Report on Marine Data Concerns and a new Track-Checking QC Code

Dr. Bradley Ballish
NOAA/NWS/NCEP/EMC/IMSG
Data Workshop 22-24 May 2018

23 May 2018

Introduction

- **NCEP needs an operational automated QC for marine data due to many location errors and stuck data problems**
- **The NRLACQC aircraft track-checking QC code, initially provided by NRL's Dr. Pat Pauley works very well with aircraft data, and NCEP is grateful for it**
- **Scientists in EMC have reported problems with the NRLACQC, that either Pat or EMC staff have fixed, which is mutually beneficial to all**
- **The NRLACQC code has large amounts of complex logic, so I decided on developing a new Marine Track-checking Quality Control (MTQC) that uses something like a computer minimization scheme to decide which observations to delete when there are track checking errors, which will be described later**
- **The new MTQC was tested with both artificial and real data and then applied to aircraft data, which allowed comparison with the NRLACQC, which helped refine it**
- **In studying the MTQC applied to both ship and buoy data it appears that we need better quality and quantity of marine data and better QC and feedback to data providers, which is a challenge to us all**
- **A number of slides are provided showing such problems**
- **Slides are given to support the possibility of having future automated ship weather reports**
- **Suggestions are made for how we could work together in the future**

Introduction with A Difficult Ship QC Example

- **The next slide shows an example that is a difficult ship QC problem for data in January 2017 and is a good introduction**
  - **From the 16th to the 20th, six observation locations were incorrectly in the Northern Hemisphere (NH) and only 3 were correctly in the SH**
  - **Three reports in a row, in time, were in the wrong hemisphere**
  - **This problem is too common for ship data likely due to errors in reporting which quadrant of the globe is correct for the location**
  - **In addition, some reports had simple location typos shown in red**
- **For marine reports that move slowly compared to aircraft and report less frequently in time, a much larger time window, on the order of 5 days, is needed to diagnose some location errors**
- **For any QC code to have skill at deleting the 3 reports in a row with wrong locations, the code needs to be of a high enough order to consider enough observations forward and backwards in time**
- **Some of the observations with wrong locations had modest analysis (AN) – background (BG) changes suggesting the analysis may have drawn slightly for bad data**
- **The following QC tests used future reports which are not available in operations**

Example of a Sequence of Reports from a Ship in January 2017

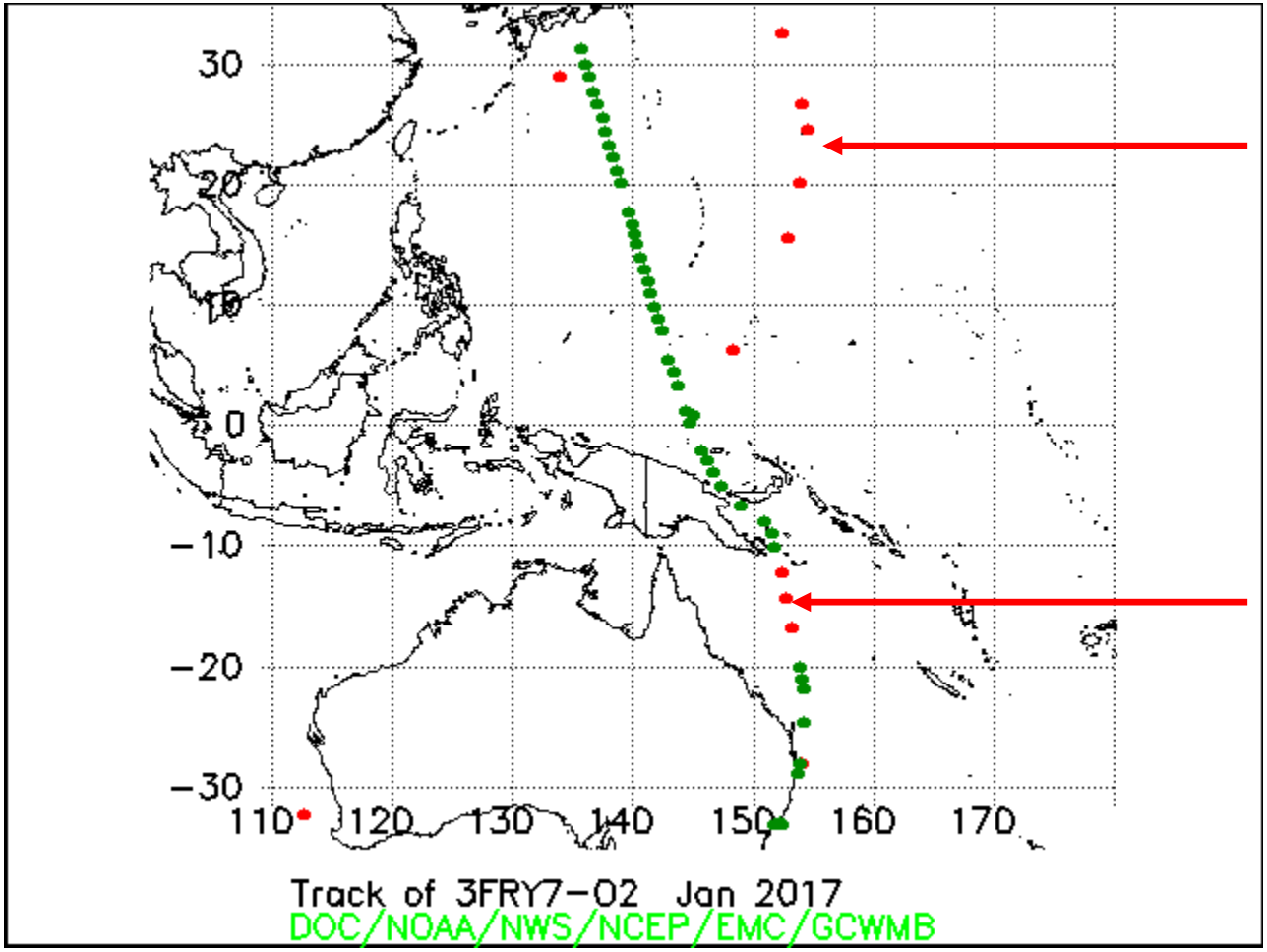
Ship	TimeInDays	Lat	Lon	Press	OB-BG	AN-BG	QM
3FRY7	13.00000	-3.90	146.60	1012.2	2.3	0.2	2
3FRY7	13.25000	-5.00	147.20	1006.7	-0.6	-0.2	2
3FRY7*	13.54167	6.30	148.20	1012.4	3.2	0.2	2
3FRY7	13.75000	-6.70	148.90	1007.7	0.2	-0.2	2
3FRY7	14.25000	-8.00	150.90	1006.3	0.1	-0.1	2
3FRY7	14.50000	-8.90	151.50	1010.6	2.0	0.3	2
3FRY7	14.75000	-10.10	151.70	1007.2	0.2	0.1	2
3FRY7	15.25000	-12.30	152.30	1009.0	1.8	0.2	2
3FRY7	15.75000	-14.40	152.70	1009.3	1.0	-0.1	2
3FRY7*	16.00000	15.50	152.90	1012.2	0.6	0.1	2
3FRY7	16.29167	-16.80	153.10	1010.2	0.4	-0.3	2
3FRY7*	17.00000	20.10	153.80	1015.2	-0.3	0.0	2
3FRY7*	18.00000	24.50	154.40	1016.2	-5.1	0.0	2
3FRY7*	18.50000	26.70	154.00	1012.2	-11.5	0.1	10
3FRY7	18.75000	-28.00	153.90	1007.0	-1.8	-0.3	2
3FRY7*	19.00000	29.00	133.80	1013.2	-5.5	-0.3	2
3FRY7*	20.00000	-32.30	112.70	1002.2	-11.6	0.1	10
3FRY7*	20.50000	32.60	152.40	999.8	-7.0	-1.5	1
3FRY7	20.83333	-33.00	152.20	1007.7	0.4	0.1	2

