3.5 COMPARISON OF AIRCRAFT AND RADIOSONDE TEMPERATURE BIASES AT NCEP

Bradley Ballish[1] and V. Krishna Kumar
NOAA/NWS/NCEP Central Operations, Camp Springs, MD 20746

1. INTRODUCTION

It is well known that wind and temperature data from radiosondes and aircraft are extremely important for Numerical Weather Prediction (NWP) models due to the models showing greater sensitivity to them compared to other data types in the data assimilation systems. Zapotocny et al. (2000) has shown that with the NCEP ETA model (now referred to as the North American Mesoscale (NAM) model), winds from ACARS (Aircraft Communication Addressing and Reporting System) data were more important than those from radiosondes. The ACARS temperature data were next in importance after those from radiosondes. Besides the importance of this data for forecasting, they are important for forecast verification. In addition, they are vital for studying climate change from NCEP/NCAR, ECMWF, NASA and other reanalyses data sets. Furthermore, radiosonde temperatures are used in calibrating satellite radiances (Reale, 2005). In this paper, we show that there are significant differences in radiosonde and aircraft temperature biases especially around 250 hPa. This raises the fundamental question of what is the "truth" in observations. Therefore in this paper, we address some of the key factors that explain some of the bias differences.

2. TEMPERATURE BIAS COMPARISONS

Statistics of the difference between temperature observations with the first guess from the NCEP Global Data Assimilation System (GDAS) (Parrish and Derber 1992) were examined. All temperatures passing the gross check by the analysis are considered for this study. The gross check rejects any observations that differ from the guess by large limits. The limits depend on the type of observation and pressure. At 250 hPa, the gross limit for radiosonde temperature is 12.0 degrees and 10.0 for automated aircraft. Fig.1 and 2 show respectively the monthly average bias (observation minus guess) for 00Z and 12Z from July 2002 to June 2005 in the pressure range of 300 to 200 hPa for radiosondes, ACARS, AMDAR (Aircraft Meteorological Data Relay) and AIREP data. Note that the monthly average radiosonde bias is always negative, while the three types of aircraft data have positive biases. The biases at 06Z and 18Z show similar results (not shown). It is remarkable that the radiosonde temperature bias in this pressure range shows persistent negative values for every model run for cycles 00Z and 12Z for the three year time period that was investigated. Fig. 3 shows observational counts of the data types studied in Fig. 1 and 2. The ACARS data shows roughly 3 times as many observations as the radiosondes but their horizontal coverage is not as uniform. Fig 4 displays the vertical structure of the average biases of radiosondes over the contiguous United States (SONDU), globe (SONDG) as well as ACARS, AMDAR and AIREP for 00Z July 2004. Here all data are interpolated to the nearest mandatory pressure level. Other monthly averages have been examined but are not shown. Typically, the three types of aircraft data show warm biases from 300 to 200 hPa, while the radiosondes exhibit opposite cold biases there. Below 300 hPa the AIREP counts are so low as to make their bias unreliable. The ACARS and AMDAR biases tend to get smaller and stay positive below 300 hPa but show cold bias near 1000 hPa for some months.

Collocation statistics between ACARS temperatures and radiosondes also corroborated the biases that were found, but the computations were considered of limited utility in part due to the non availability of high resolution 6 second data from the radiosondes that are not available to NCEP operations. In addition, collocations were

---

1 Corresponding author address: Dr. Bradley Ballish, NOAA/NWS/NCEP Central Operations, Room: 307, World Weather Building, 5200 Auth Road, Camp Springs, MD 20746; e-mail:
Bradley.Ballish@noaa.gov

low in number, would not sample all aircraft types at all pressures, were mainly over the US 48 states and still required vertical interpolations on the order of 20 hPa.

3. RADIOSONDE TEMPERATURE BIAS ANALYSIS

Investigation of the possible cause of the cold bias shown by radiosondes around 250 hPa, revealed that the NCEP radiation correction (Collins, 1999), made the cold bias colder compared to the guess. Fig. 5 and 6 show respectively the global monthly average bias of radiosondes temperatures versus the guess at 250 hPa for 00Z and 12Z from July 2004 to June 2005. The raw data compared to the guess is shown as "RAWMG", while after any radiation correction at NCEP, the bias is shown as "RADMG". Note that for all months the raw data has a cold bias compared to the guess. The radiation correction makes the cold bias colder. The analysis is colder than the guess as shown as "ANLMG". The analysis cooling is influenced by many factors. Even though the temperature observation errors assigned for radiosondes are 20% larger than the observation errors for automated aircraft around 250 hPa, which would imply a smaller weight given to radiosondes than to aircraft observations, the analysis results in net cooling on average.

This cold temperature bias in the guess is also a common problem for other operational weather prediction centers as illustrated in the NCEP web site of Suranjana Saha, [Available at http://wwwt.emc.ncep.noaa.gov/gmb/ssaha/]. Other centers show similar cold biases around 250 hPa compared to radiosondes with the NCEP radiation correction, with warm biases around 150 hPa. Note that radiation correction can only explain a part of this cold bias.

Further investigation showed a larger cold bias at 250 hPa for the Chinese radiosondes (not shown) that was made larger by the radiation correction. Fig. 7 presents a 12 month period of monthly averages of temperature differences at 250 hPa between 00Z and 12Z for Chinese radiosondes. Thus the diurnal differences in the model guess with those of the radiosonde temperatures with and without the radiation correction are compared. Note that the raw data "RAWD" shows diurnally averaged temperature differences similar to the guess "GESD". The radiation correction makes the corrected temperature's diurnal differences bigger than the guess as shown with "RADD". NCEP's QC team suspected that the Chinese radiosondes were already corrected at the site which was later confirmed by the Chinese NMC. They implemented radiation corrections in January 2001 (Y. Zhang 2005, personal communication). Consequently, the NCEP radiation correction for Chinese radiosondes was turned off as of August 2005, except above 50 hPa.

NCEP also noticed large problems with the radiation correction for the US RS80 Vaisalla radiosondes in the stratosphere as reported by Redder et al. (2003). Fortunately this problem is relatively small around 250 hPa where we are most interested. However, there is a possibility of these stratospheric errors impacting the analysis around 250 hPa.

4. AIRCRAFT TEMPERATURE BIAS ANALYSIS

As the WMO lead center for aircraft data, NCEP maintains a web site [Available at http://www.nco.ncep.noaa.gov/pmb/qap/] with monthly reports on ACARS and AMDAR data for current as well as the past 12 months. These reports show some aircraft with very warm biases to the guess. As pointed out by Moninger et al. (2003), these automated aircraft reports can have abnormally warm temperatures due to debris in their temperature sensing tubes. These aircraft units with very warm biases make the overall bias warmer, but it was suspected that there may be additional factors for the warm bias. As the lead center, NCEP has access to the encryption algorithm used by Air Radio Inc. (ARINC) to keep the public from knowing what airline or actual aircraft the encrypted IDs correspond to. By knowing the real tail numbers for the ACARS units, this study utilized data from the web site http://www.landings.com/ to identify the type of aircraft that each unit belonged to for over 98% of the whole ACARS fleet. The reliability of the "landings.com" website seems accurate when compared with the earlier incomplete information that NCEP received from ARINC. Identification of 35 different types of aircraft was possible such as Boeing 767-322 and McDonnell Douglas units MD-11 etc. The web site did not provide any information on temperature sensors of various aircraft types that may be important. NCEP received some information from United Airlines (J.

