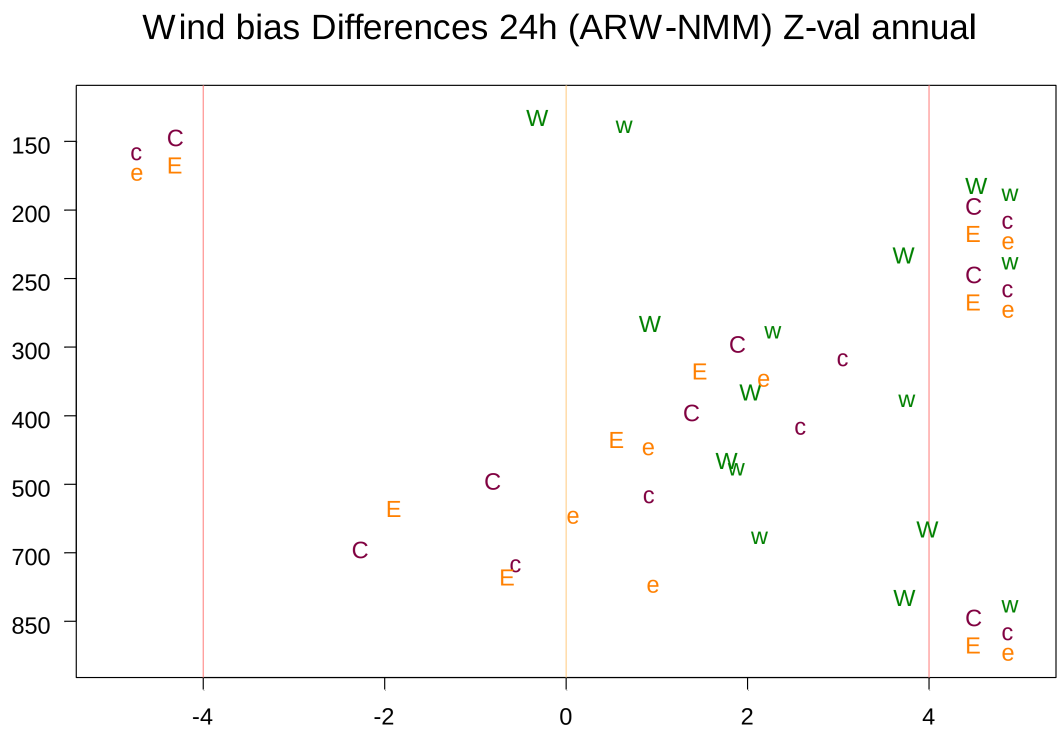
Bias for Sonde Vector Wind – Forecast Hour 24



Caption: The differences in the absolute values of wind bias are presented for three regions: Conus (C for physics 1 and c for physics 2), Conus West (W for physics 1 and w for physics 2)and Conus East (E for physics 1 and e for physics 2). Model bias for each level may be either positive or negative. The absolute value of the bias means that the results can be interpreted more directly. If the magnitude of the bias of ARW is larger, the results will be on the right side of the plot. If the magnitude of the bias of NMM is larger, the results will be on the left side of the plot. The accompanying error bars represent a 95% confidence limit on the difference of the wind bias. Results are presented for eight pressure levels, with the three regions slightly offset vertically for clarity.
(For comparison: see 24h aircraft bias on page 29)

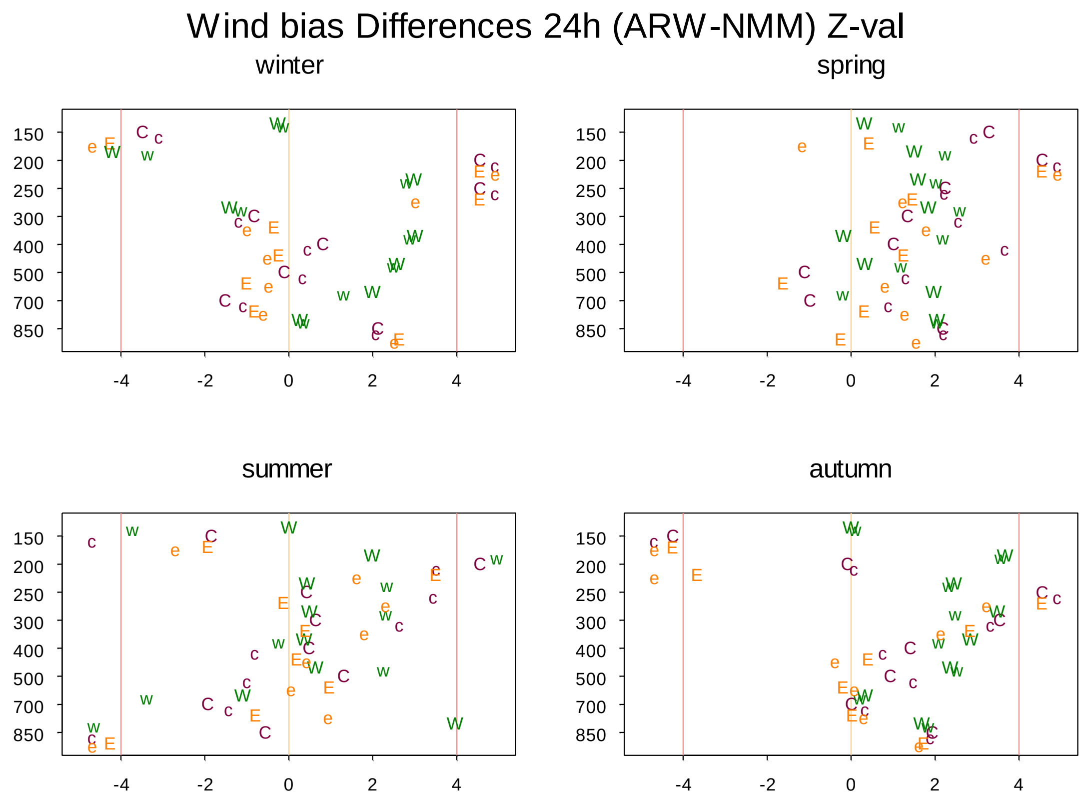


Caption: The differences in the absolute values of wind bias are presented for three regions: Conus (C for physics 1 and c for physics 2), Conus West (W for physics 1 and w for physics 2)and Conus East (E for physics 1 and e for physics 2). Model bias for each level may be either positive or negative. The absolute value of the bias means that the results can be interpreted more directly. If the magnitude of the bias of ARW is larger, the results will be on the right side of the plot. If the magnitude of the bias of NMM is larger, the results will be on the left side of the plot. The accompanying error bars represent a 95% confidence limit on the difference of the wind bias. Results are presented for eight pressure levels, with the three regions slightly offset vertically for clarity.



Capital letters are for phase 1; Weatherhead and Noonan Mon Jul 24 17:57:33 MST 2006

Caption: The number of standard errors represents how strong the statistical significance is between ARW bias and NMM bias. The values are presented for three regions: Conus (C for physics 1 and c for physics 2), Conus West (W for physics 1 and w for physics 2)and Conus East (E for physics 1 and e for physics 2). When the significance is more than four standard errors, either positively or negatively, the symbol for that region is listed in the end region. If there is little significant difference between the two models, the symbol for that region will appear near zero. If the statistical analysis indicates that it is highly likely that the bias of ARW is larger than NMM and the symbol will appear more to the right. Similarly, if the statistical analysis indicates that it is highly likely that the bias of NMM is larger than ARM, the symbol will appear on the left. The actual difference in bias may be very small. This plot only indicates how confident we are that the difference is significant in a statistical sense.



Caption: The number of standard errors represents how strong the statistical significance is between ARW bias and NMM bias. The values are presented for three regions: Conus (C for physics 1 and c for physics 2), Conus West (W for physics 1 and w for physics 2)and Conus East (E for physics 1 and e for physics 2). When the significance is more than four standard errors, either positively or negatively, the symbol for that region is listed in the end region. If there is little significant difference between the two models, the symbol for that region will appear near zero. If the statistical analysis indicates that it is highly likely that the bias of ARW is larger than NMM and the symbol will appear more to the right. Similarly, if the statistical analysis indicates that it is highly likely that the bias of NMM is larger than ARM, the symbol will appear on the left. The actual difference in bias may be very small. This plot only indicates how confident we are that the difference is significant in a statistical sense.

Bias for Sonde Vector Wind / Conus / Forecast Hour 24

ANNUAL

	ARW.ph1	NMM.ph1	ARW.ph2	NMM.ph2	ph1.diff	ph2.diff	# of Obs
P150	0.08704	0.20012	0.14922	0.27434	-0.11308	-0.12512	67.91186
sigma	0.08479	0.08350	0.08612	0.08435	0.01741	0.01771	232.00000
p-val	1.02660	2.39674	1.73258	3.25255	-6.49657	-7.06693	232.00000
P200	-0.43261	-0.22173	-0.47909	-0.29640	0.21088	0.18269	69.55404
sigma	0.07367	0.07896	0.08073	0.08544	0.01745	0.01771	232.00000
p-val	-5.87215	-2.80814	-5.93448	-3.46894	12.08378	10.31644	232.00000
P250	-0.50433	-0.40964	-0.60260	-0.48937	0.09469	0.11323	71.70724
sigma	0.06914	0.06800	0.07245	0.07536	0.01556	0.01584	232.00000
p-val	-7.29402	-6.02441	-8.31770	-6.49350	6.08426	7.14701	232.00000
P300	-0.42535	-0.39762	-0.48510	-0.43321	0.02773	0.05189	72.05456
sigma	0.05694	0.05793	0.05687	0.05966	0.01467	0.01701	232.00000
p-val	-7.47076	-6.86379	-8.52931	-7.26136	1.89064	3.05067	232.00000
P400	-0.31882	-0.29975	-0.33675	-0.29934	0.01907	0.03741	73.82162
sigma	0.05222	0.05062	0.05090	0.05043	0.01377	0.01448	232.00000
p-val	-6.10589	-5.92160	-6.61612	-5.93614	1.38545	2.58368	232.00000
P500	-0.21001	-0.22094	-0.24543	-0.23150	-0.01093	0.01393	72.79853
sigma	0.05245	0.05272	0.04999	0.05114	0.01356	0.01524	232.00000
p-val	-4.00368	-4.19087	-4.90939	-4.52699	-0.80628	0.91393	232.00000
P700	-0.09897	-0.13162	-0.14731	-0.15488	-0.03265	-0.00756	73.14376
sigma	0.03912	0.03939	0.03835	0.03830	0.01439	0.01361	232.00000
p-val	-2.52965	-3.34117	-3.84165	-4.04408	-2.26852	-0.55571	232.00000
P800	0.27222	0.14892	0.18570	0.06983	0.12330	0.11587	67.79748
sigma	0.04340	0.04534	0.04824	0.05405	0.01610	0.01770	232.00000
p-val	6.27204	3.28429	3.84927	1.29203	7.65609	6.54423	232.00000

Appendix 4: Summary of Wind Bias using Aircraft Data

Wind results are compared here for both the ARW and NMM core with 2 sets of physics, Phase-1 and Phase-2. Three regions: CONUS, CONUS-West, and CONUS-East along with two forecast hours (12h and 24h) are analyzed separately for comparison.

The wind bias is generally consistent from 850mb through 200mb with NMM showing a smaller magnitude of bias for these levels. At 150mb, the ARW model shows a smaller magnitude of bias. These results are somewhat seasonally dependent, particular for the behavior above 300mb. For Winter and Autumn, the ARW model shows a much smaller magnitude of bias than the NMM model.

This section consists of two subsections. The first subsection addresses the 12 hour forecasts while the second subsection addresses 24 hour forecasts. Initialization values are addressed in Appendix 7. Each subsection begins with a series of summary plots followed by tables of values. The tables are organized by region: CONUS, CONUS West, and CONUS East. For each region, The first table, labeled “annual,” gives results using all of the data; this table is followed by four tables giving the results when using only data for each of the four seasons.

We calculate a mean derived from daily (0Z and 12Z) bias calculations, we calculate both the Standard Error (sigma) and a Pair-Wise Comparison (Z-val) on our results. In columns 5 and 6, we show the results of a comparison made on bias, sigma, and Z-val for both cores with the same physics packages.

The following tables summarize results of bias calculations performed on the wind values generated for the WRF comparison. A description of each column is as follows:

Column 1:   The first cell of the first column describes which season is examined: Annual means all available data are used (7-15-05 thru 4-25-06); Winter indicates data from 1-16-06 thru 2-15-06; Spring indicates data from 3-25-06 thru 4-25-06; Summer indicates data from 7-15-05 thru 8-15-05; and Autumn indicates data from 11-1-05 thru 11-30-05. The remainder of Column 1 helps describe the data and are grouped to describe the results for each pressure level. Eight pressure levels are examined: 150mb, 200mb, 250mb, 300mb, 400mb; 500mb, 700mb, 850mb. For each pressure level there are three rows of information in the tables. The first row gives the mean estimate; the second row gives the standard error of the estimate and the third row gives the Z-value which is the mean divided by the standard error. More information on the statistical techniques used are available in the main document and appendices 13 and 14.

Column 2: ARW.ph1   The mean of the ARW bias using Phase-1 physics. 00Z and 12Z forecasts were used. Only forecasts for which both ARW and NMM bias values for Phase-1 physics were available were included. Below the mean estimate is the standard error on the estimate; below the standard error is the associated Z-value obtained by dividing the mean by the standard deviation. This is calculated and recorded for all eight pressure levels.

Column 3: NMM.ph1   The mean of the NMM bias using Phase-1 physics. 00Z and 12Z forecasts were used. Only forecasts for which both ARW and NMM bias values for Phase-1 physics were available were included. Below the mean estimate is the standard error on the estimate; below the standard error is the associated Z-value obtained by dividing the mean by the standard deviation. This is calculated and recorded for all eight pressure levels.

Column 4: ARW.ph2 The mean of the ARW bias using Phase-2 physics. 00Z and 12Z forecasts were used. Only forecasts for which both ARW and NMM bias values for Phase-2 physics were available were included. Below the mean estimate is the standard error on the estimate; below the standard error is the associated Z-value obtained by dividing the mean by the standard deviation. This is calculated and recorded for all eight pressure levels.

Column 5: NMM.ph2 The mean of the NMM bias using Phase-2 physics. 00Z and 12Z forecasts were used. Only forecasts for which both ARW and NMM bias values for Phase-2 physics were available were included. Below the mean estimate is the standard error on the estimate; below the standard error is the associated Z-value obtained by dividing the mean by the standard deviation. This is calculated and recorded for all eight pressure levels.

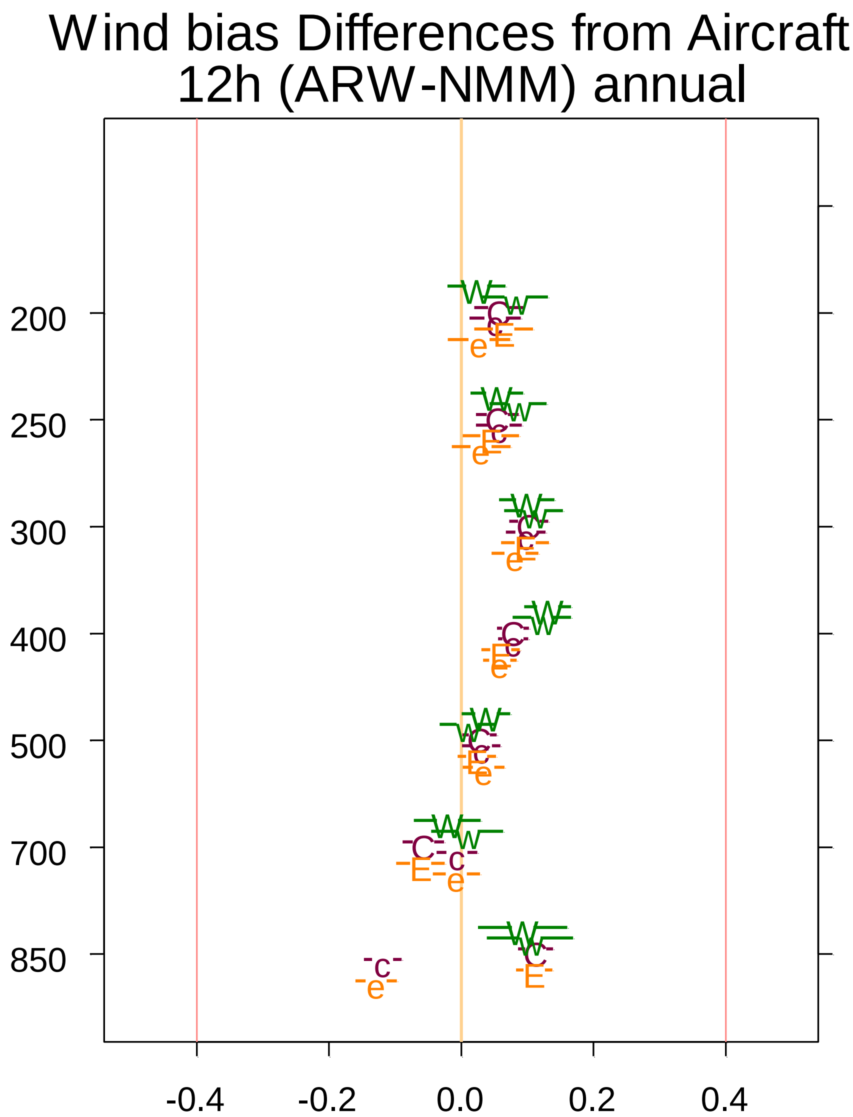
Column 6: Ph1.diff The mean of the pairwise differences of ARW bias and NMM bias using Phase-1 physics. Below this mean is the calculated standard error in the mean; below the standard error is the associated Z-value obtained by dividing the mean by the standard deviation.

Column 7: Ph2.diff The mean of the pairwise differences of ARW bias and NMM bias using Phase-2 physics. Below this mean is the calculated standard error in the mean; below the standard error is the associated Z-value obtained by dividing the mean by the standard deviation.

Column 8: # of Obs The total number of observations used in deriving the results in the prior columns. The first number is the number of forecasts used for the calculations. The second and third numbers are the average number of observations used to derive the individual forecast values for Phase-1 and Phase-2 respectively.

Bias for Aircraft Vector Wind – Forecast Hour 12

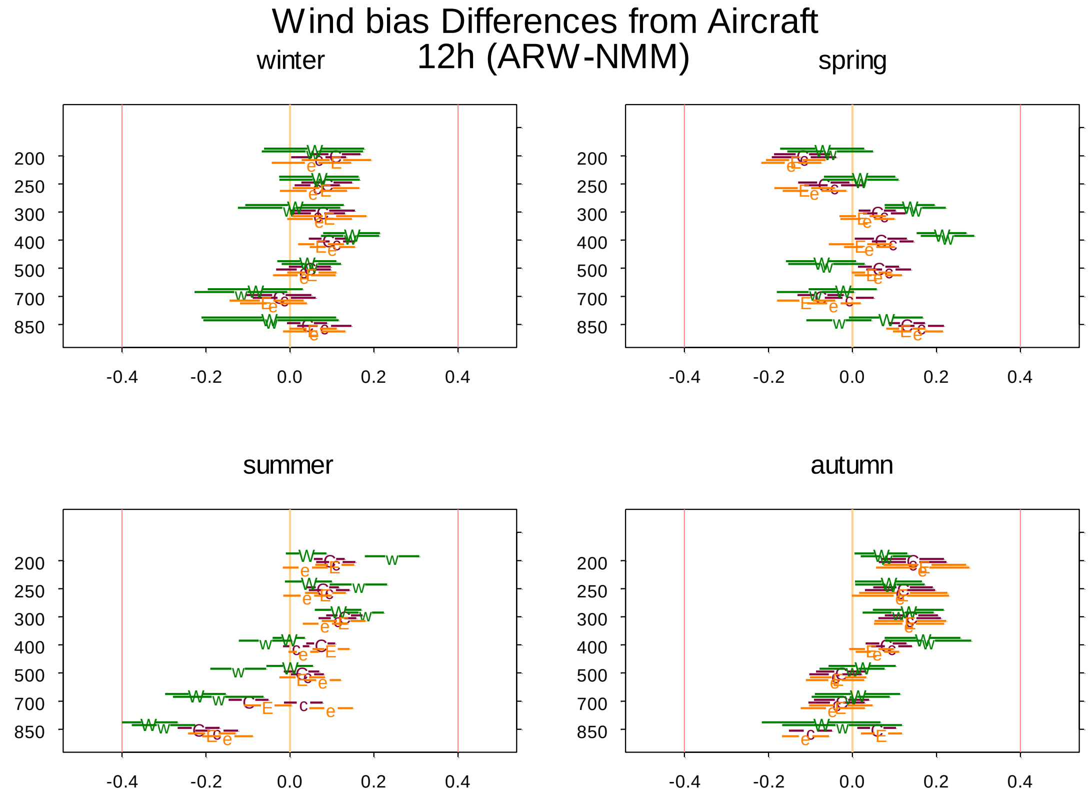
Similar calculations using sonde data show that for 200mb and below, the difference results give similar conclusions. The highest level reported for the aircraft results is 200-250mb, however sondes report up to 150mb. At 150mb, the sondes show a bias difference indicating that ARW has a smaller absolute value of bias.



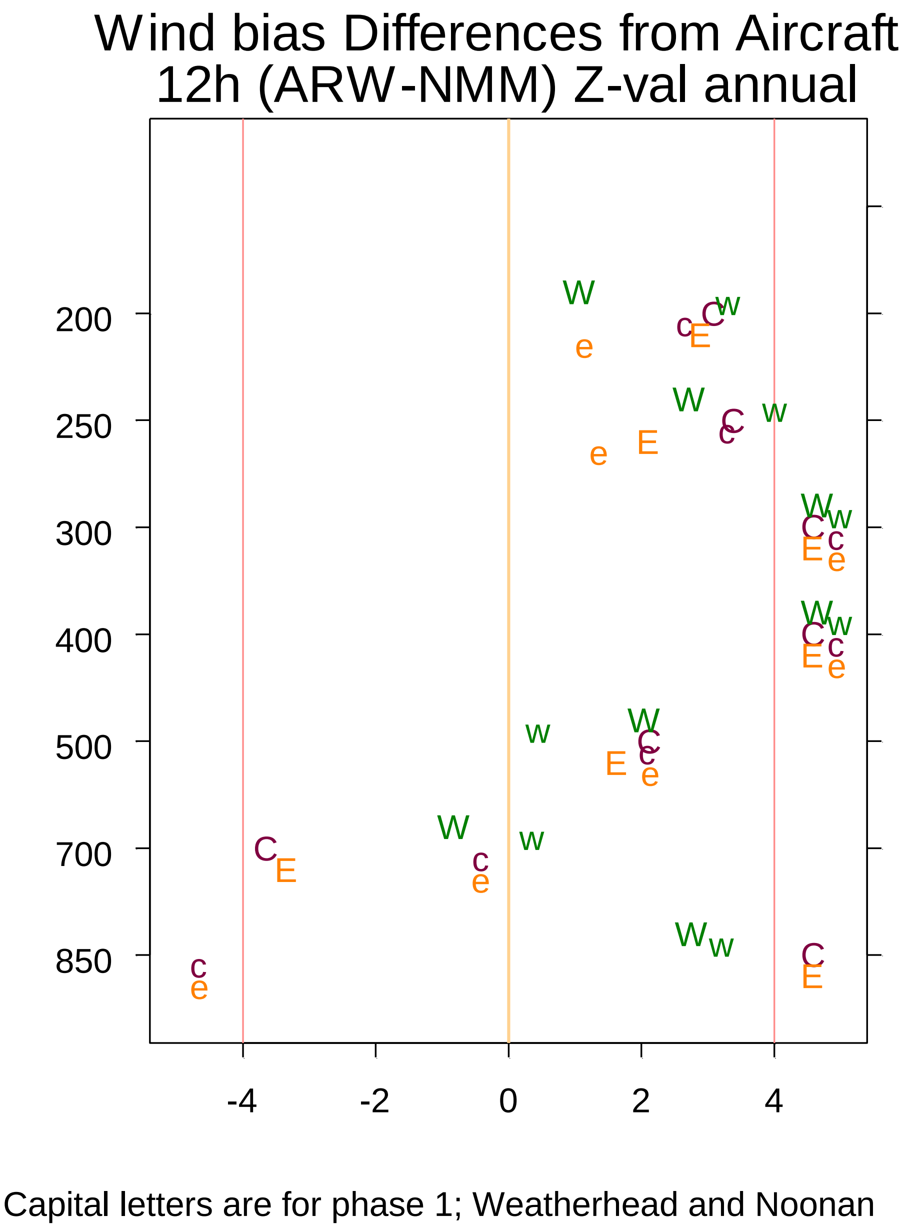
Capital letters are for phase 1; Weatherhead and Noonan

Caption: The differences in the absolute values of wind bias are presented for three regions: Conus (C for physics 1 and c for physics 2), Conus West (W for physics 1 and w for physics 2)and Conus East (E for physics 1 and e for physics 2). Model bias for each level may be either positive or negative. The absolute value of the bias means that the results can be interpreted more directly. If the magnitude of the bias of ARW is larger, the results will be on the right side of the plot. If the magnitude of the bias of NMM is larger, the results will be on the left side of the plot. The accompanying error bars represent a 95% confidence limit on the difference of the wind bias. Results are presented for seven pressure bins, with the three regions slightly offset vertically for clarity.

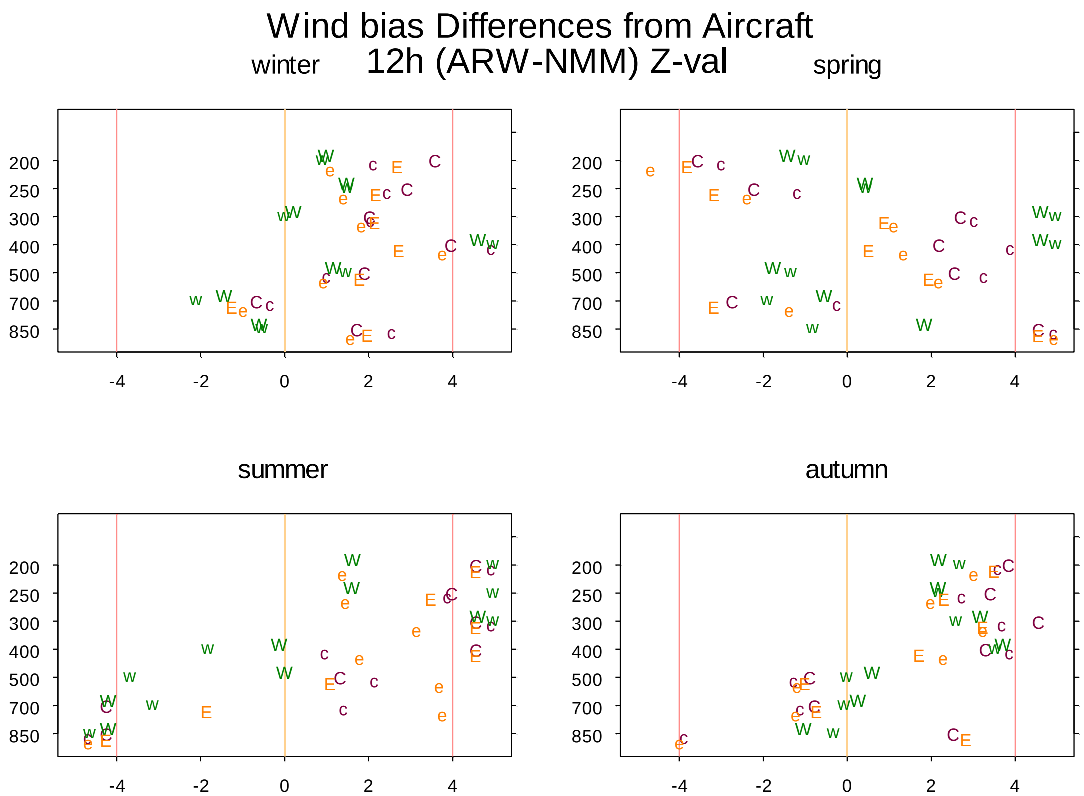
(For comparison: see 12h sonde bias on page 12)



Caption: The differences in the absolute values of wind bias are presented for three regions: Conus (C for physics 1 and c for physics 2), Conus West (W for physics 1 and w for physics 2)and Conus East (E for physics 1 and e for physics 2). Model bias for each level may be either positive or negative. The absolute value of the bias means that the results can be interpreted more directly. If the magnitude of the bias of ARW is larger, the results will be on the right side of the plot. If the magnitude of the bias of NMM is larger, the results will be on the left side of the plot. The accompanying error bars represent a 95% confidence limit on the difference of the wind bias. Results are presented for seven pressure bins, with the three regions slightly offset vertically for clarity.



Caption: The number of standard errors represents how strong the statistical significance is between ARW bias and NMM bias. The values are presented for three regions: Conus (C for physics 1 and c for physics 2), Conus West (W for physics 1 and w for physics 2)and Conus East (E for physics 1 and e for physics 2). When the significance is more than four standard errors, either positively or negatively, the symbol for that region is listed in the end region. If there is little significant difference between the two models, the symbol for that region will appear near zero. If the statistical analysis indicates that it is highly likely that the bias of ARW is larger than NMM and the symbol will appear more to the right. Similarly, if the statistical analysis indicates that it is highly likely that the bias of NMM is larger than ARM, the symbol will appear on the left. The actual difference in bias may be very small. This plot only indicates how confident we are that the difference is significant in a statistical sense.



