COMET PARTNERS PROJECT **FINAL REPORT**: S09-81065

PI: Dr. Steven M. Lazarus, Florida Institute of Technology

Project Personnel: Steve Levine (FIT), Dr. Pablo Santos (NWS-Miami), Dave Sharp (NWS-Melbourne)

Project Overview/Highlights:

The primary goal of this project has been met with the creation of both updated static and a directionally dependent reject/accept lists for the RTMA wind QC/QA (Florida region only). The project has produced a Master's thesis for Mr. Steve Levine who presented his work at NCEP in April this year. Mr. Levine is scheduled to defend his thesis in the Fall and is currently being courted by NCEP for a position following his graduation.

SECTION 1: Objectives and Accomplishments

Methodology/FIT

As previously reported (6 month report) the *Florida region* was somewhat arbitrarily divided into 5 mesoscale areas based on latitude and longitude. The regions were defined to allow for a smaller area over which to compare observations and networks in an effort to minimize the impact of regional weather conditions on the QC/QA statistics across the state. A full year of observations including downscaled RUC forecasts at the observation locations (i.e., first guess), station and network metadata, and RTMA quality control flags from the NCEP GSI were all inserted into the FIT mySQL database. From these data, various statistical tables have been created and, as a result, recommendations for changes to the present RTMA accept lists have been made to NCEP. The statistical tables generated via the database include:

1. Daily statistics for each observation-day for each station, including average and standard deviation of wind speed and direction, RMSE of wind speed, RMSE of wind direction, bias of wind speed, correlation coefficient of wind speed and a flag to identify whether observations which were part of the sample should be assimilated by the RTMA or not.
2. Wind direction statistics. This table contains Station ID, directional bin information, and the same statistics as described in the table of daily statistics above. The bin width is 45^{o}.
3. A second companion table (to #2 above) contains the highest and average absolute and

normalized RMSE of wind speed for ASOS stations in each region, for each wind direction bin.

4. Two additional directionally dependent tables were also generated using hourly statistics.

Because so few mesonet sites were included in the original use list, it was decided that in order for any new buddy check to work, the list of acceptable stations had to be expanded. This was accomplished by comparing daily statistics of mesonet sites with ASOS sites in the same region. The assumption here is that the ASOS stations are the 'gold standard' of surface data. There are enough ASOS sites in each region so they can be used as a reference. For each day, the mean and maximum daily RMSE (both absolute and normalized) from each network and region (including ASOS sites) were calculated (bullet 1 above). A station that had a lower daily RMSE than either the mean or maximum of the ASOS stations in the same region is re-flagged. Separate flags were considered for normalized and absolute RMSE. When a station's daily flag value was changed, all observations occurring on that day were changed to the new value. Stations that were flagged as acceptable for at least 50% of the days and which 2 or more observations were available, were then passed on to a second (hypothesis) test to fine-tune the static accept list. Two Z-tests were conducted to ensure that the difference between the daily mean observed wind speed of the mesonet station in question and the daily average observed wind speed of the nearest METAR site had an average difference of less than 1.0 ms^{-1}. Only days with all 24 hourly observations for both the mesonet and the nearby METAR site were used. Fifty days from this set were then randomly selected for each station to use as the sample.

A systematic low bias with respect to wind speed has been detected among most mesonet sites (Benjamin et al. 2007a; M. Pondeca personal communication). As a result, the RTMA now rejects a significant number of mesonet wind observations with only a few 'well-sited' stations on the national list presently included in the RTMA (see Table 1). While this ensures that biased mesonet data is not assimilated, it is highly likely that at least some quality data is also being excluded. Figure 1 indicates that the bias issue exists for our region (Florida) as well. Note the bias histogram is centered near -1.0

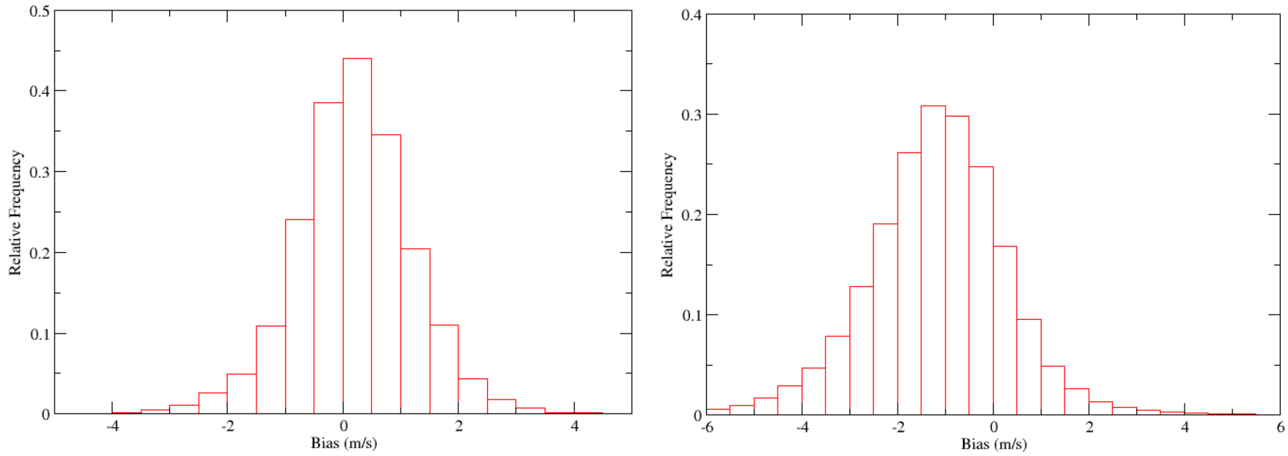

Fig. 1. Relative frequency histogram of wind speed bias (observed minus background, ms^{-1}) for a.) ASOS sites only and, b.) all observations pre-quality control.

ms^{-1} when 'all' observations (Fig. 1b) are included versus the ASOS stations only where the bias is minimal (Fig. 1a). A complete breakdown of bias by individual network is given in the Appendix.