McQuay 2005, personal communication) showing which tail numbers of United had one of the three different temperature sensors used by United ACARS units. This did not seem to be a significant factor, but may be significant for some other aircrafts.

There are other factors that may explain temperature biases for aircraft (J. Stickland 2005, personal communication). These include the temperature probe's design and exposure, how the correction is calculated for the large dynamic heating and computer processing of the data. In studying the major ACARS aircraft types using the above web site, it was found that each type is almost always used by just one airline and has a limited time range for registration dates. It could be then that each specific type of aircraft has constant factors that affect the temperature measurements.

Detailed statistics available at NCEP on ACARS temperatures compared to the NCEP guess were analyzed for dependence on a number of factors including: the airline, the type of aircraft, pressure, the aircraft Phase of Flight (POF) and the time of day. Our analysis found some dependence of the temperature bias on the POF as reported in the collocation studies by Schwartz and Benjamin (1995) and Mamrosh et al. (2002). Fig. 8 shows vertical structure of the temperature bias for different POF, interpolated to the nearest mandatory pressure level for all of January 2005. The POF are shown as "DSNT" for descent, "ASNT" for ascent, "LEVL" for level and "MISS" for the POF being missing. The data counts related to this plot are shown in Fig. 9. Note that at 700 hPa and below, the counts for level POF are so low that the biases are not reliable. In Fig.8 notice that below roughly 400 hPa the descent reports appear warmer than other types. At 600 hPa and above, the ascent reports appear warmer compared to DSNT, MISS and LEVL. Note that for 300 hPa and up, the counts for missing POF are largest. Since it will later be shown that specific aircraft types are important factors in temperature bias and since some types of aircraft report only select POF, the above bias study was repeated using a select group of aircraft. This group was based on aircraft types that reported descent, ascent and level phases of flight which included types 737-522, 757-24APF, 767-322, 767-34AF, A300F4, A310-203, MD-10 and MD-11F comprising about 335 different aircraft units. Fig. 10 shows the vertical dependence of the bias for different POF for this select group. From roughly 700 hPa to 925 hPa, the ascent and descent POF show similar biases. Around 1000 hPa, the LEVL reports are warmer. From about 500 to 300 hPa, the ascent reports are clearly warmer. This indicates that the POF is an important factor in temperature bias even amongst units that report all POF. Fig. 11 shows the data counts corresponding to Fig. 10. The counts show a maximum around 250 hPa for level POF for the select types displayed. Studying the temperature biases for different POF for all types of aircraft is beyond the scope of this report, but may be needed for some applications. This study used the POF reported by the aircraft and did not attempt to deduce the POF based on the time history of altitudes.

Next investigation focused on how temperature biases varied with specific aircraft types. It was found that the bias did not vary much based on major aircraft types like 757, 767, Airbus etc. However, the bias varies considerably based on specific types as shown in Fig 12. This plot shows temperature biases for select types for a 12 month period for 300 hPa and above. This selection was made to show a significant spread in bias for types with sizable counts that are shown in Fig. 13. Fig. 14 is the same as Fig. 12, except it was for types with the largest counts not shown in Fig. 12. Fig. 15 is the same as Fig. 13, except the counts are for the types in Fig. 14. Note that the biases tend to vary smoothly from month to month not exhibiting obvious seasonal variability. Thus using these biases in some sort of bias correction in a GDAS package may be a very worthwhile strategy to pursue. Since Fig. 9 shows that only a small percentage of reports above 300 hPa report as ascent or descent, the biases shown in Figures 12 and 14 should not require any modification for the POF. Also note that since some units can have very warm biases that are not characteristic of the group, the biases in Fig. 12 and 14 did not include any units that were beyond three standard deviations away from the mean biases of the group. Fig. 16 shows how the biases of aircraft types in Fig. 12 vary with pressure for 00Z in January 2005. The biases do vary with pressure and may show different characteristics for different POF, times of day and year.

CONCLUSIONS

The large consistent and contrasting temperature bias around 250 hPa between radiosonde (cold) and aircraft (warm) temperatures and the NCEP guess can be partially explained by the cold bias (radiosonde) being aggravated by the NCEP radiation correction and warm bias (aircraft) by a combination of factors that are dependent on specific aircraft types. Since the ACARS data show a large difference in temperature biases based on specific aircraft types, it is difficult to assume that the aircraft temperatures represent the truth. It would be very valuable for all aircraft manufactures to explain why the biases vary so much with specific aircraft types. Since it will be a long time before changes are made to the aircraft temperature measurements, it may be useful to apply some type of bias correction to these data prior to executing the NWP analysis system.

If these differences in aircraft biases can be explained, it would be beneficial in the future to design new aircraft with automated temperature sensors to be more consistent and closer to the true temperature. It would be also helpful for all automated aircraft reports to report the POF as that is a factor in temperature bias.

For future work, we expect to expand this work to include a study of AMDAR temperature biases as a function of aircraft type and the phase of flight. This would also be useful for AIREPS but that would be challenging since AIREPS do not report aircraft tail numbers and in some cases may involve different aircraft types with similar flight IDS. We plan to carry out detailed analysis and forecast experiments with the Global Forecast System (GFS) model to elucidate the impacts of bias corrected temperatures of radiosonde and various aircraft observations. This study only focused on the radiosonde temperature bias versus aircraft temperature bias but more work needs to be done using winds from both platforms.

REFERENCES

Collins, W., 1999: Determination of new adjustment tables in order to bring radiosonde temperature and height measurements from different sonde types into relative agreement [Available online at http://www.emc.ncep.noaa.gov/mmb/papers/collins/new_tables/new_tables.html]

Mamrosh, R., R. Baker and T. Jirikowic, 2002: A comparison of ACARS WVSS and NWS radiosonde temperature and moisture Data, Proceedings of the Sixth Symposium on Integrated Observing Systems. AMS Symp. Integr. Obs. Syst. Vol. 6.