Red numbers are typos

Blue shows quadrant N-S errors

The * indicates data deleted by new 3rd or 4th order QC

This plot shows the Marine QC results from the new second Order QC

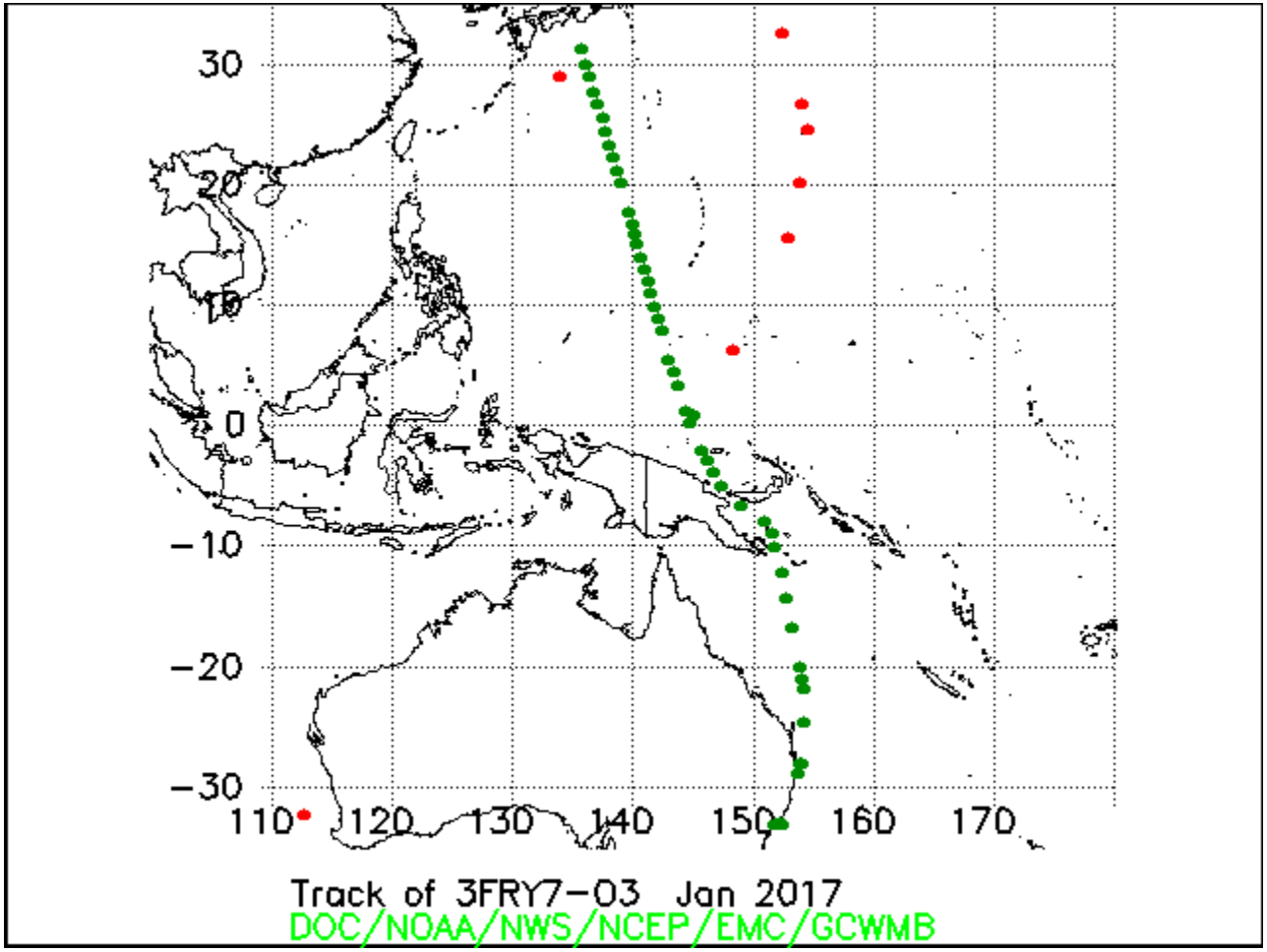


These OBs were in the wrong hemisphere with times belonging down below

The QC deleted good data due to OBs nearby in time but in the NH

The red dots are deleted reports, while the green passed the new QC

This plot shows good QC results from the new 3rd or 4th Order QC



The QC here appears completely correct

The red dots are deleted reports, while the green passed the new QC

New Marine QC Code Logic

- **First the data are sorted by ID and time such as sorts on 3FRY7 `13.25000, etc.`**
- **Since ships can and do change directions and speeds for many reasons, high travel speed is the most reliable QC check**
- **When an impossible speed of travel is found going from OB J to J+1, then using diagnostics on 2*N +1 OBs about the error:**
  - **Using from Obs J+1 –N to J+1 +N where N is the order of QC**
    - **These limits are restricted so each OB is between 1 and the total number of OBs**
  - **For each OB in this group, a measure R of the track-checking problem is calculated with each OB in the group one at a time compared with all other OBs in the group to see which OB gives the largest measure R**
  - **R(K) is the number of inter observational high speed jumps in travel with OB K compared to the others, and then decreased by .001* times the RMS differences in adjacent travel velocity vectors without using OB K**
  - **R is dominated by high travel speed checks, but the smoothness of the adjacent travel velocities gives better choices for deletes**
  - **After finding which OB gives the largest value of R, then that OB is deleted**
  - **Further iterations are performed if needed with just passing OBs from the past iteration**
  - **Diagnostics are made before and after QC in part to check on the QC**
  - **Many tests were run with real and artificial data with known errors added**
- **I do not fully understand how to best define R, which I call the Rychtar Number in honor of NOAA Port Officer Paula Rychtar**

New Marine QC Code Logic (Continued)

- **Using higher order seems to be better for more complicated errors, but does not require code rewrites or lengthy code**
  - **For example, 4th order QC has successfully deleted groups of 4 OBs in a row timewise at wrong locations**
- **To count as an inter observation high speed of travel, the speed must be at least 30 mph, Ingleby uses 15 m/sec.**
- **To avoid excessive deletion of minor errors, the high speed must also be over a distance of at least 50 miles – what is best?**
- **The code removes each deleted observation before further QC tests, but records each action and uses 4 iterations**
- **Further work is needed towards deleting large parts of the data if the QC has not worked adequately after its iterations are done**
- **Note some data QC decisions could change over time, so that data that passed QC in the last few days, could be deleted with new data being used in future runs**
- **The code is currently not using OB-BG values to help with QC decisions, but that is likely to be used in the future**
- **Recent tests suggest the code needs to be tested using quadrant corrections for east-west or north-south errors**
- **Additional QC is performed for reports with both suspect speeds and heading changes based on shared info from Bruce Ingleby**

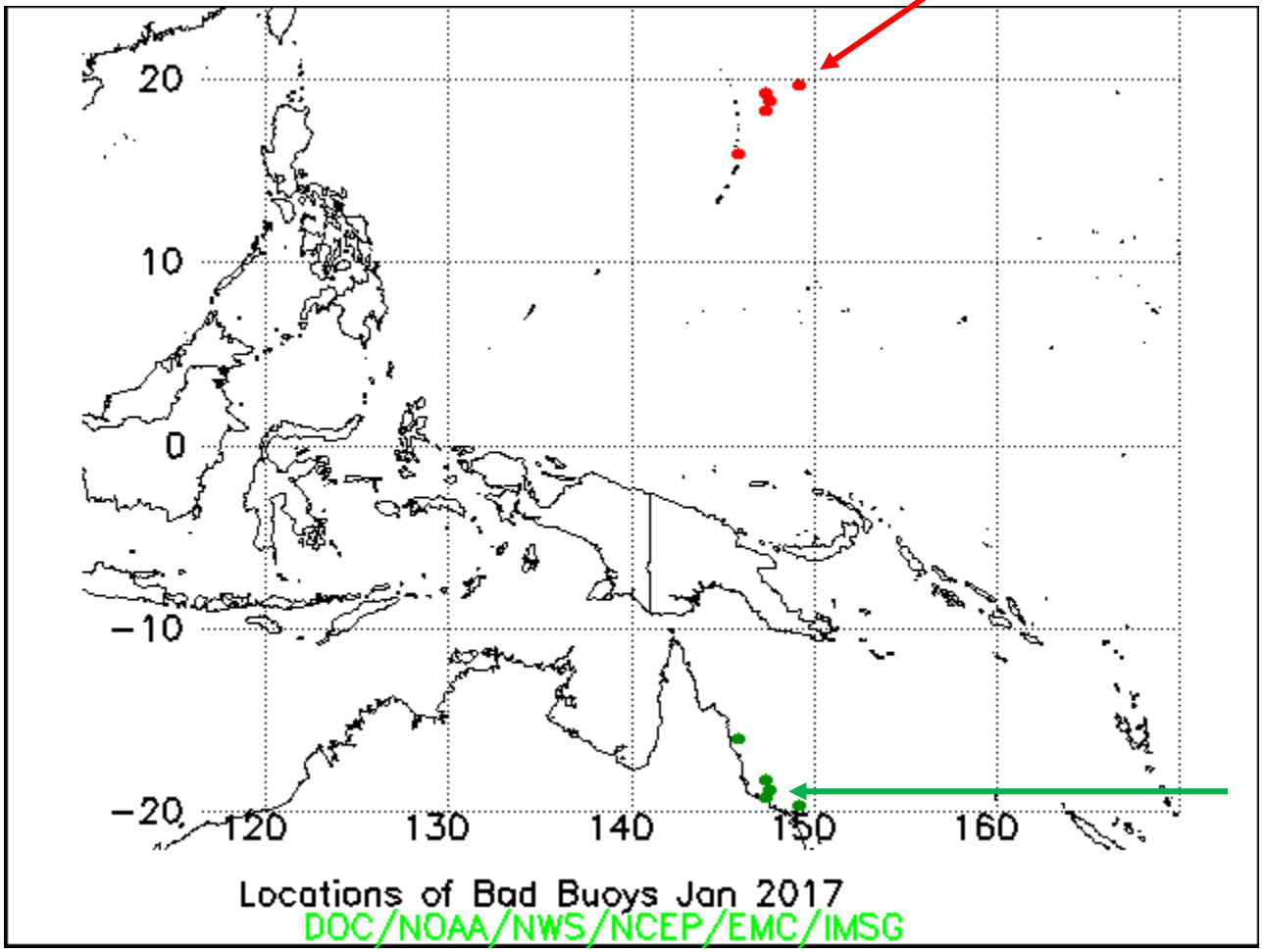
New Marine QC Code Logic (Continued)

- **In many tests with fake data with a wide variety of errors introduced as well as with tests with real data, the 4th order QC worked very well**
- **QC errors were hard to find, and examples of the errors will be emphasized here**
- **In some cases of QC errors there were more bad reports than good, for which any QC is not likely to work well**
- **It appears that other cases of QC errors were partly due to large differences in time between reports coupled with errors**
  - **Maybe changes in the Rychtar number R to include something like adding a factor for distance travelled with the adjacent travel velocity vector differences may help to give better answers for difficult QC cases**

A Serious Case of a Group of Buoys With Prolonged Location Errors

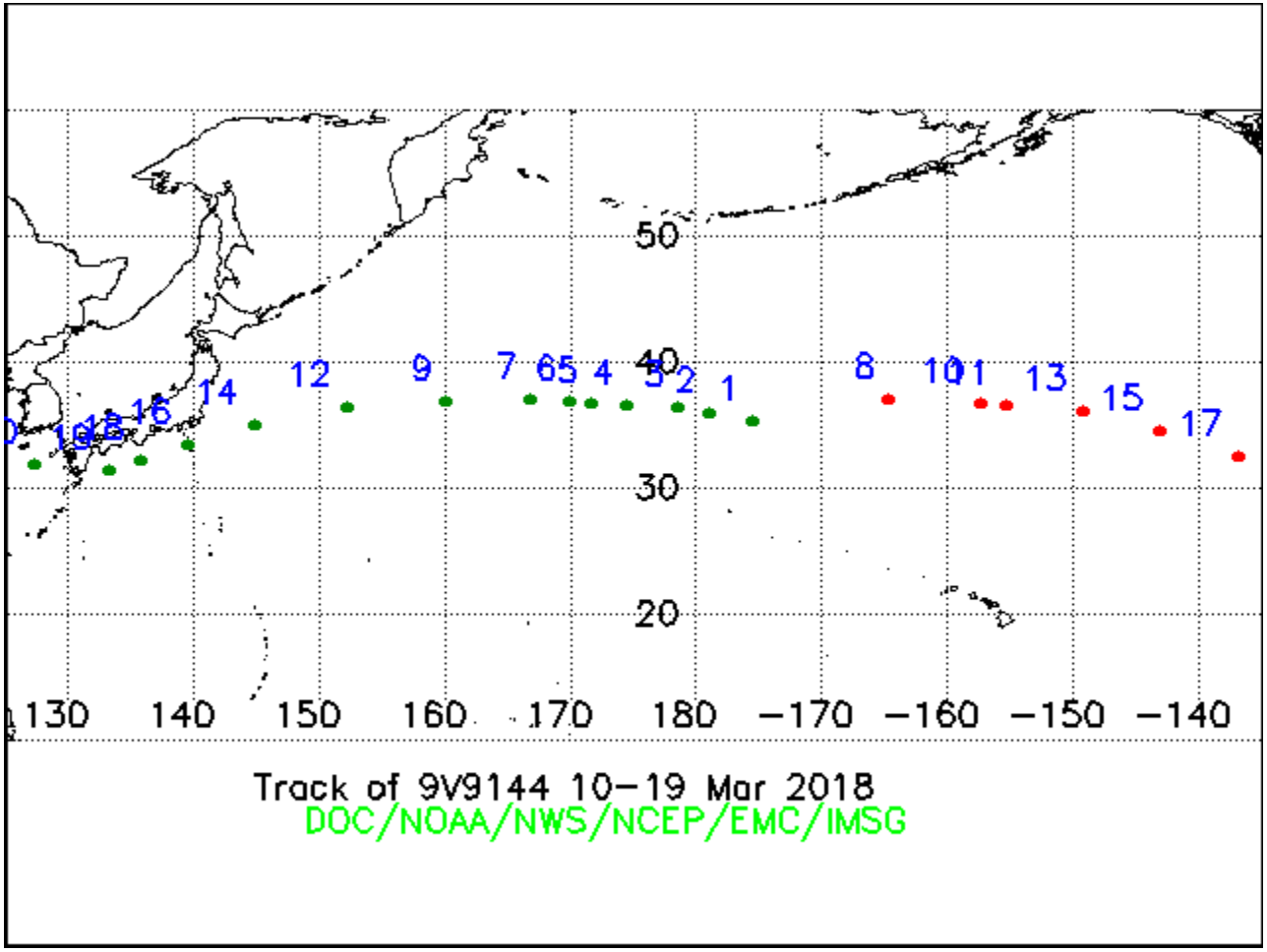
- **The next slide shows a serious problem where 5 buoys were reporting at wrong locations in the northern hemisphere (NH) for 6 months until mid day on the 16th of January 2017 when they jumped to mirrored locations in the southern hemisphere (SH)**
  - **The data in the NH had poor fits to the model BG, while the data in the SH had good fits**
  - **The buoy IDs were AGRF, GBGR, DVRF, HRRF and MYRF**
  - **EMC's Chris Hill reported the data had headers indicating the data were put on the GTS by Australia, which is consistent with their correct locations in the SH**
  - **Herve Benichou of France reported this problem was there from April 2016 to Jan 2017, and it is quite likely this caused negative impact on NCEP analyses and forecasts**
  - **ECMWF's Ersagon reported these buoys were on their reject-list in January**
  - **Many of these reports in the NH were passing QC at NCEP, with some indication that the analysis was drawing for some for the bad data**
- **It is a problem that either the data monitoring centers did not succeed in reporting that these buoys had suspect data to the data providers more quickly or the providers acted too slowly**
- **Now that NCEP's GFS has a 4DVAR ensemble system that can draw more for marine data, such problem sites need to be on their reject-list**
- **The new QC code can only delete some of the groups of reports with wrong locations, adding OB-BG info could help**