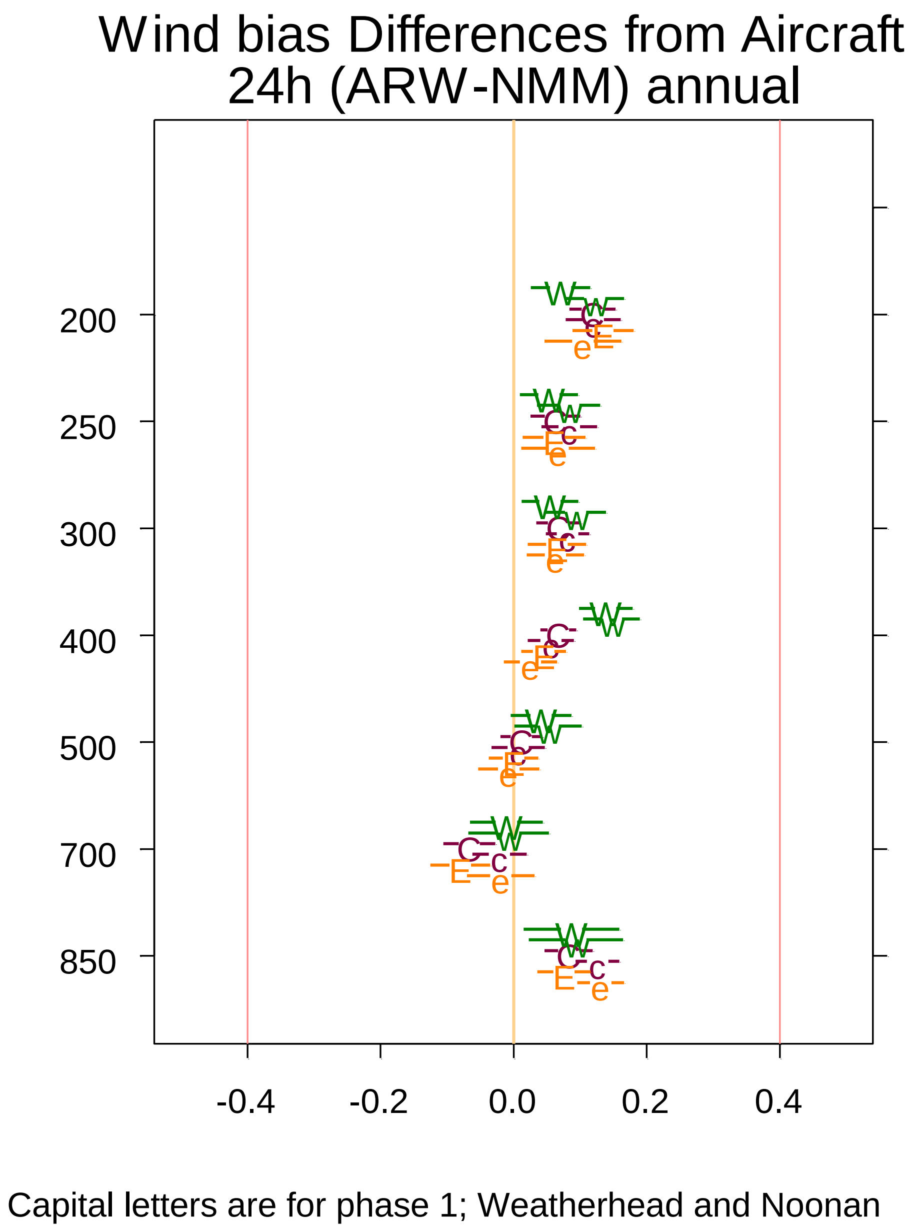
Caption: The number of standard errors represents how strong the statistical significance is between ARW bias and NMM bias. The values are presented for three regions: Conus (C for physics 1 and c for physics 2), Conus West (W for physics 1 and w for physics 2)and Conus East (E for physics 1 and e for physics 2). When the significance is more than four standard errors, either positively or negatively, the symbol for that region is listed in the end region. If there is little significant difference between the two models, the symbol for that region will appear near zero. If the statistical analysis indicates that it is highly likely that the bias of ARW is larger than NMM and the symbol will appear more to the right. Similarly, if the statistical analysis indicates that it is highly likely that the bias of NMM is larger than ARM, the symbol will appear on the left. The actual difference in bias may be very small. This plot only indicates how confident we are that the difference is significant in a statistical sense.

Bias for Aircraft Vector Wind / Conus / Forecast Hour 12

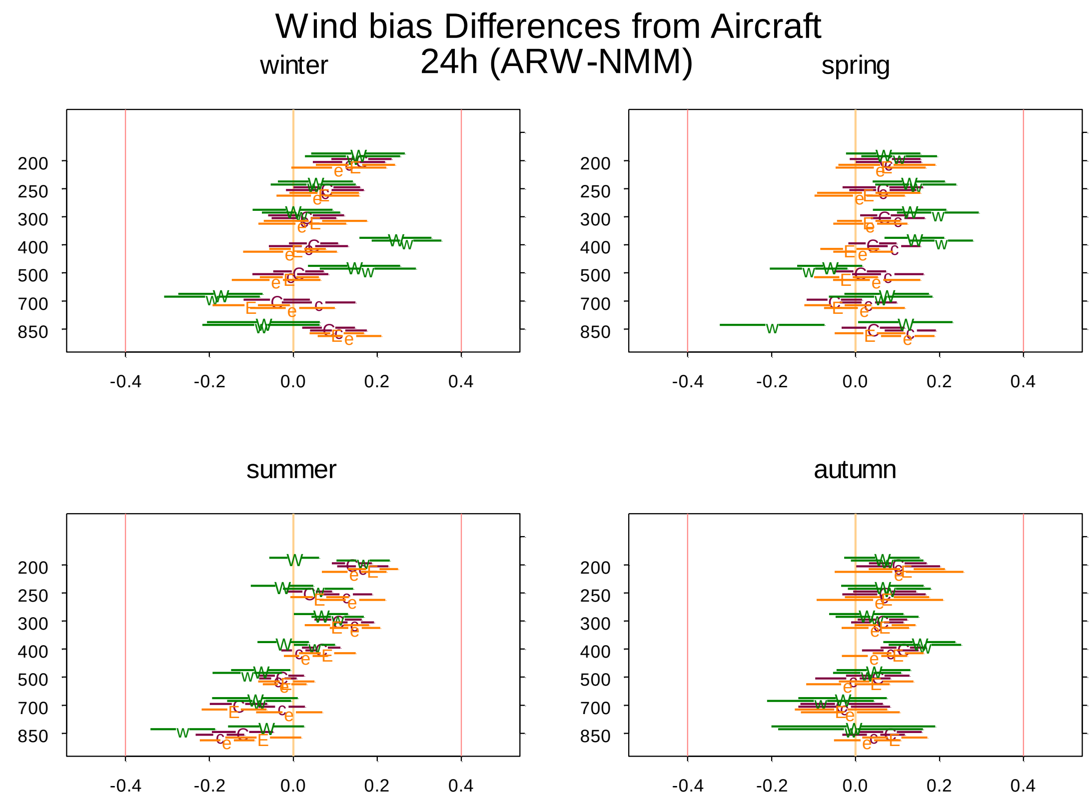
ANNUAL

	ARW.ph1	NMM.ph1	ARW.ph2	NMM.ph2	ph1.diff	ph2.diff	# of Obs
P200-250	-1.46112	-1.40429	-1.60300	-1.55175	0.05683	0.05126	1219.35789
sigma	0.08265	0.09020	0.08181	0.09143	0.01844	0.01933	232.00000
p-val	-17.67830	-15.56862	-19.59350	-16.97204	3.08221	2.65122	232.00000
P250-300	-1.32745	-1.27274	-1.43659	-1.37952	0.05472	0.05707	604.24000
sigma	0.07476	0.07762	0.07889	0.08298	0.01619	0.01735	232.00000
p-val	-17.75732	-16.39727	-18.20929	-16.62465	3.37866	3.29014	232.00000
P300-400	-0.75495	-0.65289	-0.72265	-0.62500	0.10206	0.09765	408.07789
sigma	0.06426	0.06790	0.06644	0.06969	0.01480	0.01503	232.00000
p-val	-11.74800	-9.61502	-10.87589	-8.96804	6.89408	6.49869	232.00000
P400-550	-0.48170	-0.40368	-0.43337	-0.35509	0.07802	0.07828	664.47789
sigma	0.05067	0.05018	0.05057	0.05076	0.01207	0.01142	232.00000
p-val	-9.50682	-8.04505	-8.56902	-6.99523	6.46134	6.85364	232.00000
P550-700	-0.31434	-0.28680	-0.30852	-0.27819	0.02755	0.03032	813.80632
sigma	0.04654	0.04772	0.04560	0.04752	0.01302	0.01452	232.00000
p-val	-6.75439	-6.01035	-6.76610	-5.85400	2.11614	2.08800	232.00000
P700-850	-0.24031	-0.29744	-0.30820	-0.31466	-0.05713	-0.00646	667.31789
sigma	0.04497	0.04839	0.04557	0.04750	0.01564	0.01538	232.00000
p-val	-5.34400	-6.14689	-6.76387	-6.62491	-3.65208	-0.42032	232.00000
P850-1000	0.15273	0.04014	0.00570	-0.12457	0.11259	-0.11887	936.29158
sigma	0.06463	0.06656	0.06666	0.06641	0.01324	0.01427	232.00000
p-val	2.36315	0.60308	0.08552	-1.87577	8.50138	-8.33244	232.00000

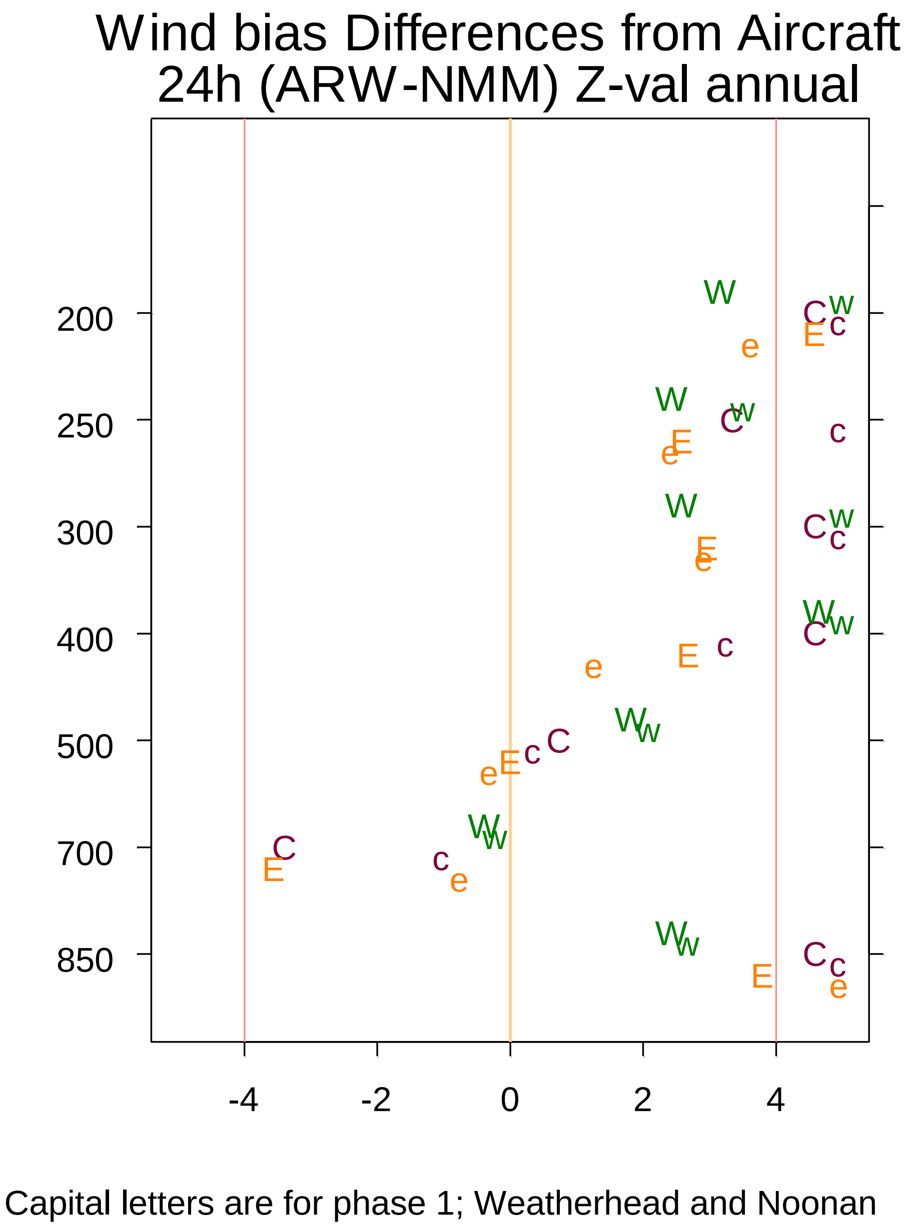
Bias for Aircraft Vector Wind – Forecast Hour 24



Caption: The differences in the absolute values of wind bias are presented for three regions: Conus (C for physics 1 and c for physics 2), Conus West (W for physics 1 and w for physics 2)and Conus East (E for physics 1 and e for physics 2). Model bias for each level may be either positive or negative. The absolute value of the bias means that the results can be interpreted more directly. If the magnitude of the bias of ARW is larger, the results will be on the right side of the plot. If the magnitude of the bias of NMM is larger, the results will be on the left side of the plot. The accompanying error bars represent a 95% confidence limit on the difference of the wind bias. Results are presented for seven pressure bins, with the three regions slightly offset vertically for clarity.
(For comparison: see 24h sonde bias on page 17)

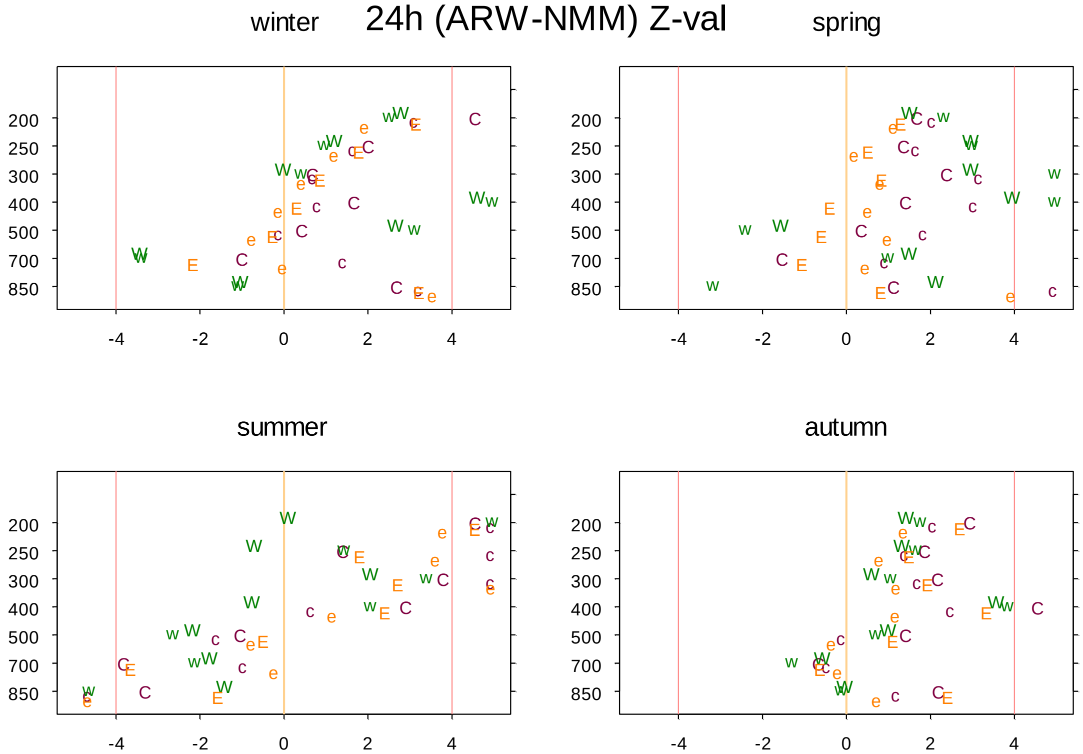


Caption: The differences in the absolute values of wind bias are presented for three regions: Conus (C for physics 1 and c for physics 2), Conus West (W for physics 1 and w for physics 2)and Conus East (E for physics 1 and e for physics 2). Model bias for each level may be either positive or negative. The absolute value of the bias means that the results can be interpreted more directly. If the magnitude of the bias of ARW is larger, the results will be on the right side of the plot. If the magnitude of the bias of NMM is larger, the results will be on the left side of the plot. The accompanying error bars represent a 95% confidence limit on the difference of the wind bias. Results are presented for seven pressure bins, with the three regions slightly offset vertically for clarity.



Caption: The number of standard errors represents how strong the statistical significance is between ARW bias and NMM bias. The values are presented for three regions: Conus (C for physics 1 and c for physics 2), Conus West (W for physics 1 and w for physics 2)and Conus East (E for physics 1 and e for physics 2). When the significance is more than four standard errors, either positively or negatively, the symbol for that region is listed in the end region. If there is little significant difference between the two models, the symbol for that region will appear near zero. If the statistical analysis indicates that it is highly likely that the bias of ARW is larger than NMM and the symbol will appear more to the right. Similarly, if the statistical analysis indicates that it is highly likely that the bias of NMM is larger than ARM, the symbol will appear on the left. The actual difference in bias may be very small. This plot only indicates how confident we are that the difference is significant in a statistical sense.

Wind bias Differences from Aircraft
24h (ARW-NMM) Z-val



Caption: The number of standard errors represents how strong the statistical significance is between ARW bias and NMM bias. The values are presented for three regions: Conus (C for physics 1 and c for physics 2), Conus West (W for physics 1 and w for physics 2)and Conus East (E for physics 1 and e for physics 2). When the significance is more than four standard errors, either positively or negatively, the symbol for that region is listed in the end region. If there is little significant difference between the two models, the symbol for that region will appear near zero. If the statistical analysis indicates that it is highly likely that the bias of ARW is larger than NMM and the symbol will appear more to the right. Similarly, if the statistical analysis indicates that it is highly likely that the bias of NMM is larger than ARM, the symbol will appear on the left. The actual difference in bias may be very small. This plot only indicates how confident we are that the difference is significant in a statistical sense.

Bias for Aircraft Vector Wind / Conus / Forecast Hour 24

ANNUAL

	ARW.ph1	NMM.ph1	ARW.ph2	NMM.ph2	ph1.diff	ph2.diff	# of Obs
P200-250	-1.70631	-1.58778	-1.77987	-1.66030	0.11852	0.11957	1212.70074
sigma	0.10481	0.10638	0.10600	0.10832	0.01728	0.02057	231.00000
p-val	-16.28079	-14.92564	-16.79125	-15.32714	6.86077	5.81194	231.00000
P250-300	-1.45979	-1.39709	-1.56030	-1.47669	0.06270	0.08360	606.03793
sigma	0.10501	0.10580	0.10727	0.10919	0.01878	0.02087	231.00000
p-val	-13.90152	-13.20494	-14.54610	-13.52439	3.33814	4.00602	231.00000
P300-400	-0.73977	-0.67209	-0.78804	-0.70704	0.06768	0.08100	410.30032
sigma	0.09305	0.09591	0.09453	0.09864	0.01672	0.01633	231.00000
p-val	-7.95005	-7.00774	-8.33644	-7.16762	4.04825	4.96125	231.00000
P400-550	-0.29441	-0.22720	-0.26844	-0.21243	0.06720	0.05601	666.47945
sigma	0.07130	0.07120	0.06958	0.06930	0.01350	0.01729	231.00000
p-val	-4.12934	-3.19094	-3.85811	-3.06550	4.97737	3.23843	231.00000
P550-700	-0.25899	-0.24750	-0.22442	-0.21764	0.01149	0.00678	819.10221
sigma	0.05735	0.05882	0.05702	0.06266	0.01571	0.02005	231.00000
p-val	-4.51576	-4.20757	-3.93579	-3.47348	0.73161	0.33805	231.00000
P700-850	-0.14750	-0.21397	-0.17159	-0.19286	-0.06647	-0.02127	671.24025
sigma	0.05046	0.05385	0.06031	0.06097	0.01956	0.02040	231.00000
p-val	-2.92280	-3.97365	-2.84493	-3.16333	-3.39872	-1.04293	231.00000
P850-1000	0.32248	0.23974	0.19330	0.06710	0.08274	0.12620	943.09378
sigma	0.07023	0.07268	0.07655	0.07843	0.01804	0.01641	231.00000
p-val	4.59170	3.29835	2.52500	0.85559	4.58594	7.68917	231.00000

Appendix 5: Summary of Wind RMSE Using Sonde Data

Wind results are compared here for both the ARW and NMM core with 2 sets of physics, Phase-1 and Phase-2. Three regions: CONUS, CONUS-West, and CONUS-East along with two forecast hours (12h and 24h) are analyzed separately for comparison.

For both the 12h and 24h forecasts, NMM exhibits a smaller RMSE for the lowest part of the atmosphere (between 850mb and 300m). Above 300mb, ARW has a smaller RMSE for all regions and all physics packages. This difference is most clear for the 12h forecast and becomes less clear with the 24h forecast, although both forecasts show similar patterns. These patterns are repeated seasonally, although at 850mb in Autumn and Winter, ARW has a smaller RMSE. These results are independent of physics package.

This section consists of two subsections. The first subsection addresses the 12 hour forecasts while the second subsection addresses 24 hour forecasts. Initialization values are addressed in Appendix 7. Each subsection begins with a series of summary plots followed by tables of values. The tables are organized by region: CONUS, CONUS West, and CONUS East. For each region, The first table, labeled "annual," gives results using all of the data; this table is followed by four tables giving the results when using only data for each of the four seasons.

We calculate a mean derived from daily (0Z and 12Z) Root Mean Square Error (RMSE) calculations, we calculate both the Standard Error (sigma) and a Pair-Wise Comparison (Z-val) on our results. In columns 5 and 6, we show the results of a comparison made on RMSE, sigma, and Z-val for both cores with the same physics packages.

The following tables summarize results of RMSE calculations performed on the wind values generated for the WRF comparison. A description of each column is as follows:

Column 1: The first cell of the first column describes which season is examined: Annual means all available data are used (7-15-05 thru 4-25-06); Winter indicates data from 1-16-06 thru 2-15-06; Spring indicates data from 3-25-06 thru 4-25-06; Summer indicates data from 7-15-05 thru 8-15-05; and Autumn indicates data from 11-1-05 thru 11-30-05. The remainder of Column 1 helps describe the data and are grouped to describe the results for each pressure level. Eight pressure levels are examined: 150mb, 200mb, 250mb, 300mb, 400mb; 500mb, 700mb, 850mb. For each pressure level there are three rows of information in the tables. The first row gives the mean estimate; the second row gives the standard error of the estimate and the third row gives the Z-value which is the mean divided by the standard error. More information on the statistical techniques used are available in the main document and appendices 13 and 14.

Column 2: ARW.ph1 The mean of the ARW RMSE using Phase-1 physics. 00Z and 12Z forecasts were used. Only forecasts for which both ARW and NMM RMSE values for Phase-1 physics were available were included. Below the mean estimate is the standard error on the estimate; below the standard error is the associated Z-value obtained by dividing the mean by the standard deviation. This is calculated and recorded for all eight pressure levels.

Column 3: NMM.ph1 The mean of the NMM RMSE using Phase-1 physics. 00Z and 12Z forecasts were used. Only forecasts for which both ARW and NMM RMSE values for Phase-1 physics were available were included. Below the mean estimate is the standard error on the estimate; below the standard error is the associated Z-value obtained by dividing the mean by the standard deviation. This is calculated and recorded for all eight pressure levels.

Column 4: ARW.ph2 The mean of the ARW RMSE using Phase-2 physics. 00Z and 12Z forecasts were used. Only forecasts for which both ARW and NMM RMSE values for Phase-2 physics were available were included. Below the mean estimate is the standard error on the estimate; below the standard error is the associated Z-value obtained by dividing the mean by the standard deviation. This is calculated and recorded for all eight pressure levels.

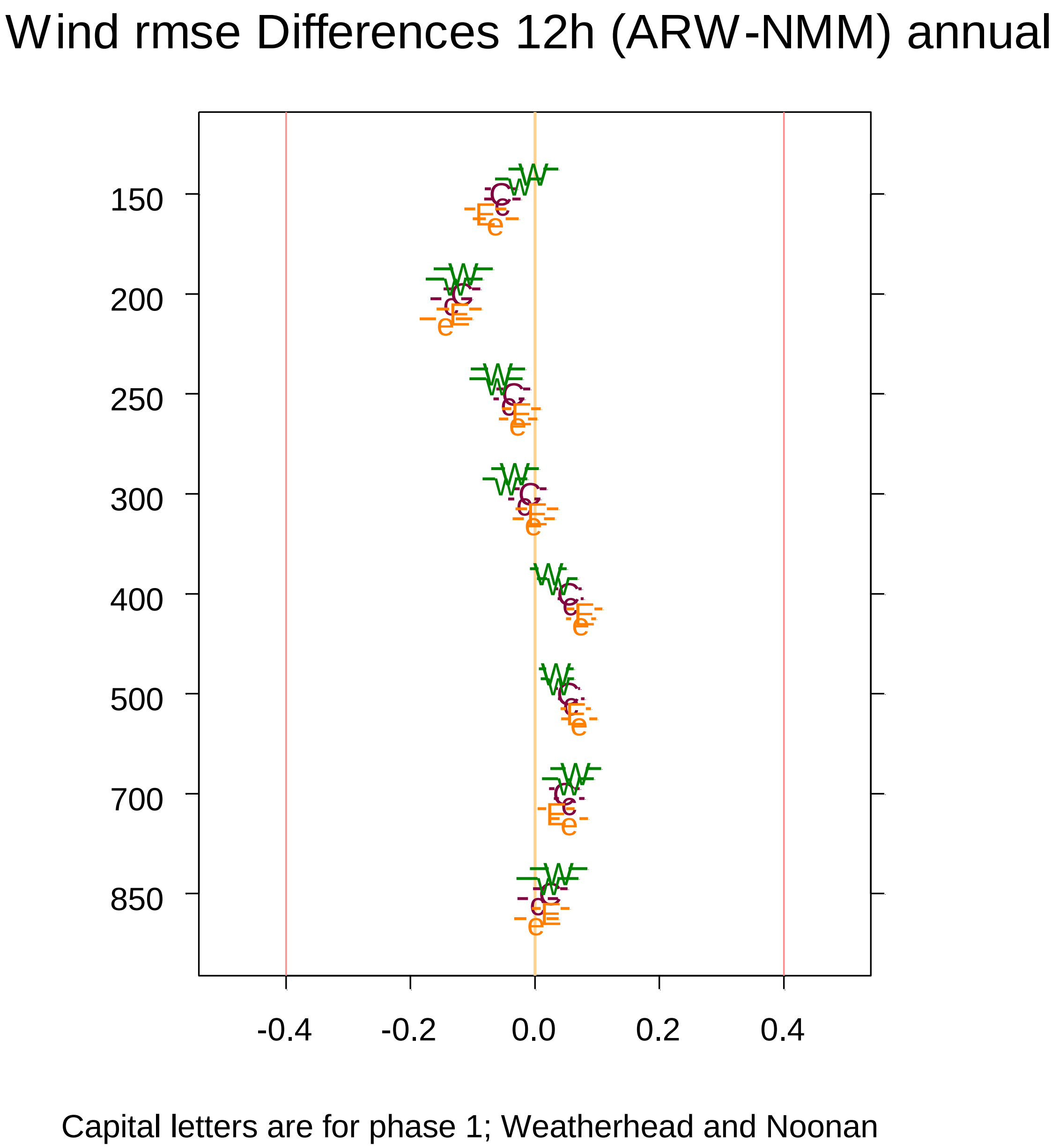
Column 5: NMM.ph2 The mean of the NMM RMSE using Phase-2 physics. 00Z and 12Z forecasts were used. Only forecasts for which both ARW and NMM RMSE values for Phase-2 physics were available were included. Below the mean estimate is the standard error on the estimate; below the standard error is the associated Z-value obtained by dividing the mean by the standard deviation. This is calculated and recorded for all eight pressure levels.

Column 6: Ph1.diff The mean of the pairwise differences of ARW RMSE and NMM RMSE using Phase-1 physics. Below this mean is the calculated standard error in the mean; below the standard error is the associated Z-value obtained by dividing the mean by the standard deviation.

Column 7: Ph2.diff The mean of the pairwise differences of ARW RMSE and NMM RMSE using Phase-2 physics. Below this mean is the calculated standard error in the mean; below the standard error is the associated Z-value obtained by dividing the mean by the standard deviation.

Column 8: # of Obs The total number of observations used in deriving the results in the prior columns. The first number is the number of forecasts used for the calculations. The second and third numbers are the average number of observations used to derive the individual forecast values for Phase-1 and Phase-2 respectively.

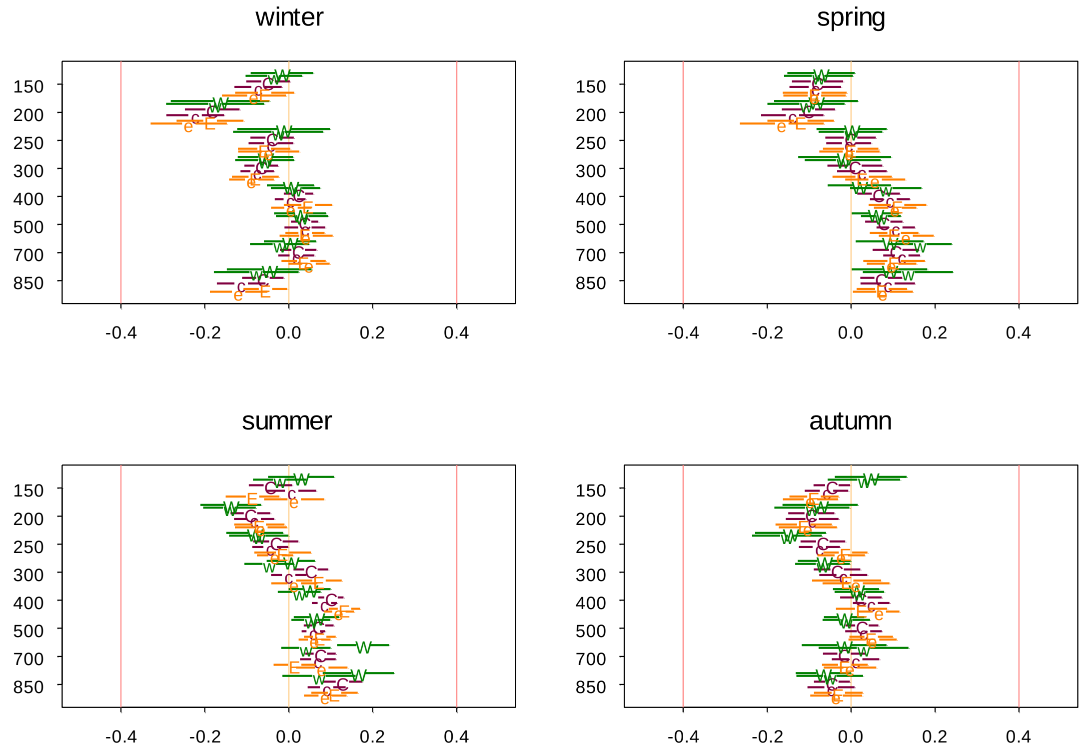
RMSE for Sonde Vector Wind – Forecast Hour 12



Caption: The differences in wind rmse are presented for three regions: Conus (C for physics 1 and c for physics 2), Conus West (W for physics 1 and w for physics 2)and Conus East (E for physics 1 and e for physics 2). Model rmse is always positive. If the rmse of ARW is larger, the results will be on the right side of the plot. If the rmse of NMM is larger, the results will be on the left side of the plot. The accompanying error bars represent a 95% confidence limit on the difference of the wind rmse. Results are presented for eight pressure levels, with the three regions slightly offset vertically for clarity.