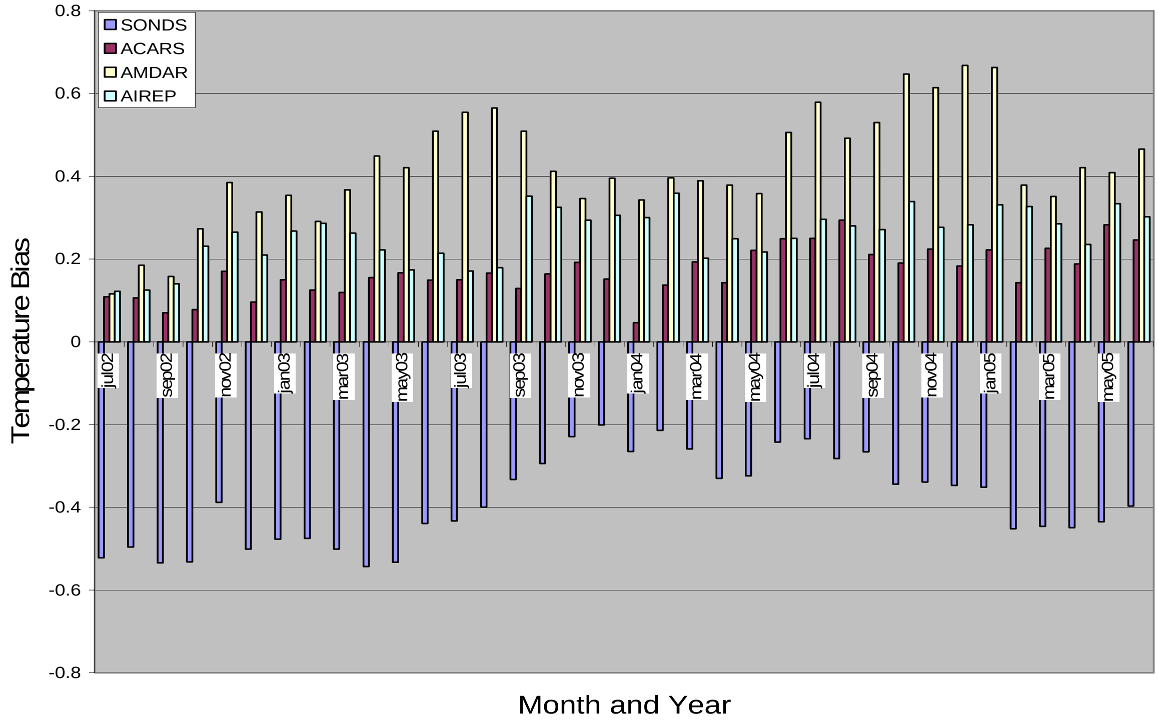
Parrish, D. and J. Derber, 1992: The National Meteorological Center's spectral statistical-interpolation analysis system. Mon. Wea. Rev., **120**, 1747-1763.

Reale, A.L., M. E. Petty and C. E. Brown, 2005: Operational polar satellite validation and evaluation systems. Technical Proceedings of the 21st International Conference on Interactive Information Processing Systems (IIPS) for Meteorology, Oceanography and Hydrology, 85th AMS Annual Meeting, San Diego, Ca., 9-13 Jan 2005.

Redder, C., J. Luers, and R. Eskridge, 2003: Unexplained discontinuity in the U. S. radiosonde temperature data. Part II: Stratosphere. J. Atmos. and Oceanic Techno., **21**, 1133-1144.

Schwartz, B. and S.G. Benjamin, 1995: A comparison of temperature and wind measurements from ACARS-Equipped aircraft and rawinsondes. Wea. Forecasting, **10**, 528-544.

Zapotocyny, T.H. S.J. Nieman, W.P. Menzel, J.P. Nelson III, J.A.Jung, E. Rogers, D.F. Parrish, G.J. DiMego, M. Baldwin and T.J. Schmit, 2000: A case study of the sensitivity of the ETA data assimilation system. Wea. Forecasting, **15**, 603-621.