Based on the bias shown in the histogram in Figure 1b, we selected the null hypothesis that the mesonet station in question would have an average bias (observed minus background) of at least 1.0 ms^{-1} lower than the closest METAR site (this necessitates running two Z-tests with differences of +/- 1.0). Using a 90% confidence level, the tests are then repeated with another randomly selected set of 50 days. Stations for which appropriate Z-values were returned on both tests were considered to have passed the test. Stations that pass both the RMSE test and Z-test were included on a permanent/static accept list. Stations that pass the RMSE test but not the Z-test <u>were not included</u> on the new accept list – but they are however further examined using a directionally dependent consistency test described next.

Two additional quality control tests were developed whereby the wind observations are stratified by placing them into eight 45-degree directional bins (the background wind direction is not considered here). The bins were sorted by station with bulk statistics – the same statistics described in the previous section – were calculated for each bin and placed in a new mySQL table (see methodology above).

A flagging procedure was developed similar to the one used in the previous test. The wind speed innovations and RMSEs were computed for each directional bin and station. This test was especially relevant for stations which only marginally passed the RMSE test and had a relatively significant number of observation days flagged as acceptable, but not enough to be put on the new accept list. As in the previously described statistical tests, mesonet sites were compared to ASOS sites in the same region. Those which had an directional bin RMSE lower than that of the nearby ASOS stations were flagged as acceptable (there are actually two separate flags defined here, one for normalized RMSE and one for absolute RMSE). Stations that were not on the original accept list, but whose directional RMSE for that day is lower than mean RMSE for accepted ASOS stations in that region for that day and bin were then examined individually and subjected to a bin-dependent Z-test. For this Z-test, 100 observations are randomly selected from each directional bin and compared with that of the nearest METAR station for the same time and day. Although we are sorting using wind direction, the statistical tests are applied to wind speed. Here, we attempt to remove times for which the winds are light and variable prior to applying the Z-test. To do this, a directional consistency test is first performed to determine whether the observed wind direction at the METAR site was within +/-22.5^{o} of the direction of the observed wind at the mesonet site. If not, the particular observation was not included in the Z-test. This criterion generally resulted in the exclusion of between 30 and 40 of the 100 randomly selected observations. Since a sample size of 30 is widely considered acceptable for purposes of such a Z-test, the test is assumed to be fairly accurate.

The Z-tests are performed on the subset of observations for which the observed direction at the METAR and mesonet sites are within 22.5^{o} of each other. As previously described, the null hypothesis for the first test is that the average difference between the observed wind speed at the METAR site and mesonet site is at least 1.0 ms^{-1} while for the second test, the difference threshold selected are values less than or equal to -1.0 ms^{-1}. *Stations that pass both Z-tests were placed on a list to be used only when the observed wind direction is within range of the bin for which the station passed the test.* All observations which met these criteria were re-flagged to be included in any retrospective or future analyses; different flags were used to identify why a certain observation was flagged as usable.

Results/FIT

A. Permanent Accept List

All stations with available observations (647) were subjected to the initial RMSE static test described previously. 154 stations passed either the NRMSE or absolute RMSE test (152 passed the absolute RMSE test and 51 passed the normalized RMSE test, 49 passed both). All of these stations were then passed on to the more stringent Z-test.

Only 33 of the 154 stations that were subjected to the Z-test returned favorable scores (i.e., returned a bias of less than 1.0 ms^{-1}). Most of these stations contained at least some bias (see Appendix A). The mesonet sites consistently recorded a lower wind speed than the nearest METAR station. *Interestingly, all of the mesonet sites within 1.5 km of a METAR site failed the Z-test*. Of the 33 stations that did pass this QC, 8 were more than 30 km from the nearest METAR site. The stations that passed the Z-test, and were subsequently placed on the permanent accept list, are given in Table 2.