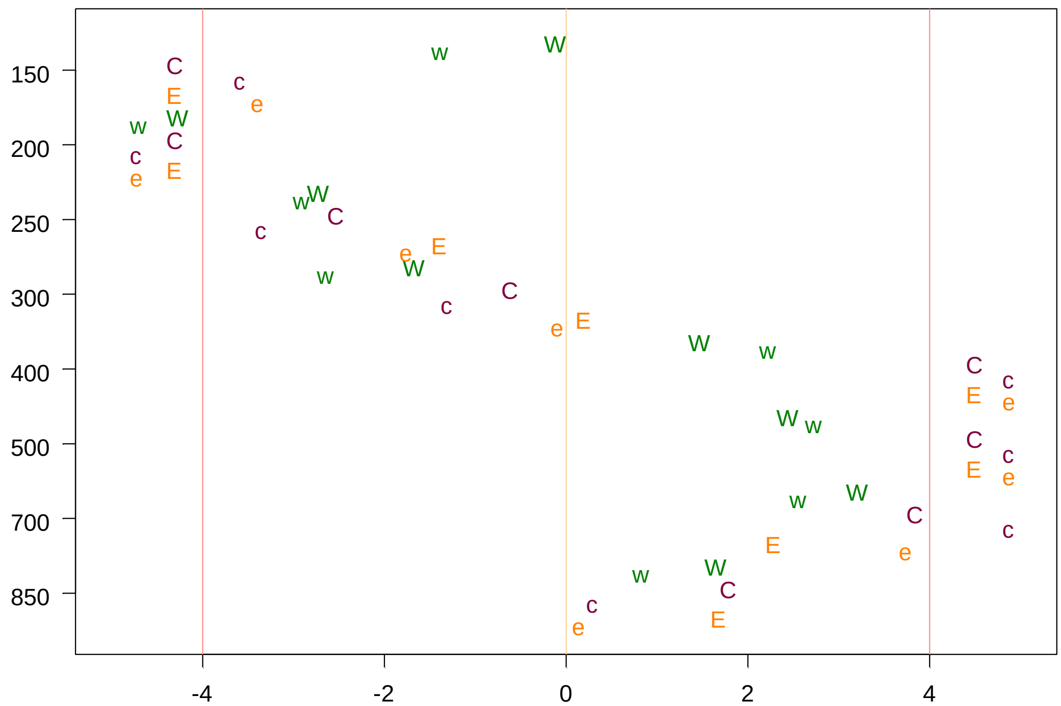
(For comaparison: see 12h aircraft RMSE on page 48)

Wind rmse Differences 12h (ARW-NMM)



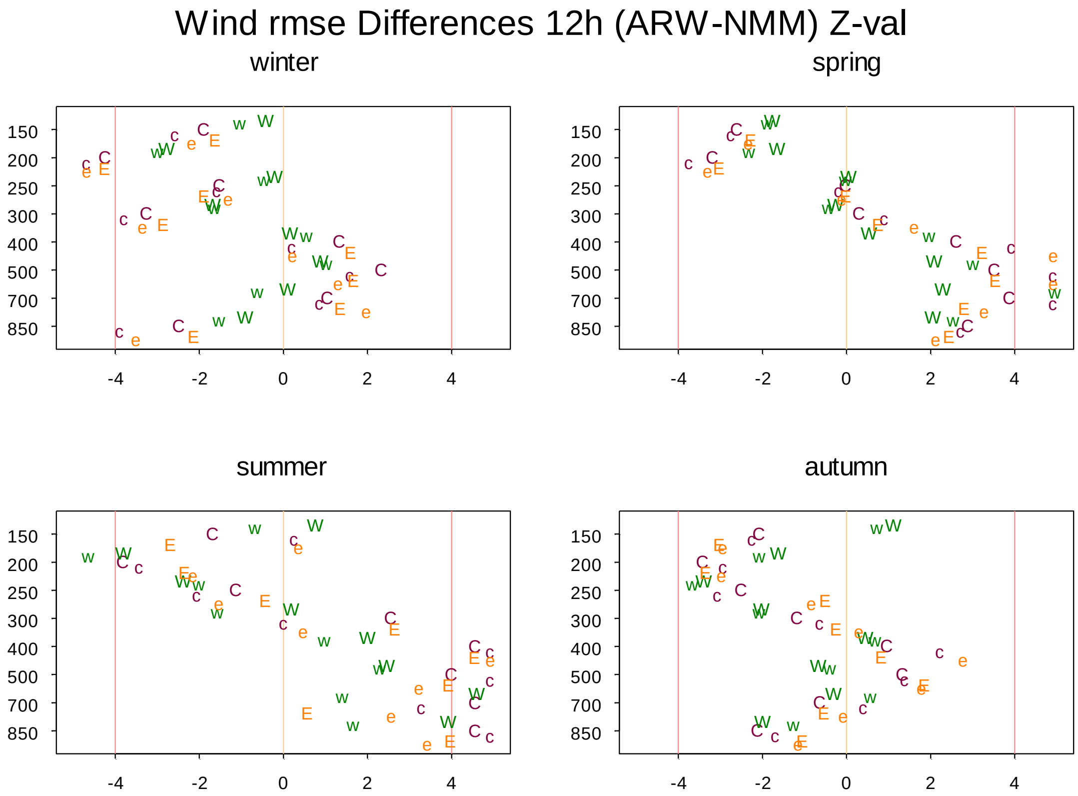
Caption: The differences in wind rmse are presented for three regions: Conus (C for physics 1 and c for physics 2), Conus West (W for physics 1 and w for physics 2)and Conus East (E for physics 1 and e for physics 2). Model rmse is always positive. If the rmse of ARW is larger, the results will be on the right side of the plot. If the rmse of NMM is larger, the results will be on the left side of the plot. The accompanying error bars represent a 95% confidence limit on the difference of the wind rmse. Results are presented for eight pressure levels, with the three regions slightly offset vertically for clarity.

Wind rmse Differences 12h (ARW-NMM) Z-val annual



Capital letters are for phase 1; Weatherhead and Noonan Mon Jul 24 00:04:12 MST 2006

Caption: The number of standard errors represents how strong the statistical significance is between ARW rmse and NMM rmse. The values are presented for three regions: Conus (C for physics 1 and c for physics 2), Conus West (W for physics 1 and w for physics 2)and Conus East (E for physics 1 and e for physics 2). When the significance is more than four standard errors, either positively or negatively, the symbol for that region is listed in the end region. If there is little significant difference between the two models, the symbol for that region will appear near zero. If the statistical analysis indicates that it is highly likely that the rmse of ARW is larger than NMM and the symbol will appear more to the right. Similarly, if the statistical analysis indicates that it is highly likely that the rmse of NMM is larger than ARM, the symbol will appear on the left. The actual difference in rmse may be very small. This plot only indicates how confident we are that the difference is significant in a statistical sense.



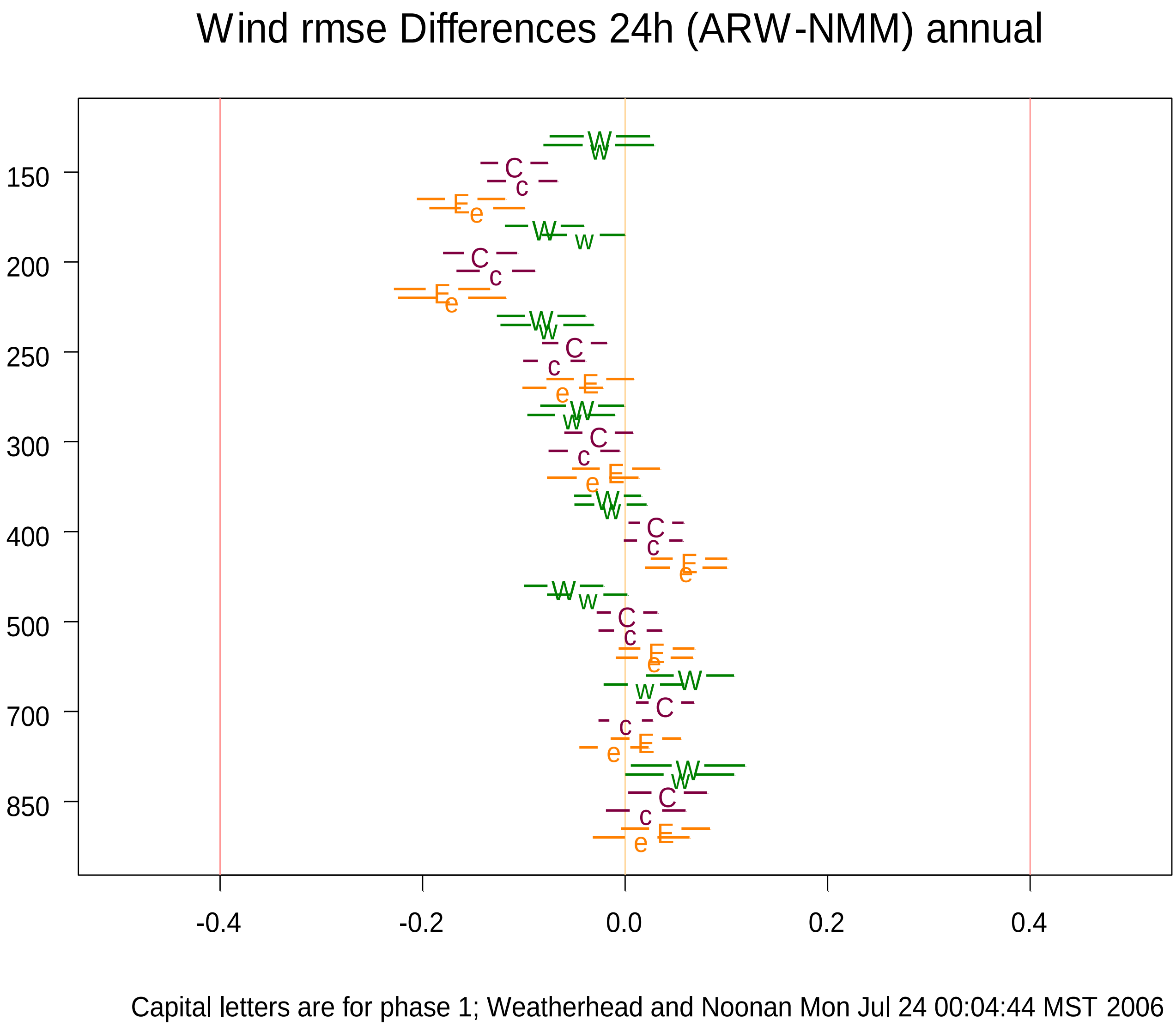
Caption: The number of standard errors represents how strong the statistical significance is between ARW rmse and NMM rmse. The values are presented for three regions: Conus (C for physics 1 and c for physics 2), Conus West (W for physics 1 and w for physics 2)and Conus East (E for physics 1 and e for physics 2). When the significance is more than four standard errors, either positively or negatively, the symbol for that region is listed in the end region. If there is little significant difference between the two models, the symbol for that region will appear near zero. If the statistical analysis indicates that it is highly likely that the rmse of ARW is larger than NMM and the symbol will appear more to the right. Similarly, if the statistical analysis indicates that it is highly likely that the rmse of NMM is larger than ARM, the symbol will appear on the left. The actual difference in rmse may be very small. This plot only indicates how confident we are that the difference is significant in a statistical sense.

RMSE for Sonde Vector Wind / Conus / Forecast Hour 12

ANNUAL

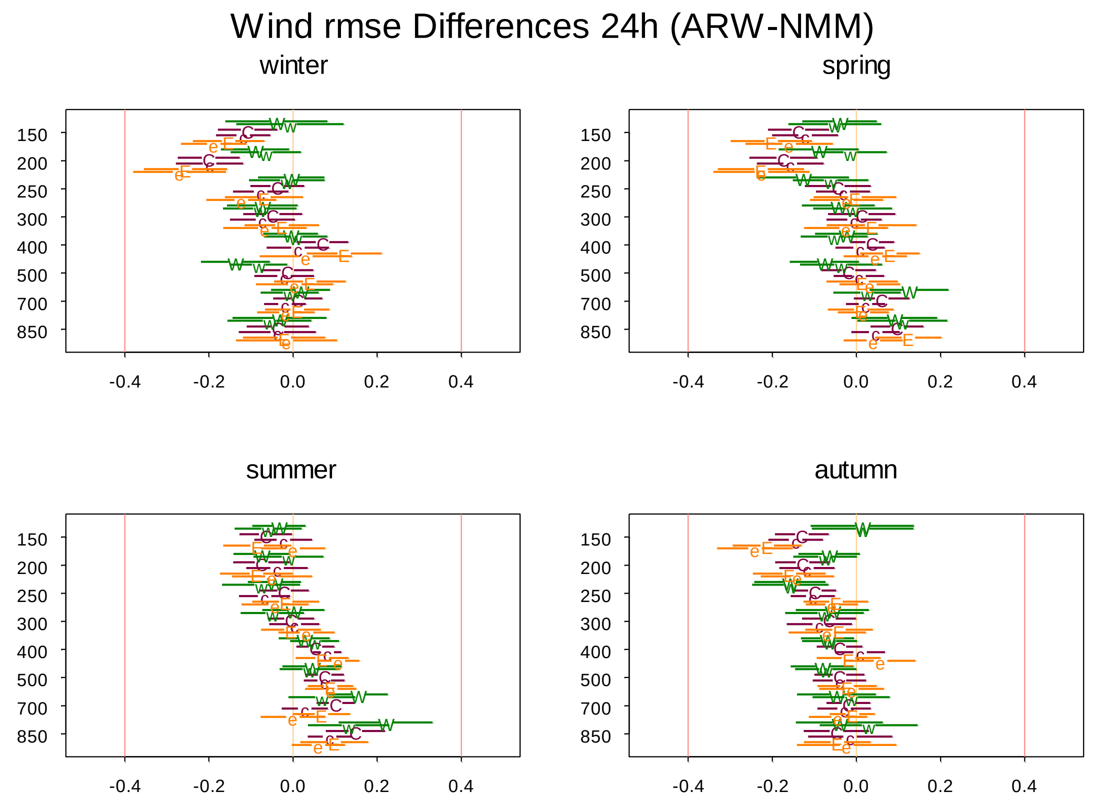
	ARW.ph1	NMM.ph1	ARW.ph2	NMM.ph2	ph1.diff	ph2.diff	# of Obs
P150	5.76002	5.81458	5.71933	5.77151	-0.05456	-0.05218	67.97377
sigma	0.09325	0.09577	0.09739	0.10171	0.01302	0.01451	232.00000
p-val	61.76881	60.71308	58.72875	56.74701	-4.19208	-3.59657	232.00000
P200	6.29986	6.41688	6.21881	6.35331	-0.11702	-0.13450	69.61385
sigma	0.07902	0.08274	0.07761	0.08407	0.01474	0.01667	232.00000
p-val	79.72206	77.55490	80.13070	75.56733	-7.93933	-8.06692	232.00000
P250	6.33551	6.37014	6.31829	6.36013	-0.03463	-0.04184	71.76810
sigma	0.12434	0.12540	0.13299	0.13169	0.01366	0.01245	232.00000
p-val	50.95308	50.79802	47.51128	48.29799	-2.53484	-3.35987	232.00000
P300	6.04231	6.05058	6.02172	6.03868	-0.00827	-0.01695	72.14900
sigma	0.12351	0.13030	0.12263	0.12559	0.01337	0.01291	232.00000
p-val	48.92244	46.43740	49.10352	48.08346	-0.61863	-1.31321	232.00000
P400	5.33949	5.28561	5.31596	5.25859	0.05389	0.05738	73.86569
sigma	0.11182	0.11431	0.11166	0.11340	0.01074	0.01040	232.00000
p-val	47.75234	46.23940	47.60983	46.37042	5.01704	5.51492	232.00000
P500	4.89171	4.83833	4.87855	4.82016	0.05338	0.05839	72.83945
sigma	0.07149	0.07452	0.07410	0.07643	0.00964	0.01063	232.00000
p-val	68.42943	64.92861	65.83932	63.06721	5.53871	5.49131	232.00000
P700	4.54878	4.50133	4.51893	4.46397	0.04745	0.05496	73.26128
sigma	0.05660	0.05850	0.05954	0.06033	0.01237	0.01234	232.00000
p-val	80.36737	76.95184	75.89654	73.99398	3.83772	4.45495	232.00000
P800	4.38317	4.35839	4.42960	4.42489	0.02478	0.00471	67.86464
sigma	0.04364	0.04681	0.04615	0.04944	0.01390	0.01635	232.00000
p-val	100.44679	93.11234	95.98822	89.49923	1.78282	0.28797	232.00000

RMSE for Sonde Vector Wind – Forecast Hour 24

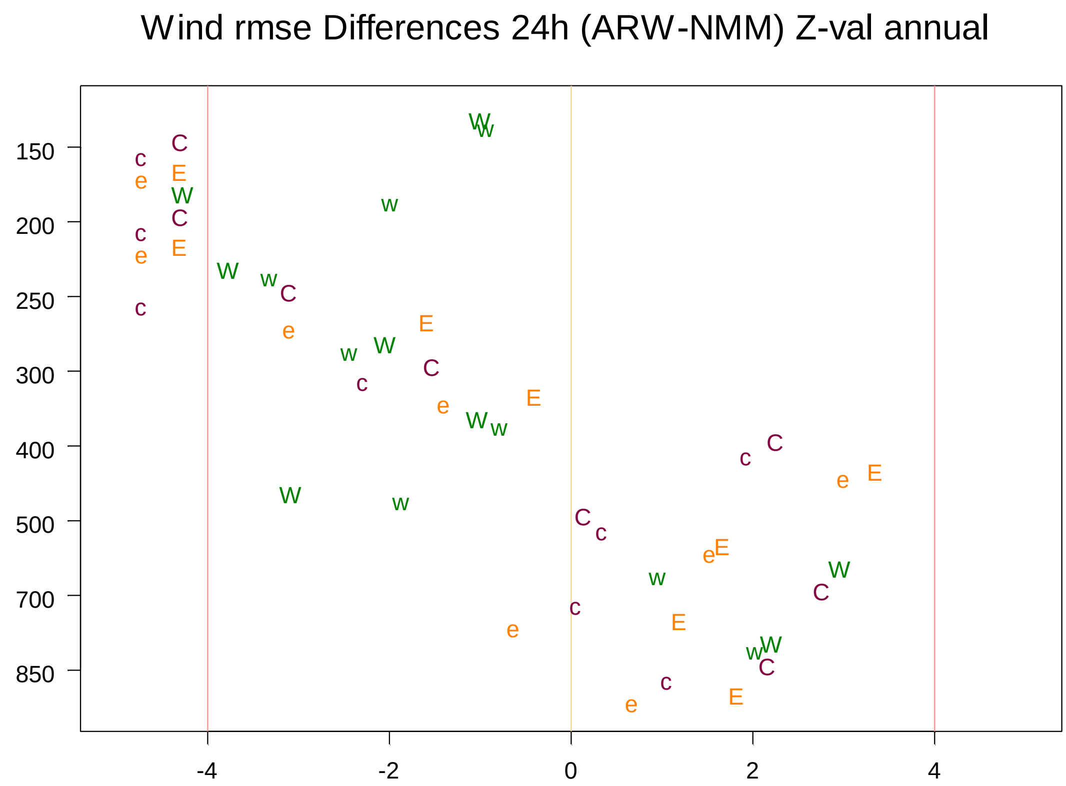


Caption: The differences in wind rmse are presented for three regions: Conus (C for physics 1 and c for physics 2), Conus West (W for physics 1 and w for physics 2)and Conus East (E for physics 1 and e for physics 2). Model rmse is always positive. If the rmse of ARW is larger, the results will be on the right side of the plot. If the rmse of NMM is larger, the results will be on the left side of the plot. The accompanying error bars represent a 95% confidence limit on the difference of the wind rmse. Results are presented for eight pressure levels, with the three regions slightly offset vertically for clarity.

(For comparison: see 24h aircraft RMSE on page 53)



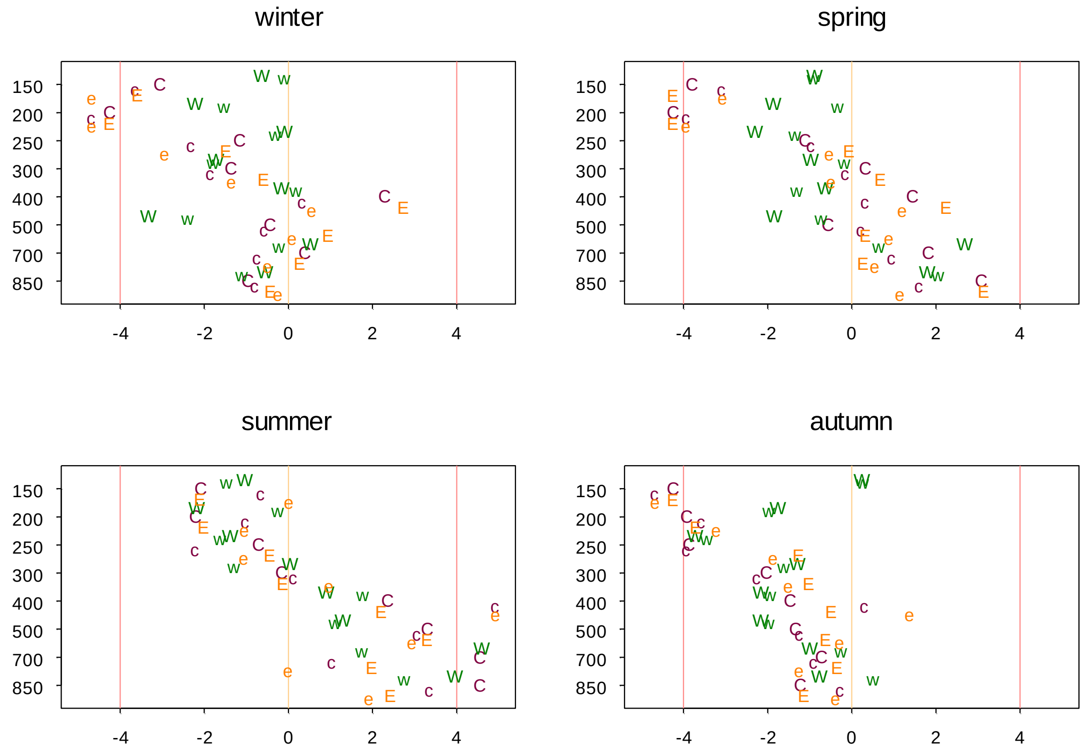
Caption: The differences in wind rmse are presented for three regions: Conus (C for physics 1 and c for physics 2), Conus West (W for physics 1 and w for physics 2)and Conus East (E for physics 1 and e for physics 2). Model rmse is always positive. If the rmse of ARW is larger, the results will be on the right side of the plot. If the rmse of NMM is larger, the results will be on the left side of the plot. The accompanying error bars represent a 95% confidence limit on the difference of the wind rmse. Results are presented for eight pressure levels, with the three regions slightly offset vertically for clarity.



Capital letters are for phase 1; Weatherhead and Noonan Mon Jul 24 00:05:46 MST 2006

Caption: The number of standard errors represents how strong the statistical significance is between ARW rmse and NMM rmse. The values are presented for three regions: Conus (C for physics 1 and c for physics 2), Conus West (W for physics 1 and w for physics 2)and Conus East (E for physics 1 and e for physics 2). When the significance is more than four standard errors, either positively or negatively, the symbol for that region is listed in the end region. If there is little significant difference between the two models, the symbol for that region will appear near zero. If the statistical analysis indicates that it is highly likely that the rmse of ARW is larger than NMM and the symbol will appear more to the right. Similarly, if the statistical analysis indicates that it is highly likely that the rmse of NMM is larger than ARM, the symbol will appear on the left. The actual difference in rmse may be very small. This plot only indicates how confident we are that the difference is significant in a statistical sense.

Wind rmse Differences 24h (ARW-NMM) Z-val



Caption: The number of standard errors represents how strong the statistical significance is between ARW rmse and NMM rmse. The values are presented for three regions: Conus (C for physics 1 and c for physics 2), Conus West (W for physics 1 and w for physics 2)and Conus East (E for physics 1 and e for physics 2). When the significance is more than four standard errors, either positively or negatively, the symbol for that region is listed in the end region. If there is little significant difference between the two models, the symbol for that region will appear near zero. If the statistical analysis indicates that it is highly likely that the rmse of ARW is larger than NMM and the symbol will appear more to the right. Similarly, if the statistical analysis indicates that it is highly likely that the rmse of NMM is larger than ARM, the symbol will appear on the left. The actual difference in rmse may be very small. This plot only indicates how confident we are that the difference is significant in a statistical sense.

RMSE for Sonde Vector Wind / Conus / Forecast Hour 24

ANNUAL

	ARW.ph1	NMM.ph1	ARW.ph2	NMM.ph2	ph1.diff	ph2.diff	# of Obs
P150	6.12744	6.23684	6.17358	6.27504	-0.10940	-0.10146	67.91186
sigma	0.10245	0.10395	0.10169	0.10309	0.01669	0.01735	232.00000
p-val	59.80674	60.00094	60.70763	60.86771	-6.55526	-5.84789	232.00000
P200	7.11178	7.25492	7.10607	7.23359	-0.14314	-0.12753	69.55404
sigma	0.09139	0.09701	0.09505	0.10139	0.01831	0.01943	232.00000
p-val	77.81683	74.78900	74.76280	71.34377	-7.81581	-6.56280	232.00000
P250	7.65155	7.70145	7.63434	7.70420	-0.04991	-0.06986	71.70724
sigma	0.15948	0.16098	0.16598	0.16541	0.01605	0.01534	232.00000
p-val	47.97763	47.84191	45.99545	46.57682	-3.10974	-4.55391	232.00000
P300	7.47166	7.49773	7.50318	7.54359	-0.02607	-0.04041	72.05456
sigma	0.19721	0.19914	0.19634	0.19730	0.01696	0.01759	232.00000
p-val	37.88721	37.64979	38.21505	38.23493	-1.53669	-2.29783	232.00000
P400	6.49079	6.46018	6.51170	6.48384	0.03061	0.02786	73.82162
sigma	0.18510	0.18342	0.18560	0.18767	0.01362	0.01449	232.00000
p-val	35.06652	35.22151	35.08536	34.54973	2.24662	1.92224	232.00000
P500	5.67329	5.67133	5.68493	5.67971	0.00196	0.00522	72.79853
sigma	0.11888	0.12365	0.12326	0.13107	0.01499	0.01570	232.00000
p-val	47.72417	45.86695	46.12221	43.33499	0.13067	0.33229	232.00000
P700	5.09312	5.05367	5.06037	5.05975	0.03945	0.00062	73.14376
sigma	0.06226	0.06555	0.06590	0.06912	0.01433	0.01342	232.00000
p-val	81.80357	77.09451	76.79026	73.20091	2.75359	0.04622	232.00000
P800	4.90816	4.86612	4.94883	4.92825	0.04204	0.02058	67.79748
sigma	0.05699	0.05868	0.05803	0.06344	0.01949	0.01964	232.00000
p-val	86.12862	82.92554	85.28472	77.68939	2.15694	1.04759	232.00000

Appendix 6: Summary of Wind RMSE Using Aircraft Data

Wind results are compared here for both the ARW and NMM core with 2 sets of physics, Phase-1 and Phase-2. Three regions: CONUS, CONUS-West, and CONUS-East along with two forecast hours (12h and 24h) are analyzed separately for comparison.

For both the 12h and 24h forecasts, NMM exhibits a smaller RMSE for the lowest part of the atmosphere (between 850mb and 300m). Above 300mb, ARW has a smaller RMSE for all regions and all physics packages. This difference is most clear for the 12h forecast and becomes less clear with the 24h forecast, although both forecasts show similar patterns. These patterns are repeated seasonally, although at 850mb in Autumn and Winter, ARW has a smaller RMSE. These results are independent of physics package.

This section consists of two subsections. The first subsection addresses the 12 hour forecasts while the second subsection addresses 24 hour forecasts. Initialization values are addressed in Appendix 7. Each subsection begins with a series of summary plots followed by tables of values. The tables are organized by region: CONUS, CONUS West, and CONUS East. For each region, The first table, labeled "annual," gives results using all of the data; this table is followed by four tables giving the results when using only data for each of the four seasons.

We calculate a mean derived from daily (0Z and 12Z) Root Mean Square Error (RMSE) calculations, we calculate both the Standard Error (sigma) and a Pair-Wise Comparison (Z-val) on our results. In columns 5 and 6, we show the results of a comparison made on RMSE, sigma, and Z-val for both cores with the same physics packages.

The following tables summarize results of RMSE calculations performed on the wind values generated for the WRF comparison. A description of each column is as follows:

Column 1: The first cell of the first column describes which season is examined: Annual means all available data are used (7-15-05 thru 4-25-06); Winter indicates data from 1-16-06 thru 2-15-06; Spring indicates data from 3-25-06 thru 4-25-06; Summer indicates data from 7-15-05 thru 8-15-05; and Autumn indicates data from 11-1-05 thru 11-30-05. The remainder of Column 1 helps describe the data and are grouped to describe the results for each pressure level. Eight pressure levels are examined: 150mb, 200mb, 250mb, 300mb, 400mb; 500mb, 700mb, 850mb. For each pressure level there are three rows of information in the tables. The first row gives the mean estimate; the second row gives the standard error of the estimate and the third row gives the Z-value which is the mean divided by the standard error. More information on the statistical techniques used are available in the main document and appendices 13 and 14.

Column 2: ARW.ph1 The mean of the ARW RMSE using Phase-1 physics. 00Z and 12Z forecasts were used. Only forecasts for which both ARW and NMM RMSE values for Phase-1 physics were available were included. Below the mean estimate is the standard error on the estimate; below the standard error is the associated Z-value obtained by dividing the mean by the standard deviation. This is calculated and recorded for all eight pressure levels.

Column 3: NMM.ph1 The mean of the NMM RMSE using Phase-1 physics. 00Z and 12Z forecasts were used. Only forecasts for which both ARW and NMM RMSE values for Phase-1 physics were available were included. Below the mean estimate is the standard error on the estimate; below the standard error is the associated Z-value obtained by dividing the mean by the standard deviation. This is calculated and recorded for all eight pressure levels.

Column 4: ARW.ph2 The mean of the ARW RMSE using Phase-2 physics. 00Z and 12Z forecasts were used. Only forecasts for which both ARW and NMM RMSE values for Phase-2 physics were available were included. Below the mean estimate is the standard error on the estimate; below the standard error is the associated Z-value obtained by dividing the mean by the standard deviation. This is calculated and recorded for all eight pressure levels.