Fi.g 1. Temperature Biases 300 to 200 hPa 00Z

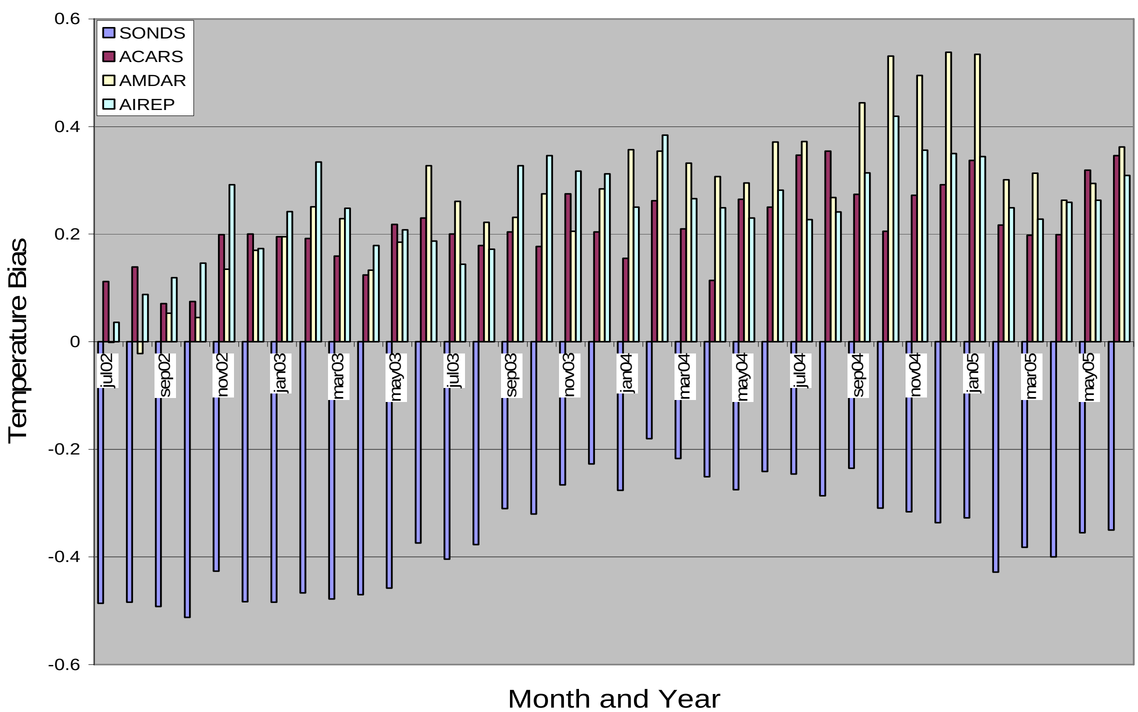


Fig 2. Temperature Biases 300 to 200 hPa 12Z

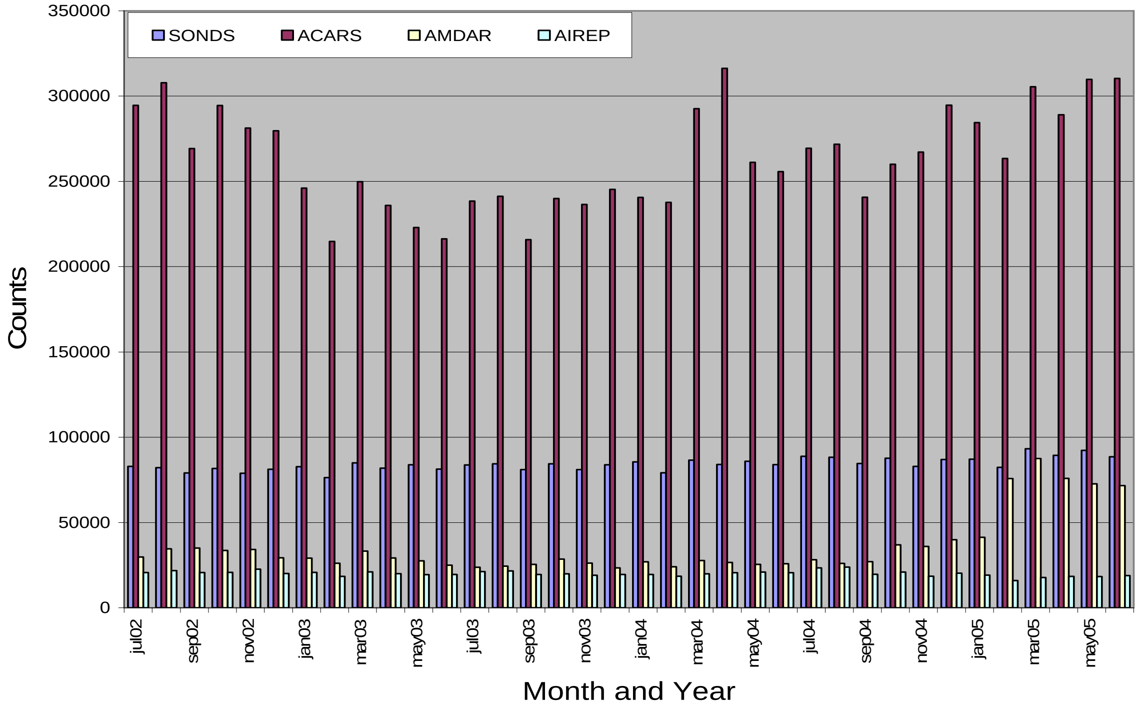


Fig 3. Temperature Counts 300 to 200 hPa 00Z

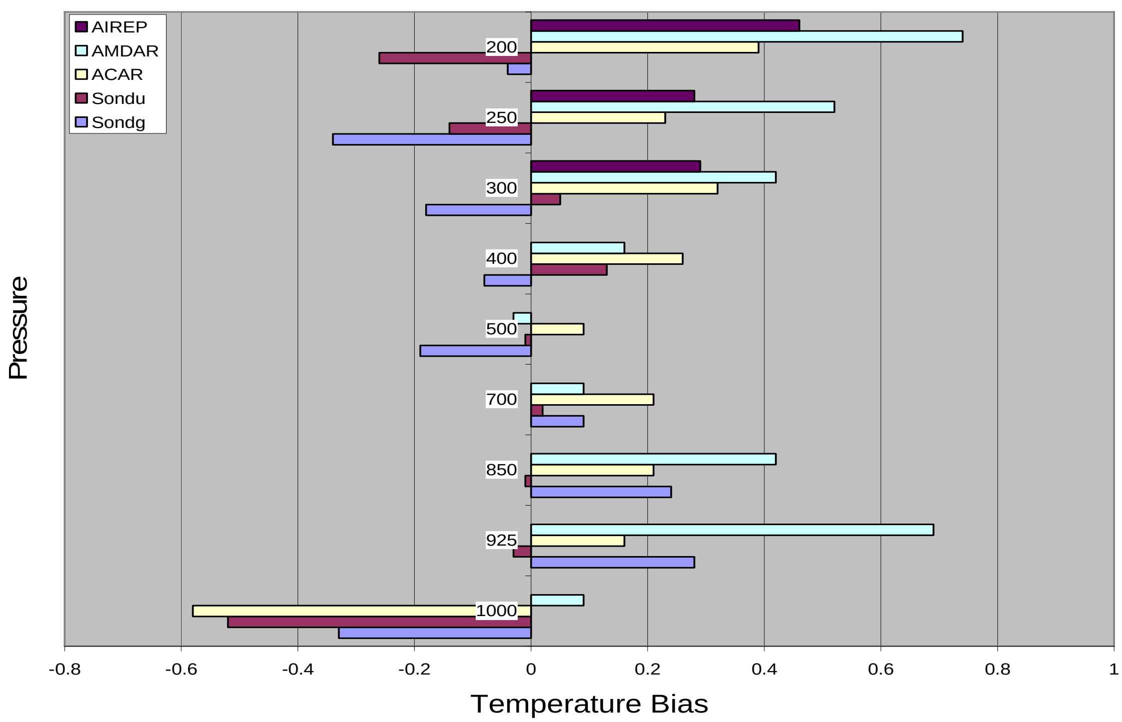


Fig 4. Temperature Biases 00Z July 2004

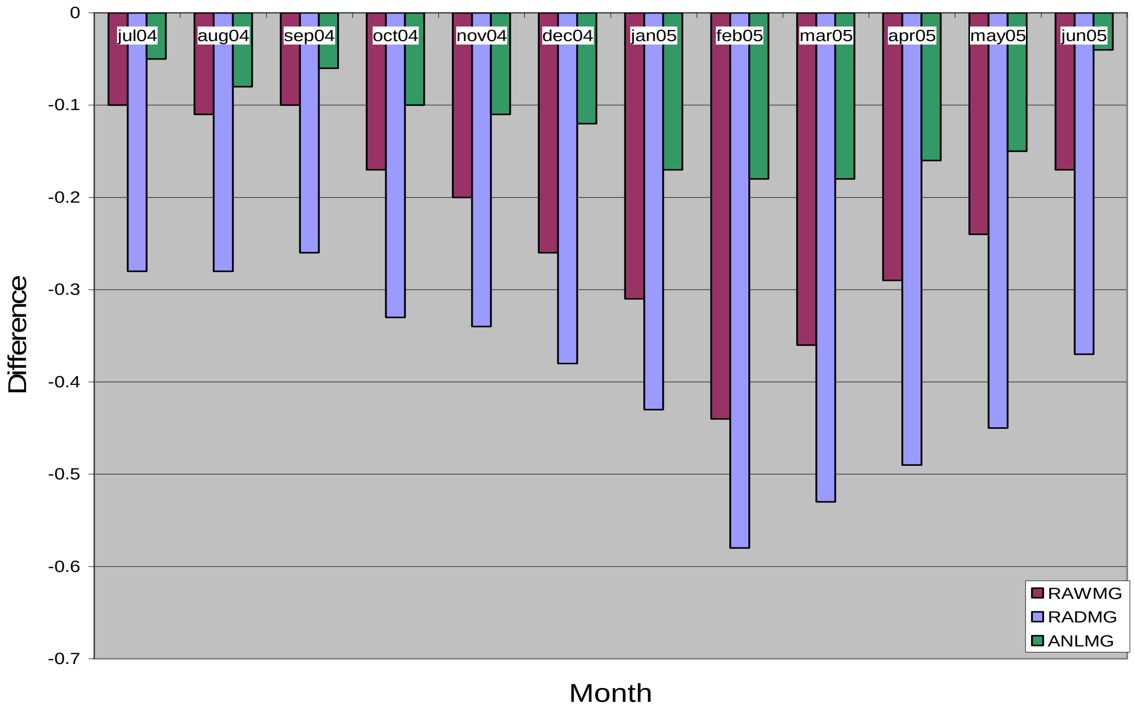


Fig 5. Monthly Average Temperature Differences Versus Guess 00Z 250 hPa All Sondes