These OBs were at wrong
Locations for several months



Correct
Locations

This plot shows the Marine QC results from the latest QC for Ship 9V9144
The blue numbers show the time order of the observations



The red dots are good rejections of OBs in wrong quadrant

The plots often show more data than the text listings, and green dots show OBs that passed later full QC and red failed

Sequence of Reports from Ship 9V9144 in Mar 2018

Quality Control Diags Before QC - Unit 9V9144

TimenDays	Lat	Lon	DelTH	Dist	DIR	SPD
9.95833	35.30	184.40	999.9	999.9	999	999.9
10.41667	35.90	181.10	11.0	189.9	103	17.3
10.79167	36.40	178.50	9.0	149.1	103	16.6
11.33333	36.60	174.40	13.0	228.1	93	17.5
11.83333	36.80	171.60	12.0	155.7	95	13.0
12.00000	36.90	169.90	4.0	94.2	94	23.6
12.41667	37.00	166.80	10.0	171.3	92	17.1
12.87500	37.10	195.20	11.0	1560.0	270	141.8
13.37500	36.90	160.00	12.0	1930.9	90	160.9
13.75000	36.70	202.60	9.0	2336.5	270	259.6
14.04167	36.60	204.70	7.0	116.6	273	16.7
14.45833	36.40	152.30	10.0	2872.3	90	287.2
14.91667	36.10	210.80	11.0	3206.5	270	291.5
15.54167	35.10	144.90	15.0	3627.2	89	241.8
15.87500	34.50	216.80	8.0	3982.4	271	497.8
16.54167	33.40	139.50	16.0	4312.0	89	269.5
16.95833	32.50	223.20	10.0	4704.5	271	470.5
17.12500	32.20	135.80	4.0	4933.9	90	1233.4
17.58333	31.40	133.30	11.0	156.8	69	14.3
18.50000	31.90	127.40	22.0	348.7	96	15.8

There are many reports in wrong quadrant E VS W that result in high travel speeds (SPD) in MPH and big direction changes

DelTH is time in hours to last OB
Dist is distance in miles to last OB
DIR is direction of travel from last OB with 270 being west to east movement
SPD is the speed of travel in MPH from last OB
999 means missing

For operational use times like 18.5000 will be replaced by the Julian Date with .50000 added on. This will allow QC across changing months and years

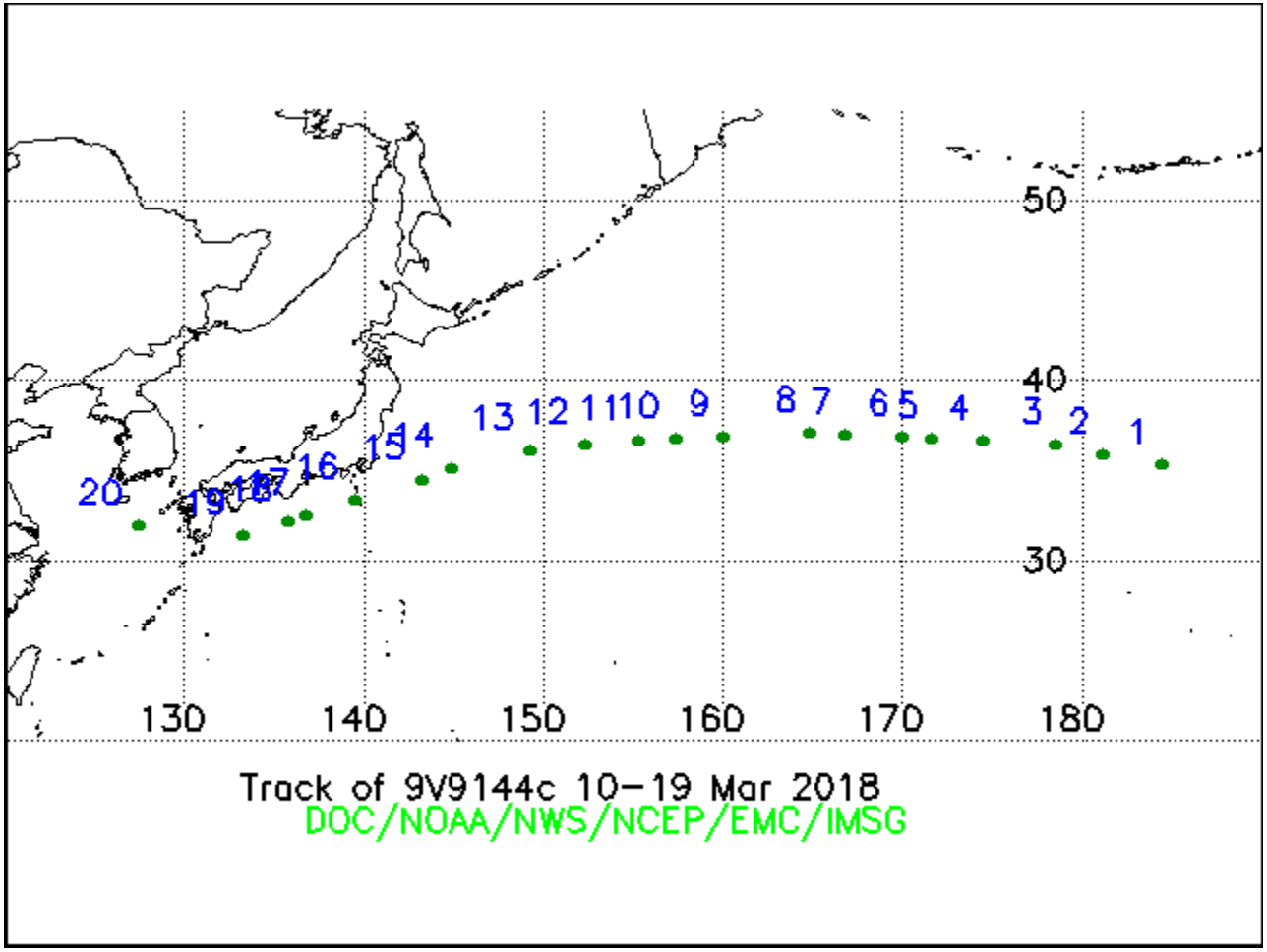
Sequence of Reports from Ship 9V9144 in Mar 2018

Quality Control Diags After QC - Unit 9V9144

TimenDays	Lat	Lon	DelTH	Dist	DIR	SPD
9.95833	35.30	184.40	999.9	999.9	999	999.9
10.41667	35.90	181.10	11.0	189.9	103	17.3
10.79167	36.40	178.50	9.0	149.1	103	16.6
11.33333	36.60	174.40	13.0	228.1	93	17.5
11.83333	36.80	171.60	12.0	155.7	95	13.0
12.00000	36.90	169.90	4.0	94.2	94	23.6
12.41667	37.00	166.80	10.0	171.3	92	17.1
13.37500	36.90	160.00	23.0	375.4	89	16.3
14.45833	36.40	152.30	26.0	428.1	85	16.5
15.54167	35.10	144.90	26.0	424.4	78	16.3
16.54167	33.40	139.50	24.0	329.9	69	13.7
17.12500	32.20	135.80	14.0	230.3	69	16.4
17.58333	31.40	133.30	11.0	156.8	69	14.3
18.50000	31.90	127.40	22.0	348.7	96	15.8

After QC the speeds and directions look fine, but many reports are gone

This plot shows the Marine QC Results After Quadrant Correction for Ship 9V9144



After quadrant corrections, the locations look fine and no data were rejected

The plots often show more data than the text listings, and green dots show OBs that passed later full QC and red failed