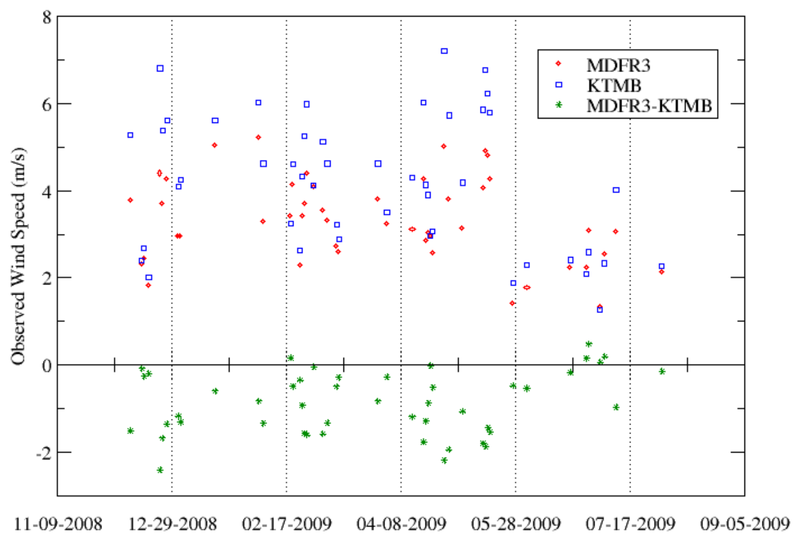

Fig. 2. Time series of daily average wind speeds (ms^{-1}) and the difference (mesonet – METAR, filled green circles) between two proximity weather stations: KTMB (open blue squares), a METAR site at the airport and MDFR3 (filled red circles), a weather station set up by the local fire department at the airport fire station.

Fig. 3. Kendall-Tamiami Executive Airport, showing location of METAR site KTMB (white circle near center of airport) and AWS site MDFR4 (red circle on east side of airport).

An interesting example of two nearby stations that record different values is shown in Figure 2. The figure shows the daily average wind speeds at two stations located at Kendall-Tamiami Executive Airport in Miami. KTMB is the airport ASOS site, while a nearby AWS site, MDFR3 is located at the airport fire station; the siting is shown in Figure 3. Figure 2 shows a small (on the order of 1.0 ms^{-1}) but fairly consistent bias in daily average wind speed between the two stations. Similar biases were found in other stations that are near (< 5.0 km) ASOS sites and are fairly consistent. The aerial photograph in Figure 3 indicates that the ASOS sites are generally located in open areas near airport runways whereas the nearby mesonet sites are located in less than ideal wind monitoring locations (i.e., sheltered/obstructed by buildings and trees).

B. Directionally Dependent Accept List

Of the 782 directional bins tested, 151 (19.3%) bins passed the Z-test, and another 46 (5.9%) nearly passed. Many of those that ‘passed’ resulted from the same station which produced successful Z-tests for multiple directional bins. The station-pairs and number of directional bins in which they passed are shown in Table 3. *A number of the stations that failed the daily Z-test subsequently passed the* *directional Z-test for at least one bin.* As an example, one of these stations, UMLF1 (from the Florida Mesonet) is examined in more detail here (additional stations are discussed in more detail in Mr. Levine's thesis). The station itself appears to be well-sited – in an open field. However a small bias is evident when the daily average wind speeds of the two stations are compared (Fig. 5). We show the breakdown by directional bin in Appendix B (we are in the process of running Z-tests by direction bin). However, it is clear that the bias varies depending on wind direction. Here, as the wind direction changes from easterly to southerly, the bias apparently increases (Fig. B.1). This appears to be consistent with Fig. 4 which clearly suggests that the land surface (vegetation) differ between the two – especially for southeast to northwest fetch.

Figure 4. Location of Florida Mesonet station (UMLF1, green circle) and a nearby METAR site (KLEE, white circle). The two stations are about 21 km apart. Also shown is the mesonet station heading with respect to the METAR site.

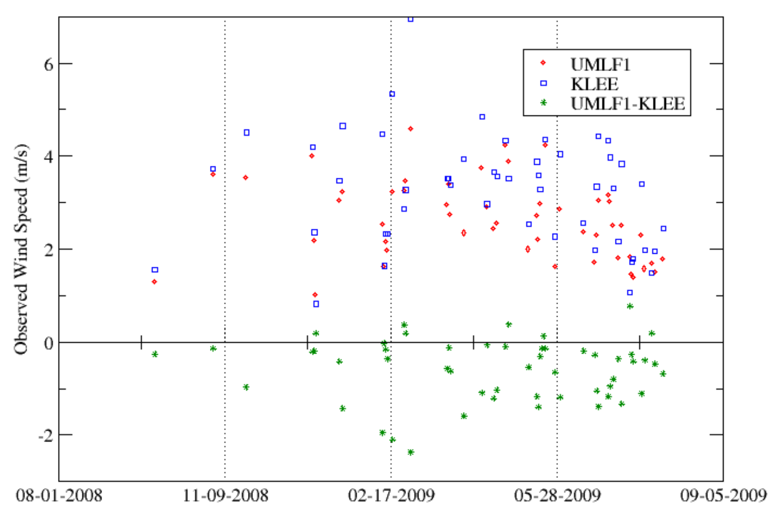

Fig. 5. As in Figure 2 but for stations UMLF1 and KLEE.

SECTION 2: Related Accomplishments

FIT

The PI and graduate student Steve Levin visited NCEP in April – Mr. Levine gave a presentation on his work and, as a result, has been arranging for a position post-graduation.

NWS/NCEP

The updated static accept lists for the region studied here have been passed along to NCEP for operational use in the RTMA.