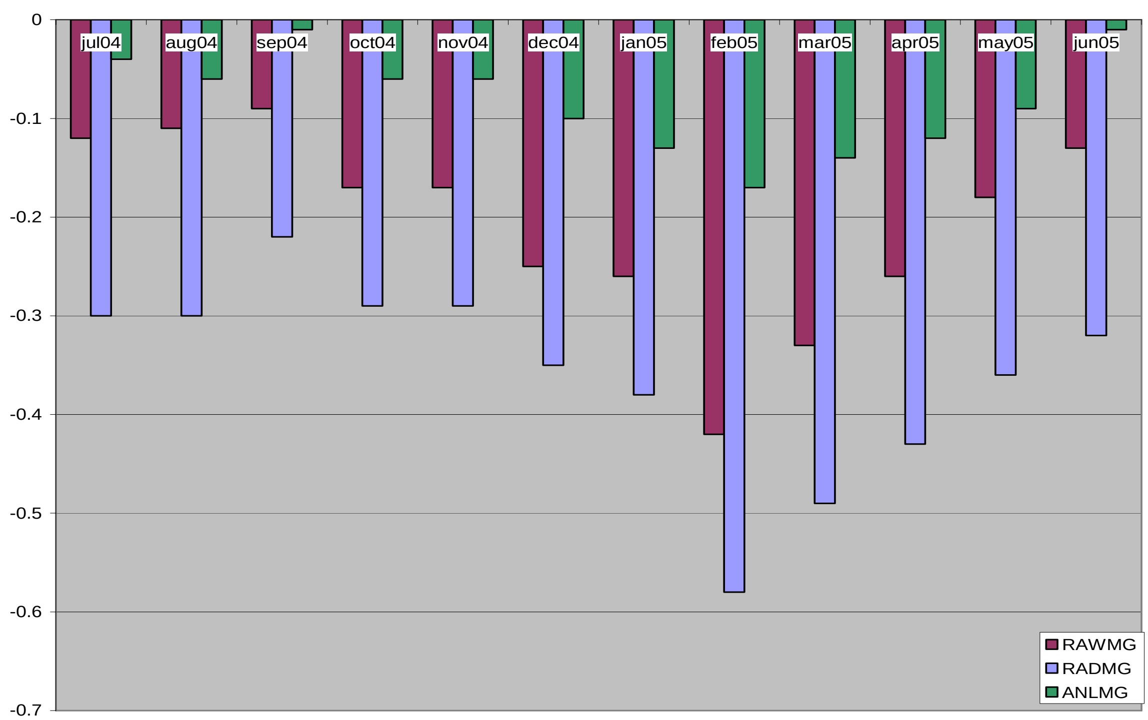


Fig 6. Monthly Average Temperature Differences Versus Guess 12Z 250 hPa All Sondes