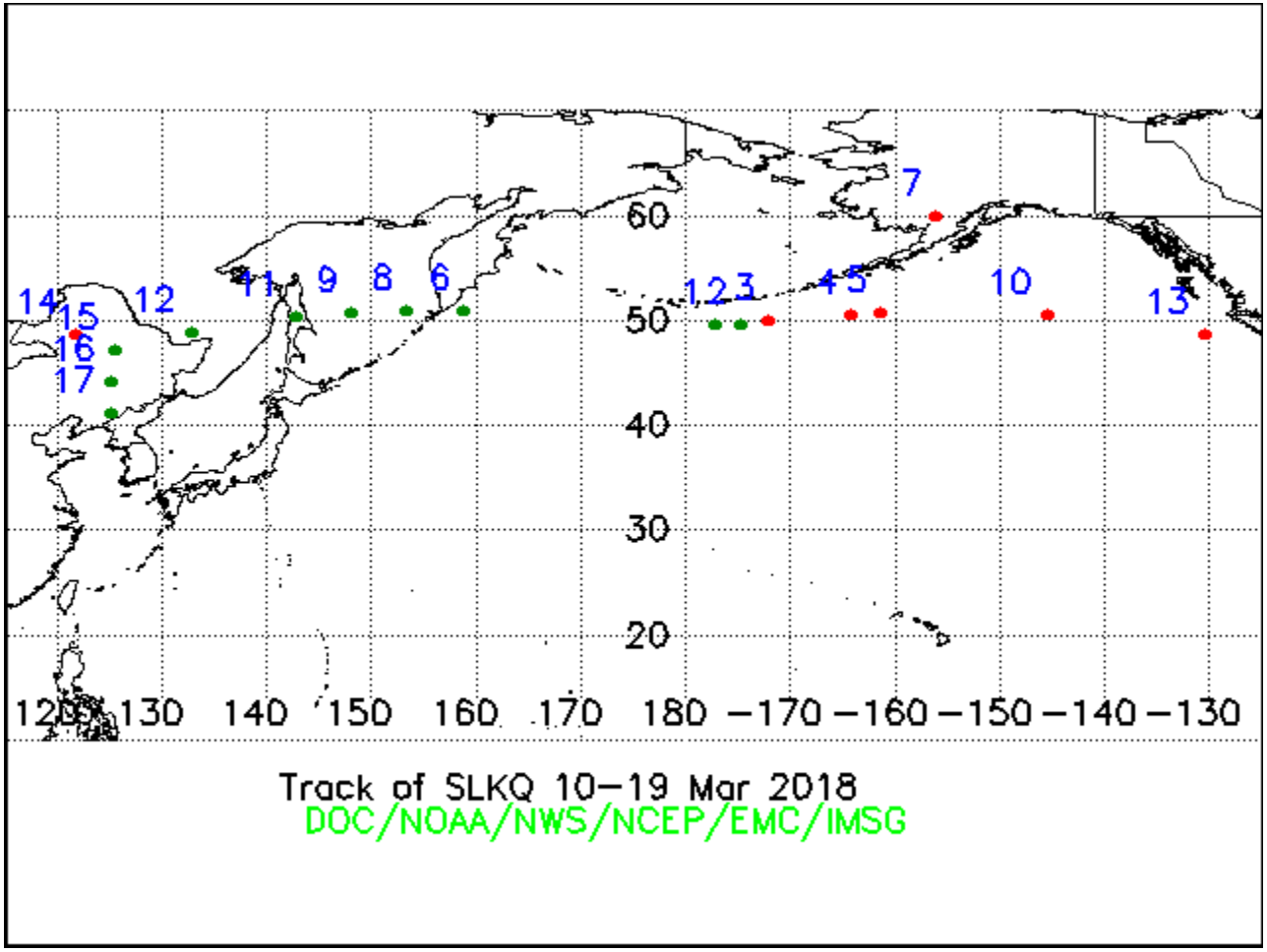
Sequence of Reports from Ship 9V9144 in Mar 2018

Quality Control Diags After Quad Cor - Unit 9V9144

TimenDays	Lat	Lon	DelTH	Dist	DIR	SPD
9.95833	35.30	184.40	999.9	999.9	999	999.9
10.41667	35.90	181.10	11.0	189.9	103	17.3
10.79167	36.40	178.50	9.0	149.1	103	16.6
11.33333	36.60	174.40	13.0	228.1	93	17.5
11.83333	36.80	171.60	12.0	155.7	95	13.0
12.00000	36.90	169.90	4.0	94.2	94	23.6
12.41667	37.00	166.80	10.0	171.3	92	17.1
12.87500	37.10	164.80	11.0	110.5	94	10.0
13.37500	36.90	160.00	12.0	265.2	87	22.1
13.75000	36.70	157.40	9.0	144.5	85	16.1
14.04167	36.60	155.30	7.0	116.6	87	16.7
14.45833	36.40	152.30	10.0	167.2	85	16.7
14.91667	36.10	149.20	11.0	173.9	83	15.8
15.54167	35.10	144.90	15.0	251.2	74	16.7
15.87500	34.50	143.20	8.0	105.0	67	13.1
16.54167	33.40	139.50	16.0	225.2	70	14.1
16.95833	32.50	136.80	10.0	168.4	68	16.8
17.12500	32.20	135.80	4.0	61.9	70	15.5
17.58333	31.40	133.30	11.0	156.8	69	14.3
18.50000	31.90	127.40	22.0	348.7	96	15.8

The speeds and directions after quadrant corrections look fine, and none were rejected by the QC

This plot shows the Marine QC results from the latest QC for Ship SLKQ

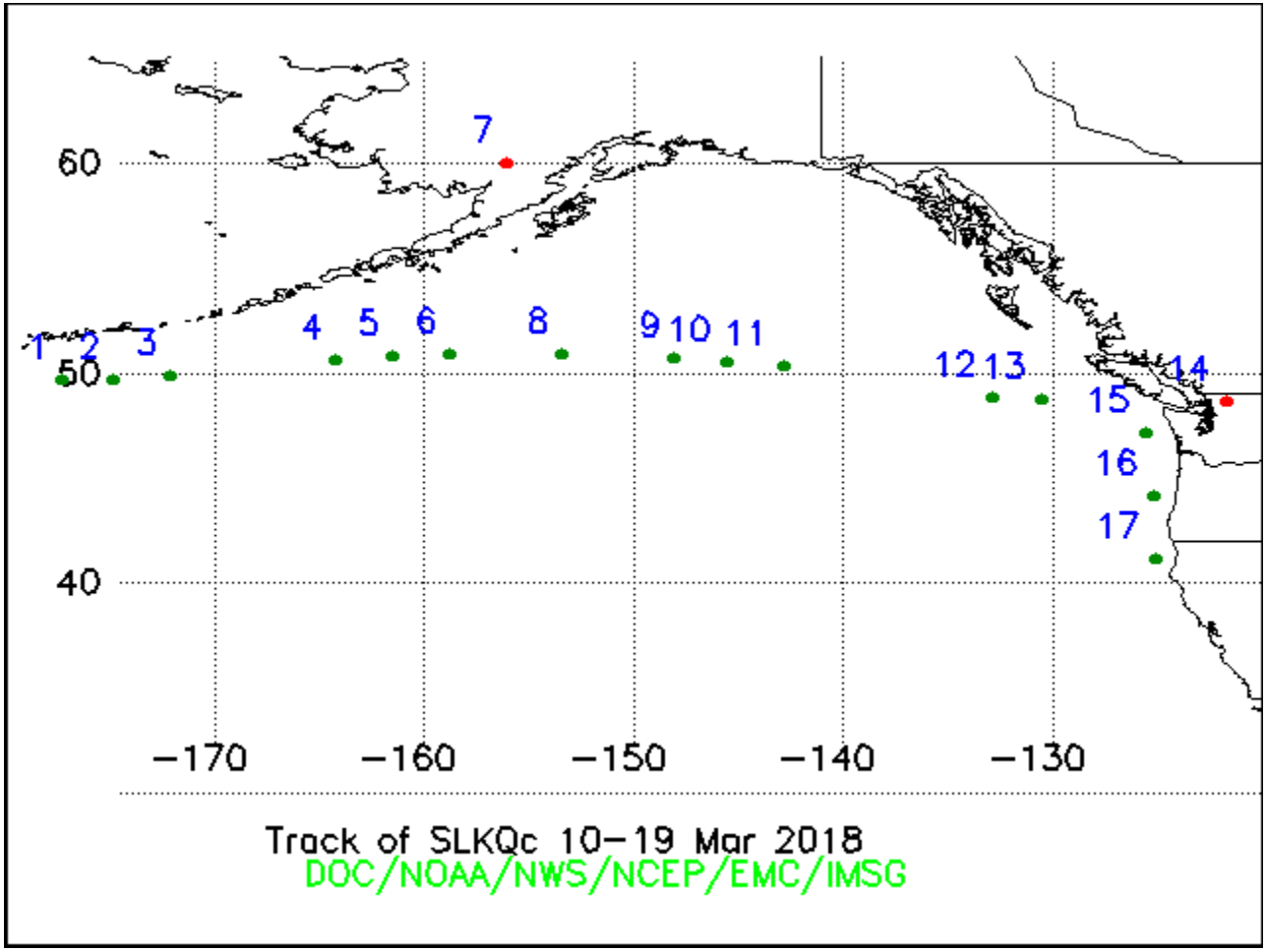


Most red dots are wrong rejections of OBs in a good quadrant, and far west green dots are OBs that are wrong and should be east of 180

Many OBs to the far left are over land and fail earlier QC

This example has so many errors that the best automated solution could be to delete all of the reports

This plot shows the Marine QC results from the latest QC
After Quadrant Correction for Ship SLKQ



Two red dots are good rejections of OBs with simple typos. The green dots look fine

The plots often show more data than the text listings, and green dots show OBs that passed later full QC and red failed

Sequence of Reports from Ship A8RW4 in Sep 2017

Quality Control Diags After Cycling QC - Unit A8RW4

TimenDays	Lat	Lon	DelTH	Dist	DIR	SPD
4.75000	8.10	132.60	6.0	136.0	315	22.7
5.50000	4.40	136.50	18.0	370.2	314	20.6
6.25000	0.50	140.60	18.0	390.7	314	21.7
6.58333	-1.00	142.40	8.0	161.9	310	20.2
6.75000	-1.80	143.30	4.0	83.2	312	20.8
8.00000	7.40	149.90	30.0	781.6	216	26.1
8.75000	-12.10	152.20	18.0	1356.4	353	75.4
9.00000	-13.90	152.60	6.0	127.2	348	21.2
9.25000	-15.70	152.90	6.0	126.0	351	21.0

Quality Control Diags After Final QC - Unit A8RW4

TimenDays	Lat	Lon	DelTH	Dist	DIR	SPD
4.75000	8.10	132.60	6.0	136.0	315	22.7
5.50000	4.40	136.50	18.0	370.2	314	20.6
6.25000	0.50	140.60	18.0	390.7	314	21.7
6.58333	-1.00	142.40	8.0	161.9	310	20.2
6.75000	-1.80	143.30	4.0	83.2	312	20.8
8.75000	-12.10	152.20	48.0	936.9	319	19.5
9.00000	-13.90	152.60	6.0	127.2	348	21.2
9.25000	-15.70	152.90	6.0	126.0	351	21.0