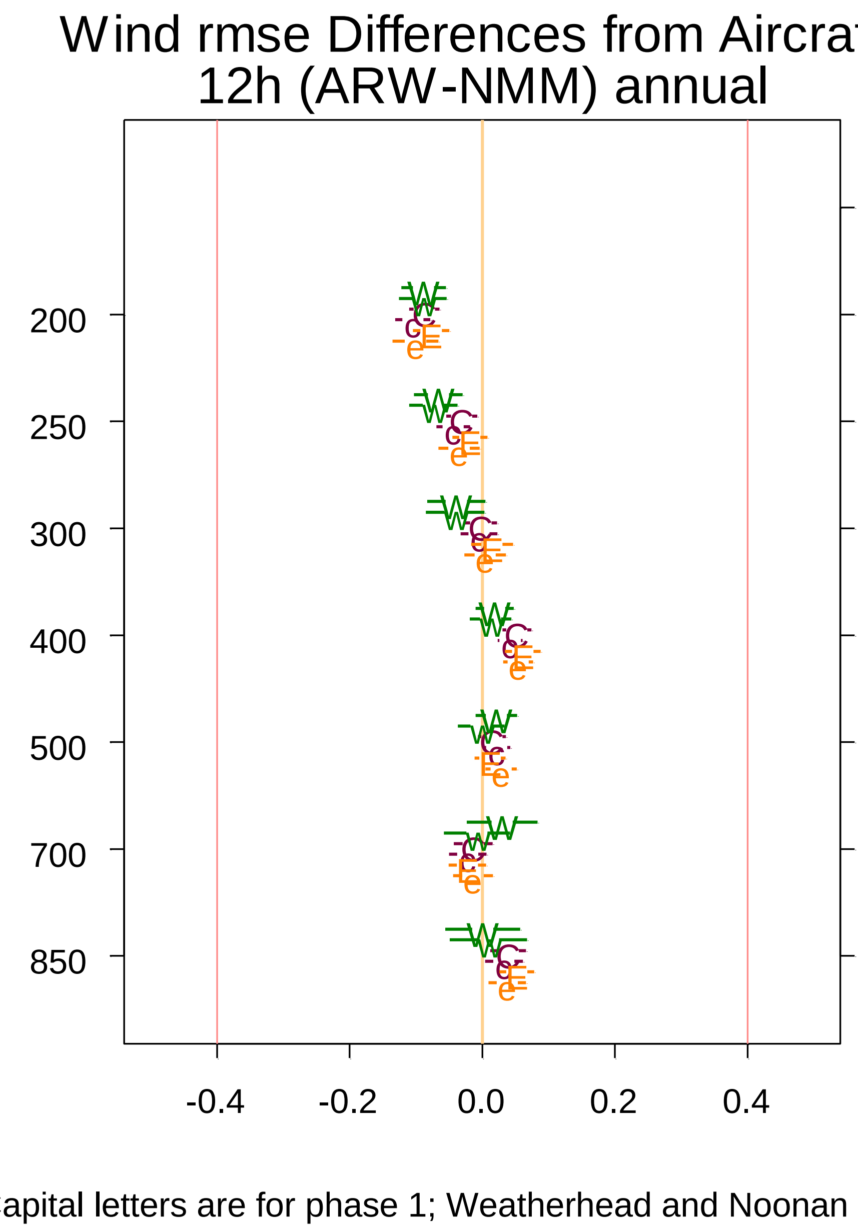
Column 5: NMM.ph2 The mean of the NMM RMSE using Phase-2 physics. 00Z and 12Z forecasts were used. Only forecasts for which both ARW and NMM RMSE values for Phase-2 physics were available were included. Below the mean estimate is the standard error on the estimate; below the standard error is the associated Z-value obtained by dividing the mean by the standard deviation. This is calculated and recorded for all eight pressure levels.

Column 6: Ph1.diff The mean of the pairwise differences of ARW RMSE and NMM RMSE using Phase-1 physics. Below this mean is the calculated standard error in the mean; below the standard error is the associated Z-value obtained by dividing the mean by the standard deviation.

Column 7: Ph2.diff The mean of the pairwise differences of ARW RMSE and NMM RMSE using Phase-2 physics. Below this mean is the calculated standard error in the mean; below the standard error is the associated Z-value obtained by dividing the mean by the standard deviation.

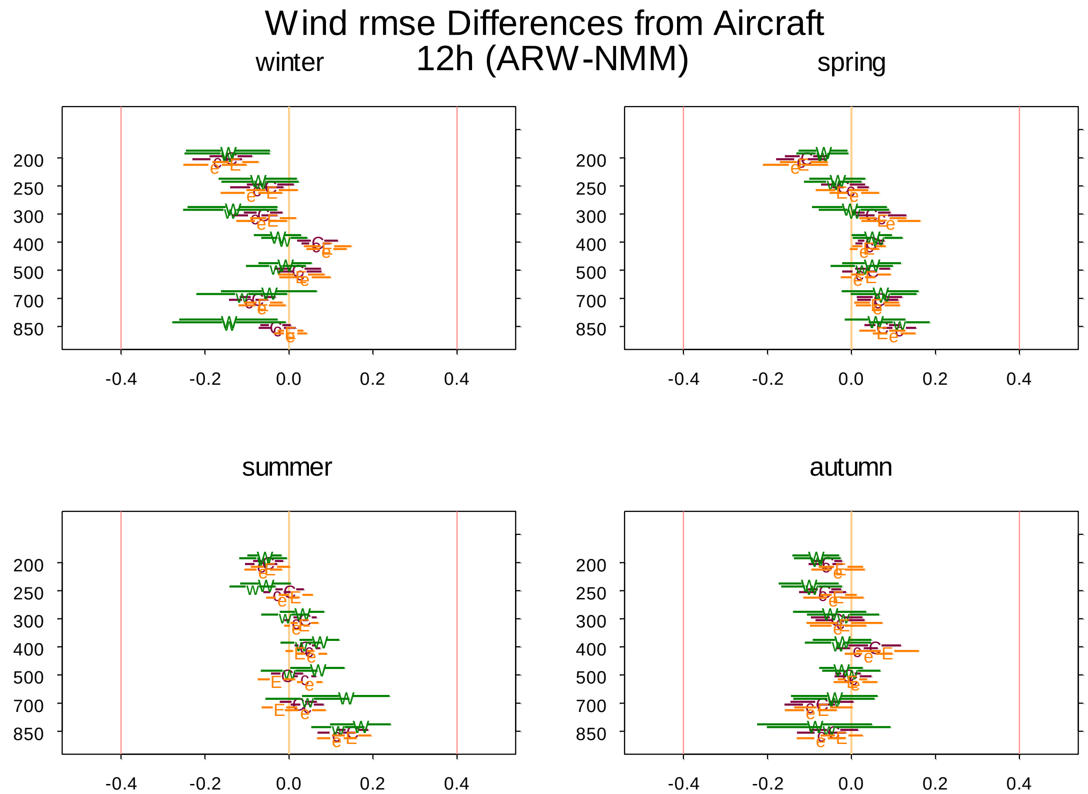
Column 8: # of Obs The total number of observations used in deriving the results in the prior columns. The first number is the number of forecasts used for the calculations. The second and third numbers are the average number of observations used to derive the individual forecast values for Phase-1 and Phase-2 respectively.

RMSE for Aircraft Vector Wind – Forecast Hour 12

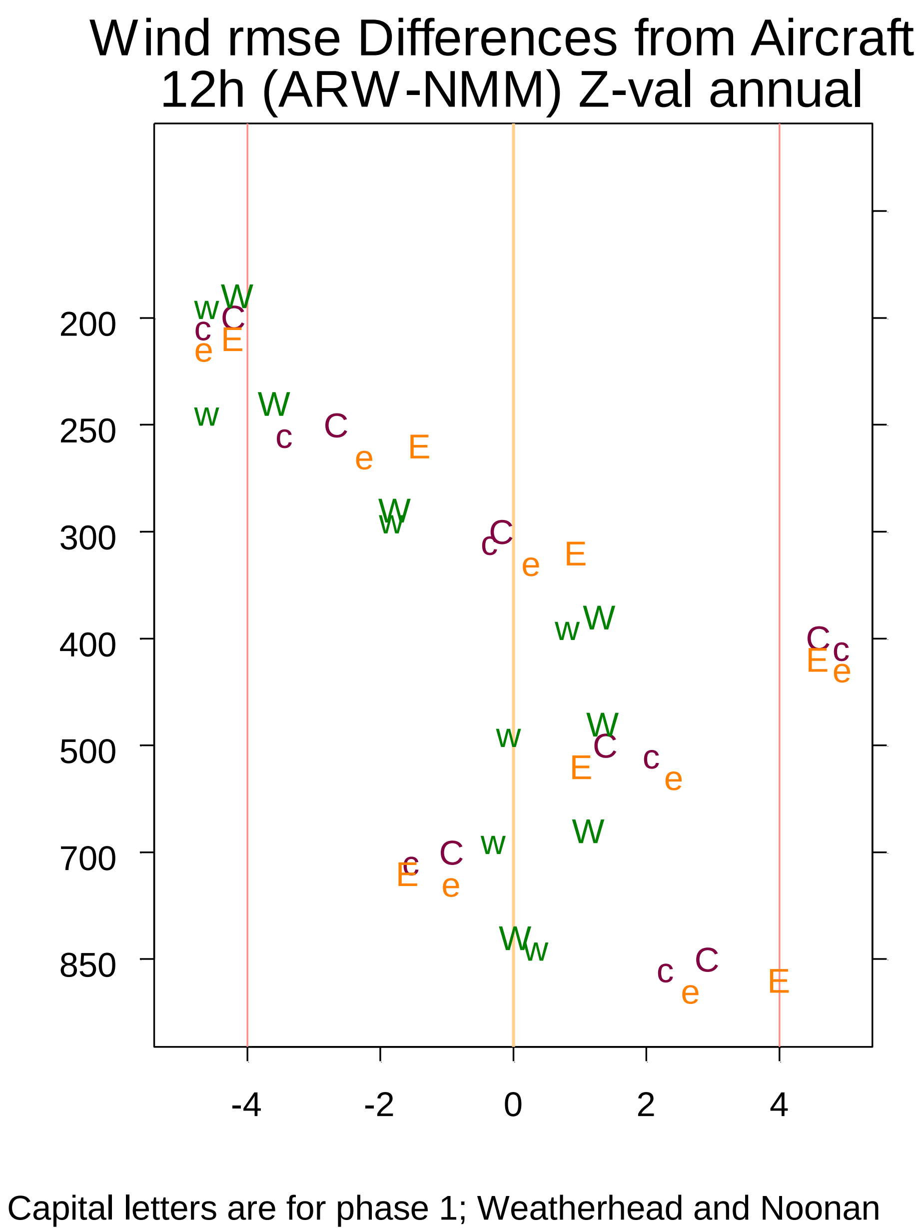


Caption: The differences in wind rmse are presented for three regions: Conus (C for physics 1 and c for physics 2), Conus West (W for physics 1 and w for physics 2)and Conus East (E for physics 1 and e for physics 2). Model rmse is always positive. If the rmse of ARW is larger, the results will be on the right side of the plot. If the rmse of NMM is larger, the results will be on the left side of the plot. The accompanying error bars represent a 95% confidence limit on the difference of the wind rmse. Results are presented for seven pressure bins, with the three regions slightly offset vertically for clarity.

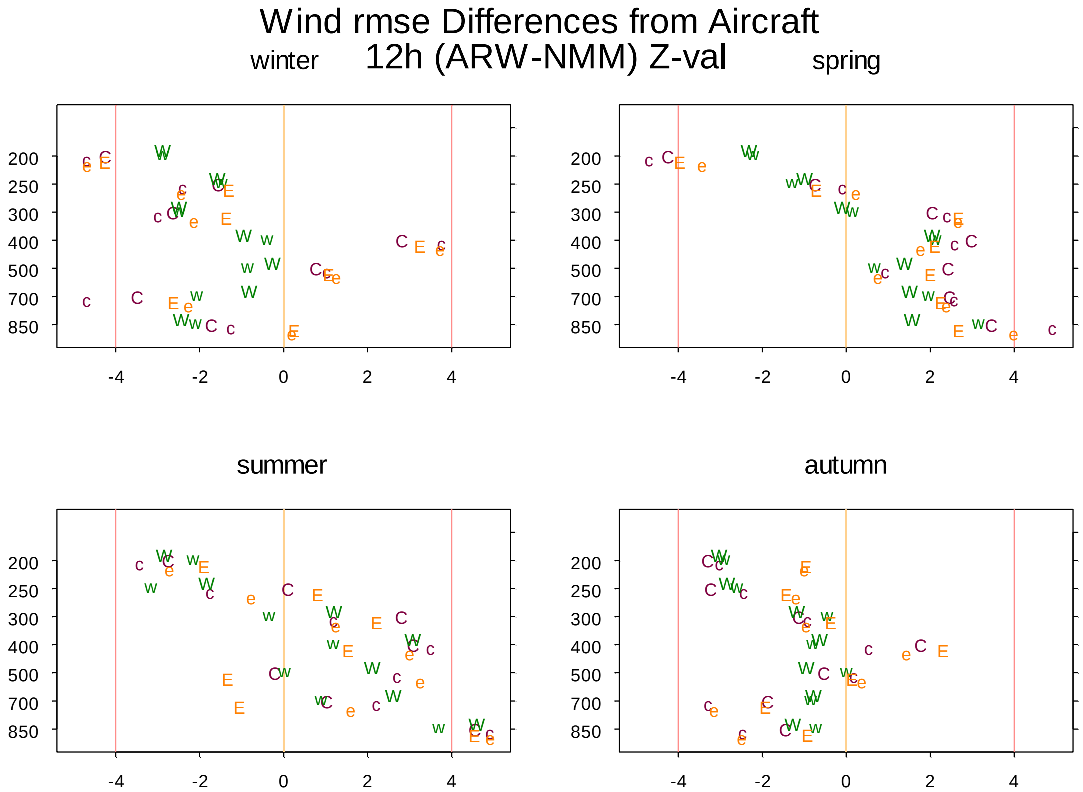
(For comparison: see 12h sonde RMSE on page 36)



Caption: The differences in wind rmse are presented for three regions: Conus (C for physics 1 and c for physics 2), Conus West (W for physics 1 and w for physics 2)and Conus East (E for physics 1 and e for physics 2). Model rmse is always positive. If the rmse of ARW is larger, the results will be on the right side of the plot. If the rmse of NMM is larger, the results will be on the left side of the plot. The accompanying error bars represent a 95% confidence limit on the difference of the wind rmse. Results are presented for seven pressure bins, with the three regions slightly offset vertically for clarity.



Caption: The number of standard errors represents how strong the statistical significance is between ARW rmse and NMM rmse. The values are presented for three regions: Conus (C for physics 1 and c for physics 2), Conus West (W for physics 1 and w for physics 2)and Conus East (E for physics 1 and e for physics 2). When the significance is more than four standard errors, either positively or negatively, the symbol for that region is listed in the end region. If there is little significant difference between the two models, the symbol for that region will appear near zero. If the statistical analysis indicates that it is highly likely that the rmse of ARW is larger than NMM and the symbol will appear more to the right. Similarly, if the statistical analysis indicates that it is highly likely that the rmse of NMM is larger than ARM, the symbol will appear on the left. The actual difference in rmse may be very small. This plot only indicates how confident we are that the difference is significant in a statistical sense.



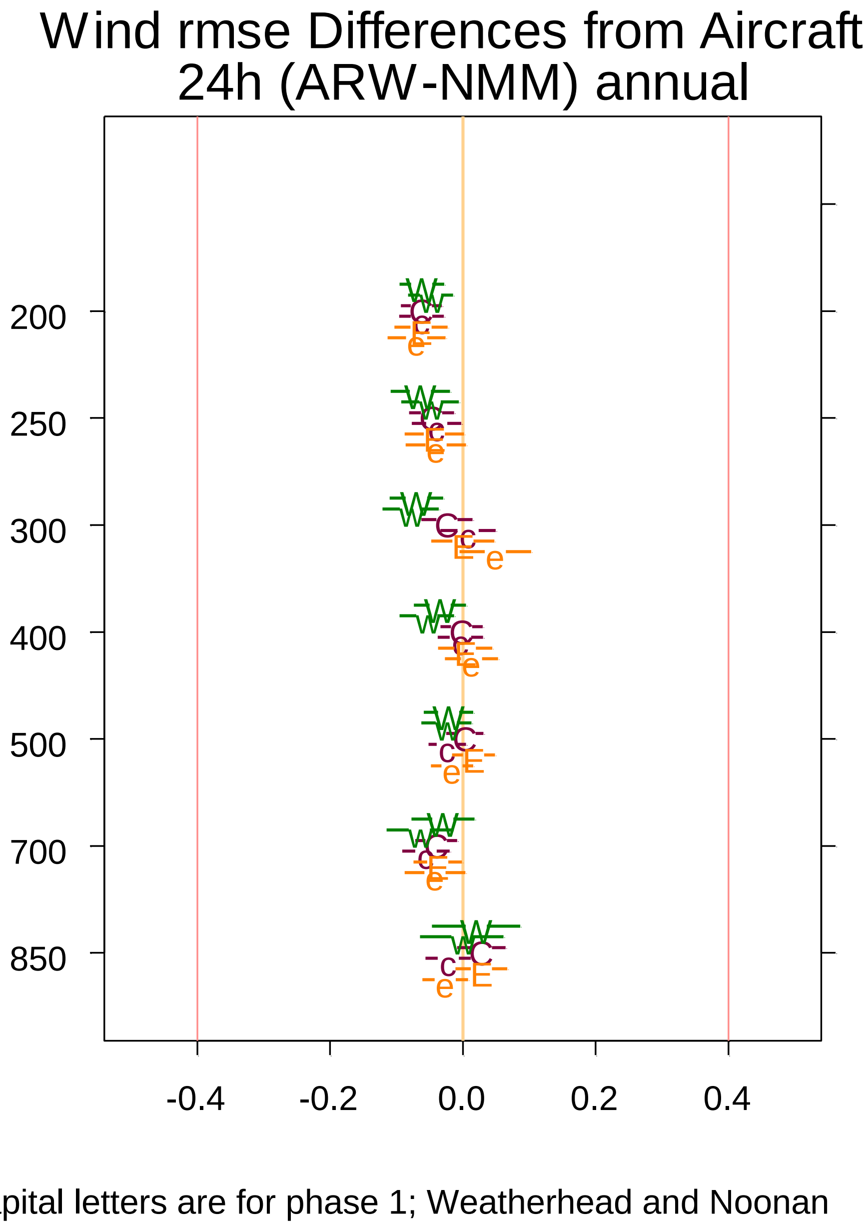
Caption: The number of standard errors represents how strong the statistical significance is between ARW rmse and NMM rmse. The values are presented for three regions: Conus (C for physics 1 and c for physics 2), Conus West (W for physics 1 and w for physics 2)and Conus East (E for physics 1 and e for physics 2). When the significance is more than four standard errors, either positively or negatively, the symbol for that region is listed in the end region. If there is little significant difference between the two models, the symbol for that region will appear near zero. If the statistical analysis indicates that it is highly likely that the rmse of ARW is larger than NMM and the symbol will appear more to the right. Similarly, if the statistical analysis indicates that it is highly likely that the rmse of NMM is larger than ARM, the symbol will appear on the left. The actual difference in rmse may be very small. This plot only indicates how confident we are that the difference is significant in a statistical sense.

RMSE for Aircraft Vector Wind / Conus / Forecast Hour 12

ANNUAL

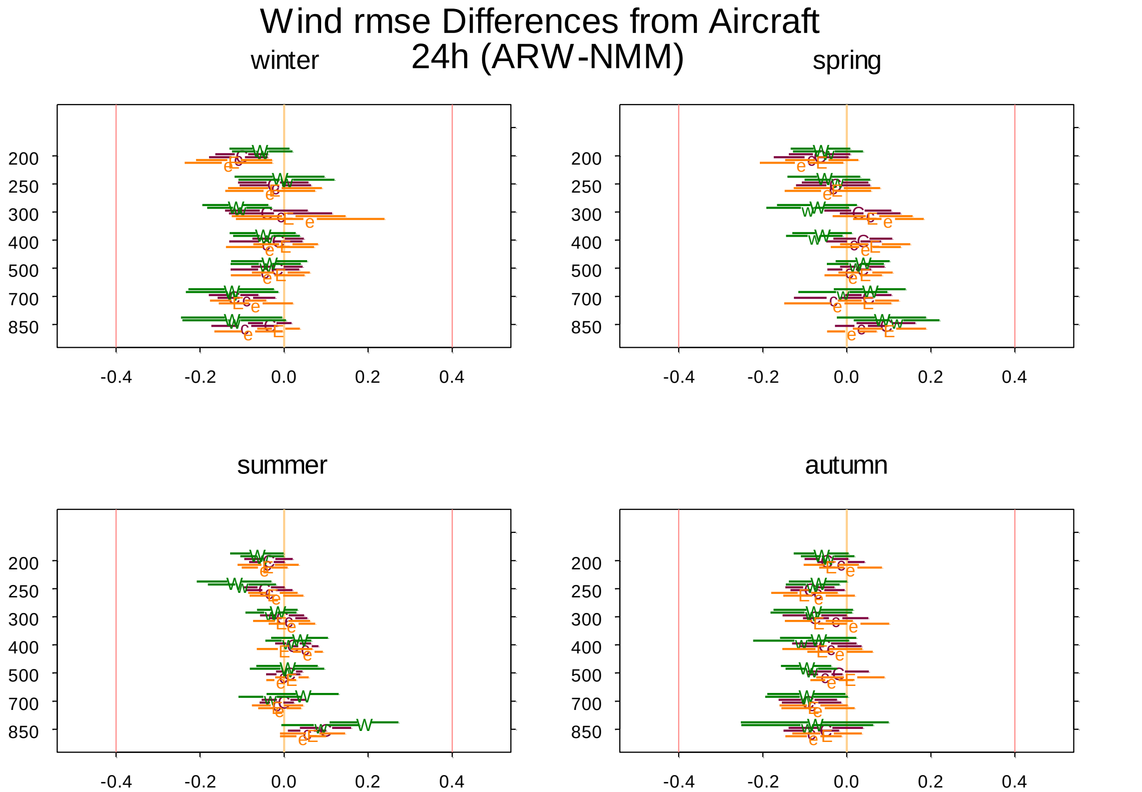
	ARW.ph1	NMM.ph1	ARW.ph2	NMM.ph2	ph1.diff	ph2.diff	# of Obs
P200-250	6.77915	6.86650	6.73485	6.83954	-0.08735	-0.10469	1219.35789
sigma	0.09717	0.10096	0.10319	0.11097	0.01138	0.01317	232.00000
p-val	69.76783	68.00998	65.26378	61.63168	-7.67785	-7.94957	232.00000
P250-300	6.91422	6.94529	6.87907	6.92290	-0.03108	-0.04383	604.24000
sigma	0.10136	0.10300	0.10396	0.10734	0.01167	0.01273	232.00000
p-val	68.21228	67.42696	66.17099	64.49781	-2.66406	-3.44412	232.00000
P300-400	6.20083	6.20301	6.15727	6.16226	-0.00219	-0.00500	408.07789
sigma	0.11442	0.11793	0.11867	0.12201	0.01225	0.01388	232.00000
p-val	54.19331	52.59768	51.88618	50.50489	-0.17848	-0.35987	232.00000
P400-550	5.12216	5.06995	5.10111	5.05908	0.05221	0.04203	664.47789
sigma	0.08871	0.08574	0.08906	0.08713	0.01097	0.00931	232.00000
p-val	57.73790	59.13168	57.28003	58.06342	4.75844	4.51610	232.00000
P550-700	4.74331	4.72883	4.73326	4.71162	0.01448	0.02164	813.80632
sigma	0.06298	0.06264	0.06896	0.06881	0.01046	0.01041	232.00000
p-val	75.31231	75.49736	68.63420	68.47375	1.38452	2.07846	232.00000
P700-850	4.81703	4.83055	4.78752	4.80926	-0.01352	-0.02174	667.31789
sigma	0.05763	0.05944	0.05927	0.06225	0.01459	0.01412	232.00000
p-val	83.58481	81.27022	80.76890	77.25233	-0.92663	-1.53965	232.00000
P850-1000	4.70381	4.66480	4.67917	4.64656	0.03901	0.03261	936.29158
sigma	0.06872	0.07252	0.06924	0.07330	0.01339	0.01427	232.00000
p-val	68.45341	64.32608	67.57983	63.39392	2.91398	2.28475	232.00000

RMSE for Vector Wind – Forecast Hour 24



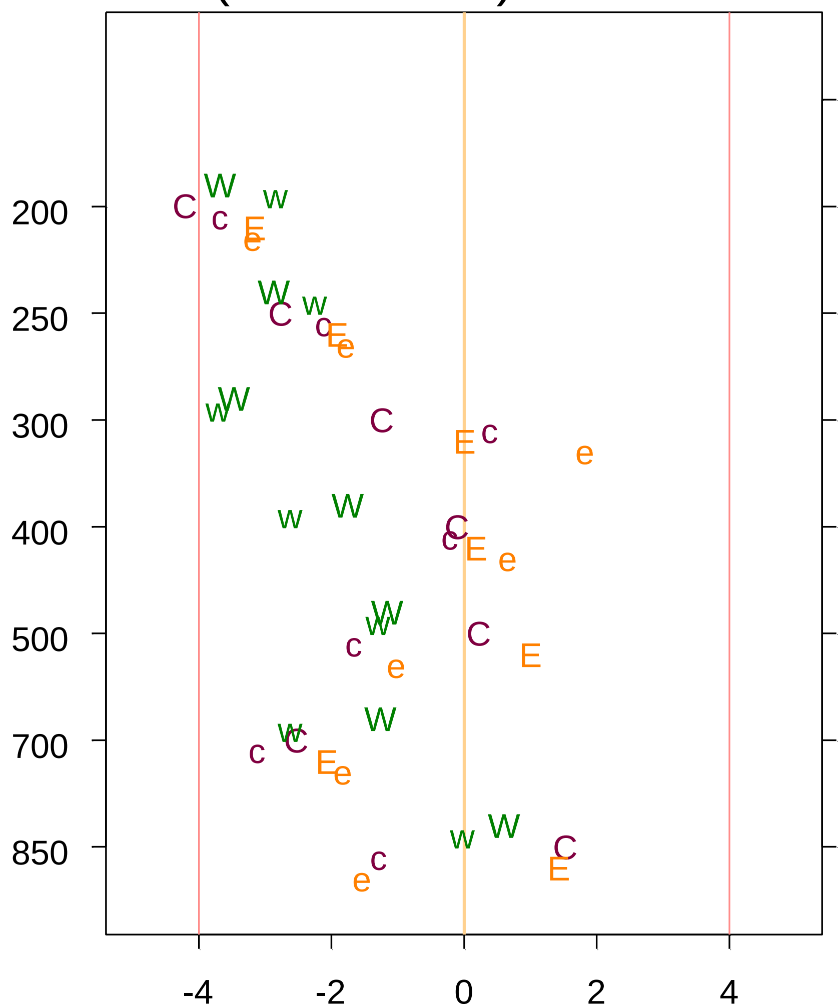
Caption: The differences in wind rmse are presented for three regions: Conus (C for physics 1 and c for physics 2), Conus West (W for physics 1 and w for physics 2)and Conus East (E for physics 1 and e for physics 2). Model rmse is always positive. If the rmse of ARW is larger, the results will be on the right side of the plot. If the rmse of NMM is larger, the results will be on the left side of the plot. The accompanying error bars represent a 95% confidence limit on the difference of the wind rmse. Results are presented for seven pressure bins, with the three regions slightly offset vertically for clarity.

(For comparison: see 24h sonde RMSE on page 41)



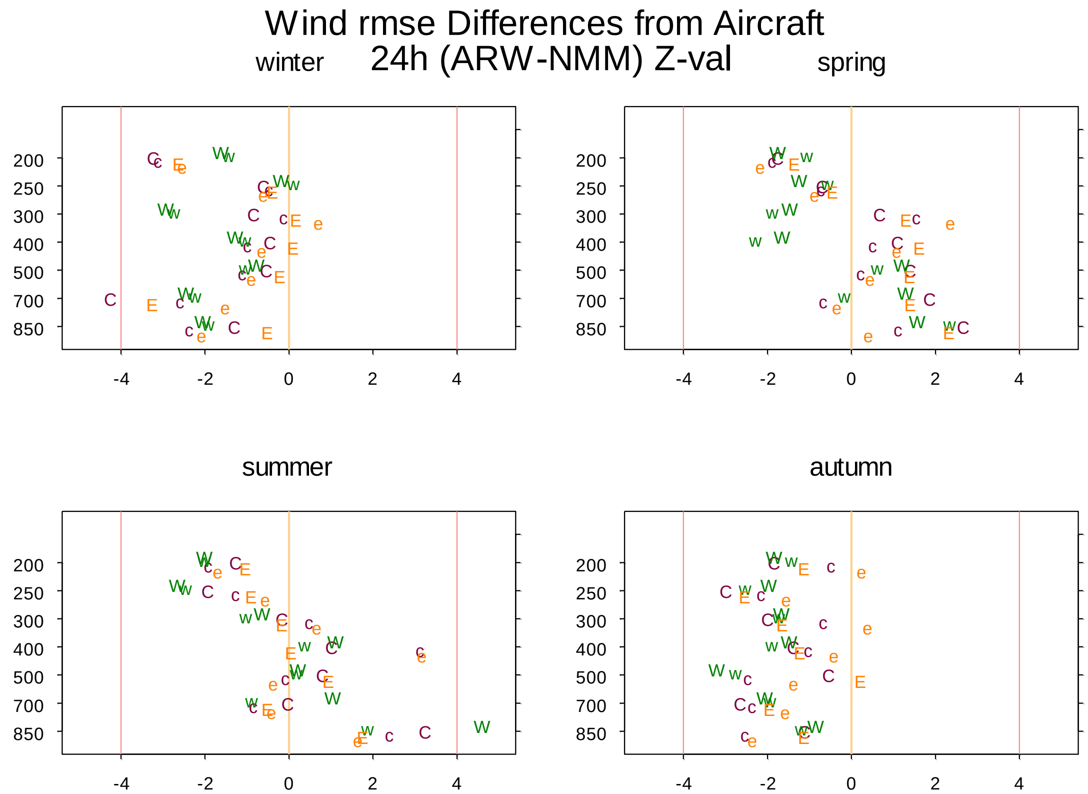
Caption: The differences in wind rmse are presented for three regions: Conus (C for physics 1 and c for physics 2), Conus West (W for physics 1 and w for physics 2)and Conus East (E for physics 1 and e for physics 2). Model rmse is always positive. If the rmse of ARW is larger, the results will be on the right side of the plot. If the rmse of NMM is larger, the results will be on the left side of the plot. The accompanying error bars represent a 95% confidence limit on the difference of the wind rmse. Results are presented for seven pressure bins, with the three regions slightly offset vertically for clarity.

Wind rmse Differences from Aircraft 24h (ARW-NMM) Z-val annual



Capital letters are for phase 1; Weatherhead and Noonan

Caption: The number of standard errors represents how strong the statistical significance is between ARW rmse and NMM rmse. The values are presented for three regions: Conus (C for physics 1 and c for physics 2), Conus West (W for physics 1 and w for physics 2)and Conus East (E for physics 1 and e for physics 2). When the significance is more than four standard errors, either positively or negatively, the symbol for that region is listed in the end region. If there is little significant difference between the two models, the symbol for that region will appear near zero. If the statistical analysis indicates that it is highly likely that the rmse of ARW is larger than NMM and the symbol will appear more to the right. Similarly, if the statistical analysis indicates that it is highly likely that the rmse of NMM is larger than ARM, the symbol will appear on the left. The actual difference in rmse may be very small. This plot only indicates how confident we are that the difference is significant in a statistical sense.



Caption: The number of standard errors represents how strong the statistical significance is between ARW rmse and NMM rmse. The values are presented for three regions: Conus (C for physics 1 and c for physics 2), Conus West (W for physics 1 and w for physics 2)and Conus East (E for physics 1 and e for physics 2). When the significance is more than four standard errors, either positively or negatively, the symbol for that region is listed in the end region. If there is little significant difference between the two models, the symbol for that region will appear near zero. If the statistical analysis indicates that it is highly likely that the rmse of ARW is larger than NMM and the symbol will appear more to the right. Similarly, if the statistical analysis indicates that it is highly likely that the rmse of NMM is larger than ARM, the symbol will appear on the left. The actual difference in rmse may be very small. This plot only indicates how confident we are that the difference is significant in a statistical sense.