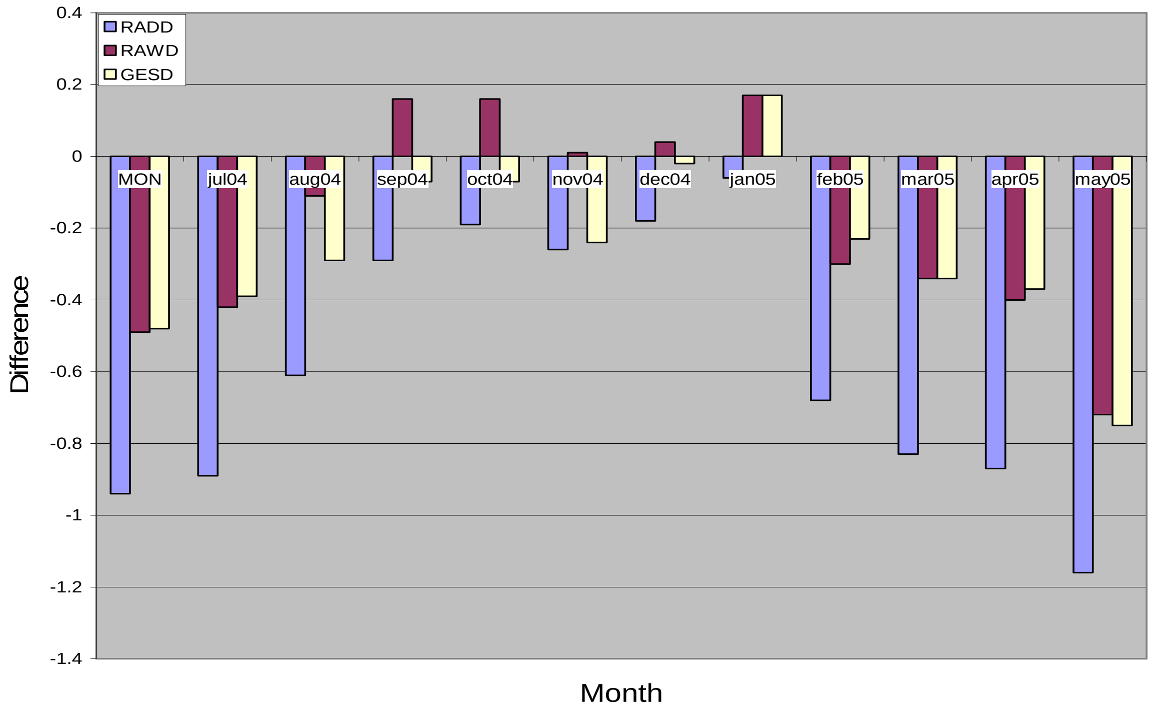


Fig 7. Monthly Average Temperature Differences 00Z – 12Z Chinese Sondes

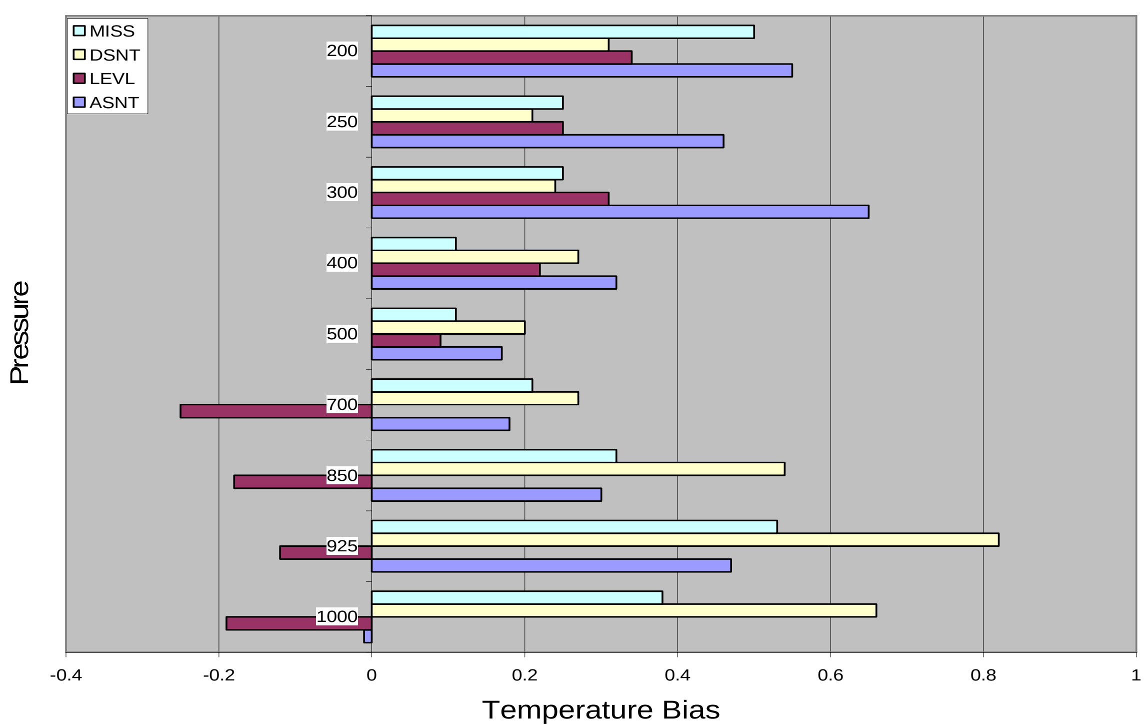


Fig 8. ACAR Temperature Biases by POF January 2005 All Types

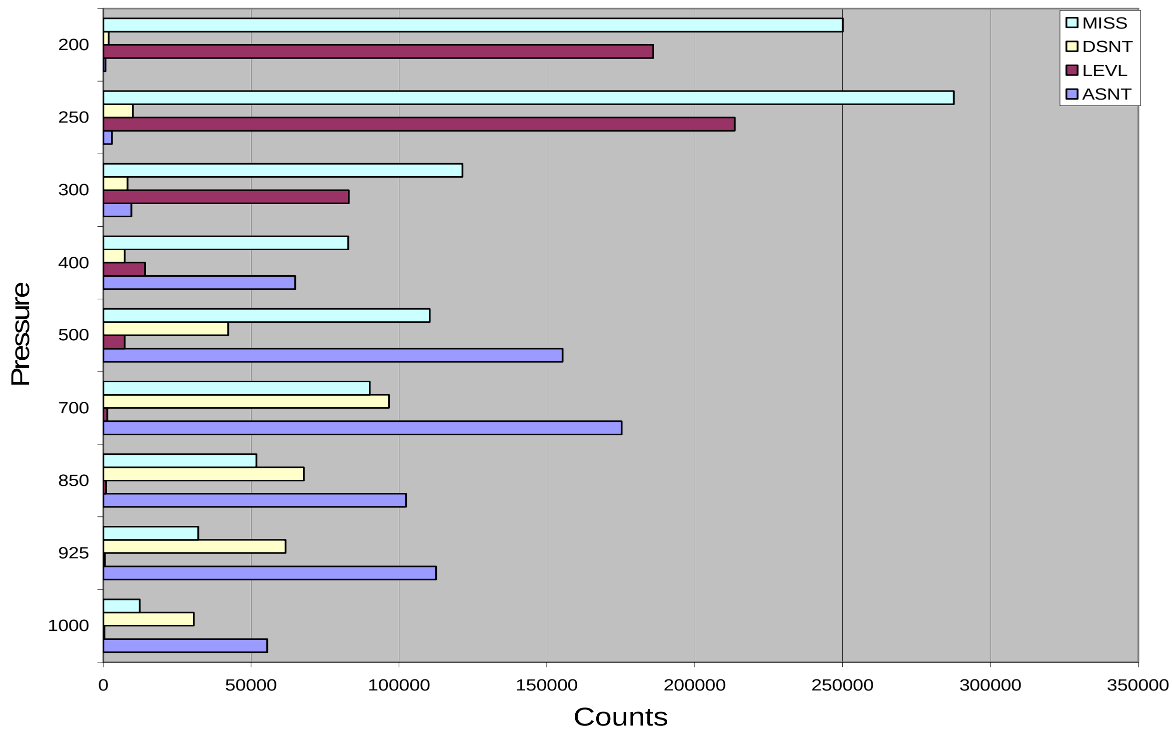


Fig 9. ACAR Temperature Counts by POF January 2005 All Types

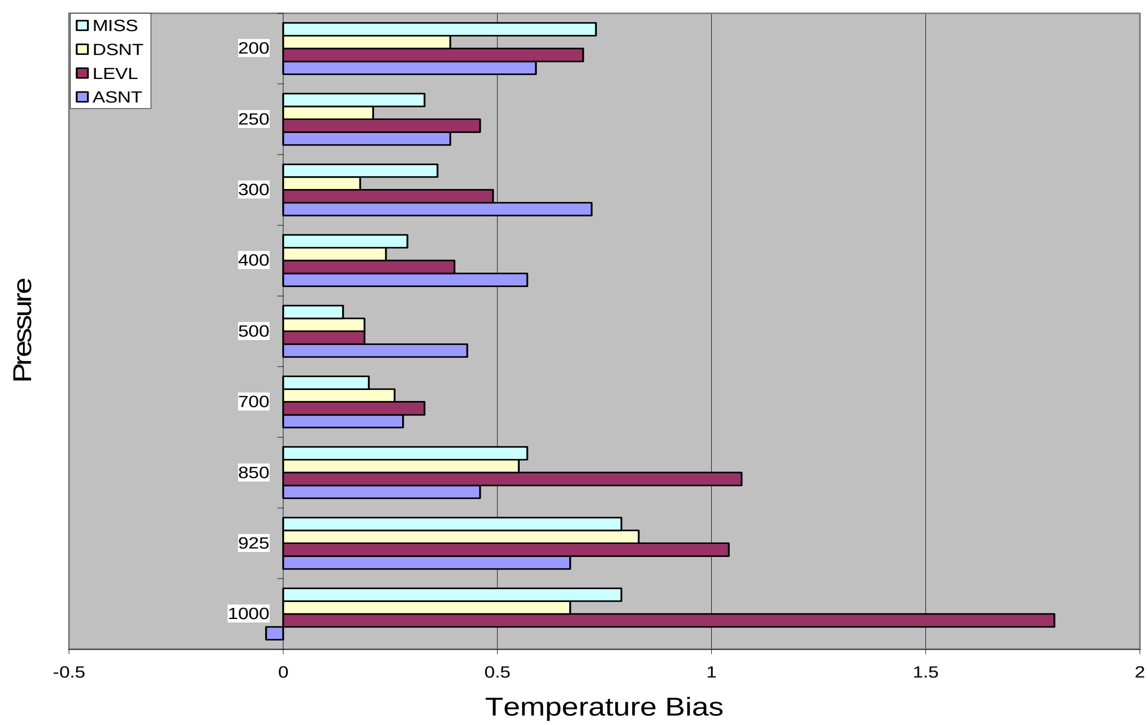