The report at 7.40N is wrong, but passes QC until OB 18 hours later is used

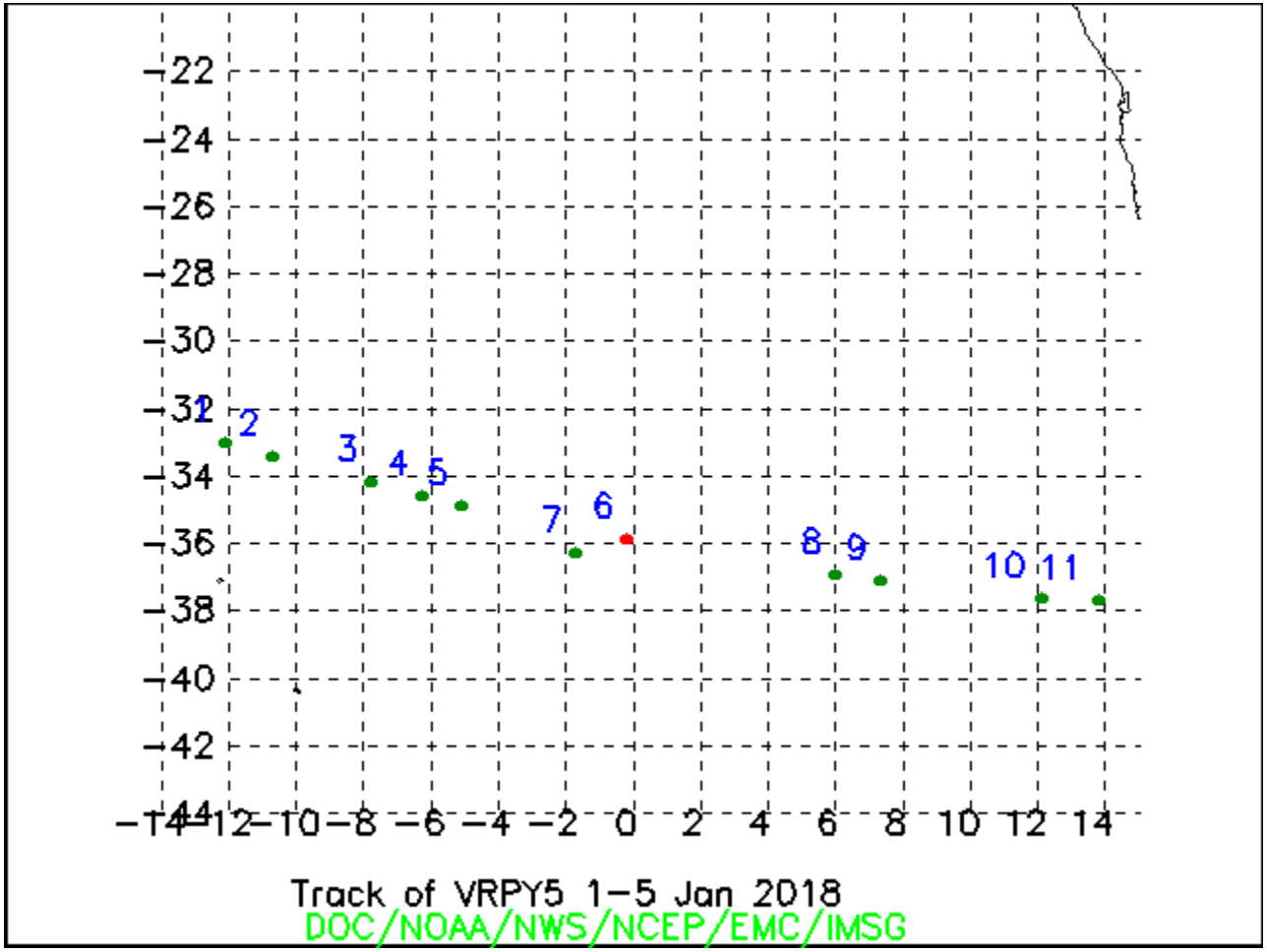
Hence QC results can change with new data

The 00Z report at 7.4N Had a speed of 26.1 from The previous OB but caused a high speed, 75.4 mph to the next OB – and passed the early QC for 00Z on the 8th which does not have the the next OB to show a speed problem

After QC the speeds and directions look good

The OB at 8.00000 causes a large direction change over a good distance, but not a high speed. The angle QC logic can delete this OB without future data due to suspect heading and speed

This plot shows the Marine QC results from the QC for Ship VRPY5

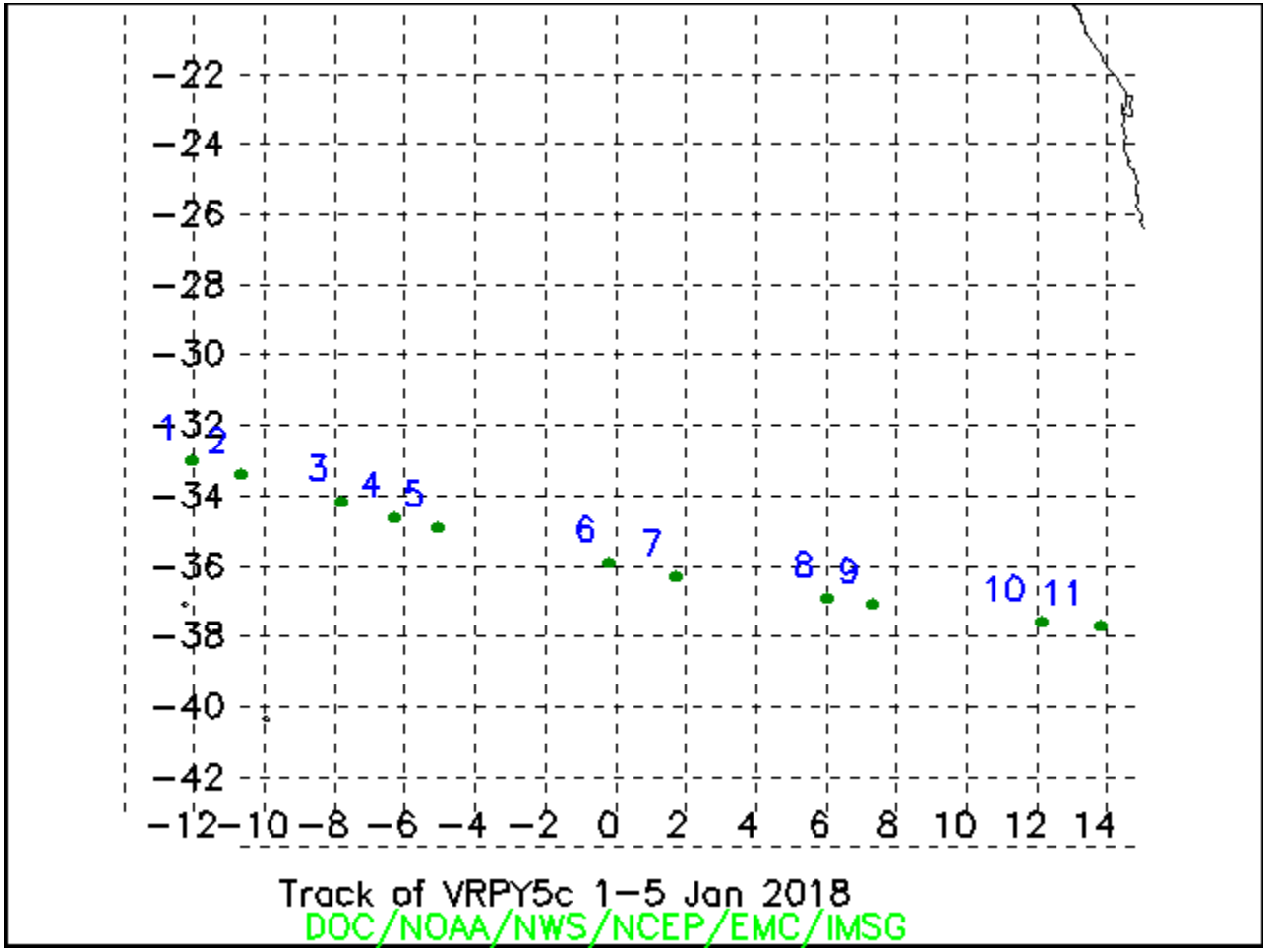


OB 6 red dot is a wrong rejection of a good OB, due to OB 7 error and large time increments

OB 7 at 1.7W should be 1.7E and OB 6 was correctly located

The plots often show more data than the text listings, and green dots show OBs that passed later full QC and red failed

This Plot Shows the Results After Quadrant Correction for Ship VRPY5



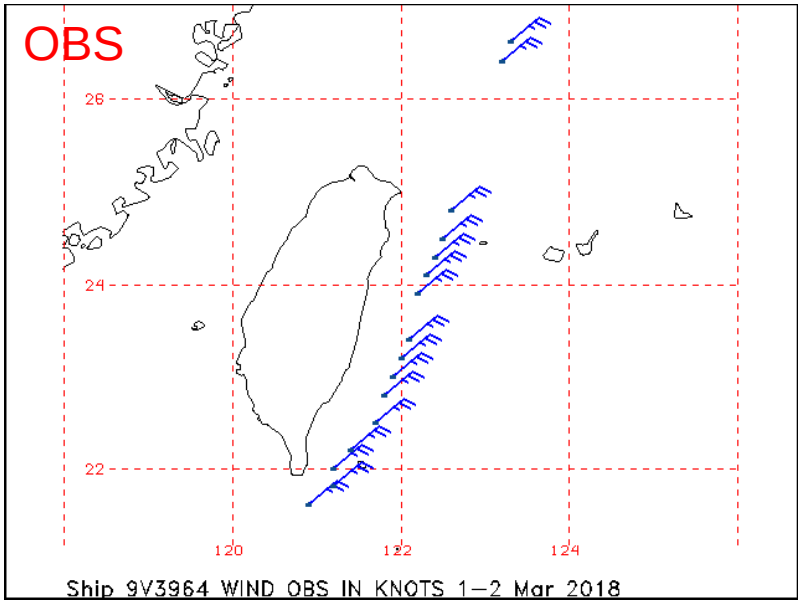
The speeds and directions look fine and none were rejected after OB 7 was moved from 1.7W to 1.7E

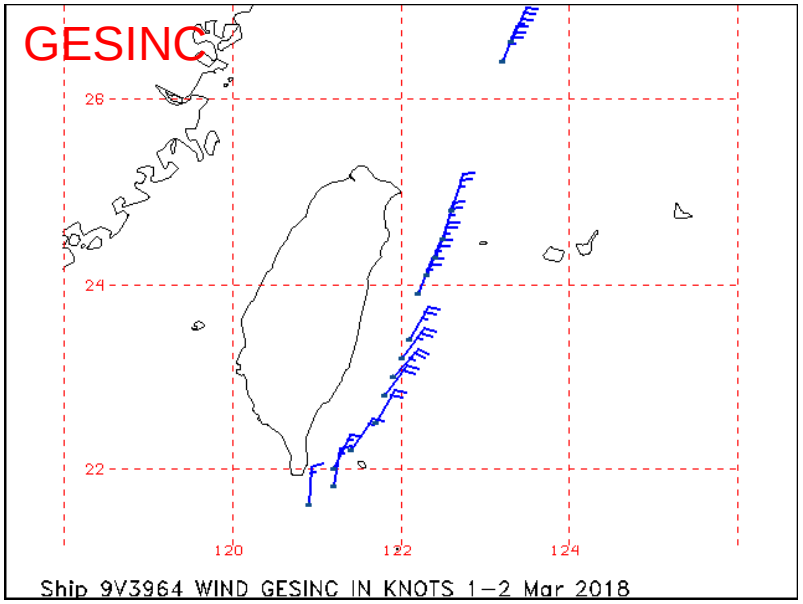
The plots often show more data than the text listings, and green dots show OBs that passed later full QC and red failed