RMSE for Aircraft Vector Wind / Conus / Forecast Hour 24

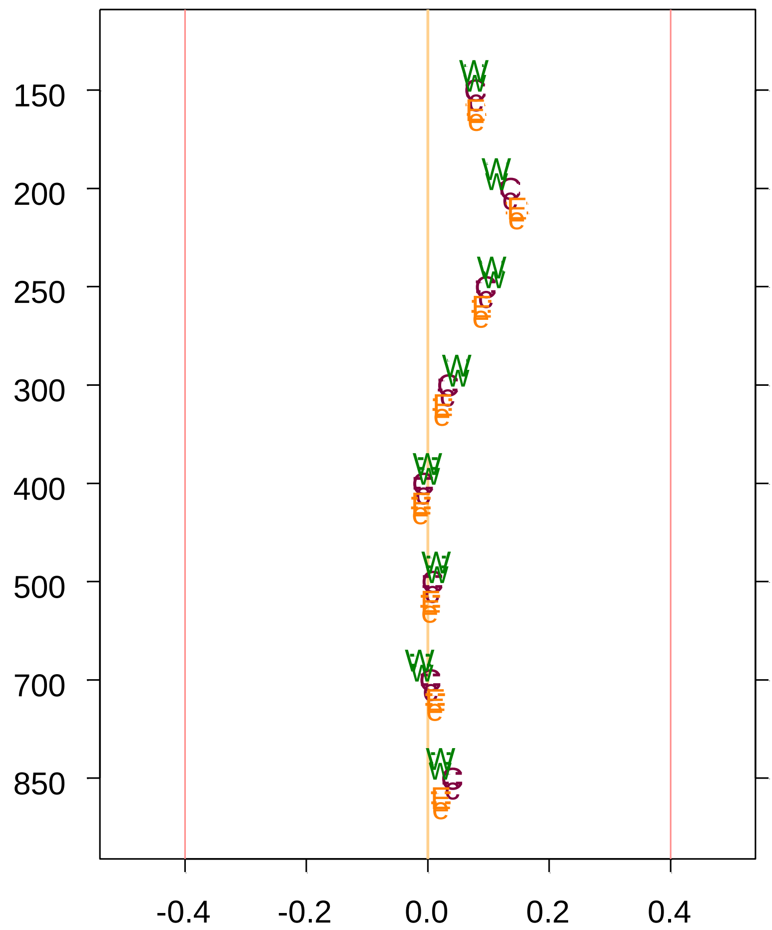
ANNUAL	ARW.ph1	NMM.ph1	ARW.ph2	NMM.ph2	ph1.diff	ph2.diff	# of Obs
P200-250	7.84392	7.90638	7.82557	7.88774	-0.06246	-0.06217	1212.70074
sigma	0.13513	0.13678	0.13865	0.14123	0.01549	0.01689	231.00000
p-val	58.04591	57.80274	56.44186	55.84897	-4.03169	-3.68057	231.00000
P250-300	8.24293	8.28997	8.24908	8.28854	-0.04704	-0.03946	606.03793
sigma	0.17348	0.17309	0.17956	0.17869	0.01699	0.01861	231.00000
p-val	47.51435	47.89422	45.94161	46.38507	-2.76821	-2.12058	231.00000
P300-400	7.41672	7.44062	7.47947	7.47146	-0.02390	0.00801	410.30032
sigma	0.17601	0.17861	0.17966	0.18295	0.01929	0.02075	231.00000
p-val	42.13904	41.65884	41.63218	40.83949	-1.23909	0.38596	231.00000
P400-550	5.83314	5.83484	5.86426	5.86792	-0.00169	-0.00365	666.47945
sigma	0.12229	0.12299	0.12842	0.13008	0.01592	0.01710	231.00000
p-val	47.69983	47.44082	45.66297	45.11096	-0.10632	-0.21368	231.00000
P550-700	5.29238	5.28924	5.32471	5.34819	0.00313	-0.02348	819.10221
sigma	0.08504	0.08472	0.09058	0.09205	0.01405	0.01411	231.00000
p-val	62.23637	62.42883	58.78339	58.09965	0.22307	-1.66330	231.00000
P700-850	5.21349	5.25359	5.24517	5.30071	-0.04010	-0.05554	671.24025
sigma	0.06905	0.07435	0.07779	0.08244	0.01586	0.01781	231.00000
p-val	75.50246	70.65718	67.42433	64.29975	-2.52804	-3.11849	231.00000
P850-1000	5.02624	4.99835	5.02907	5.05107	0.02788	-0.02200	943.09378
sigma	0.07693	0.08097	0.07848	0.08484	0.01823	0.01709	231.00000
p-val	65.33368	61.72771	64.08210	59.53814	1.52930	-1.28765	231.00000

Appendix 7: Discussion of Model Initialization

This appendix is presented in four sections: bias from sondes, bias from aircraft, followed by RMSE from sondes, and RMSE from aircraft.

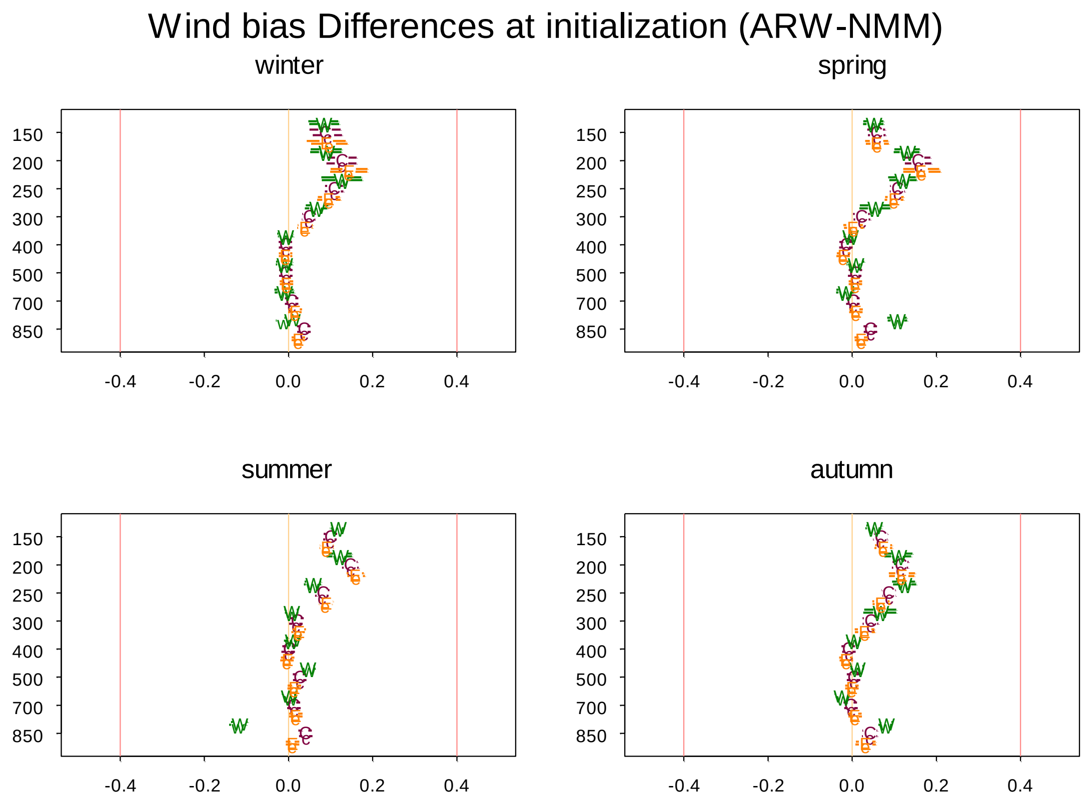
Bias for Vector Wind from sondes - (initialization)

Wind bias Differences at initialization (ARW-NMM) annual

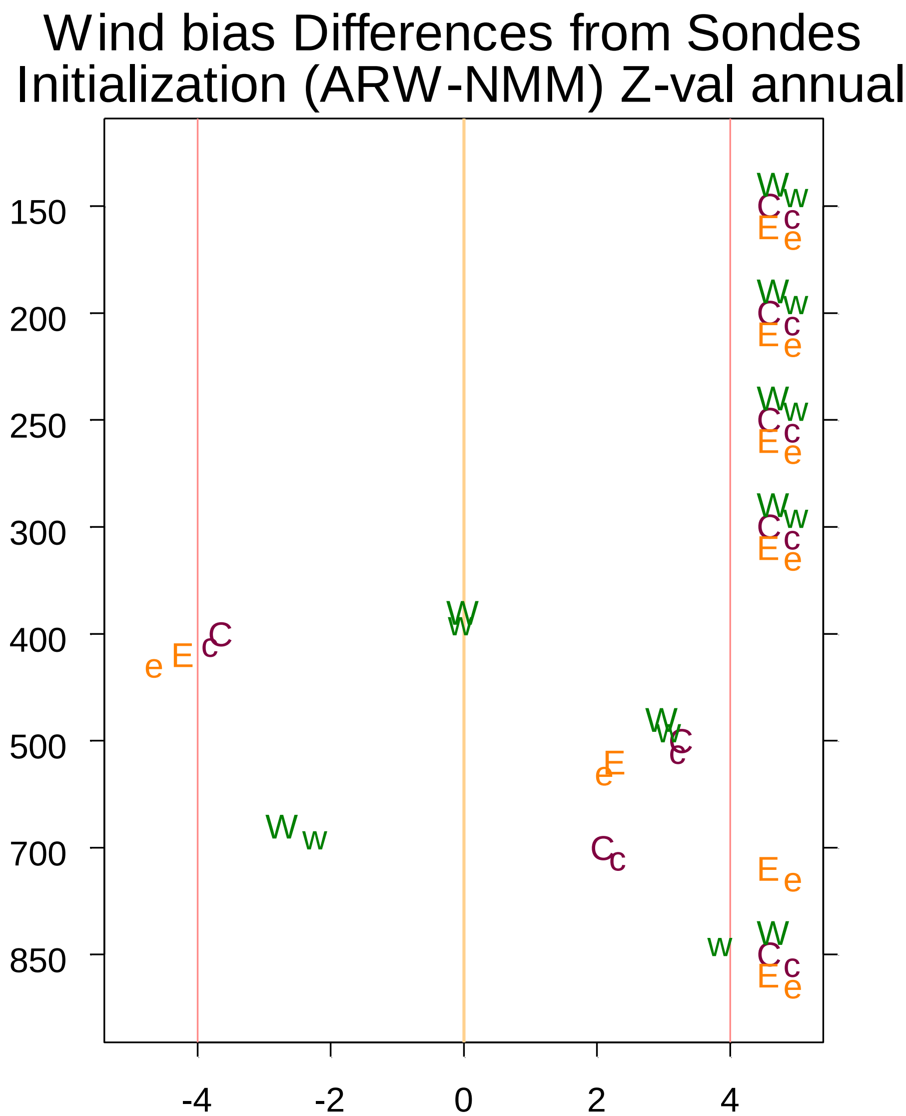


Capital letters are for phase 1; Weatherhead and Noonan

Caption: The differences in the absolute values of wind bias are presented for three regions: Conus (C for physics 1 and c for physics 2), Conus West (W for physics 1 and w for physics 2)and Conus East (E for physics 1 and e for physics 2). Model bias for each level may be either positive or negative. The absolute value of the bias means that the results can be interpreted more directly. If the magnitude of the bias of ARW is larger, the results will be on the right side of the plot. If the magnitude of the bias of NMM is larger, the results will be on the left side of the plot. The accompanying error bars represent a 95% confidence limit on the difference of the wind bias. Results are presented for eight pressure levels, with the three regions slightly offset vertically for clarity.
(For comparison: see aircraft bias at initialization on page 63)



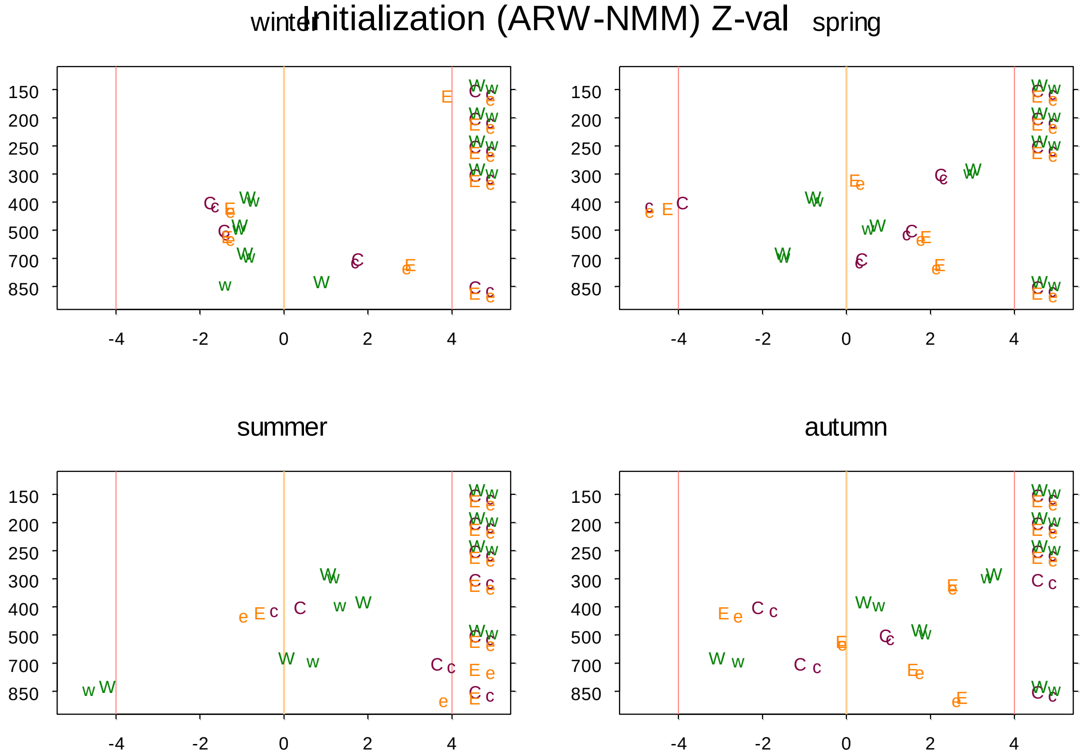
Caption: The differences in the absolute values of wind bias are presented for three regions: Conus (C for physics 1 and c for physics 2), Conus West (W for physics 1 and w for physics 2)and Conus East (E for physics 1 and e for physics 2). Model bias for each level may be either positive or negative. The absolute value of the bias means that the results can be interpreted more directly. If the magnitude of the bias of ARW is larger, the results will be on the right side of the plot. If the magnitude of the bias of NMM is larger, the results will be on the left side of the plot. The accompanying error bars represent a 95% confidence limit on the difference of the wind bias. Results are presented for eight pressure levels, with the three regions slightly offset vertically for clarity.



Capital letters are for phase 1; Weatherhead and Noonan

Caption: The number of standard errors represents how strong the statistical significance is between ARW bias and NMM bias. The values are presented for three regions: Conus (C for physics 1 and c for physics 2), Conus West (W for physics 1 and w for physics 2)and Conus East (E for physics 1 and e for physics 2). When the significance is more than four standard errors, either positively or negatively, the symbol for that region is listed in the end region. If there is little significant difference between the two models, the symbol for that region will appear near zero. If the statistical analysis indicates that it is highly likely that the bias of ARW is larger than NMM and the symbol will appear more to the right. Similarly, if the statistical analysis indicates that it is highly likely that the bias of NMM is larger than ARM, the symbol will appear on the left. The actual difference in bias may be very small. This plot only indicates how confident we are that the difference is significant in a statistical sense.

Wind bias Differences from Sondes
Initialization (ARW-NMM) Z-val



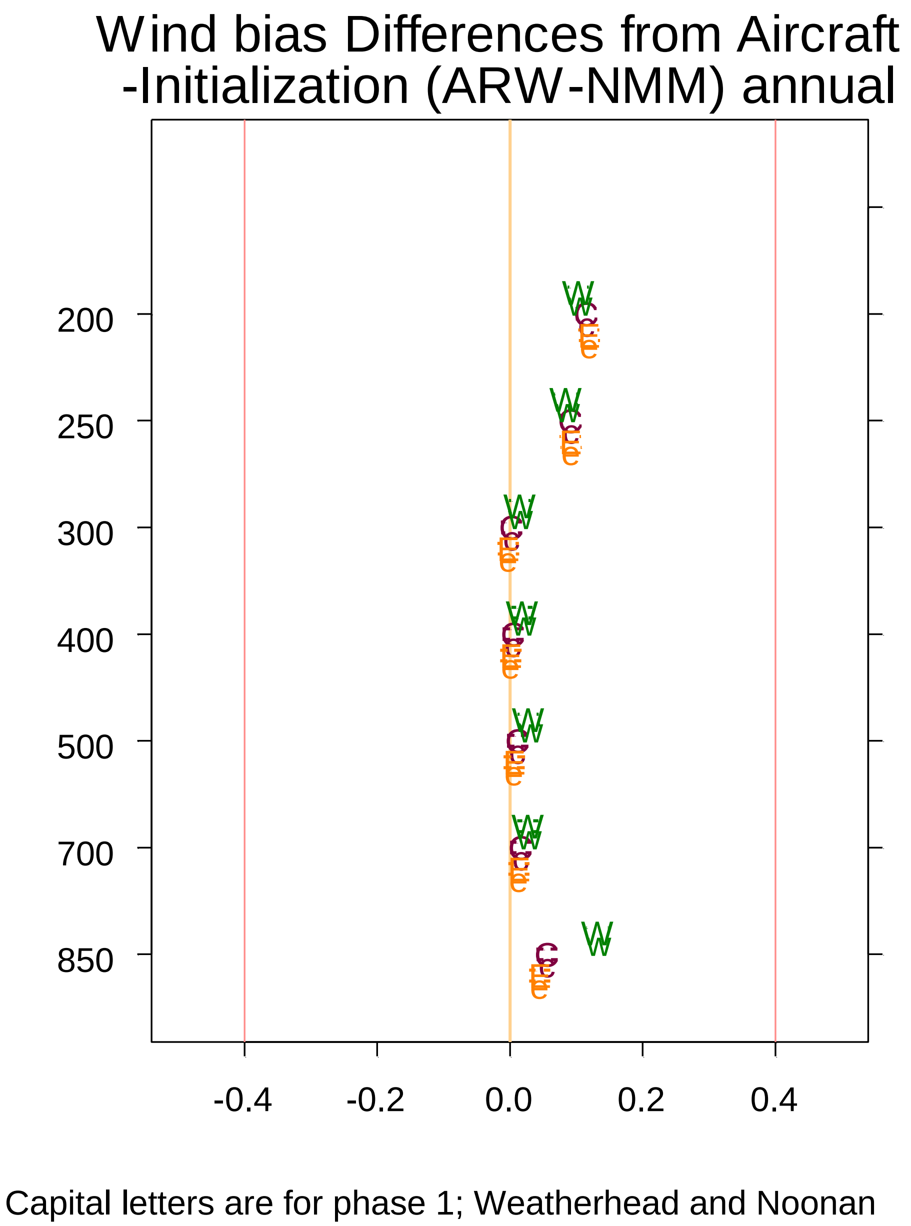
Caption: The number of standard errors represents how strong the statistical significance is between ARW bias and NMM bias. The values are presented for three regions: Conus (C for physics 1 and c for physics 2), Conus West (W for physics 1 and w for physics 2)and Conus East (E for physics 1 and e for physics 2). When the significance is more than four standard errors, either positively or negatively, the symbol for that region is listed in the end region. If there is little significant difference between the two models, the symbol for that region will appear near zero. If the statistical analysis indicates that it is highly likely that the bias of ARW is larger than NMM and the symbol will appear more to the right. Similarly, if the statistical analysis indicates that it is highly likely that the bias of NMM is larger than ARM, the symbol will appear on the left. The actual difference in bias may be very small. This plot only indicates how confident we are that the difference is significant in a statistical sense.

Bias for Sonde Vector Wind / Conus / Forecast Hour 00 (initialization)

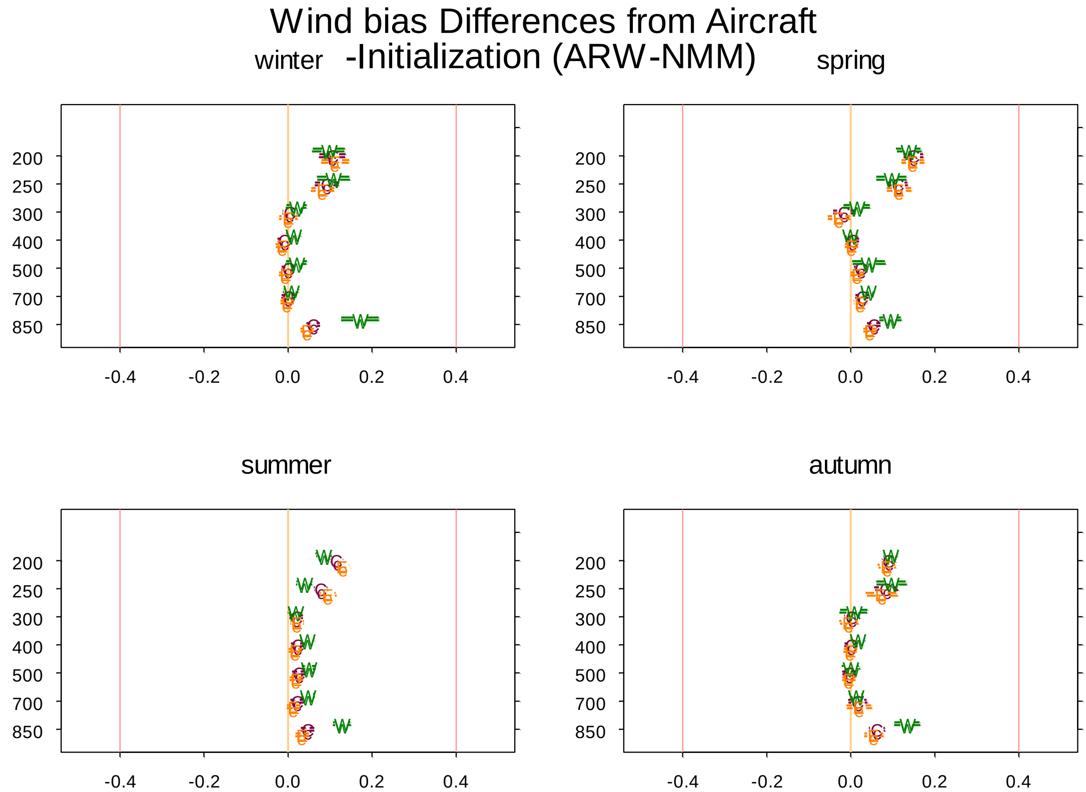
ANNUAL

	ARW.ph1	NMM.ph1	ARW.ph2	NMM.ph2	ph1.diff	ph2.diff	# of Obs
P150	-0.24917	-0.17095	-0.24741	-0.16822	0.07821	0.07918	67.99265
sigma	0.05377	0.05260	0.05526	0.05410	0.00596	0.00588	232.00000
p-val	-4.63371	-3.25036	-4.47755	-3.10956	13.11919	13.46152	232.00000
P200	-0.35102	-0.21504	-0.34525	-0.20926	0.13598	0.13599	69.58447
sigma	0.03410	0.03287	0.03408	0.03264	0.00673	0.00701	232.00000
p-val	-10.29329	-6.54231	-10.13076	-6.41108	20.19341	19.39509	232.00000
P250	-0.36811	-0.27258	-0.36426	-0.26866	0.09553	0.09560	71.77859
sigma	0.02770	0.02722	0.02820	0.02770	0.00424	0.00438	232.00000
p-val	-13.28851	-10.01245	-12.91882	-9.69853	22.51892	21.81202	232.00000
P300	-0.34307	-0.30966	-0.34617	-0.31345	0.03341	0.03273	72.24134
sigma	0.02401	0.02336	0.02461	0.02386	0.00408	0.00413	232.00000
p-val	-14.28720	-13.25795	-14.06611	-13.13837	8.19961	7.91646	232.00000
P400	-0.37735	-0.38478	-0.38128	-0.38887	-0.00743	-0.00760	73.88667
sigma	0.01864	0.01856	0.01970	0.01967	0.00203	0.00199	232.00000
p-val	-20.24722	-20.72785	-19.35628	-19.77401	-3.65293	-3.81200	232.00000
P500	-0.45781	-0.44996	-0.46204	-0.45445	0.00785	0.00759	72.86149
sigma	0.02308	0.02310	0.02356	0.02354	0.00240	0.00236	232.00000
p-val	-19.83723	-19.47597	-19.60860	-19.30674	3.26524	3.21013	232.00000
P700	-0.42771	-0.42337	-0.42702	-0.42217	0.00434	0.00485	73.29171
sigma	0.01921	0.01901	0.01974	0.01956	0.00208	0.00210	232.00000
p-val	-22.25925	-22.27589	-21.62989	-21.58738	2.08855	2.31272	232.00000
P800	-0.18089	-0.14071	-0.18258	-0.14203	0.04019	0.04055	67.90766
sigma	0.01560	0.01602	0.01611	0.01658	0.00276	0.00285	232.00000
p-val	-11.59615	-8.78188	-11.33525	-8.56525	14.54399	14.24659	232.00000

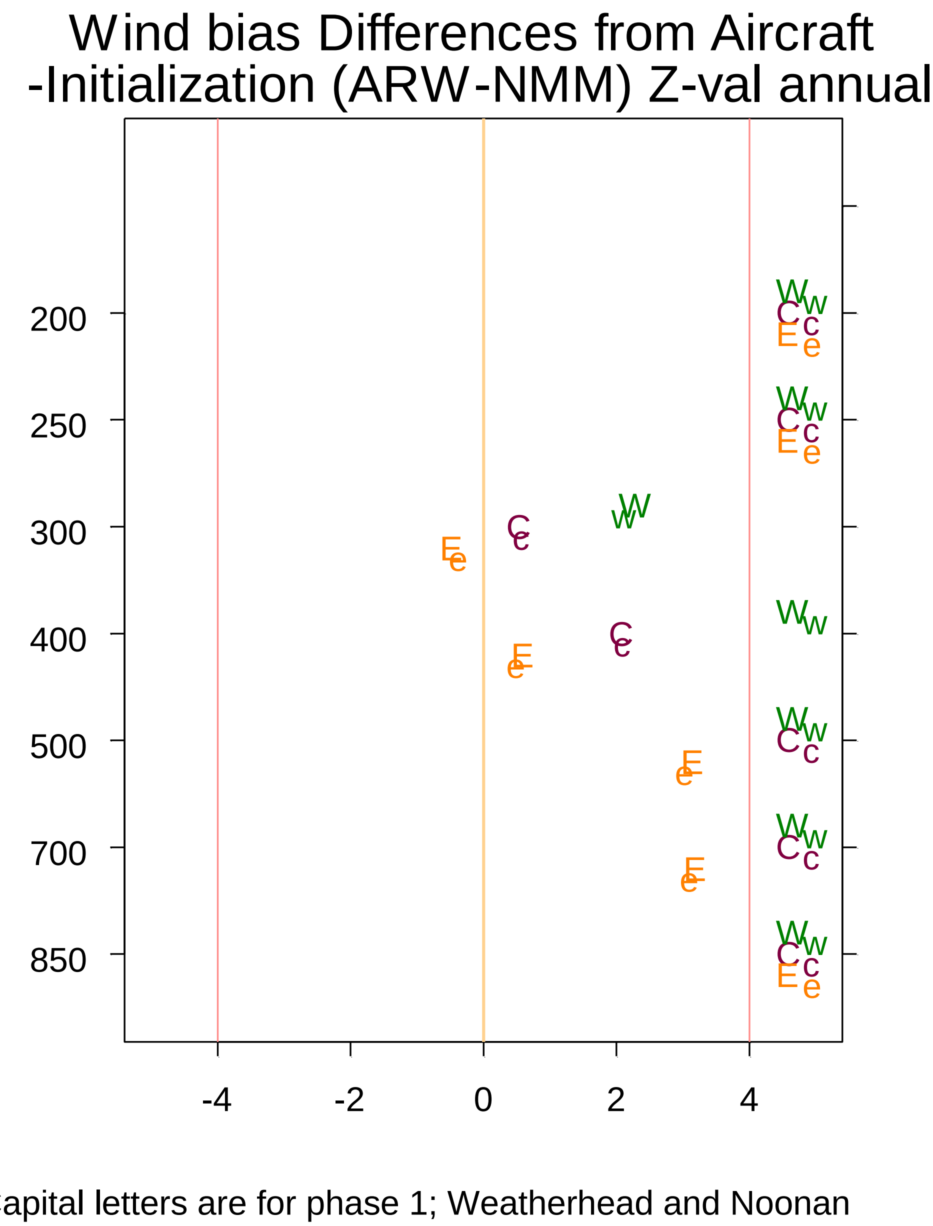
Bias for Vector Wind from aircraft - (initialization)



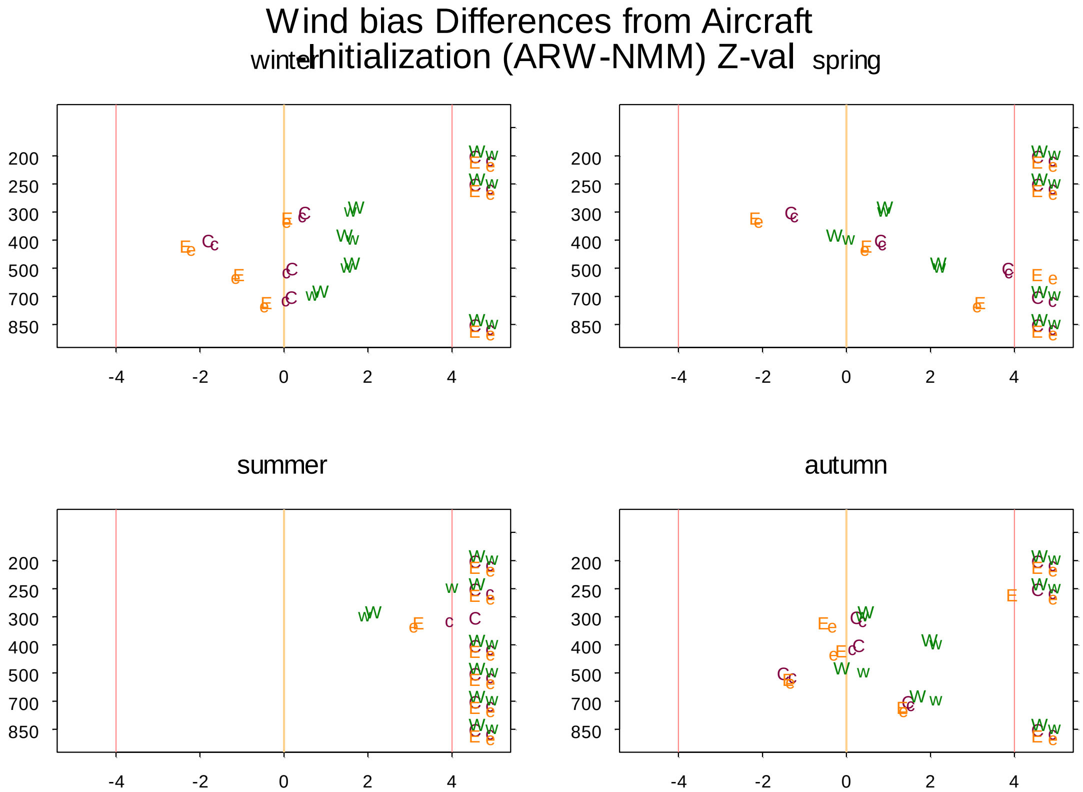
Caption: The differences in the absolute values of wind bias are presented for three regions: Conus (C for physics 1 and c for physics 2), Conus West (W for physics 1 and w for physics 2)and Conus East (E for physics 1 and e for physics 2). Model bias for each level may be either positive or negative. The absolute value of the bias means that the results can be interpreted more directly. If the magnitude of the bias of ARW is larger, the results will be on the right side of the plot. If the magnitude of the bias of NMM is larger, the results will be on the left side of the plot. The accompanying error bars represent a 95% confidence limit on the difference of the wind bias. Results are presented for seven pressure bins, with the three regions slightly offset vertically for clarity.
(For comparison: see sonde bias at initialization on page 58)



Caption: The differences in the absolute values of wind bias are presented for three regions: Conus (C for physics 1 and c for physics 2), Conus West (W for physics 1 and w for physics 2)and Conus East (E for physics 1 and e for physics 2). Model bias for each level may be either positive or negative. The absolute value of the bias means that the results can be interpreted more directly. If the magnitude of the bias of ARW is larger, the results will be on the right side of the plot. If the magnitude of the bias of NMM is larger, the results will be on the left side of the plot. The accompanying error bars represent a 95% confidence limit on the difference of the wind bias. Results are presented for seven pressure bins, with the three regions slightly offset vertically for clarity.