Fig 10. ACAR Temperature Biases by POF January 2005 Select Types

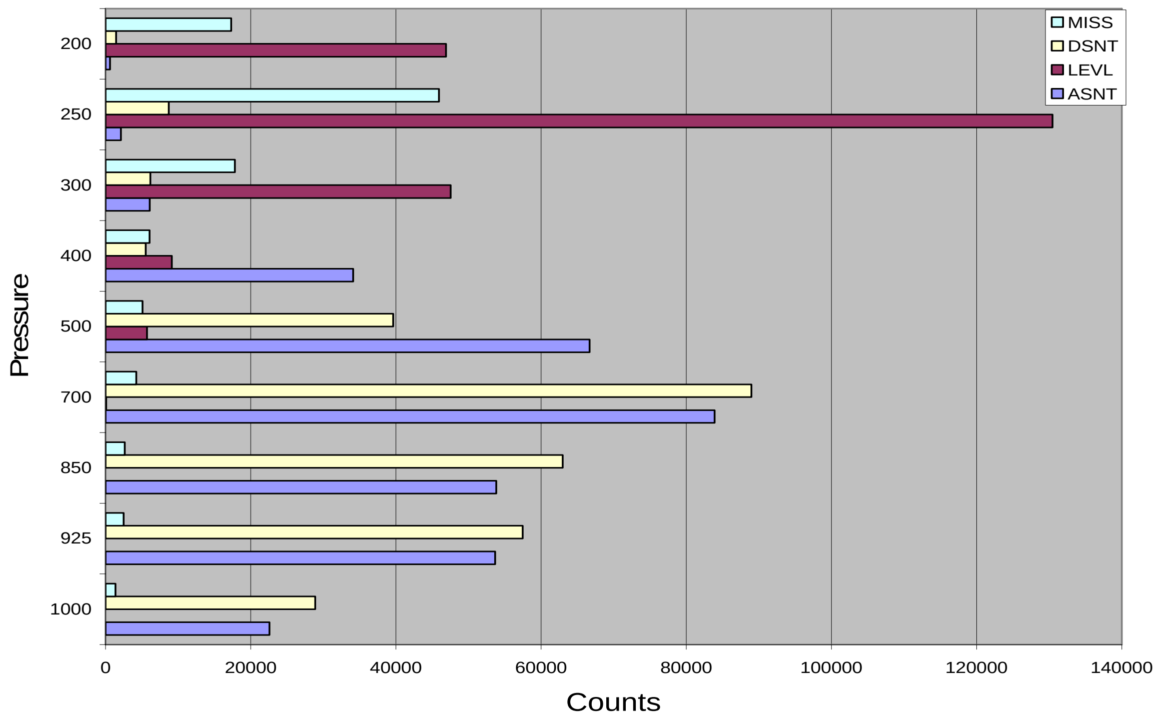


Fig 11. ACAR Temperature Counts by POF January 2005 Select Types

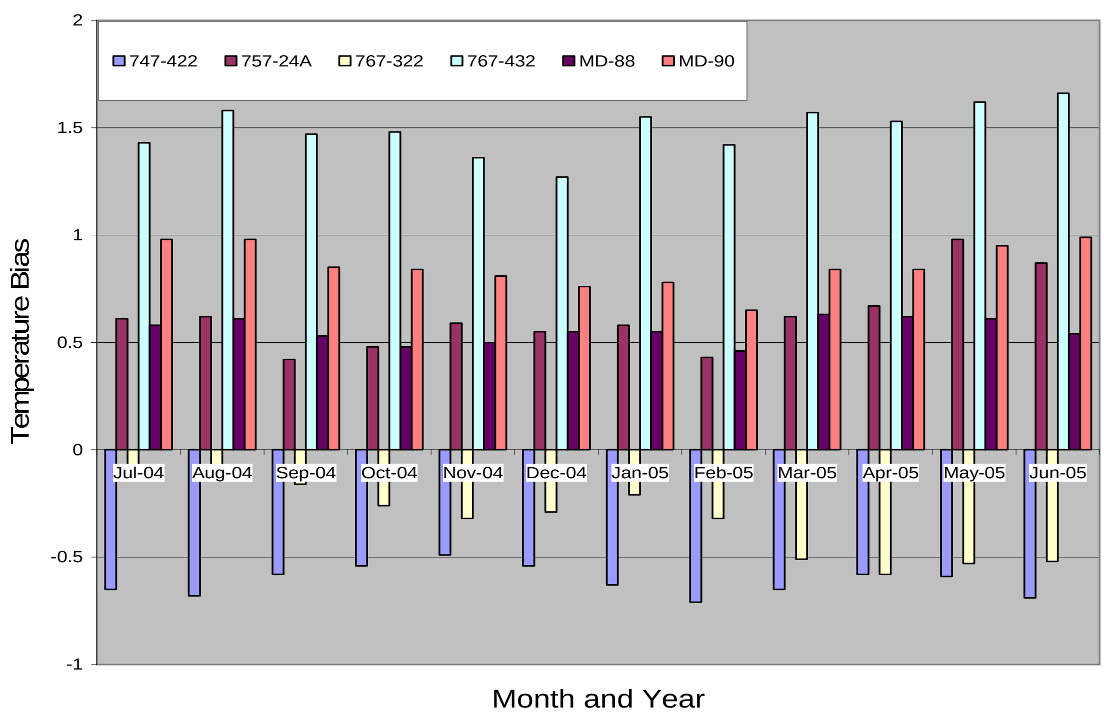


Fig 12. ACAR Temperature Biases by Aircraft Type 300 hPa and up

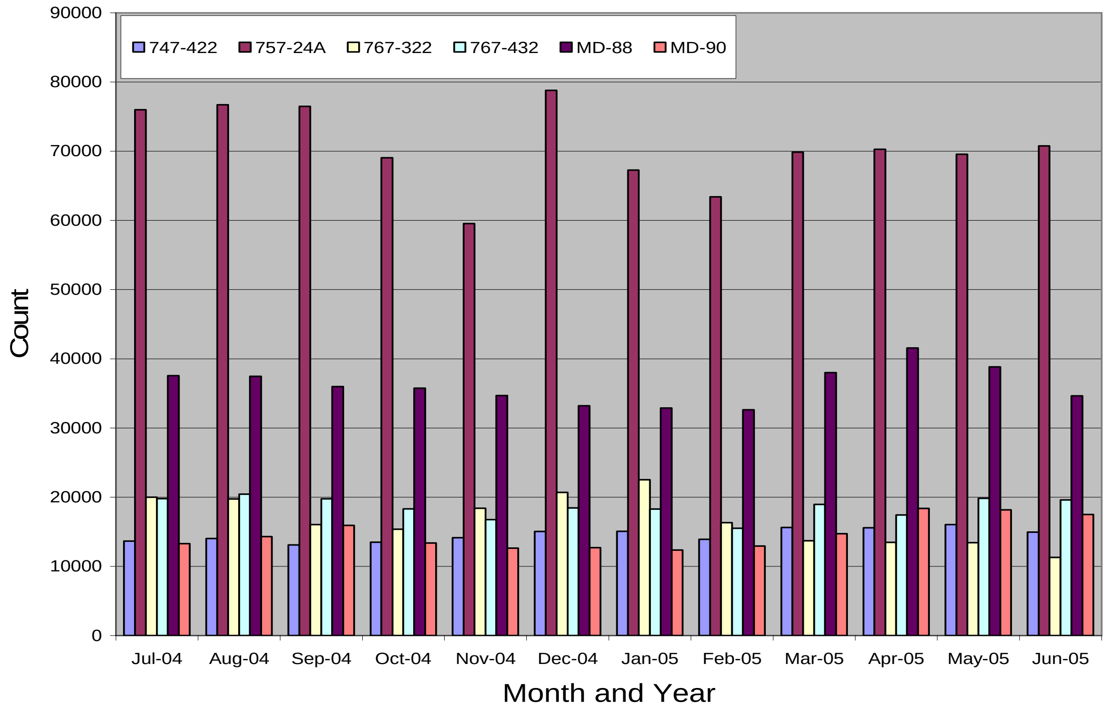