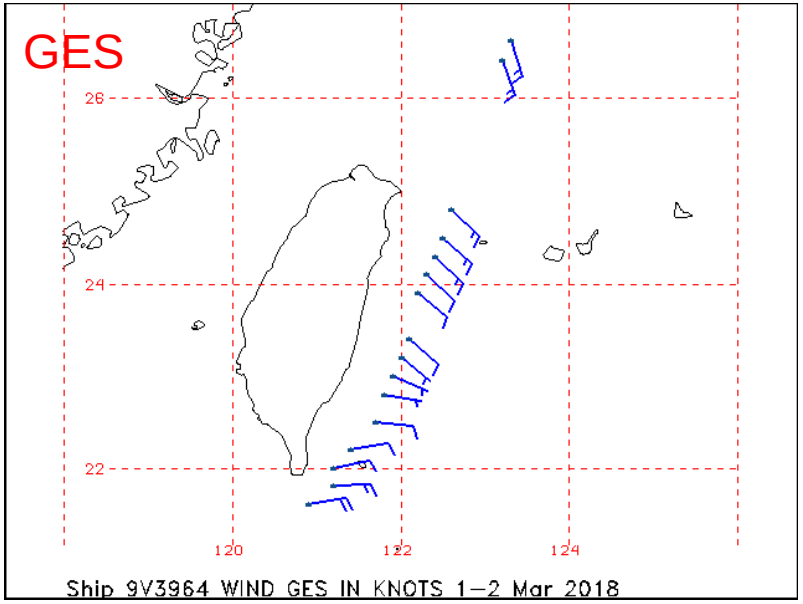
Sites with Significant Surface Pressure Biases 1-24 Apr 2018. Many of these need bias corrections, but some have large standard deviations, possibly due to typos, and should be rejected

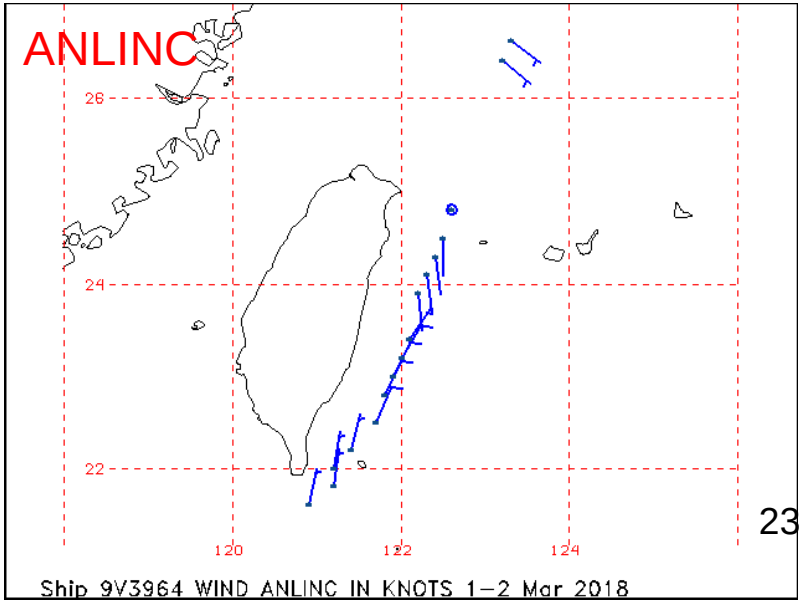
•2IYN3	3.6	3EFI	3.1	3ENG6	-4.4	3EQO	3.7
•3FCJ9	3.8	3FRT8	-3.1	5BZE2	5.4	5BZL3	4.8
•9HA3553	4.8	9HJD9	-4.0	9HYO7	8.9	9V9289	4.9
•9V9373	4.6	AUYP	5.7	AVBF	6.2	C6AB7	11.1
•C6DF6	-4.5	C6FM9	5.8	C6FN5	-4.2	C6FS9	-3.2
•C6FW9	-3.2	C6UC3	4.9	C6YA7	6.8	HPUO	-4.2
•ICJA	-5.5	J8AZ3	3.5	KGTX	-3.8	ONFN	-3.0
•ONHA	4.0	OWTW2	-4.1	OZ2049	-5.6	S6NQ	3.5
•SBPQ	-6.3	UBMO9	3.6	UCFT	-4.0	UGZM	-3.2
•V7BY3	4.0	VNSZ	-3.1	VRBF3	4.4	VRBJ5	4.6
•VRGO3	-4.4	VRGO8	-5.1	VRIB2	4.1	VRID2	4.8
•VRID6	3.8	VRJL6	3.6	VRLA6	4.1	VRLQ4	-3.2
•VRLZ4	-3.1	VRNA8	5.0	VRNR5	5.7	VRPY5	3.4
•VRPY7	-3.6	VRWN7	3.4	VTFG	-3.7	VTSG	14.2
•VTXB	8.8	VWTI	8.4	WCAJ	4.7	WDB3161	3.9
•WDC6027	-3.6	WDE4432	-4.1	WDG2803	3.0	WDI6469	3.8
•WDJ7294	-4.8	WKPM	4.2	XJBO	-4.0	YJUP4	-14.6
•ZCDK2	6.6						

Ship 9V3964 With Stuck Winds 09Z 1 Mar to 11Z 2 Mar 2018









Ships with Significant Sequences of Stuck Winds in March 2018

Ship	DirSpd	Num	THR	RMS	Ship	DirSpd	Num	THR	RMS
9HXC9	040022	6	7	10	V7UU4	170008	6	18	22
9V2782	320015	8	8	16	V7ZZ5	060030	7	6	15
9V3964	050024	21	26	23	VCBW	000000	16	15	12
9V9040	159005	6	5	13	VRDU9	050016	7	22	12
A8PQ7	270008	6	5	20	VRRB5	050024	8	11	21
C6SI6	000000	7	6	12	VRRB5	060037	11	10	14
C6VG7	050021	7	9	11	VRRB5	090037	9	8	35
C6YA5	120015	5	5	14	VRRB5	160019	11	10	15
CFN3031	000000	18	17	20	VRRB5	160034	8	12	29
CFN3031	000000	25	25	22	VRRB5	170019	6	5	11
CFN3031	000000	37	36	17	VRRB5	220012	10	9	10
PCHM	230018	5	5	11	VRRI4	050020	12	18	10

THR is the time in hours of the stuck data
RMS is the RMS OB-BG wind differences in knots
Spd is the OB speed in knots
Dir is the OB wind direction in degrees

EMC codes delete calm winds if the guess speed is at least 5 m/sec,
but note this stage of data only has the low resolution non-time interpolated prepbufr guess

Counts of Possible Stuck Data Combinations 00Z 1 Apr to 06Z 24 Apr 2018
Ps is surface pressure T is temperature Td is dew point temperature
Winds is for both direction and speed

Count	Ps	Count	PsTTd*	Count	TTd	Count	Winds
100574	1	115670	1	82576	1	99774	1
6973	2	1616	2	2688	2	4405	2
1137	3	387	3	582	3	616	3
367	4	217	4	283	4	237	4
147	5	98	5	129	5	89	5
76	6	47	6	63	6	60	6
54	7	38	7	46	7	32	7
31	8	23	8	36	8	26	8
13	9	10	9	16	9	9	9
16	10	1	10	6	10	6	10
12	11	4	11	7	11	6	11
11	12	5	12	6	12	7	12
8	13	5	13	7	13	6	13
8	14	2	14	4	14	3	14
3	15	1	32	3	15	1	15
1	17	1	33	1	16	...	...
6	18	1	36	2	20	1	197

Here **76** **6** **means 76 times there were 6 repeats of Ps in a row**

*The counts here excluded missing values

Suggestion for Automated Ship Reports

- **There is a company, MarineTraffic, that tracks locations of large numbers of commercial ships in part to avoid costly collisions, and has Automatic Identification System (AIS) data**
- **Ilias Rigopoulus of MarineTraffic shared some AIS data, but could not share reports using satellite transmission**
- **The next two graphics show counts of AIS location reports in 1x1 degree boxes for select areas for 0900 to 1500 UTC on 9 May 2018**
  - **Green dots are 1-9 reports, blue 10 to 99 and red 100 or more**
- **The plot in the Europe area had almost 1.5 million reports from 35,882 different ships, with lower counts in the Atlantic graphic**
- **There were no reports in many parts of the Atlantic far away from land due to regulations not allowing the sharing of reports by satellite**
- **For the same time period and area as the AIS Atlantic graphic, the third following graphic shows ship report locations from the GTS**
- **Since GTS ship reports have low counts and significant location errors as well as data typos, would it be possible and worthwhile to have automated ship weather reports produced?**

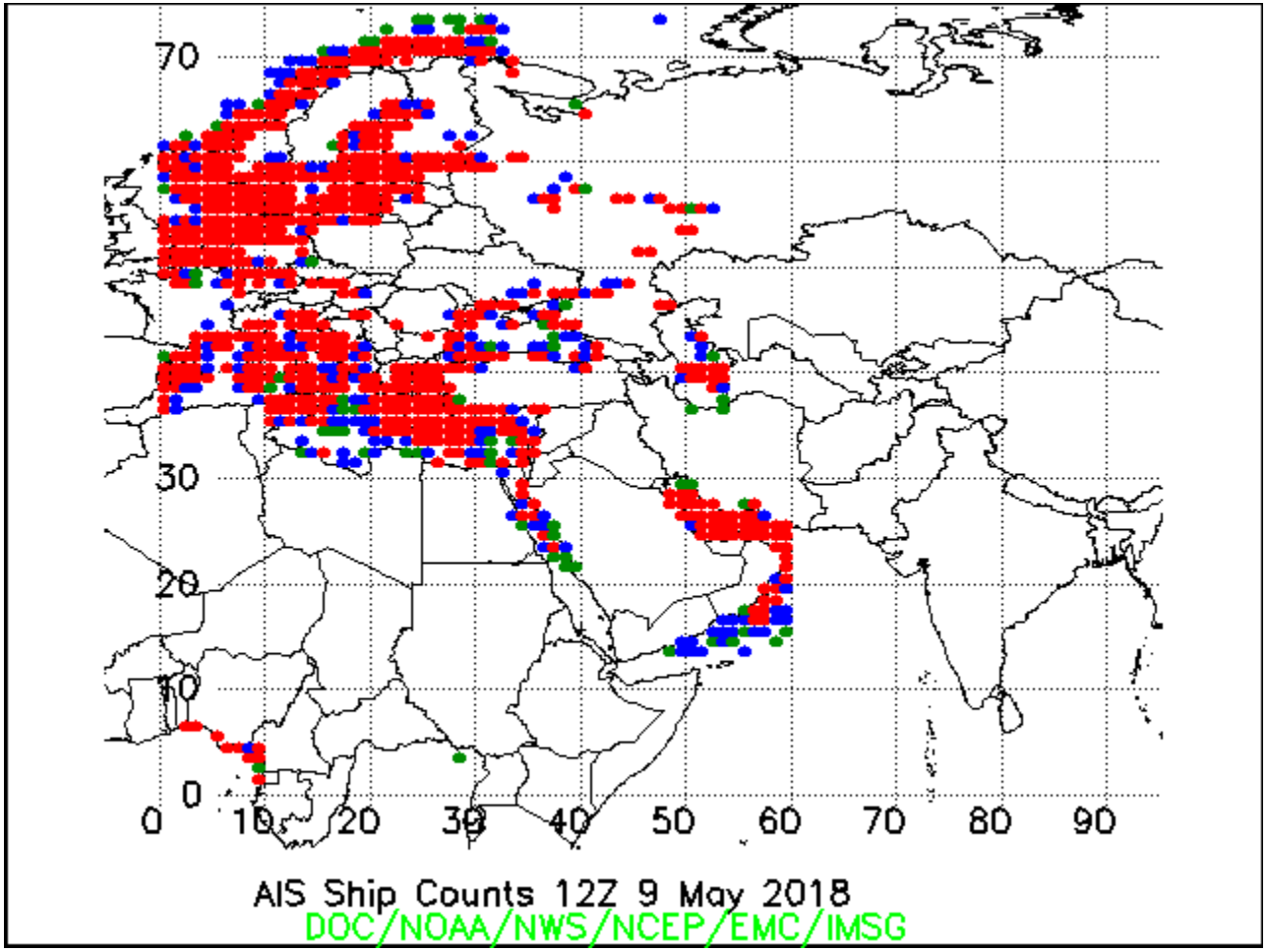
Suggestion for Automated Ship Reports Continued