Caption: The number of standard errors represents how strong the statistical significance is between ARW bias and NMM bias. The values are presented for three regions: Conus (C for physics 1 and c for physics 2), Conus West (W for physics 1 and w for physics 2)and Conus East (E for physics 1 and e for physics 2). When the significance is more than four standard errors, either positively or negatively, the symbol for that region is listed in the end region. If there is little significant difference between the two models, the symbol for that region will appear near zero. If the statistical analysis indicates that it is highly likely that the bias of ARW is larger than NMM and the symbol will appear more to the right. Similarly, if the statistical analysis indicates that it is highly likely that the bias of NMM is larger than ARM, the symbol will appear on the left. The actual difference in bias may be very small. This plot only indicates how confident we are that the difference is significant in a statistical sense.



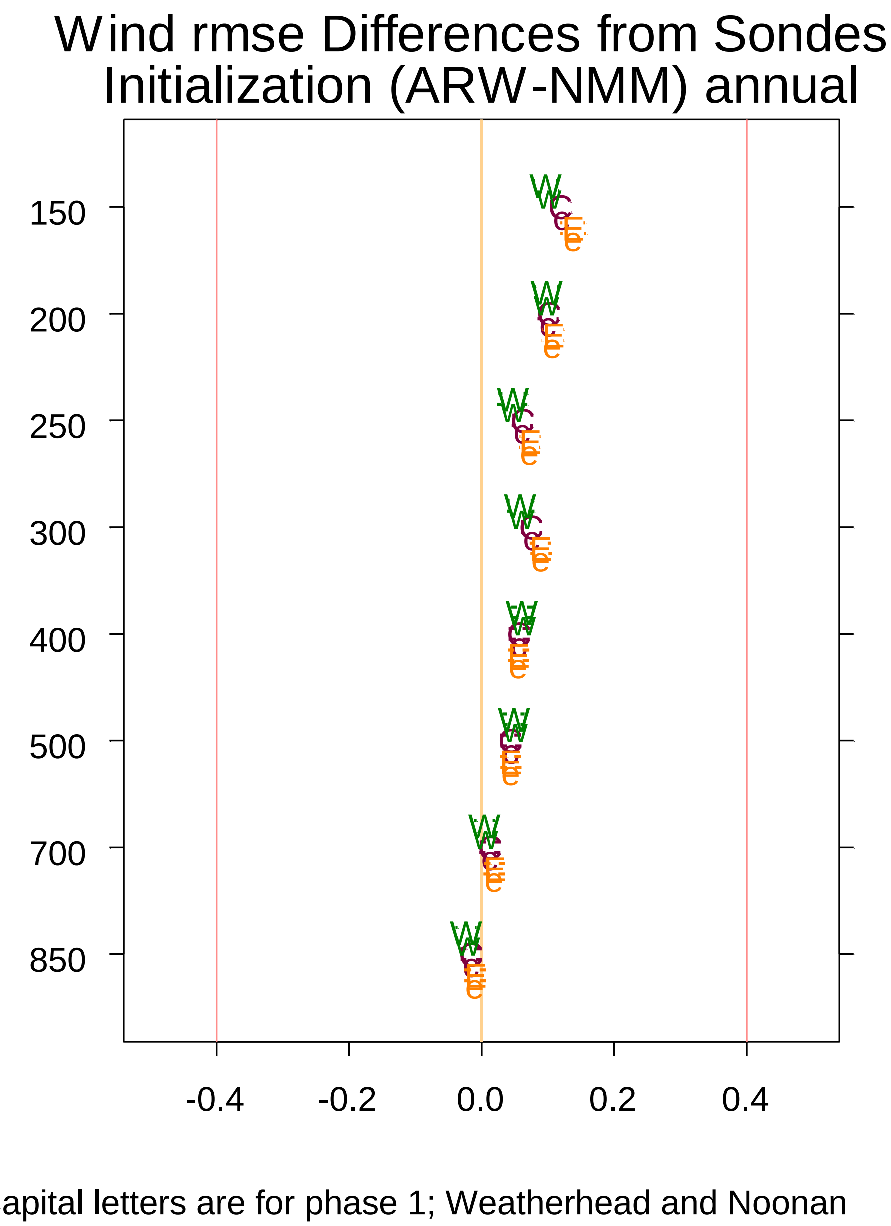
Caption: The number of standard errors represents how strong the statistical significance is between ARW bias and NMM bias. The values are presented for three regions: Conus (C for physics 1 and c for physics 2), Conus West (W for physics 1 and w for physics 2)and Conus East (E for physics 1 and e for physics 2). When the significance is more than four standard errors, either positively or negatively, the symbol for that region is listed in the end region. If there is little significant difference between the two models, the symbol for that region will appear near zero. If the statistical analysis indicates that it is highly likely that the bias of ARW is larger than NMM and the symbol will appear more to the right. Similarly, if the statistical analysis indicates that it is highly likely that the bias of NMM is larger than ARM, the symbol will appear on the left. The actual difference in bias may be very small. This plot only indicates how confident we are that the difference is significant in a statistical sense.

Bias for Aircraft Vector Wind / Conus / Forecast Hour 00 (initialization)

ANNUAL

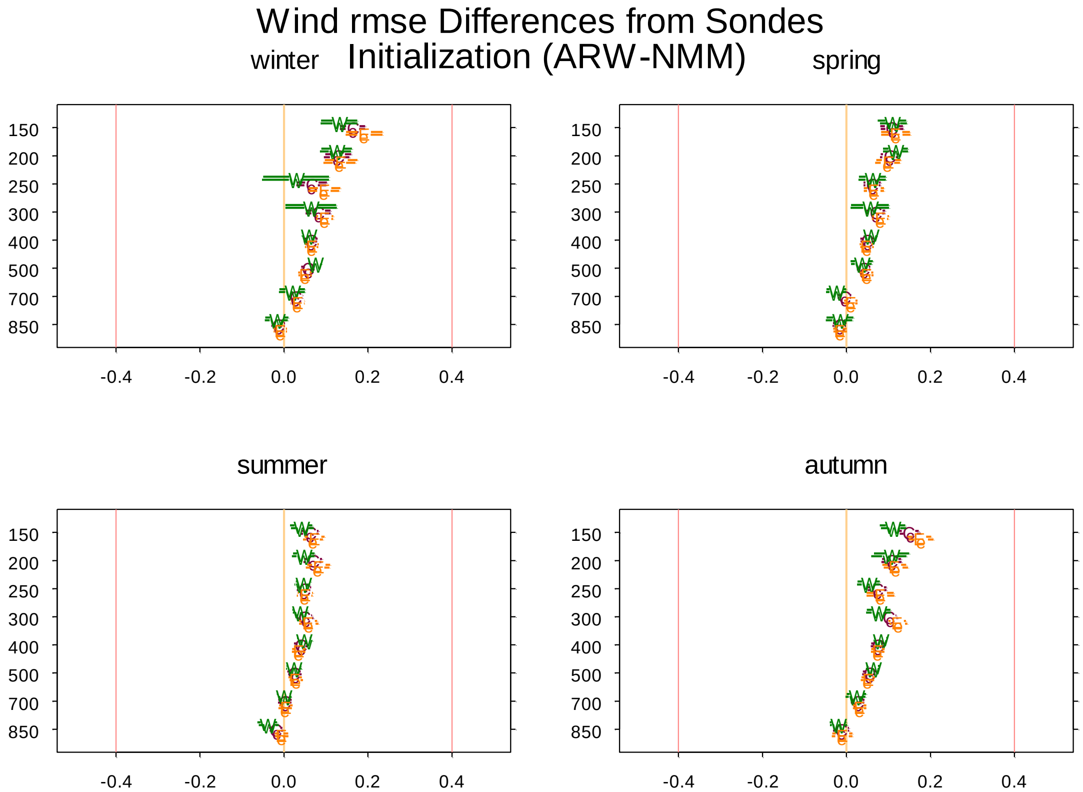
	ARW.ph1	NMM.ph1	ARW.ph2	NMM.ph2	ph1.diff	ph2.diff	# of Obs
P200-250	-1.34384	-1.22917	-1.34522	-1.22967	0.11467	0.11555	1209.51579
sigma	0.03865	0.03570	0.03946	0.03648	0.00586	0.00592	232.00000
p-val	-34.77162	-34.42621	-34.08975	-33.70654	19.56752	19.52217	232.00000
P250-300	-1.04739	-0.95652	-1.05445	-0.96218	0.09088	0.09227	600.44211
sigma	0.03233	0.03210	0.03345	0.03319	0.00590	0.00601	232.00000
p-val	-32.39381	-29.79430	-31.51955	-28.99135	15.39259	15.35537	232.00000
P300-400	-0.55811	-0.55569	-0.55907	-0.55638	0.00242	0.00269	409.02421
sigma	0.03143	0.03079	0.03188	0.03125	0.00459	0.00473	232.00000
p-val	-17.75694	-18.04645	-17.53660	-17.80596	0.52866	0.56813	232.00000
P400-550	-0.40890	-0.40401	-0.41015	-0.40525	0.00489	0.00490	664.30632
sigma	0.02423	0.02432	0.02406	0.02417	0.00236	0.00235	232.00000
p-val	-16.87270	-16.61454	-17.05012	-16.76617	2.06934	2.08538	232.00000
P550-700	-0.39603	-0.38446	-0.39388	-0.38213	0.01158	0.01174	814.02526
sigma	0.02256	0.02291	0.02319	0.02350	0.00241	0.00245	232.00000
p-val	-17.55291	-16.77971	-16.98712	-16.25855	4.79544	4.80179	232.00000
P700-850	-0.40564	-0.38933	-0.41251	-0.39605	0.01631	0.01645	666.38211
sigma	0.02036	0.02046	0.02044	0.02055	0.00325	0.00332	232.00000
p-val	-19.92597	-19.02540	-20.17934	-19.26995	5.02497	4.95509	232.00000
P850-1000	-0.39349	-0.33772	-0.39940	-0.34338	0.05577	0.05602	935.83789
sigma	0.04771	0.04778	0.04894	0.04902	0.00275	0.00283	232.00000
p-val	-8.24717	-7.06767	-8.16101	-7.00545	20.27174	19.78968	232.00000

RMSE for Vector Wind from Sondes - (initialization)

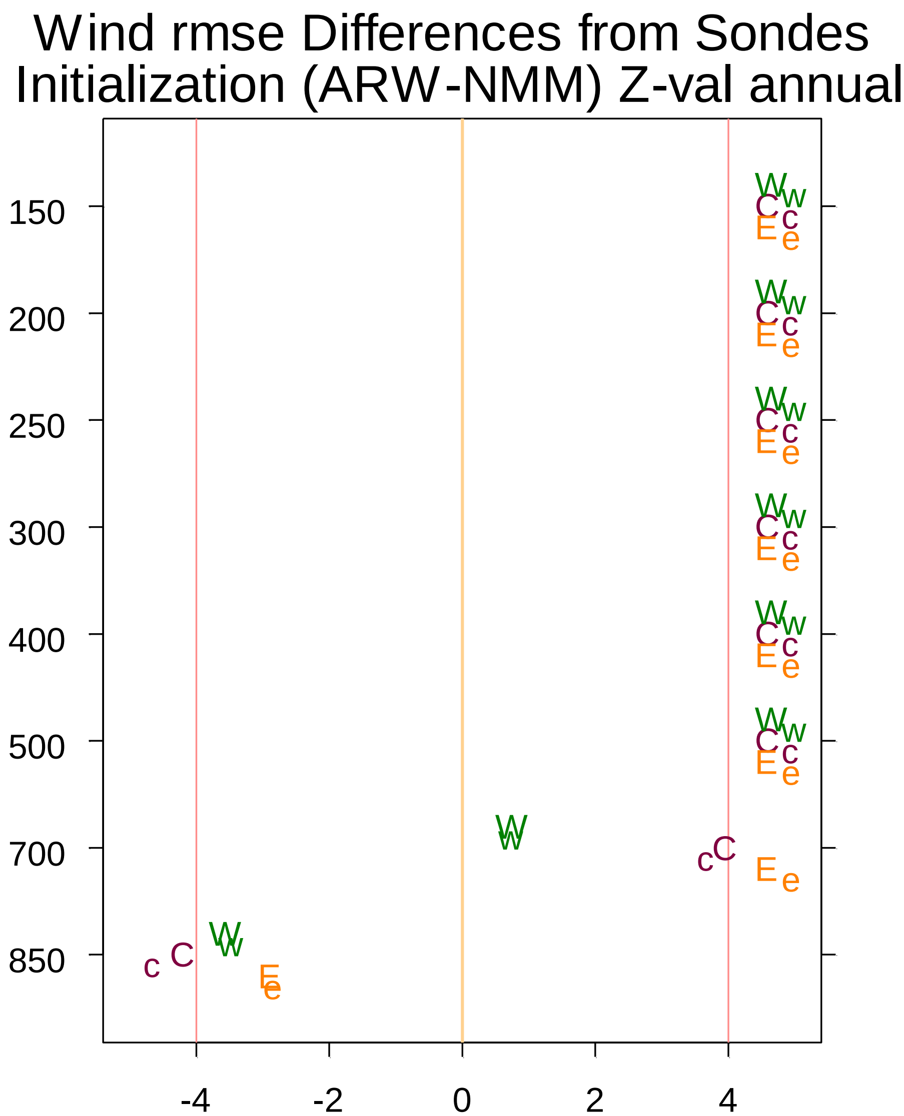


Caption: The differences in wind rmse are presented for three regions: Conus (C for physics 1 and c for physics 2), Conus West (W for physics 1 and w for physics 2)and Conus East (E for physics 1 and e for physics 2). Model rmse is always positive. If the rmse of ARW is larger, the results will be on the right side of the plot. If the rmse of NMM is larger, the results will be on the left side of the plot. The accompanying error bars represent a 95% confidence limit on the difference of the wind rmse. Results are presented for eight pressure levels, with the three regions slightly offset vertically for clarity.

(For comparison: see aircraft RMSE at initialization on page 73)

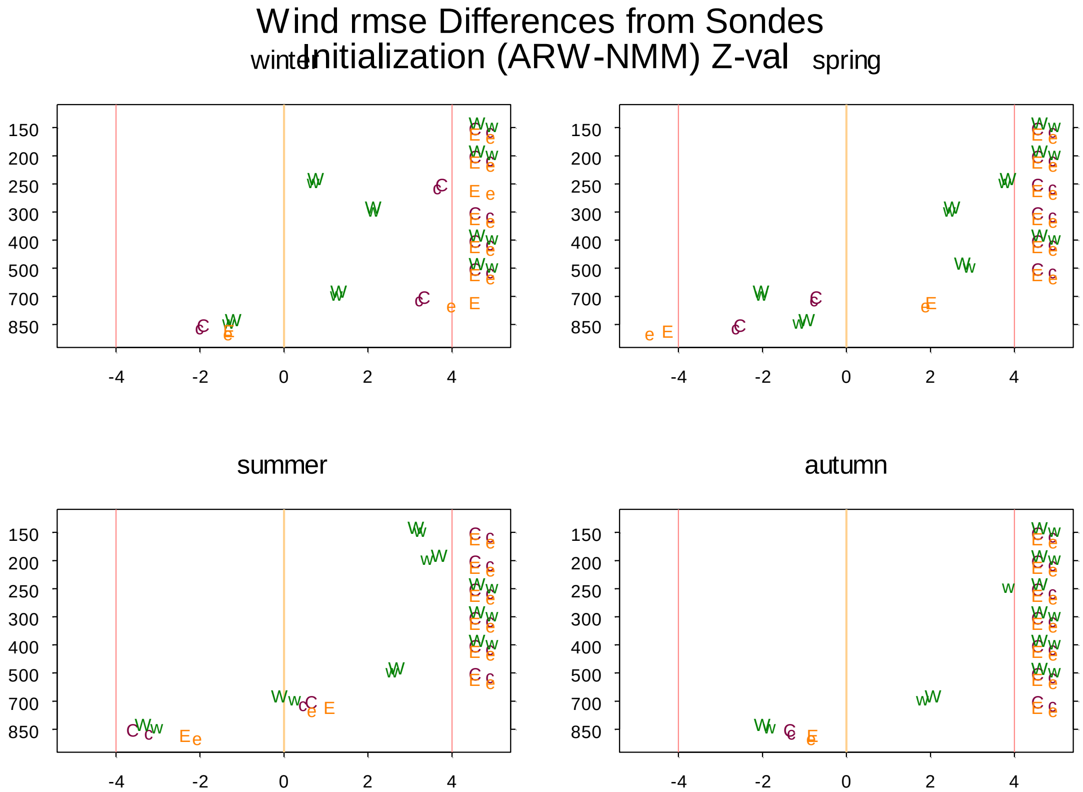


Caption: The differences in wind rmse are presented for three regions: Conus (C for physics 1 and c for physics 2), Conus West (W for physics 1 and w for physics 2)and Conus East (E for physics 1 and e for physics 2). Model rmse is always positive. If the rmse of ARW is larger, the results will be on the right side of the plot. If the rmse of NMM is larger, the results will be on the left side of the plot. The accompanying error bars represent a 95% confidence limit on the difference of the wind rmse. Results are presented for eight pressure levels, with the three regions slightly offset vertically for clarity.



Capital letters are for phase 1; Weatherhead and Noonan

Caption: The number of standard errors represents how strong the statistical significance is between ARW rmse and NMM rmse. The values are presented for three regions: Conus (C for physics 1 and c for physics 2), Conus West (W for physics 1 and w for physics 2)and Conus East (E for physics 1 and e for physics 2). When the significance is more than four standard errors, either positively or negatively, the symbol for that region is listed in the end region. If there is little significant difference between the two models, the symbol for that region will appear near zero. If the statistical analysis indicates that it is highly likely that the rmse of ARW is larger than NMM and the symbol will appear more to the right. Similarly, if the statistical analysis indicates that it is highly likely that the rmse of NMM is larger than ARM, the symbol will appear on the left. The actual difference in rmse may be very small. This plot only indicates how confident we are that the difference is significant in a statistical sense.



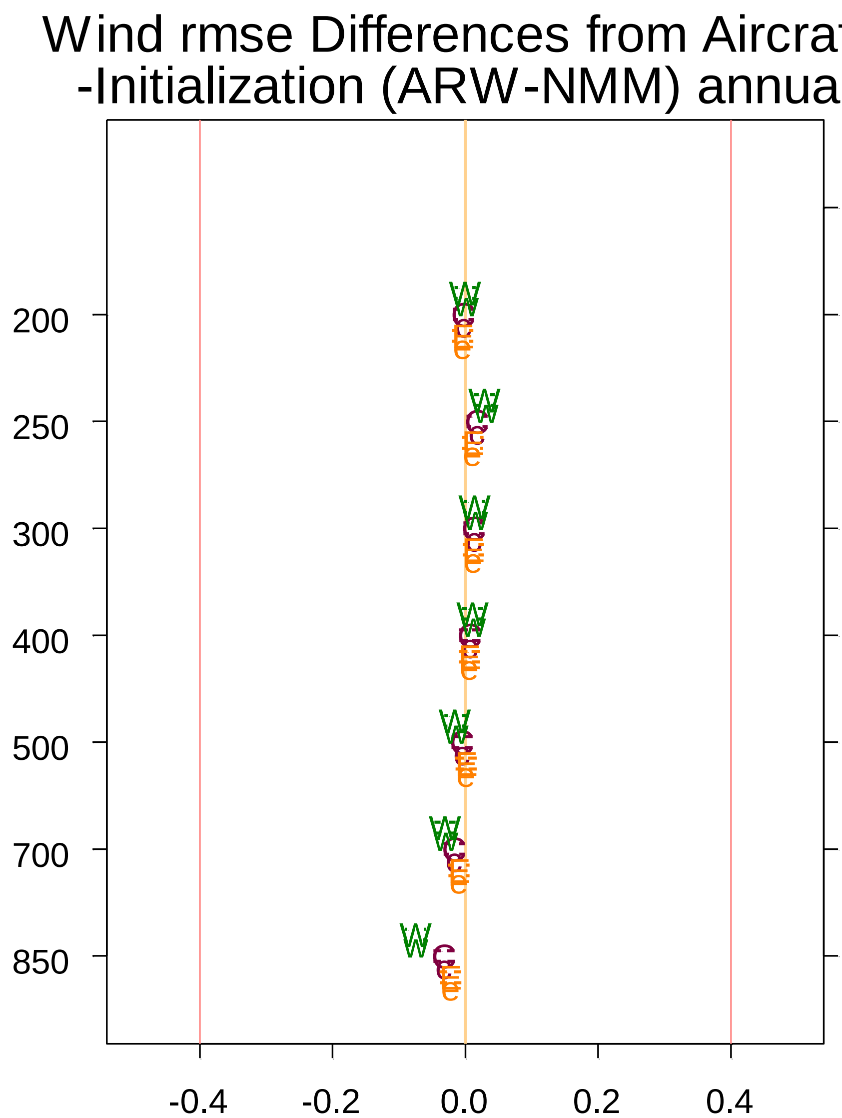
Caption: The number of standard errors represents how strong the statistical significance is between ARW rmse and NMM rmse. The values are presented for three regions: Conus (C for physics 1 and c for physics 2), Conus West (W for physics 1 and w for physics 2)and Conus East (E for physics 1 and e for physics 2). When the significance is more than four standard errors, either positively or negatively, the symbol for that region is listed in the end region. If there is little significant difference between the two models, the symbol for that region will appear near zero. If the statistical analysis indicates that it is highly likely that the rmse of ARW is larger than NMM and the symbol will appear more to the right. Similarly, if the statistical analysis indicates that it is highly likely that the rmse of NMM is larger than ARM, the symbol will appear on the left. The actual difference in rmse may be very small. This plot only indicates how confident we are that the difference is significant in a statistical sense.

RMSE for Sonde Vector Wind / Conus / Forecast Hour 00 (initialization)

ANNUAL

	ARW.ph1	NMM.ph1	ARW.ph2	NMM.ph2	ph1.diff	ph2.diff	# of Obs
P150	3.33224	3.21285	3.33243	3.21158	0.11939	0.12084	67.99265
sigma	0.09823	0.09817	0.10085	0.10092	0.00738	0.00754	232.00000
p-val	33.92219	32.72607	33.04426	31.82400	16.18598	16.02563	232.00000
P200	3.23128	3.12937	3.23029	3.12994	0.10191	0.10035	69.58447
sigma	0.05966	0.06109	0.06112	0.06264	0.00624	0.00632	232.00000
p-val	54.16550	51.22842	52.85382	49.96564	16.33076	15.89028	232.00000
P250	3.32915	3.26680	3.34221	3.28095	0.06235	0.06126	71.77859
sigma	0.08223	0.08445	0.08452	0.08681	0.00594	0.00601	232.00000
p-val	40.48396	38.68160	39.54335	37.79321	10.49598	10.18859	232.00000
P300	3.12248	3.04747	3.12652	3.05073	0.07501	0.07579	72.24134
sigma	0.06324	0.06538	0.06458	0.06700	0.00571	0.00588	232.00000
p-val	49.37147	46.61154	48.41309	45.53622	13.13120	12.88285	232.00000
P400	2.62538	2.56879	2.63104	2.57429	0.05659	0.05675	73.88667
sigma	0.04879	0.04807	0.04919	0.04856	0.00265	0.00264	232.00000
p-val	53.80635	53.43512	53.48556	53.01769	21.37372	21.51110	232.00000
P500	2.53632	2.49249	2.53654	2.49223	0.04383	0.04431	72.86149
sigma	0.04890	0.04877	0.04916	0.04914	0.00285	0.00284	232.00000
p-val	51.86288	51.10174	51.60110	50.71222	15.39142	15.60196	232.00000
P700	2.39553	2.38256	2.38821	2.37579	0.01297	0.01242	73.29171
sigma	0.04838	0.04848	0.04997	0.05008	0.00329	0.00339	232.00000
p-val	49.51492	49.14738	47.79067	47.44085	3.94447	3.66071	232.00000
P800	2.34375	2.35940	2.34990	2.36539	-0.01564	-0.01549	67.90766
sigma	0.03736	0.03825	0.03835	0.03930	0.00355	0.00361	232.00000
p-val	62.73104	61.68241	61.27025	60.19311	-4.40841	-4.29630	232.00000

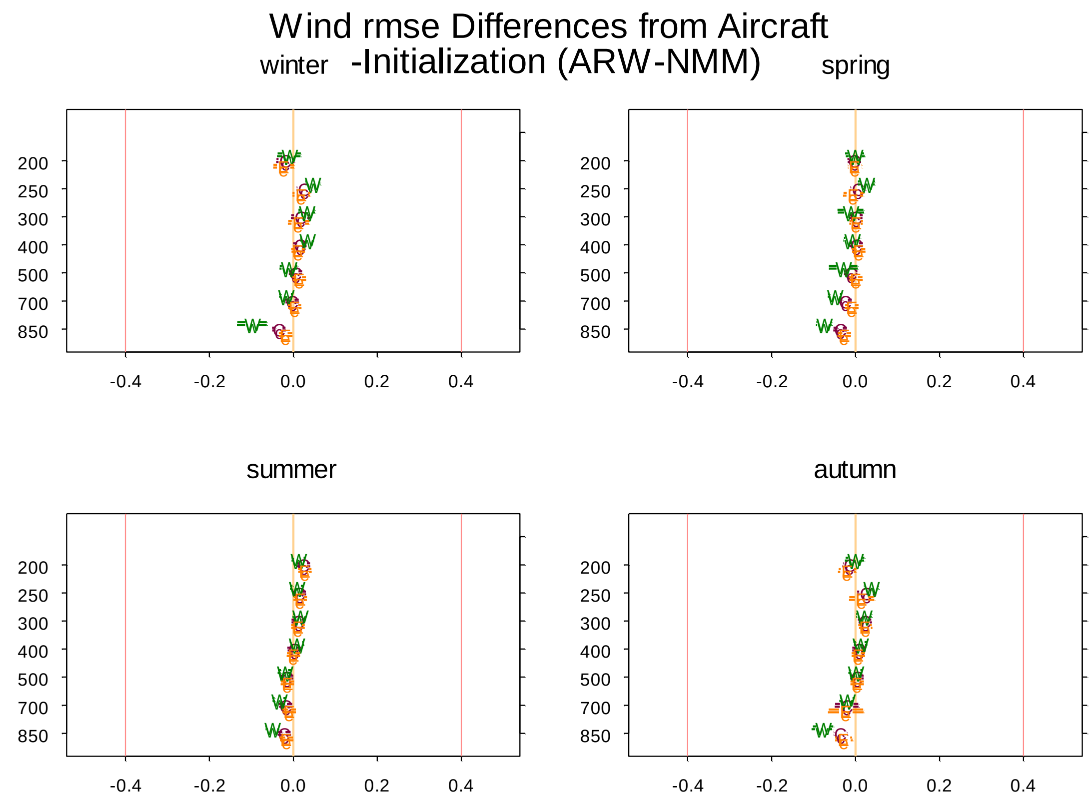
RMSE for Vector Wind from Aircraft - (initialization)



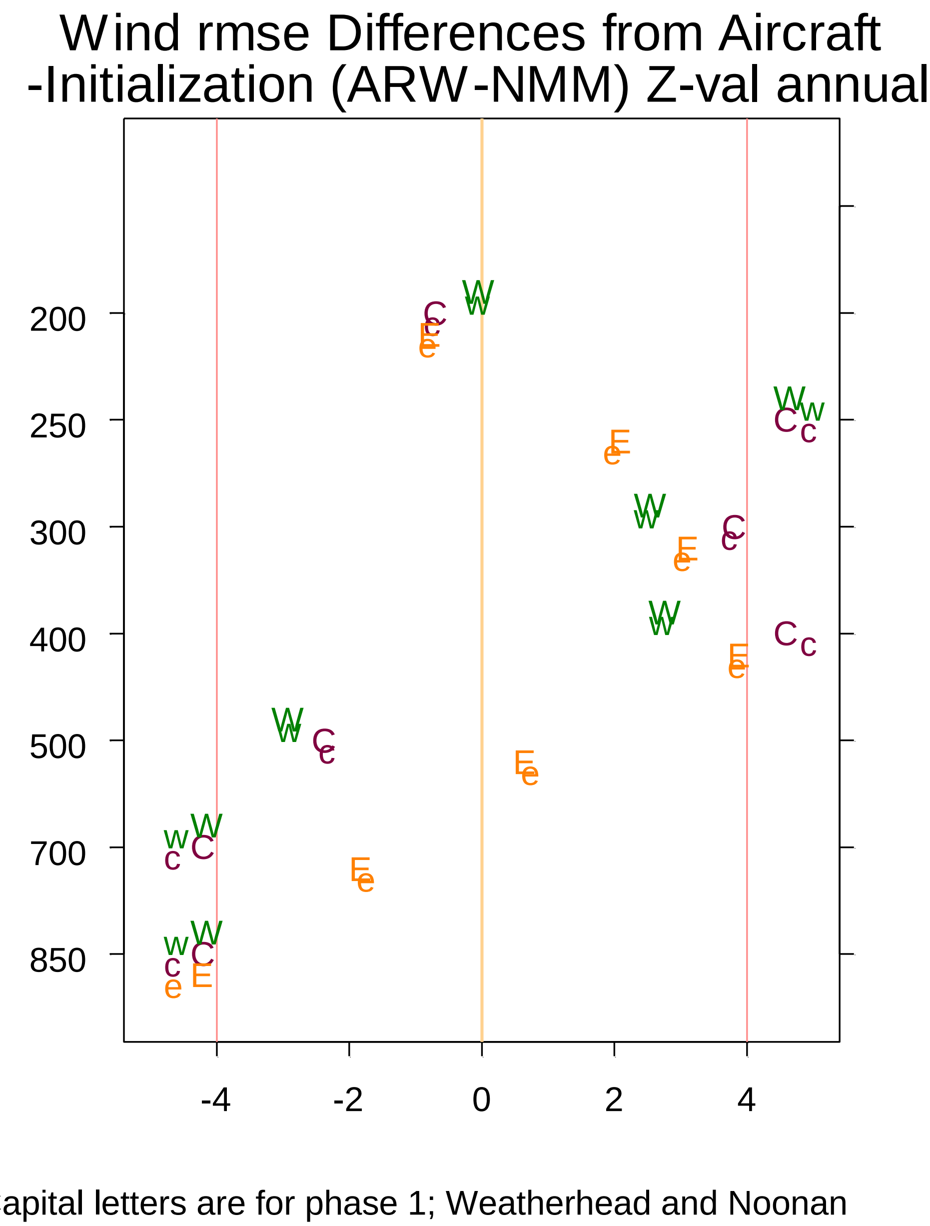
Capital letters are for phase 1; Weatherhead and Noonan

Caption: The differences in wind rmse are presented for three regions: Conus (C for physics 1 and c for physics 2), Conus West (W for physics 1 and w for physics 2)and Conus East (E for physics 1 and e for physics 2). Model rmse is always positive. If the rmse of ARW is larger, the results will be on the right side of the plot. If the rmse of NMM is larger, the results will be on the left side of the plot. The accompanying error bars represent a 95% confidence limit on the difference of the wind rmse. Results are presented for seven pressure bins, with the three regions slightly offset vertically for clarity.