Fig 13. ACAR Temperature Counts by Aircraft Type 300 hPa up

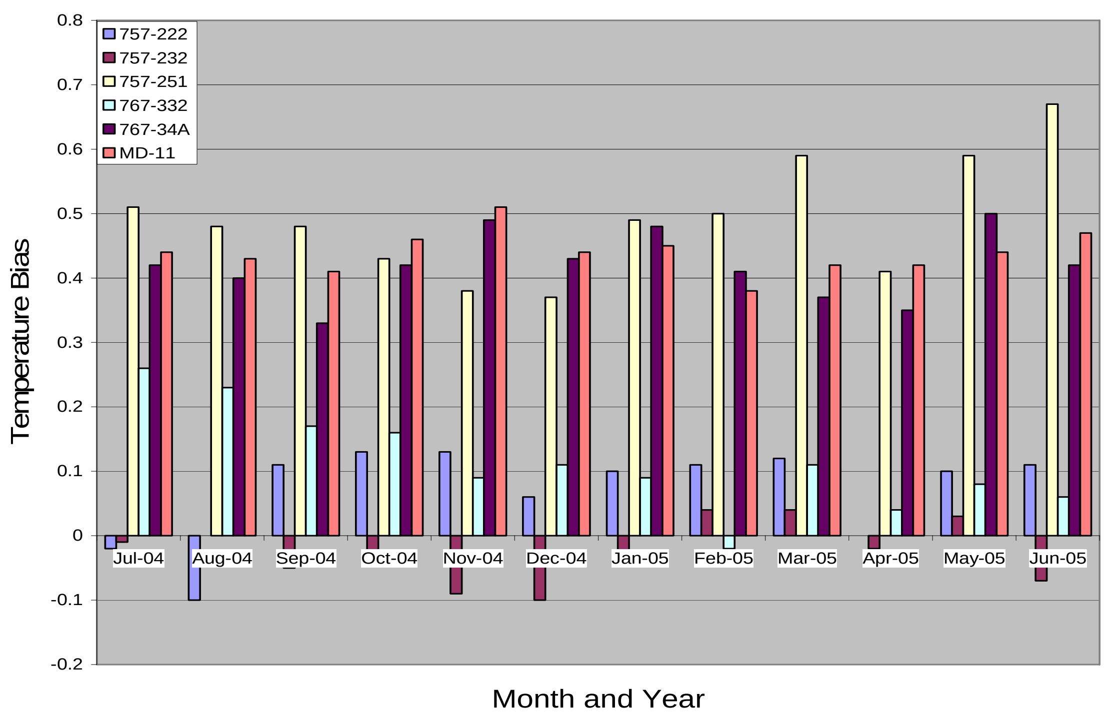


Fig 14. ACAR Temperature Biasess by Aircraft Type 300 hPa up

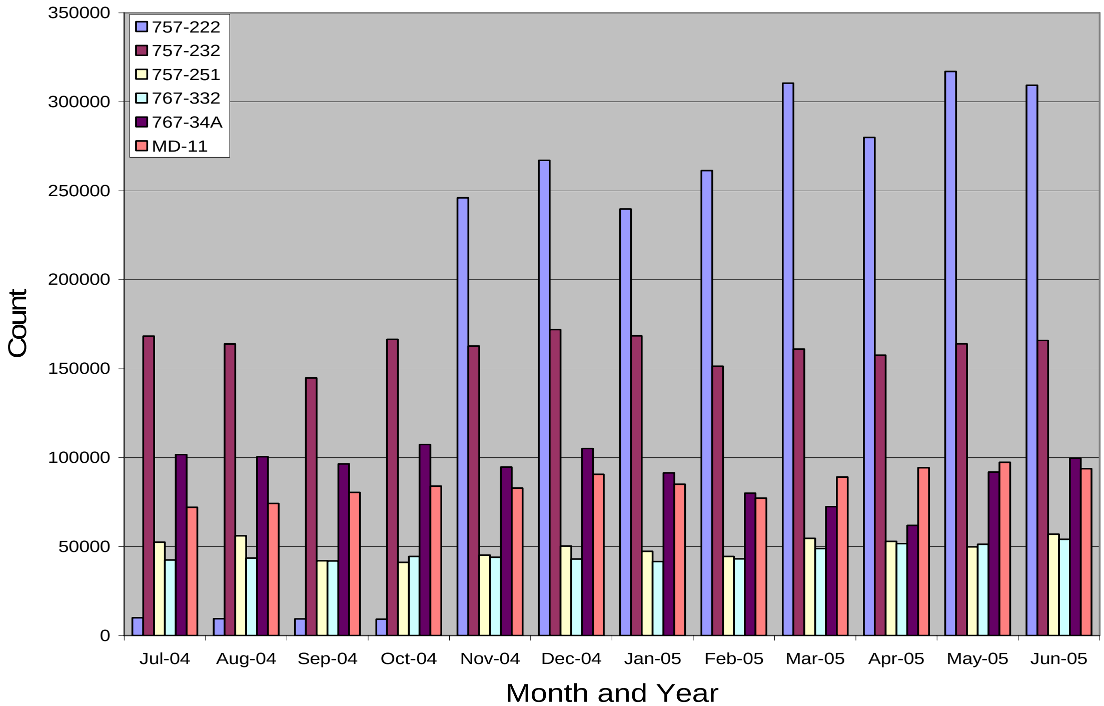


Fig 15. ACAR Temperature Counts by Aircraft Type 300 hPa up

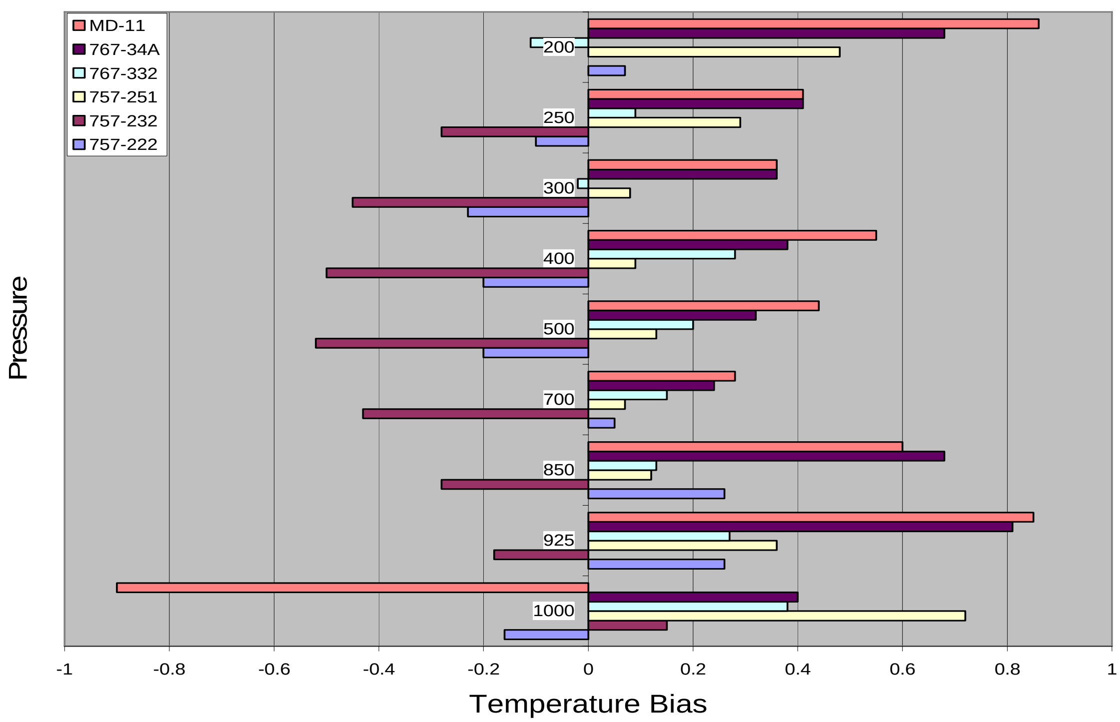


Fig 16. ACAR Temperature Biases 00Z January 2005