- **For April 2018 ship data in NCEP GDAS runs, for ships moving from point A to B that had high travel speeds**
  - **For point A to B distances of 50 miles or more, only .6% of the data were rejected**
  - **For distances of 10 miles or more, the rejection was 1%**
  - **The total number of location errors may be higher as some errors do not result in high travel speeds**
- **Due to the lack of precision in ship surface pressure, temperature, dew points, wind directions and speeds, it is difficult to estimate how many reports are spuriously stuck or constant in time**
- **Errors in location or data typos can have significant impact in data sparse areas as shown in the Atlantic GTS ship data slide**
- **Automated reports would help decrease the typos as well as possibly have more precision in the data as well as more frequent reports in time**
- **It would be useful if the reports could give the height above sea level of the ship's barometer since this can vary depending on the ship's weight with cargo or passengers**
- **It would be good to have the ship's planned trip from port A to B if relevant, which may require bufr or other format upgrades**

New EUCAWS Automated Ship Reports

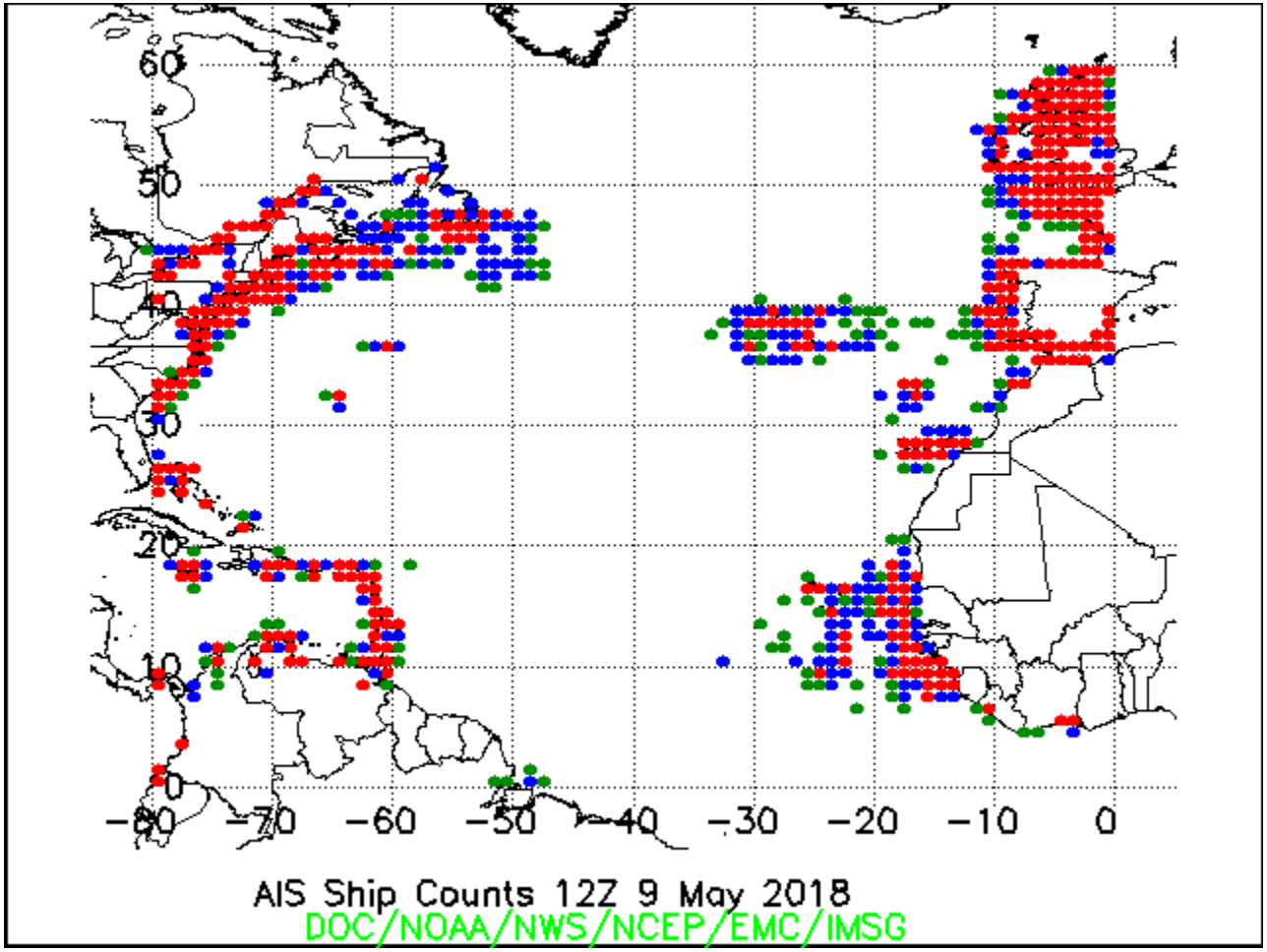
- **There has been a new project for automated ship weather reports, the Shipborne Europe Common Automated Weather Station (EUCAWS) with an EUMETNET report available at**
  - **https://www.wmo.int/pages/prog/www/IMOP/publications/IOM-125_TECO_2016/Session_2/O2(5)_Cohuet_The%20Shipborne%20European%20Common%20Automatic%20Weather%20Station%20(EUCAWS).pdf, Cohuet, J.,B., and Coauthors 2016**
  - **For OB-BG stats see** http://esurfmar.meteo.fr/qctools/last-report-list_surfmar_vos.html
- **These new data need more study over time, but the initial first look at the data is very encouraging**
  - **For 19 June 2018 there were 567 reports from 27 different ships with such data**
  - **The ship IDs ranged from EUCDE03 to EUCFR07 – Countries FR and DE?**
  - **They are said to report pressure reduced to sea level in tenths millibars**
  - **Their winds directions are to the nearest 5 degrees compared to the old 10**
  - **Locations are given to two decimal points compared to the old one decimal**
  - **Temperature and dew points have one decimal place**
  - **So far all reports are one hour apart in time, but that can be adjusted**
  - **The data receipt times at NCEP are very timely, just about 12 minutes after the observation time!**
- **Four slides ahead is a graphic showing the locations of these automated ship reports for 19 June 2018 processed in bufr at NCEP**

This Plot Shows Counts of AIS Location Data for 12Z 9 May 2018 on 1x1 degree boxes from 0 to 80 N and 0 to 60 E



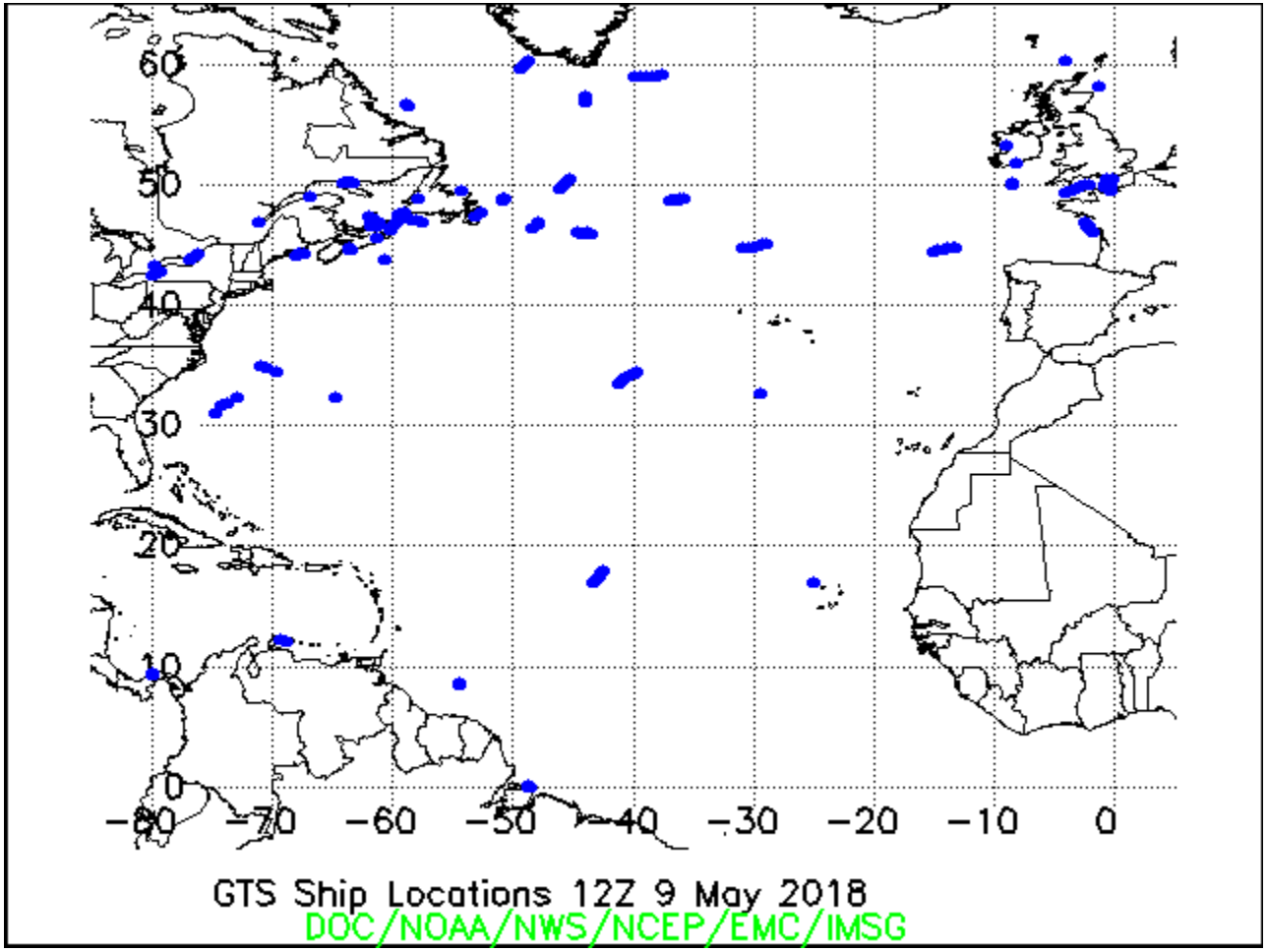
The green dots are 1-9 reports, blue 10 to 99 and red 100 or more