SECTION 3: Summary of Benefits and Current Work

FIT

The text/content for this report was lifted from the work of Mr. Steve Levine, a graduate student of the PI that has been working on the project since its inception (Mr. Levine will be graduating this fall). The work performed here (i.e., the Florida region) can be extended to cover the CONUS RTMA and perhaps extended to Alaska, Puerto Rico and Guam – where regional versions of the RTMA are running. In bi-weekly teleconferencing with NCEP (which are part of the regular RTMA committee duties of the PI), NCEP has on several occasions asked when the work can be extended to cover the entire RTMA. We believe that we have successfully demonstrated the utility of some mesonet wind observations that heretofore have not been included in the RTMA. In addition, the work marks the beginning of a more extensive effort that is tied to improving the metadata associated with the various networks as part of the National Mesonet effort (C. Marshall, personal communication). In particular, the work funded here begins to address siting issues – especially as they relate to surface roughness.

NWS/NCEP

Updated reject/accept lists (for regional winds only) based on individual station statistics as well as the examination of network bias are advancements that will most likely improve the quality of the RTMA.

SECTION 4: Presentations and Publications

The work presented herein is a portion of the graduate thesis project of Mr. Steven Levine. As previously mentioned, Mr. Levine gave a presentation at NCEP in April and will be presenting at the AMS Annual meeting in January.

SECTION 5: Summary of Problems Encountered and Issues/Questions raised

FIT

The project has been limited to the Florida region (as proposed) and to a year of data given the scope (> 5 million wind observations). Otherwise, there have been no real obstacles or problems.

NWS/NCEP

The only issue raised (by NCEP) – is whether or not we could do the same for the CONUS RTMA. Ultimately, Mr. Levine is planning on extending (completing) this work once hired by NCEP. Recent (16 August 2010) comments and a request or two from Dr. Pablo Santos (NWS Miami) follow:

"…this project is a good example of what a good partners project should be. The project was well focused and there is direct operational gain that can be attained from it."

1) The results of the SFWMD network are not surprising to me. They take good care of that network, the sites are generally well sited, and the maintenance is good.

2) Wxflow surprises me. They pay special attention to their siting and exposure. Their stations are also all coastal stations which surprises me even more. Have you shared this with Jay Titlow, their point person? He does go out of his way reaching out to us from coordinating the station siting to seeking feedback from us. Given their growing network, and that it is one of the few networks we actually get 5 minutes obs from along our coasts in our systems, I think they should be approached with this finding. It would be interesting to see their response.

3) Would you mind me sharing this with him? Don't worry, I won't do without your permission. But to see none of their observations are being used because they did not pass the test to me is a big deal.

4) Do you have the equivalent of Table 1 from the directional bin tests, in other words, what is the percentage of acceptance as a function of direction given the RMSE and Z tests for each of those networks? It is possible I misread something here but if I did please then clarify it for me.

5) What about sharing in detail the acceptance list with the WFOs so that they can use it in their LAPS static blacklists? That would be a nice side benefit of this study.

The PI also spoke with NWS Melbourne SOO via phone (13 August 2010) regarding this report. Mr. Sharp asked whether or not the MySql approach was amenable to multiple data assimilation systems (e.g., LAPS, ADAS). He also requested that we invite their IT guy (Peter Blottman) and Forecaster Matt Volkmer to Steve Levine's defense.

SECTION 6: Project Future

The following were listed in the 6 month report. Comments (highlighted in yellow) follow each bullet in this final COMET report:

- Expand the study to 12 months from 5 (from 1 August 2008 – 31 July 2009)
  Nearing Completion.
- Include additional mesonet networks
  Nearing Completion.
- Perform/publish a similar statistical analyses but more comprehensive
  Nearing Completion (i.e., Z-tests, directional bin tests).
- Publish regional accept and reject lists for the wind field
  Updated lists were recently passed along to NCEP.
- Work with NCEP to develop a directional dependent accept/reject list
  In progress.

As previously mentioned, in addition to updating the static accept list database used by the RTMA, the COMET funded project has seeded an additional effort to improve the metadata which has been identified as a significant issue with respect to the National Mesonet effort.