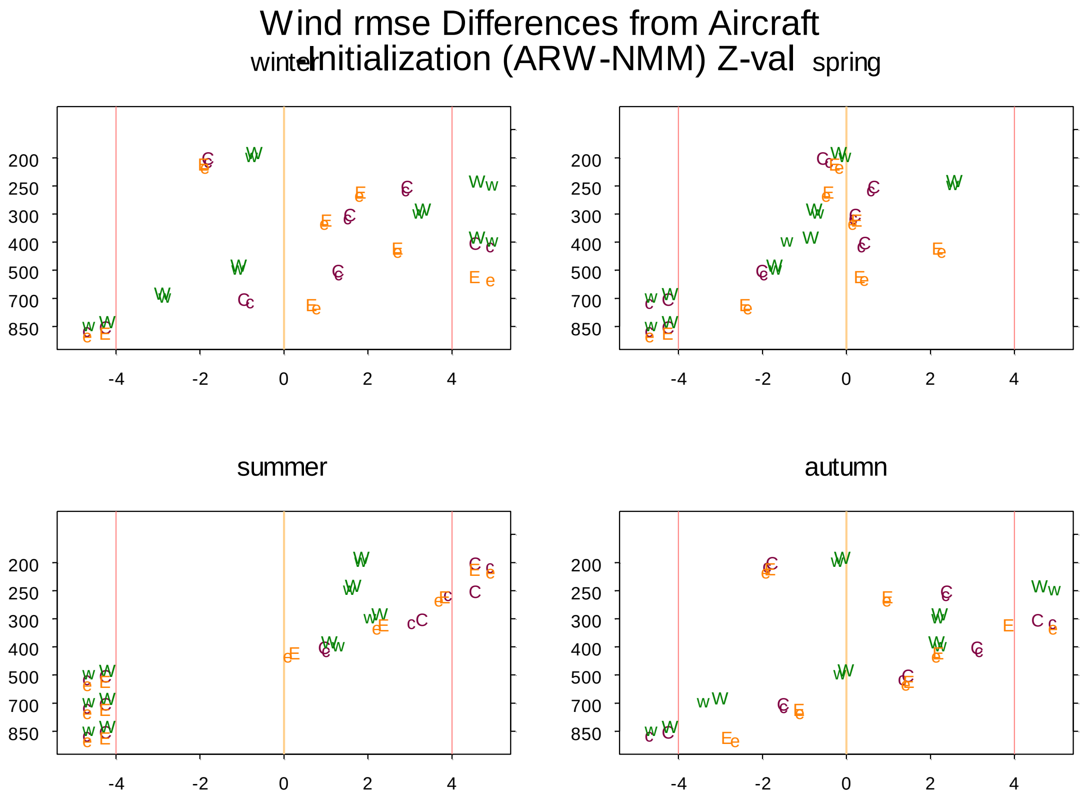
(For comparison: see sonde RMSE at initialization on page 68)



Caption: The differences in wind rmse are presented for three regions: Conus (C for physics 1 and c for physics 2), Conus West (W for physics 1 and w for physics 2)and Conus East (E for physics 1 and e for physics 2). Model rmse is always positive. If the rmse of ARW is larger, the results will be on the right side of the plot. If the rmse of NMM is larger, the results will be on the left side of the plot. The accompanying error bars represent a 95% confidence limit on the difference of the wind rmse. Results are presented for seven pressure bins, with the three regions slightly offset vertically for clarity.



Caption: The number of standard errors represents how strong the statistical significance is between ARW rmse and NMM rmse. The values are presented for three regions: Conus (C for physics 1 and c for physics 2), Conus West (W for physics 1 and w for physics 2)and Conus East (E for physics 1 and e for physics 2). When the significance is more than four standard errors, either positively or negatively, the symbol for that region is listed in the end region. If there is little significant difference between the two models, the symbol for that region will appear near zero. If the statistical analysis indicates that it is highly likely that the rmse of ARW is larger than NMM and the symbol will appear more to the right. Similarly, if the statistical analysis indicates that it is highly likely that the rmse of NMM is larger than ARM, the symbol will appear on the left. The actual difference in rmse may be very small. This plot only indicates how confident we are that the difference is significant in a statistical sense.



Caption: The number of standard errors represents how strong the statistical significance is between ARW rmse and NMM rmse. The values are presented for three regions: Conus (C for physics 1 and c for physics 2), Conus West (W for physics 1 and w for physics 2)and Conus East (E for physics 1 and e for physics 2). When the significance is more than four standard errors, either positively or negatively, the symbol for that region is listed in the end region. If there is little significant difference between the two models, the symbol for that region will appear near zero. If the statistical analysis indicates that it is highly likely that the rmse of ARW is larger than NMM and the symbol will appear more to the right. Similarly, if the statistical analysis indicates that it is highly likely that the rmse of NMM is larger than ARM, the symbol will appear on the left. The actual difference in rmse may be very small. This plot only indicates how confident we are that the difference is significant in a statistical sense.

RMSE for Aircraft Vector Wind / Conus / Forecast Hour 00 (initialization)

ANNUAL

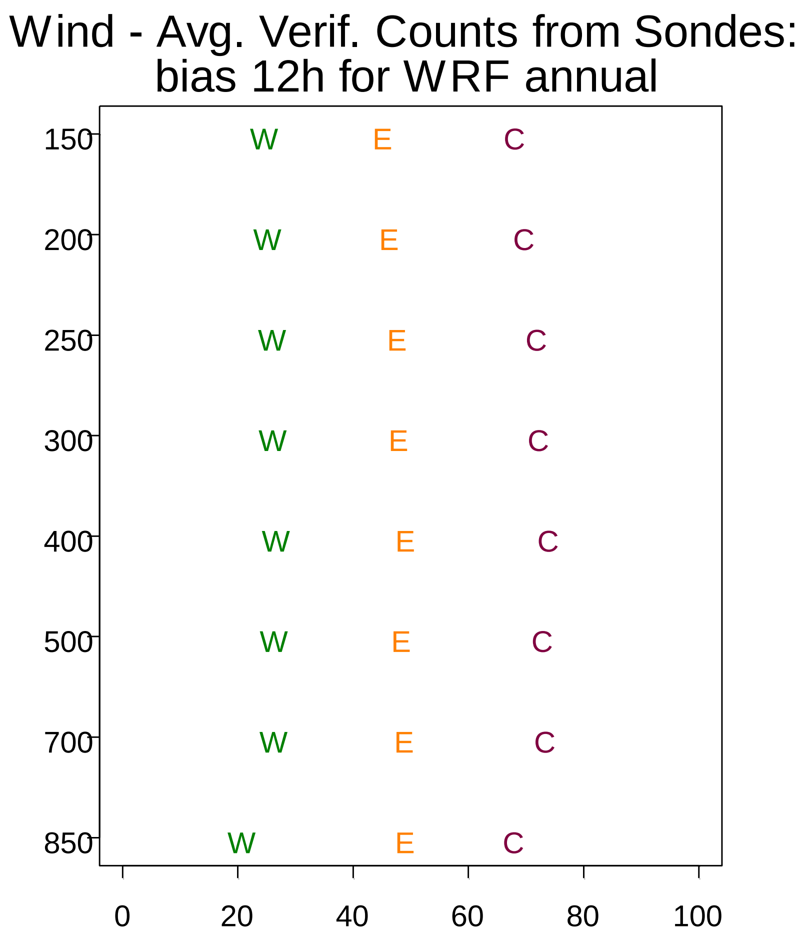
	ARW.ph1	NMM.ph1	ARW.ph2	NMM.ph2	ph1.diff	ph2.diff	# of Obs
P200-250	5.03904	5.04181	5.04095	5.04397	-0.00277	-0.00301	1209.51579
sigma	0.07157	0.07287	0.07207	0.07329	0.00395	0.00404	232.00000
p-val	70.41202	69.18820	69.94412	68.82185	-0.70023	-0.74562	232.00000
P250-300	4.92378	4.90626	4.92464	4.90704	0.01752	0.01760	600.44211
sigma	0.04991	0.04943	0.04902	0.04847	0.00412	0.00425	232.00000
p-val	98.65857	99.25624	100.45901	101.23461	4.25291	4.14215	232.00000
P300-400	4.37510	4.36214	4.37306	4.35994	0.01296	0.01313	409.02421
sigma	0.05497	0.05482	0.05483	0.05469	0.00340	0.00352	232.00000
p-val	79.59172	79.57246	79.76258	79.72334	3.80862	3.73375	232.00000
P400-550	3.50910	3.50241	3.51091	3.50415	0.00668	0.00677	664.30632
sigma	0.05169	0.05131	0.05396	0.05354	0.00154	0.00157	232.00000
p-val	67.88941	68.26502	65.06601	65.45459	4.34790	4.30009	232.00000
P550-700	3.24773	3.25263	3.24473	3.24965	-0.00489	-0.00492	814.02526
sigma	0.03507	0.03451	0.03546	0.03493	0.00205	0.00211	232.00000
p-val	92.60873	94.23835	91.51180	93.04560	-2.38235	-2.33206	232.00000
P700-850	3.44267	3.45952	3.43961	3.45638	-0.01685	-0.01678	666.38211
sigma	0.03866	0.03860	0.03868	0.03869	0.00366	0.00377	232.00000
p-val	89.04995	89.62771	88.92072	89.34115	-4.60146	-4.44763	232.00000
P850-1000	3.48912	3.52137	3.48518	3.51711	-0.03225	-0.03193	935.83789
sigma	0.06467	0.06503	0.06628	0.06665	0.00239	0.00245	232.00000
p-val	53.95499	54.15056	52.58187	52.77110	-13.49316	-13.00975	232.00000

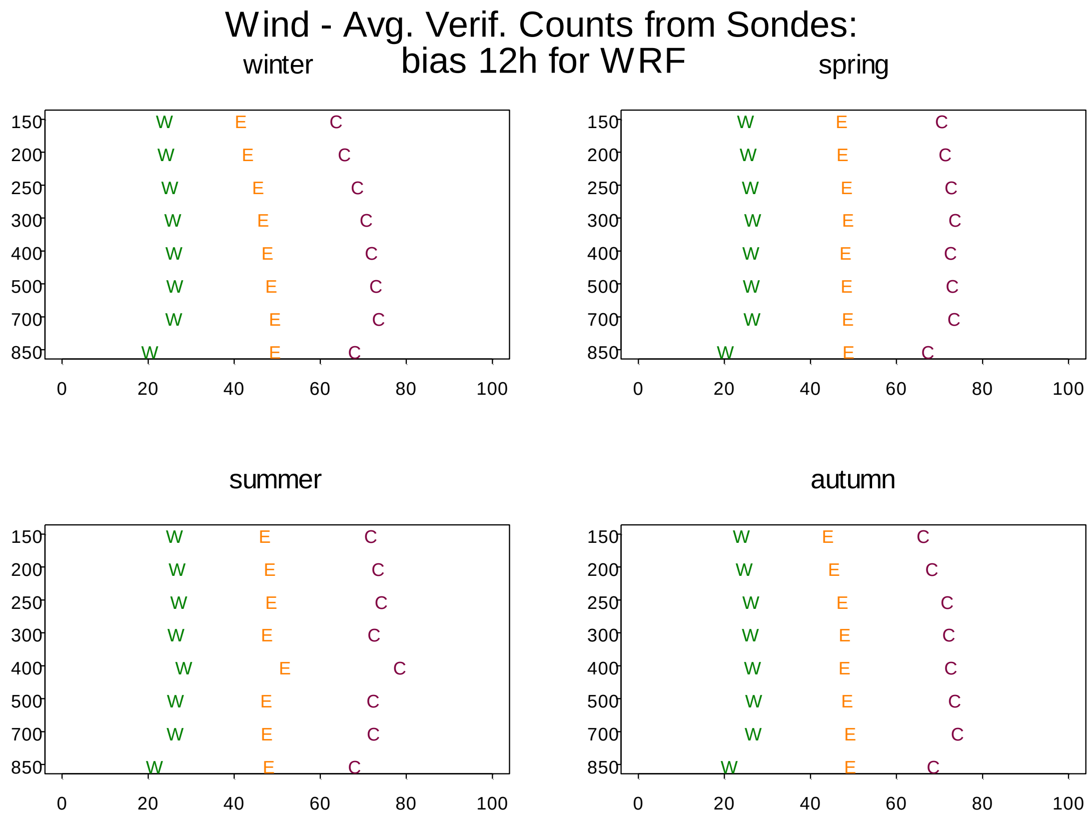
Appendix 8: Dependence of Results

One concern is that the observations have known problems for winds which may result in spurious conclusions. In particular, the possibility for wind blow off to sondes or aircraft preferentially seeking or avoiding strong winds. To examine how the results here may depend on these particular aspects of the data, we look at some of the relationships available in the data.

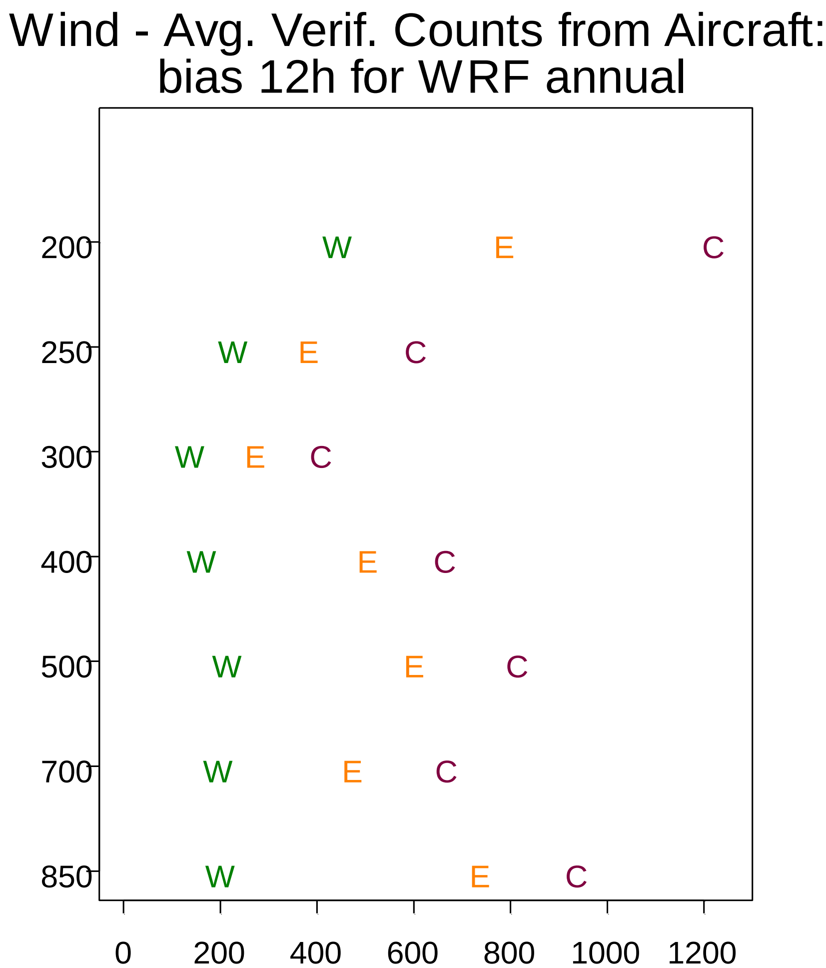
Does the sonde observation rate vary with height?

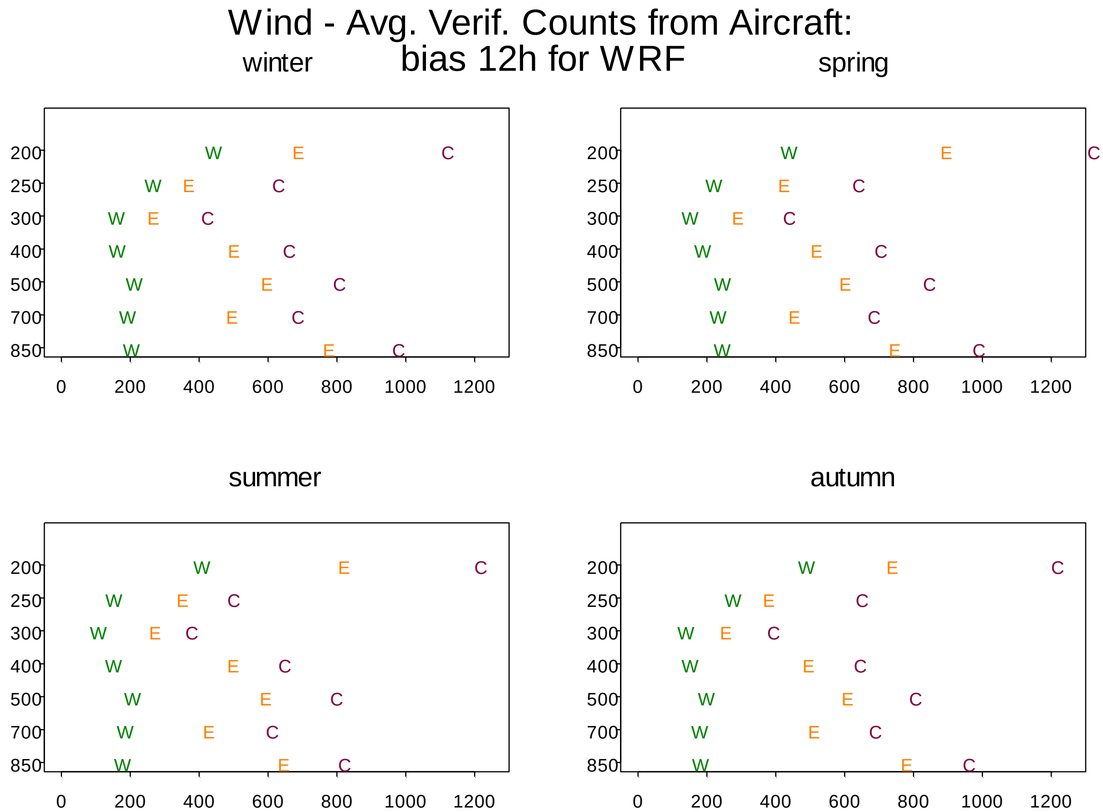
First, to test the possibility that a large number of sondes are blown out of use at high altitudes, we examine the average number of sondes used in each forecast verification as a function of height.




Wind - Avg. Verif. Counts from Sondes:
bias 12h for WRF
winter
spring
summer
autumn
150
200
250
300
400
500
700
850
0
20
40
60
80
100

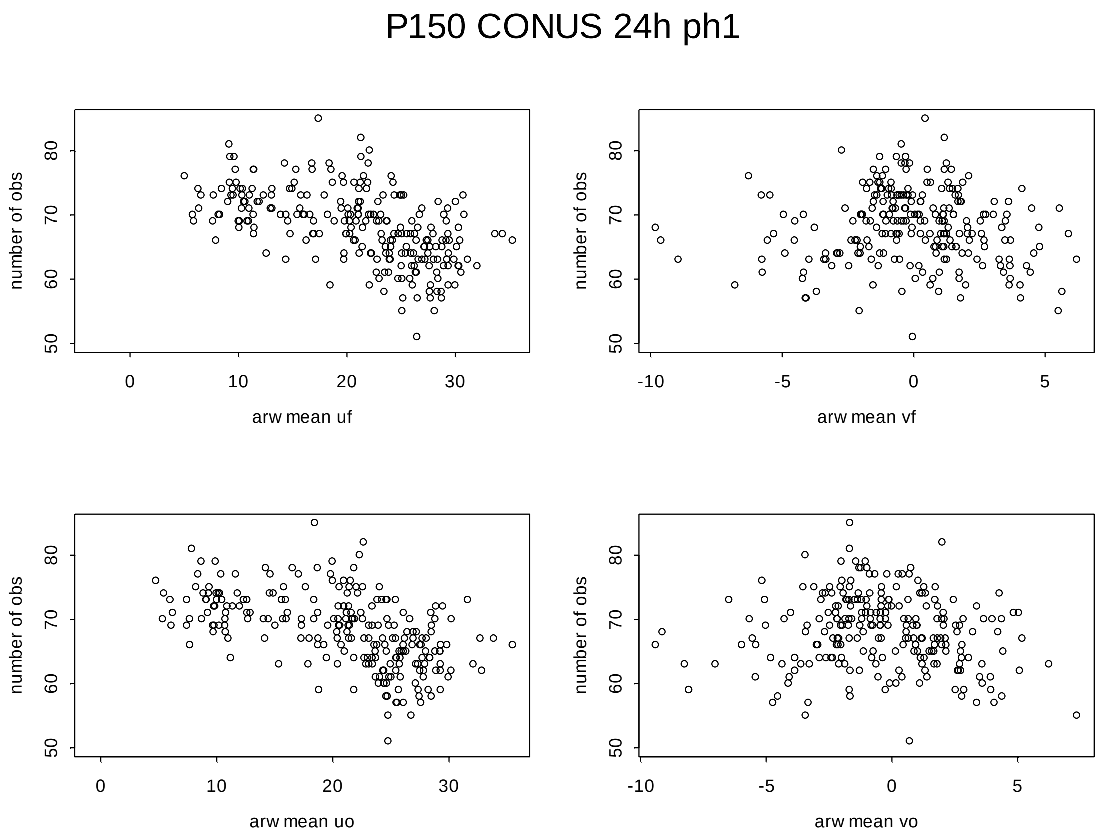
We can contrast the previous plot with the average number of observations available for each forecast from aircraft.



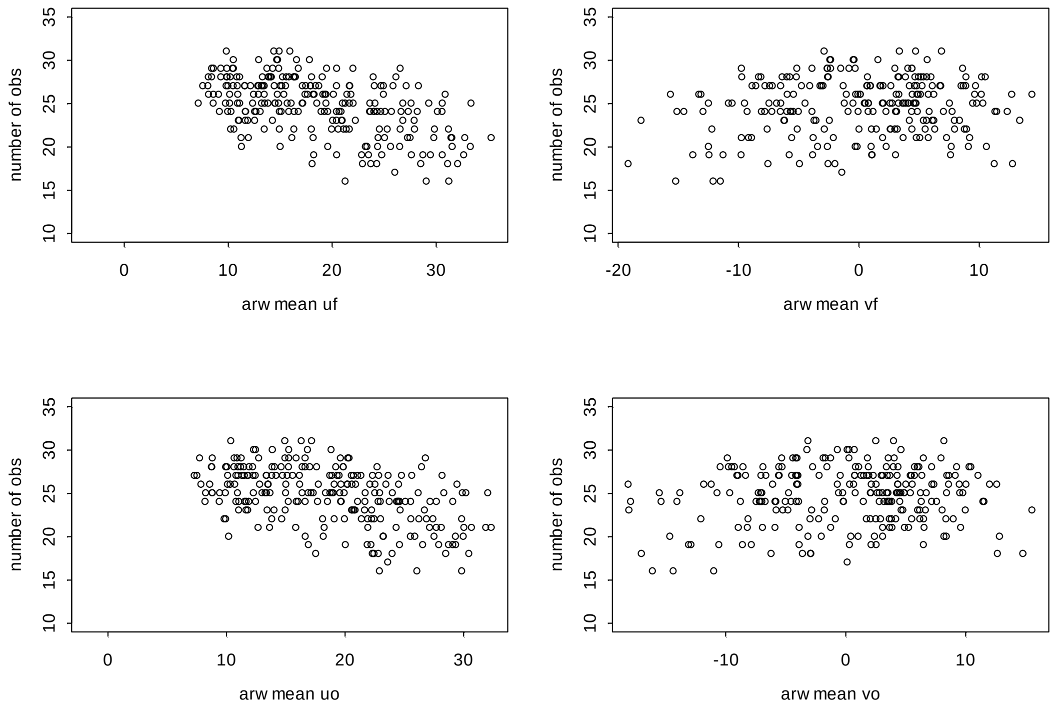


It is clear that there is some dropoff of sondes at higher altitudes. We now turn to the question of whether these dropoffs occur preferentially, for instance during high wind events. This may be better handled on a case-by-case basis, but for this study we are trying to understand if the dropoff rate is random or if it is related to the meteorological conditions.

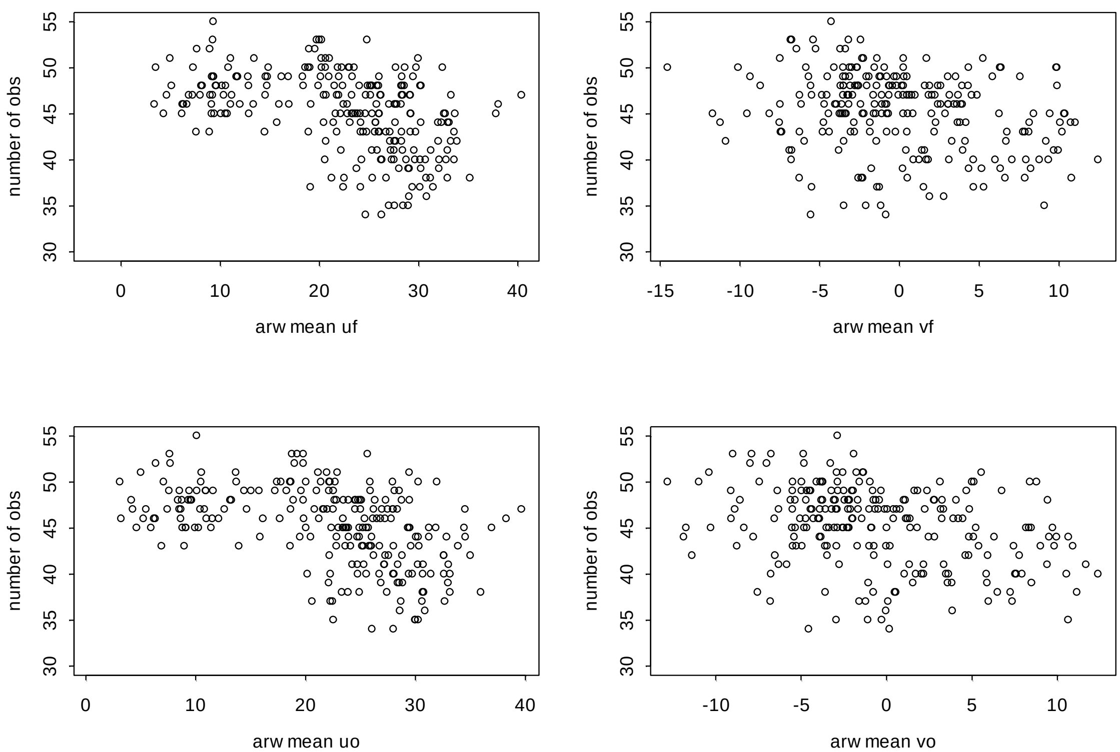
Is the sonde dropoff rate at high altitudes random?