APPENDIX A: Network Bias

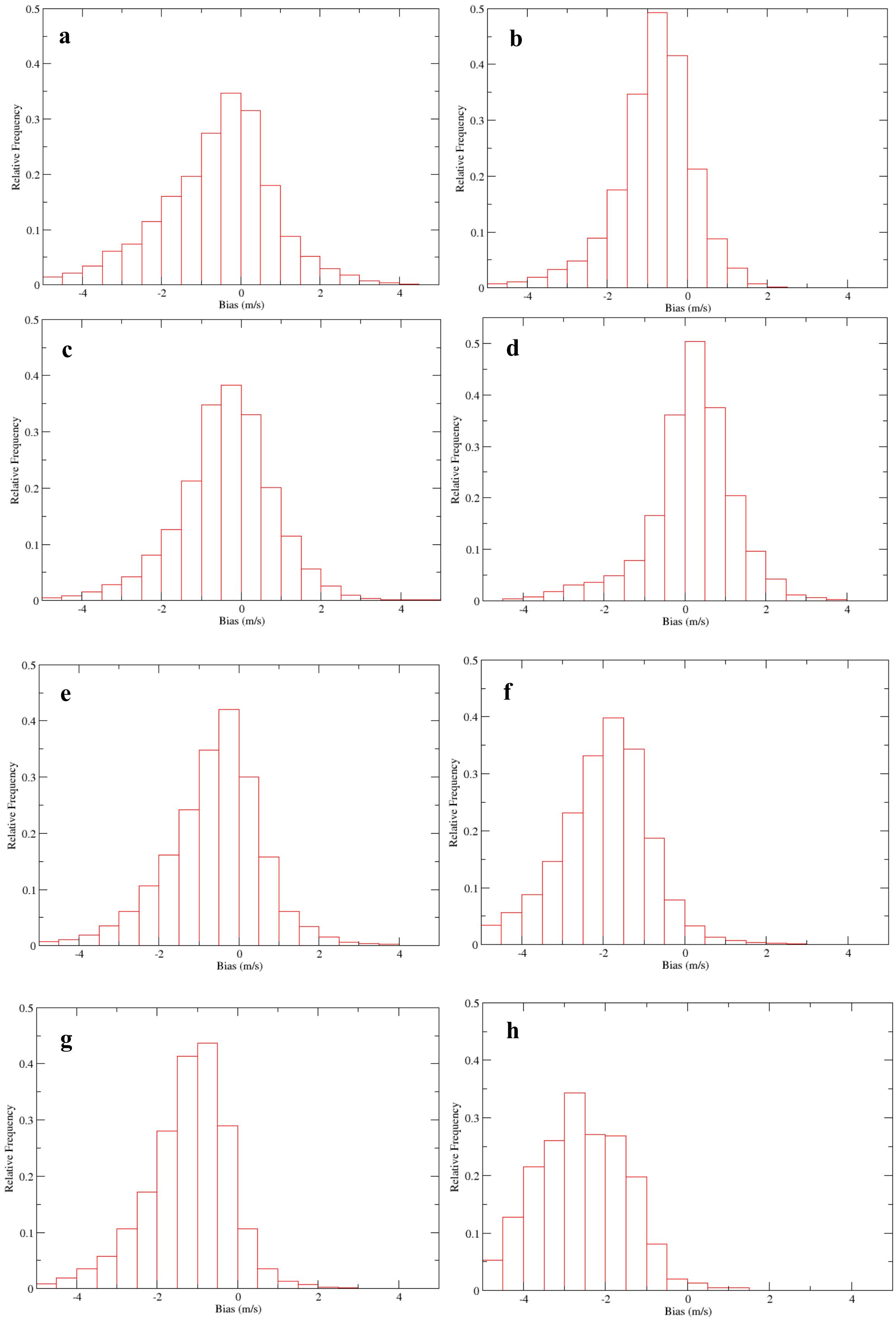

Fig. A.1 Relative frequency histogram of average daily bias for stations in: a.) National Ocean System, b.) RAWS, c.) METAR, d.) South Florida Water Management, e.) Weatherflow, f.) APRS/CWOP, AWS Convergence, and g.) Anything Weather networks.

APPENDIX B: UMLF1 vs. KLEE Directional Bias

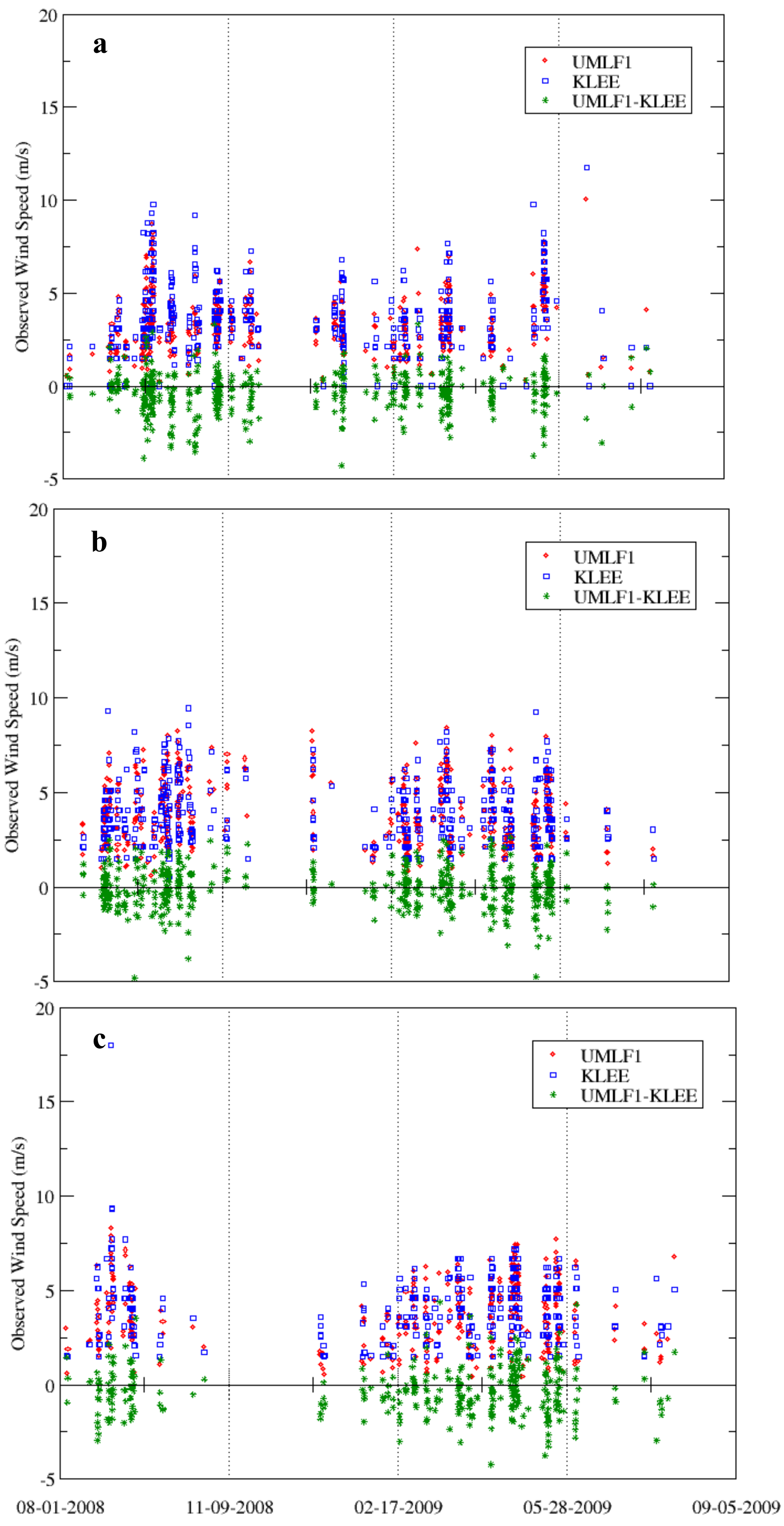

Figure B.1 Time series of wind speeds (ms^{-1}) and the difference (Mesonet – METAR, filled green circles) between two proximity weather stations: METAR site KLEE (open blue squares), and Florida Mesonet station UMLF1 (filled red circles) for directional bins a.) $0^o \leq d < 45^o$, b.) $45^o \leq d < 90^o$, and c.) $90^o \leq d < 135^o$.

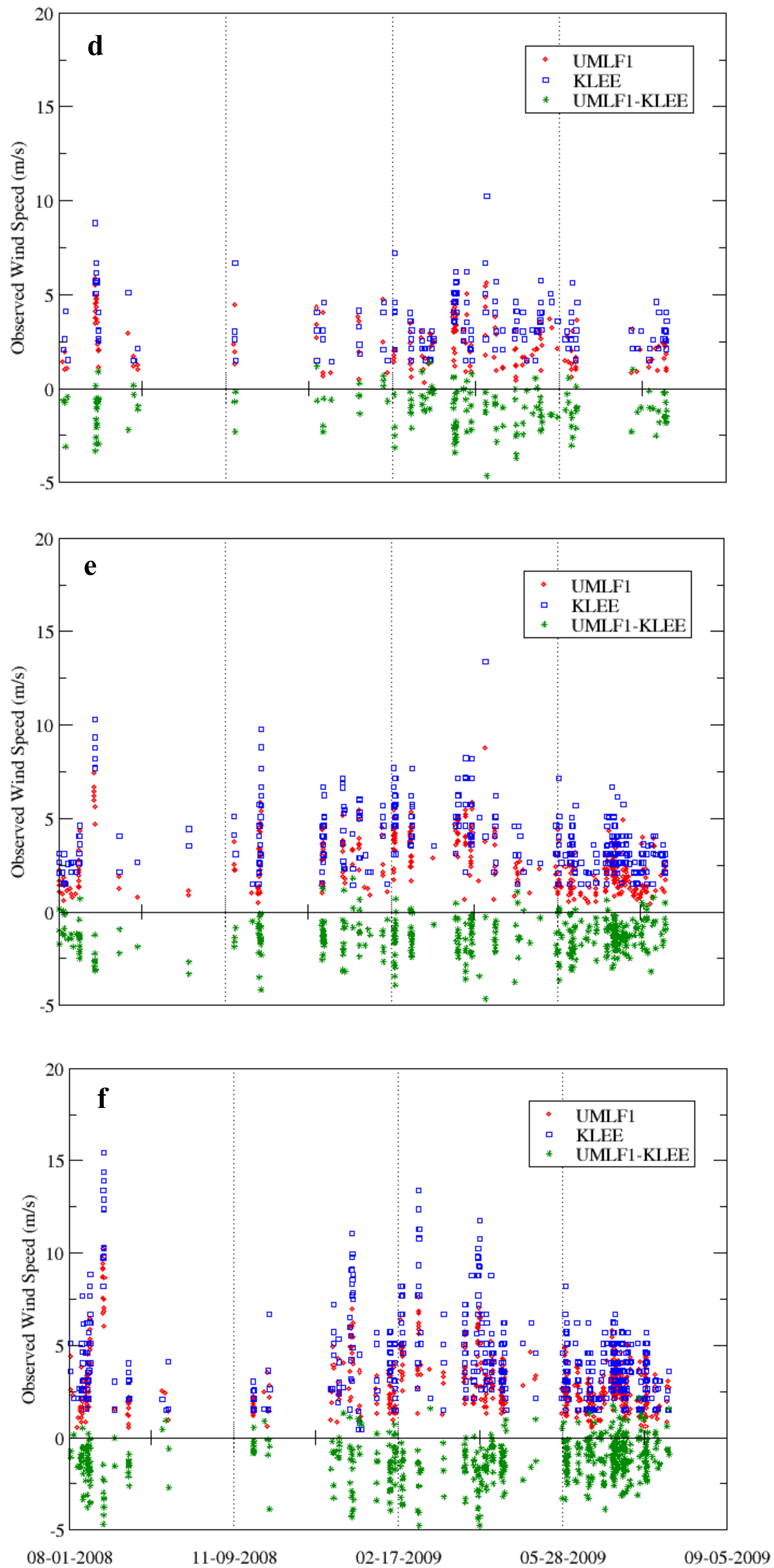

Figure B.1 d.) $135^o \leq d < 180^o$, e.) $180^o \leq d < 225^o$, and f.) $225^o \leq d < 270^o$.