P150 CONUS WEST 24h ph1

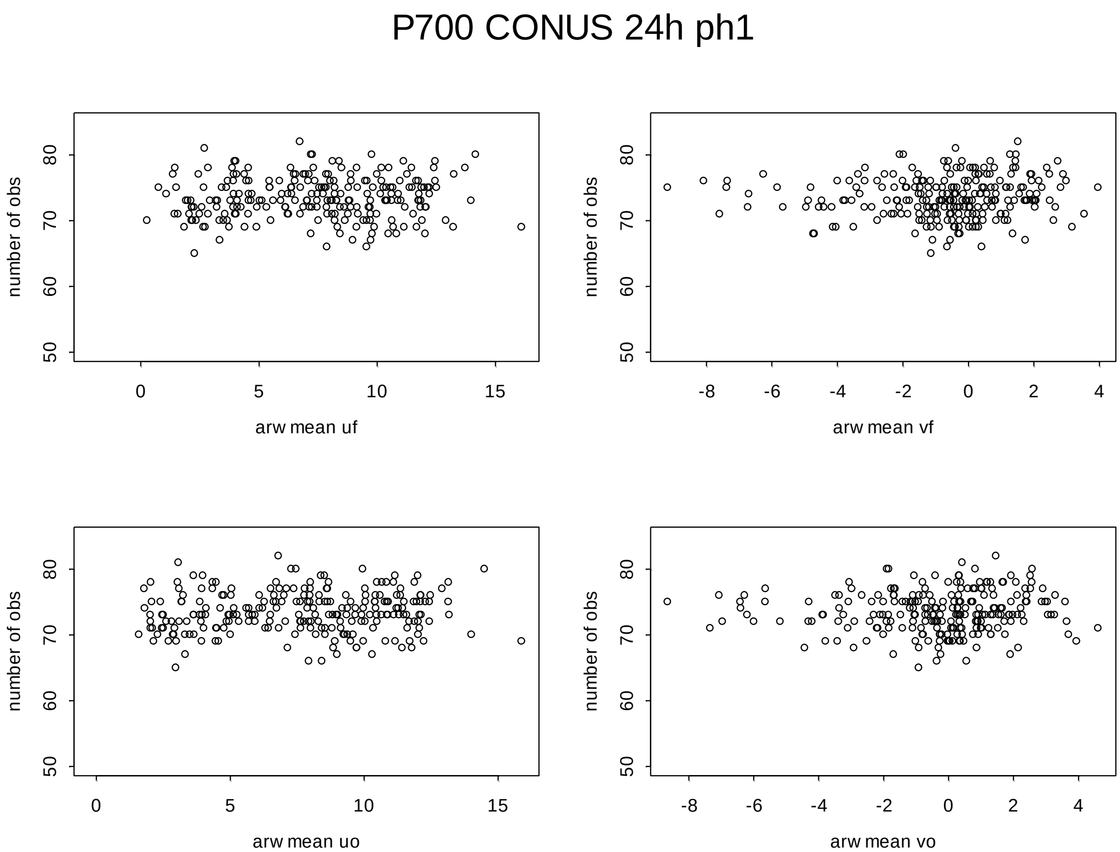


P150 CONUS EAST 24h ph1



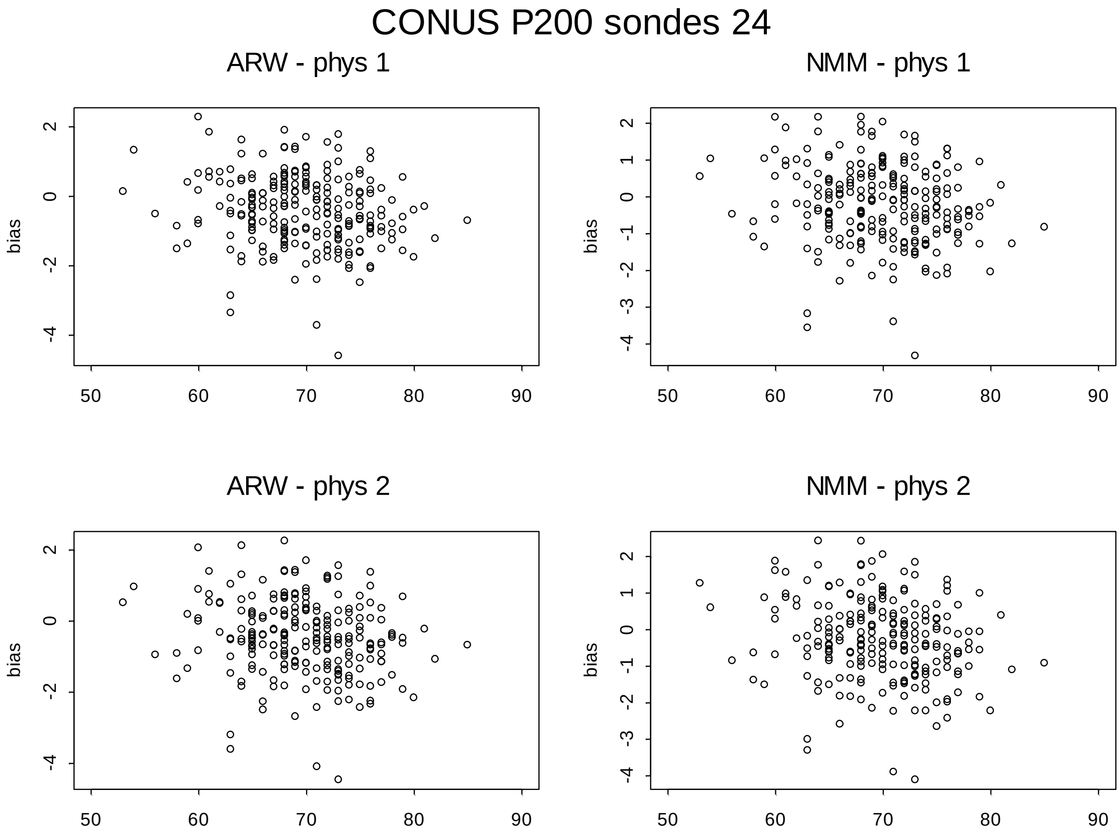
The results show a strong relationship between the east-west winds, with the days with fewer observations available for forecast verification occurring almost exclusively in the days with high average east-west winds.

It is significant to note that the relationship in the previous plots is not true for lower in the atmosphere as shown in the set of figures below.



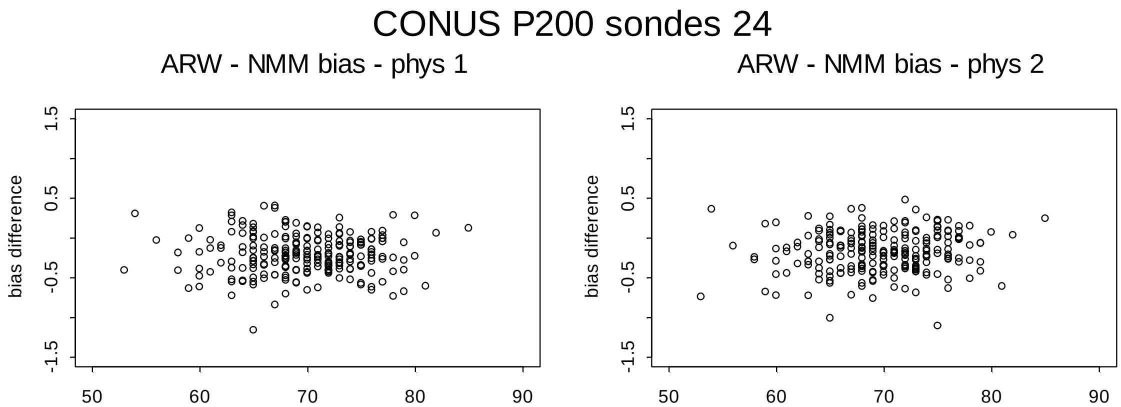
A more relevant question is whether the bias results are a function of how many observations are taken. In other words, are the blow-offs likely to be influencing the derived bias results?

Are the estimates of bias influenced by the number of sonde observations?



Are the estimates of bias differences (ARW-NMM) influenced by the number of sonde observations?

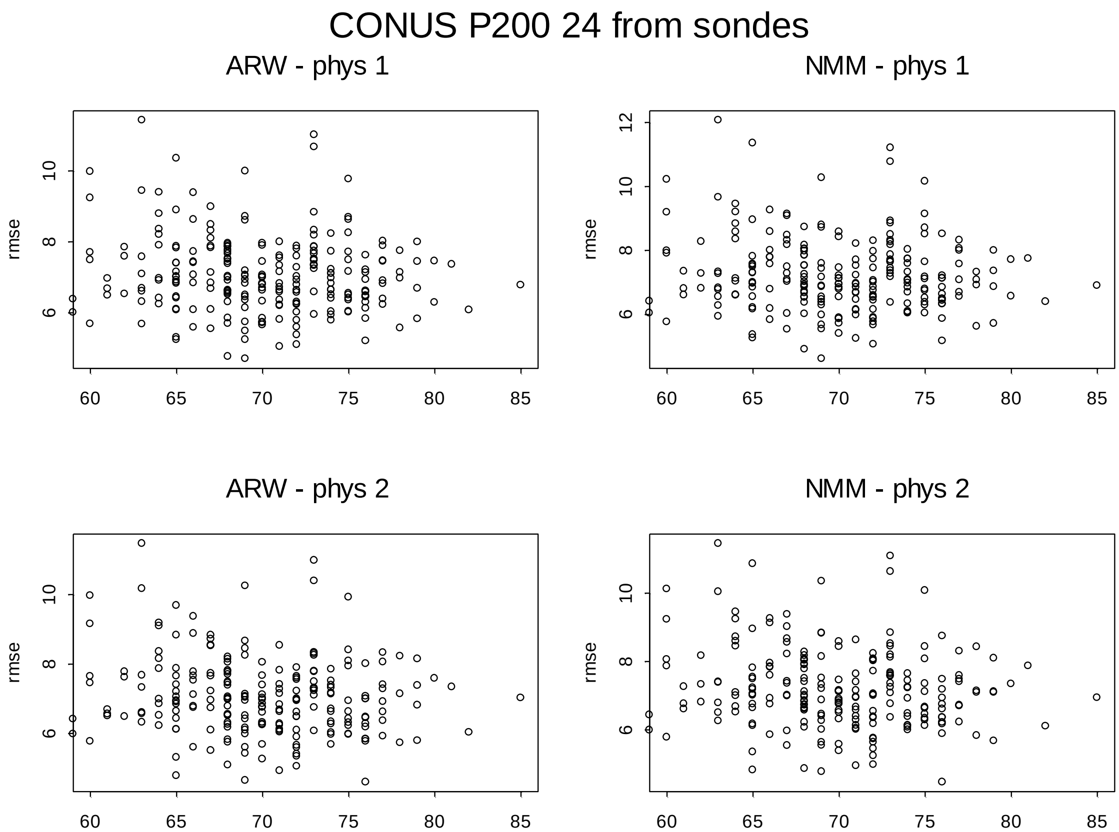
The most important question is whether the dropoffs are influencing the derived differences in bias derived from the ARW and NMM models? One way to estimate this is to see if the difference in bias varies based on how many measurements are available. The following plot shows that the difference is roughly independent of the number of measurements.



This plot shows how the bias differences for all 24 hour forecasts vary as a function of the number of measurements available for verification. The results show that the bias differences are not strongly dependent on the number of measurements.

Are the estimates of RMSE influenced by the number of sonde observations?

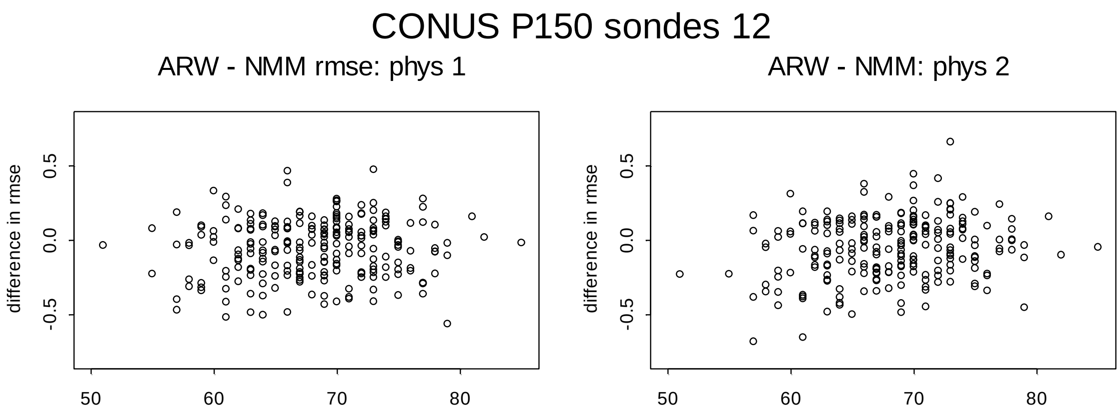
Similar questions can be asked with respect to RMSE. Are the RMSE results a function of how many sondes are available for verification?



This plot shows that in general the RMSE differences are smaller when a large number of sondes are available for observations.

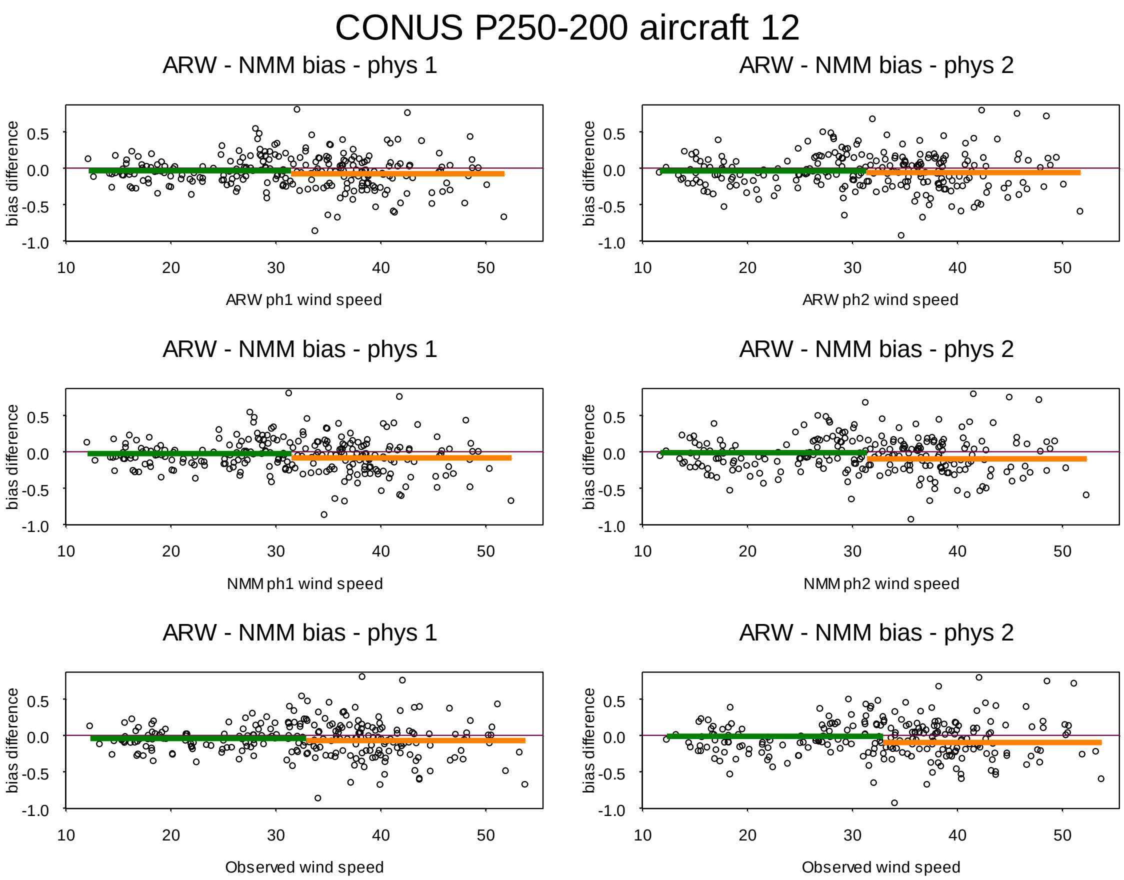
Are the estimates of RMSE differences (ARW-NMM) influenced by the number of sonde observations?

And the perhaps the most relevant question is whether the dropoff rates are influencing the differences in RMSE between the ARW and NMM cores?

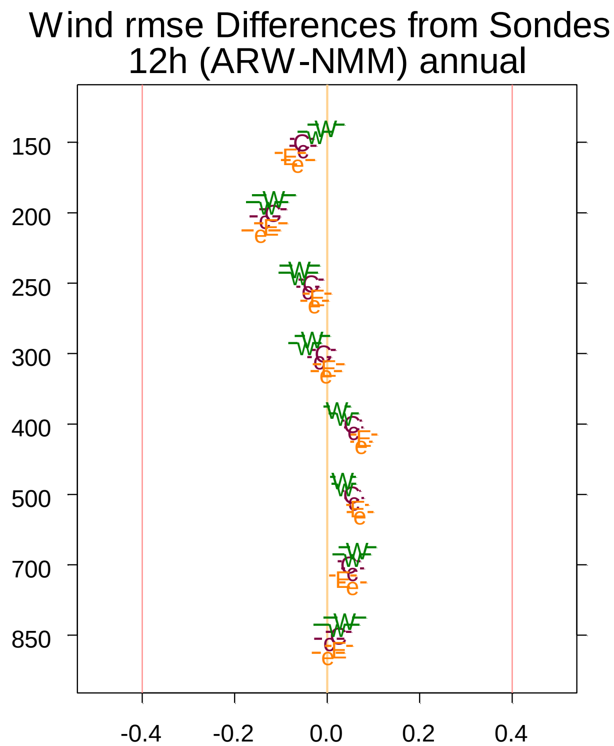


Next we look at how bias differences depend on the wind speed to determine if the models behave differently in high wind situations.

Do differences in the bias scores have any dependence on wind speed?

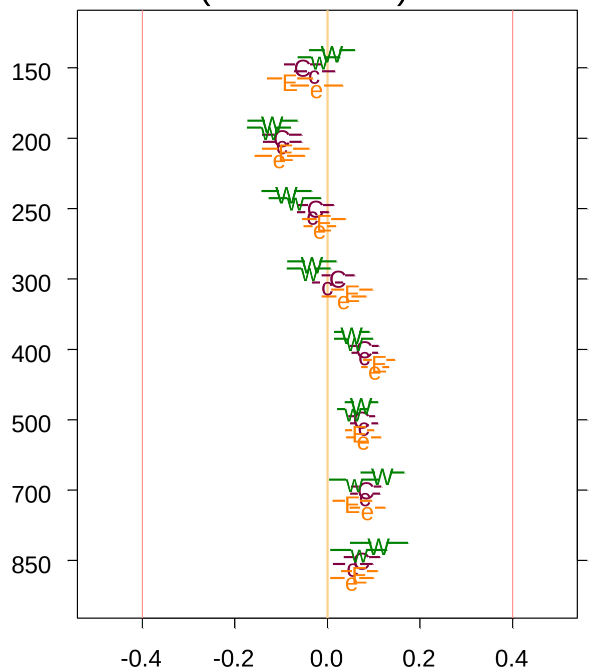


Do differences in the RMSE scores have any dependence on wind speed?

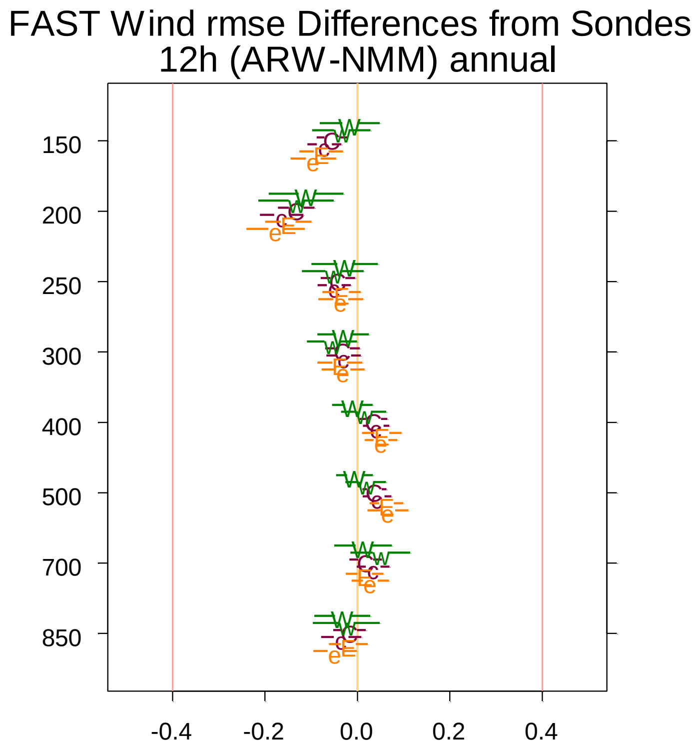


Capital letters are for phase 1; Weatherhead and Noonan

SLOW Wind rmse Differences from Sondes
12h (ARW-NMM) annual

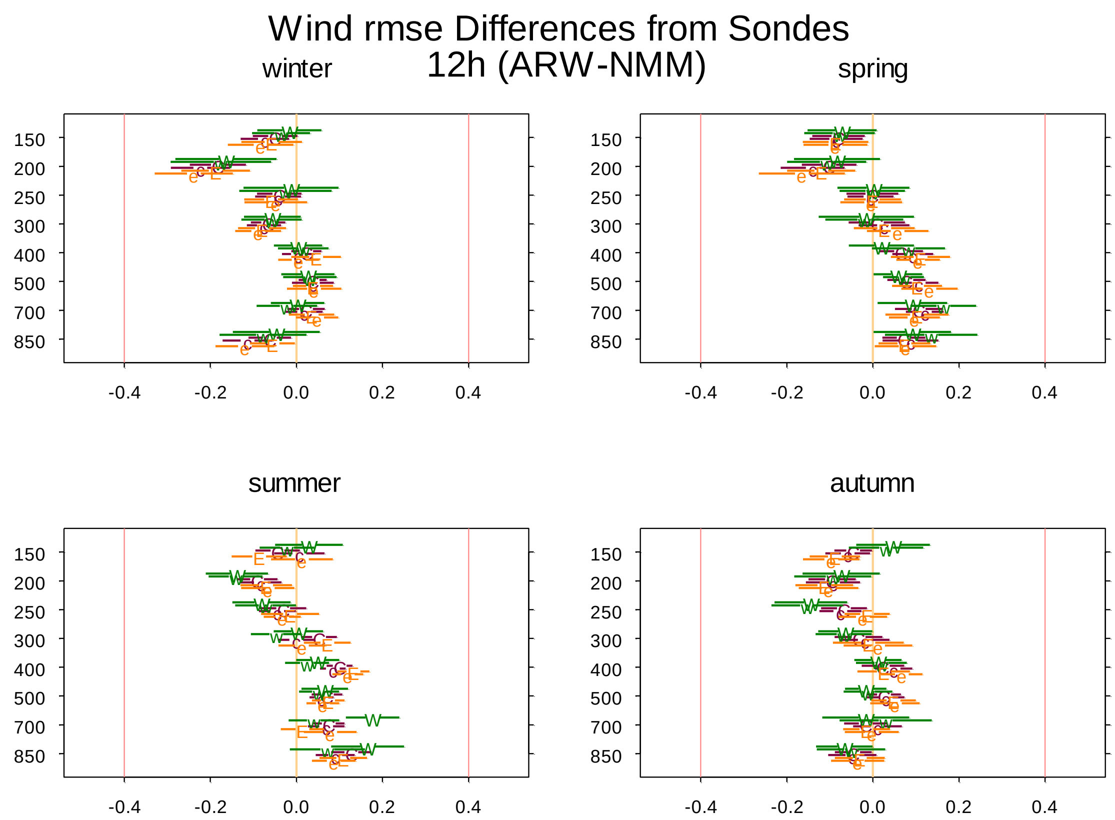


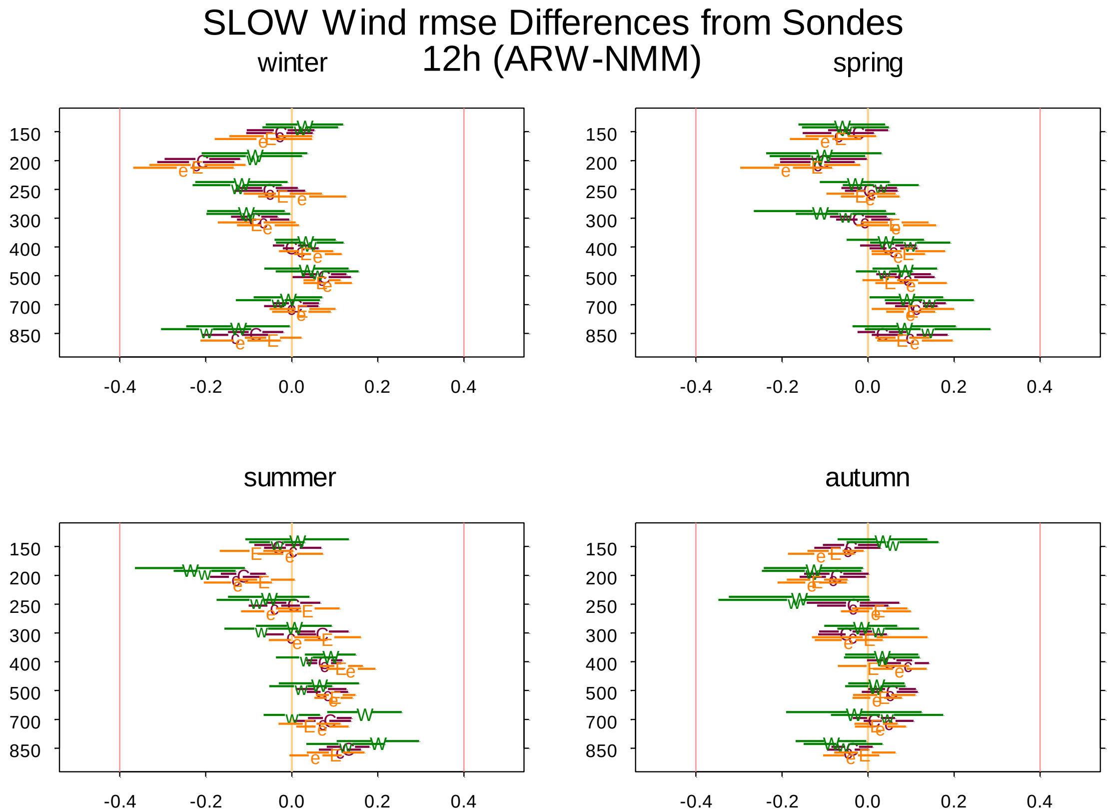
Capital letters are for phase 1; Weatherhead and Noonan

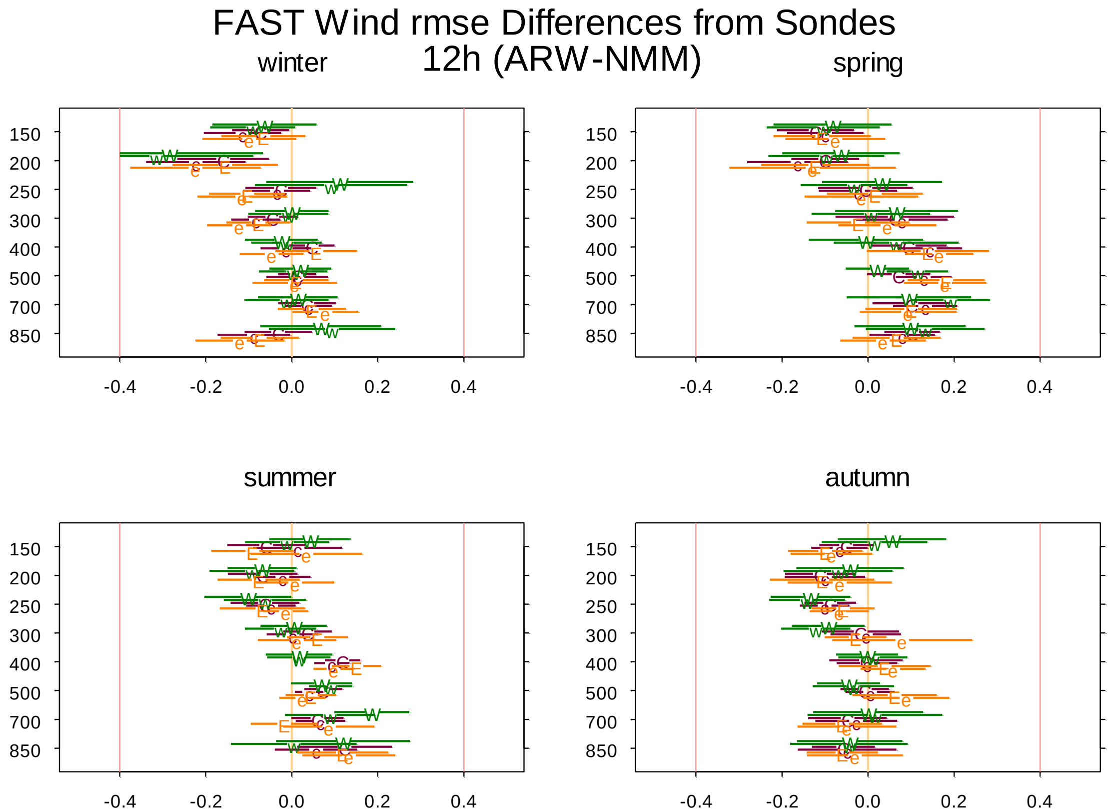


Capital letters are for phase 1; Weatherhead and Noonan

Do seasonal differences in RMSE scores have any dependence on wind speed?




SLOW Wind rmse Differences from Sondes
12h (ARW-NMM)
winter
spring
summer
autumn
150
200
250
300
400
500
700
850
-0.4
-0.2
0.0
0.2
0.4


FAST Wind rmse Differences from Sondes
12h (ARW-NMM)
winter
spring
summer
autumn
150
200
250
300
400
500
700
850
-0.4
-0.2
0.0
0.2
0.4

Appendix 9: Discussion of Statistical Assumptions

The statistical tests employed in analyzing the temperature, relative humidity and wind fields, as for all statistical tests, are based on a number of assumptions about the underlying distributions of the data. These assumptions were tested and decisions were made concerning how to address these issues.

Normal (Gaussian) distribution
The pairwise data were tested for normal distributions by comparing against a standard normal distribution. The data were found to be nearly Gaussian for many cases and showed some skewness in other cases. No corrections were made due to the non-gaussian nature of the data. Corrections would likely have a small influence on the derived bias or RMSE, but would have a larger influence on the uncertainty associated with these estimates (the error bars). Two reasons for ignoring the non-Gaussian nature were time required to make the non-Gaussian corrections and the value in having a single standard deviation associated with each estimate.

The most critical situations for the non-Gaussian behavior to cause problems is in the situation of unusual outliers. Some unusual values were identified visually. These outliers were modest, because of prescreening before the data were made available for analysis. The outliers could have some difference on the seasonally estimated values, but would have a more pronounced influence on the standard deviation of the estimated values.

Independence of the data
The pairwise data were tested for temporal independence. That is, the data were examined to determine if each forecast represented an independent test of the difference in the models. Both a Partial Autocorrelation Analysis and an Autocorrelation Analysis were performed on a sub sample of the data. The results show that the differences support an AR(1) model with time lag of 12 hours. Physically this implies that if the ARW had a larger bias than the NMM in day 20's 00Z 12 hour forecast, it would likely have a larger bias in its day 20's 12Z 12 hour forecast. However a few days later, NMM may have the advantage. In other words, the unique challenges determining which model was better did not change on the time scale in which new forecasts were made (12 hours). This redundancy in the data indicates that there were not 241 independent test of how well the ARW and NMM models perform. This lack of independence was taken into account by deriving the AR(1) coefficient independently and adjusting the confidence intervals accordingly. Taking the independence into account has a small influence on the estimated values (bias, RMSE, difference in bias and difference in RMSE), but has a direct affect on the error bars associated with these values. The effect of incorporating the data's redundancy is to increase the size of the error bars, typically by about 20-50%.

The incorporation of the temporal autocorrelation is important for planning future intercomparisons. Because redundancy reduces the amount of used information, more value is obtained from running comparisons every second or third day, if possible rather than running the comparisons every twelve hours. Defensible estimates can be made as to how many forecasts need to be compared to note a particular difference in, for instance, relative humidity bias. These calculations can help make sure that future comparisons result in meaningful and useful results.

Equal weighting.

For most upper air measurements, the bias and RMSE results are based on between 25 and 260 upper air observations, depending on region and pressure level. We can have a higher confidence that the metrics (bias or RMSE) are more accurate when more measurements are available for comparison. Currently, the difference in confidence we have between the various RMSE and bias values is not incorporated in the statistical analysis. A related issue is whether there is an important reason for the difference in observations. For instance, when few measurements are available at 100mb for winds, this could be because high winds are prevalent, thus our calculations in high wind circumstances may be less certain. The incorporation of the unequal confidence we have on each measurement of bias or RMSE would likely result in small changes in the estimates of bias and RMSE, but could have a slightly larger influence on the uncertainty limits.

Student-t distribution

For small sample sizes, the student-t distribution is important, particularly for estimating mean values and confidence intervals on those mean values. With the lack of independence of data, the effective number of data points is less than the number of collected data points (typically 241 for annual comparisons and roughly 60 for each seasonal comparison). Even assuming that the effective number of data points is 2/3 of the collected number, 40 measurements is large enough that the student-t distribution is identical to the Gaussian distribution to at least five decimal places. Therefore the student-t distribution is not used, but should give identical results as the Gaussian distribution used in this analysis.