.

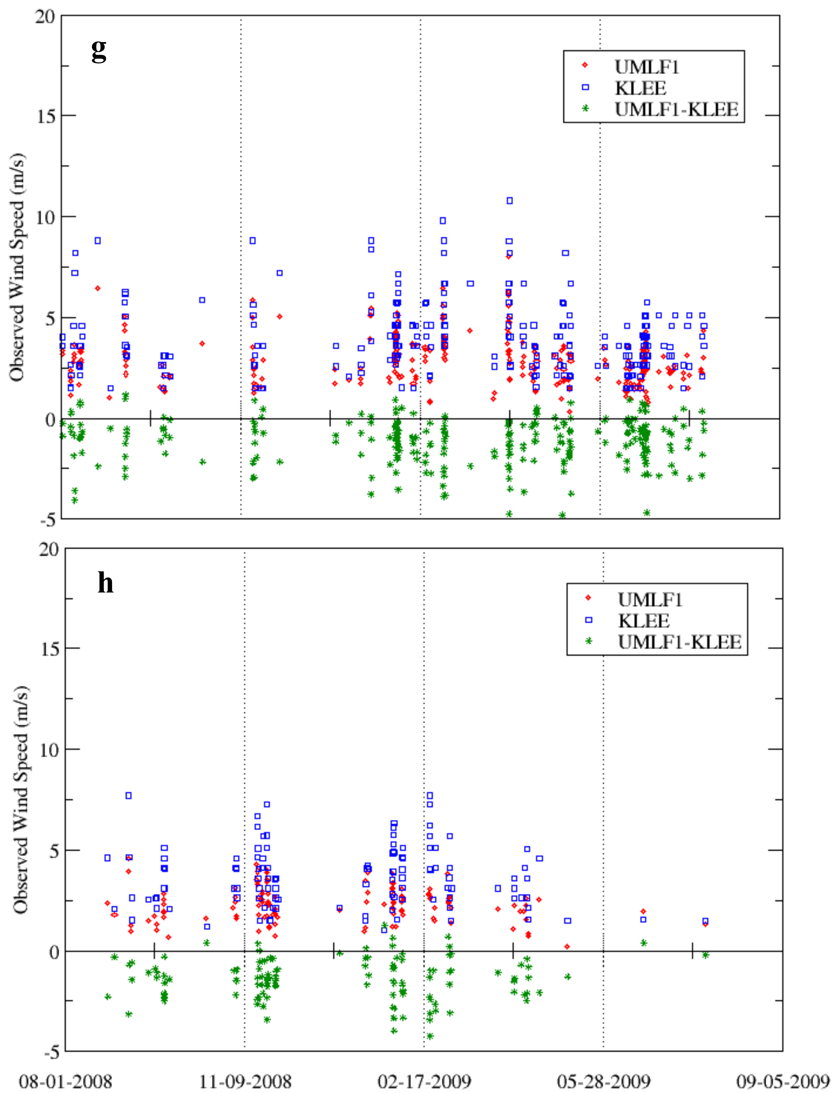

Figure B.1 g.) $270^{o} \leq d < 315^{o}$ and, h.) $315^{o} \leq d < 360^{o}$.

TABLES

Table 1: Surface observation networks, number of stations within each network, number of observations and number of observations accepted under the original RTMA quality control methods for the period 1 August 2008 – 31 July 2009. See text for details.

Provider Name	Network Code	Number Stations	# observations	Number accepted	% accepted
ARPS/CWOP Weather Network	ARPSWXNET	341	1508132	208513	13.8
Weatherflow	WXFLOW	48	262046	0	0
Florida Mesonet	Fl-Meso	35	217758	0	0
South Florida Water Management District	SFWMD	26	110753	110753	100
RAWS	RAWS	30	153309	147384	96.1
Anything Weather Network	AWX	4	20023	5287	26.4
Airport ASOS sites	ASOS	41	250463	247108	98.7
Non-ASOS METAR sites	OTHER-MTR	26	130760	129744	99.2
National Estuarine Research Reserve System	NERRS	3	816	765	93.8
National Ocean Service	NOS	18	78859	77057	97.8
Non-Federal AWOS Sites	NonFed AWOS	3	14924	6206	41.6
Weather For You	WxForYou	4	20230	0	0
Climate Reference Network	CRN	1	304	304	100
AWS Convergence Network	AWS	202	962627	250912	26.1
Synoptic 6-hour sites	SYNOP	8	9642	9562	99.2
Other	OTHR	1	3762	0	0
Total	**TOTAL**	**1119**	**3869954**	**1315380**	**34.0**

Table 2 Stations passing the expanded static test (i.e., passed both the initial RMS test as well as the Z-test). Z+/- 1 denotes test for bias of +/- 1 ms-1 respectively. 'Second' represent the same test applied to a separate/random sample taken from the same population.

Station1	Station2	Distance (km)	Heading (deg)	Z+1	Z-1	Second Z+1	Second Z-1
XNMI	KHWO	13	320	-2.81	9.83	-2.9	9.34
ALHF1	KGNV	19	135	-11.76	16.55	-9.06	12.65
XLRL	KSRQ	29	340	-2.59	10.05	1.75	6.58
XTER	KSRQ	17	173	-2.33	5.76	-2.32	7.01
XWKI	KBKV	13	115	-11.74	5.02	-13.53	5.88
XJAK	KNRB	6	80	-6.24	4.2	-7.51	5.94
XKYW	KEYW	6	115	-5.98	6.61	-7.33	6.37
MAIF	KMAI	1.5	220	-4.52	15.28	0.13	6.2
XRDY	KMCO	30	90	-4.76	13.34	-4.78	11.2
LIOF1	K40J	68	250	-6.32	8.66	-6.23	8.27
DLVL1	KDAB	5	230	-5.94	4.58	-6.54	3.4
QUIF1	KTLH	28	120	-5.84	8.94	-13.97	17.13
XURB	KMIA	10	110	-4.39	5.44	-3.66	5.44
MMSHL	KMIA	11	330	-3.62	5.03	-3.62	5.03
XTRP	KFMY	14	70	-3.09	2.14	-3.31	3.12
PCEF1	KFPR	8	20	-6.93	6.33	-5.65	5.74
C0924	KNQX	29	250	-1.71	16.01	-2.9	16.46
THVLL	KLEE	18	110	-3.72	25.63	-2.44	14.27
POPF1	KORL	24	115	-3.78	18.9	-3.46	16.21
D0905	KPGD	71	245	-2.21	16.92	-1.19	14.11
XCHL	KPGD	10	110	-2.37	5.87	-2.45	6.25
ARCAD	KPGD	35	200	-2.91	10.47	-1.77	7.42
D0832	KRSW	74	245	-3.62	8.82	-3.54	11.66
AIRGL	KRSW	74	250	-5.91	3.39	-11.18	5.61
SEBF1	KOBE	57	110	-4.59	12.93	-1.72	13.57
TRAF1	KOCF	27	190	-2.09	13.44	-2.3	15.8
BRZF1	KOCF	44	125	-7.87	4.69	-9.62	5.91
MDFR4	KOPF	16	270	-5.93	-1.7	-4.08	-0.18
XLOL	KVDF	26	140	-18	4.13	-17.74	4.54
DOVF1	KVDF	12	265	-11.93	10.89	-9.01	8.76
XGRF	KPCM	15	230	-11.24	11.8	-16.24	15.61
FROST	KBOW	32	310	-5.66	8.08	-12.26	13.9
STDF1	KHST	13	105	-2.81	13.11	-1.62	9.6

Table 3 Station-pairs that passed the directional bin test for at least 1 of the 8 directional bins. Headings (deg) give the METAR station location with respect to the Mesonet station.

Mesonet site	METAR site	Station Distance (km)	Heading (deg)	# bins passed
UMLF1	KLEE	21	240	7
XDAI	KMLB	7	350	6
MACF1	KVQQ	26	105	6
MDFR3	KTMB	1	270	6
HMSTD	KHST	10	135	6
MACF1	KVQQ	26	105	6
BLDF1	KPBI	51	90	5
FDLF1	KFLL	9	105	5
HMSHS	KHST	9	65	5
IMKF1	KRSW	32	285	5
STNF1	KSGJ	33	20	5
TAMPA	KTPF	9	200	5
XTKY	KHST	6	330	4
XFLM	KHWO	8	65	4
XCVN	KFXE	13	90	4
EPRF1	KDAB	39	95	4
C1018	KTPA	14	225	3
CLRCG	KFXE	15	115	3
FRMCB	KFMY	4	240	3
KENF1	KMLB	42	70	3
WCHUL	KTPF	75	305	3
XBCG	KFMY	42	110	3
ABS01	KPGD	68	245	2
AR589	KFPR	30	80	2
D1956	KMCF	24	315	2
FRDHV	KFMY	6	235	2
FTLSS	KFLL	6	125	2
HLLBS	KHWO	46	85	2
LKBTR	KGNV	38	170	2
LKWHL	KBOW	12	55	2
MDFR4	KOCF	1	270	2
MMLCC	KHST	6	110	2
NPLSS	KAPF	28	245	2
TTTFD	KTPF	14	205	2
VLNF1	KMCO	32	100	2
WNT16	KMCO	8	35	2
XBON	KRSW	17	15	2
XLWS	KSGJ	6	360	2
XROY	KPBI	6	100	2
XSAR	KSRQ	6	330	2
XSPR	KORL	15	150	2
XNPL	KAPF	2	90	1

XDSO	KFMY	14	225	1
XAZL	KORL	8	285	1
WNT14	KMCO	8	100	1
PMPHS	KPMP	1.5	320	1
OPLKA	KHWO	4	360	1
NPORT	KPGD	32	140	1
NPFD3	KAPF	14	215	1
NPFD1	KAPF	9	270	1
MMFMS	KMIA	3.5	160	1
MDFR2	KOCF	16	270	1
DVNHR	KHWO	12	160	1
D0479	KHST	14	85	1