This Plot Shows Counts of AIS Location Data for 12Z 9 May 2018 on 1x1 degree boxes from 0 to 60 N and 0 to 80 W

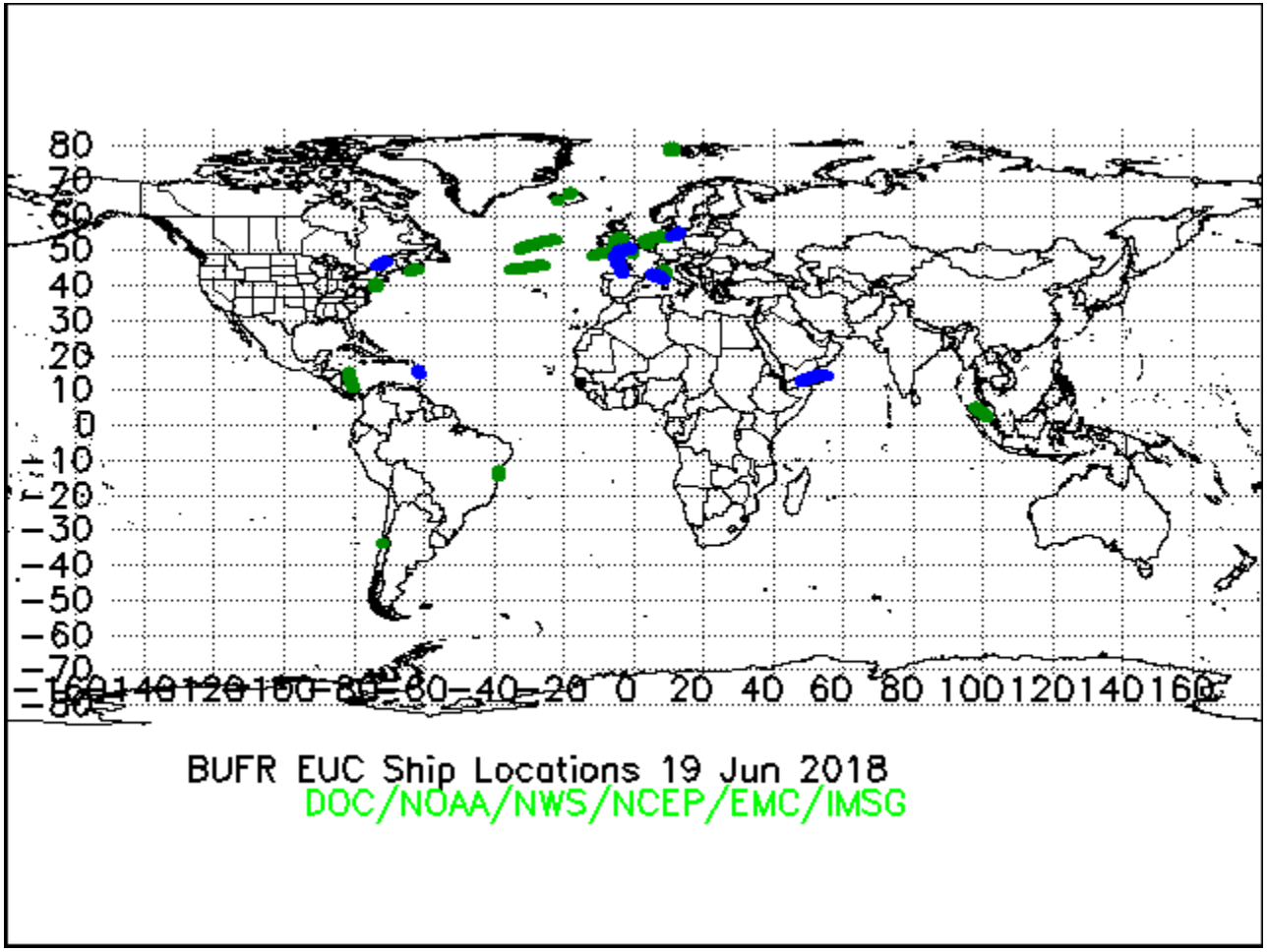


The green dots are 1-9 reports, blue 10 to 99 and red 100 or more.
The numbers and locations of ships reporting by satellite are not available

This Plot Shows Locations of GTS Ship Reports for 12Z 9 May 2018



This Plot Shows Locations of EUCAWS Ship Reports for 19 June 2018



Green dots are reports with missing winds, blue with winds

Acknowledgements

- **EMC's Chris Hill provided prepbufr files with high resolution time interpolated background and analysis values**
- **Thanks to EMC's John Derber and Jim Purser for allowing me to pursue this work**
- **ECMWF's Bruce Ingleby provided a Fortran 90 code for the UKMET ship track-checking code that was compiled and tested. It showed the value of taking QC action for the combination of suspect heading changes along with suspect travel speeds**
- **NCEP's Joe Sienkiewicz informed me about the AIS ship data**
- **Special thanks to Ilias Rigopoulus of MarineTraffic who shared some AIS data with me, but could not share reports using satellite transmission**
- **Discussions with Jim Purser and Daryl Kleist helped to decide how to describe the MTQC algorithm**
- **Thanks to Krishna Kumar for help with gempak scripts**
- **Thanks to NCO's Christine Caruso Magee who many years ago said something like "You'll never be able to develop an automated QC to handle all the errors in the ship data", which is still true and helped me to plan the new QC to handle complex errors**

The Path Forward

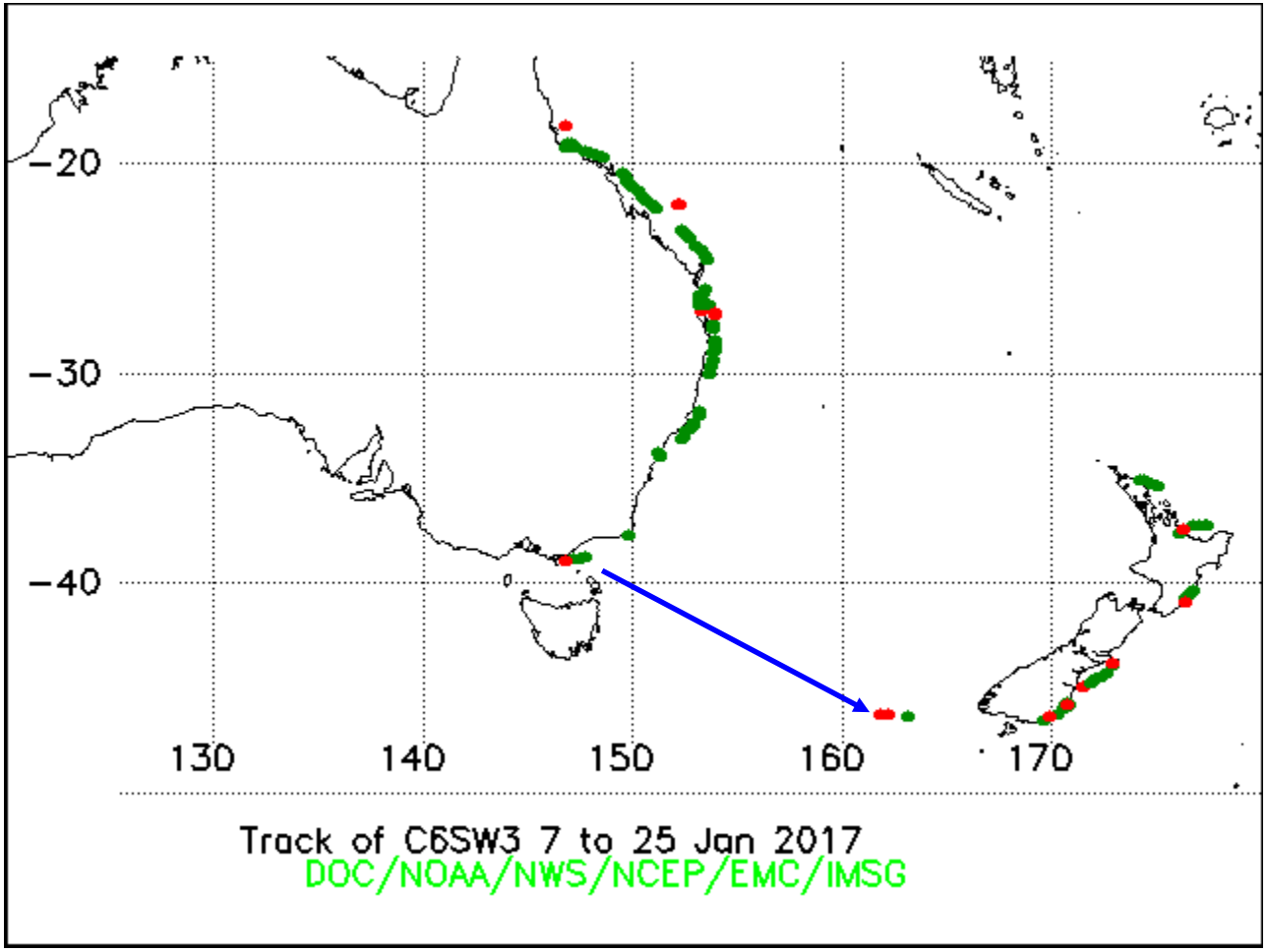
- **NCEP needs to implement the new MTQC when it is ready**
- **NCEP needs more effort on managing marine sites on its reject-list, but that takes more resources**
- **More and better feedback is needed to data suppliers**
  - **Fast data quality alerts would work well with buoy data**
- **For stuck data, what is the best way to judge if the data repeats are just luck or errors?**
- **If AIS data could include automated ship weather reports, that would make the data both more reliable as well as increasing data counts and coverage**
- **Automated ship reports from the EUCAWS or similar programs would also be useful**
- **Surface pressure bias correction of marine surface reports should be done on data passing marine QC**
- **Work is also needed for what to do with marine reports near land that can be affected by small scale orography and friction**
- **Work is needed to perform better marine QC for data use in things like reanalysis, which could have better QC using longer time periods of data including future reports**
- **Mark Ignaszewski of FNMOC may do some QC comparison tests**

Back Ground Additional Slides Follow

Ship Example of a Group Track-check Error as Well as Many Correct Zigzags

- **The next slide shows a group of reports from ship C6SW3 along the east coast of Australia that jump in 7 hours to near New Zealand moving at an impossible 130 mph**
  - **The slide suggests this may be a cruise ship stopping at different places for passengers to get on and off at along both coast lines**
  - **Many nearby observations intercompared before and after this big jump in position will show impossible high speeds**
  - **This may simply be a serious time error for many observations rather than the locations being wrong**
  - **It is difficult to compare the data quality of the two groups because the models do not capture well the real circulations near coasts and because this ship's data had many observations that were stuck (not changing with time)**
- **Some of the small scale speed violations could be due to the ship times being in hours with only one decimal place for locations**
- **With groups of reports that are self consistent but have impossible speeds needed to move from one group to the other are not handled well with the current QC codes**
- **This example had many changes in heading, which appear to be correct (ships do not travel as straight as aircraft)**

The blue arrow shows a jump in position of 911 miles in 7 hours at an impossible speed of 130 mph



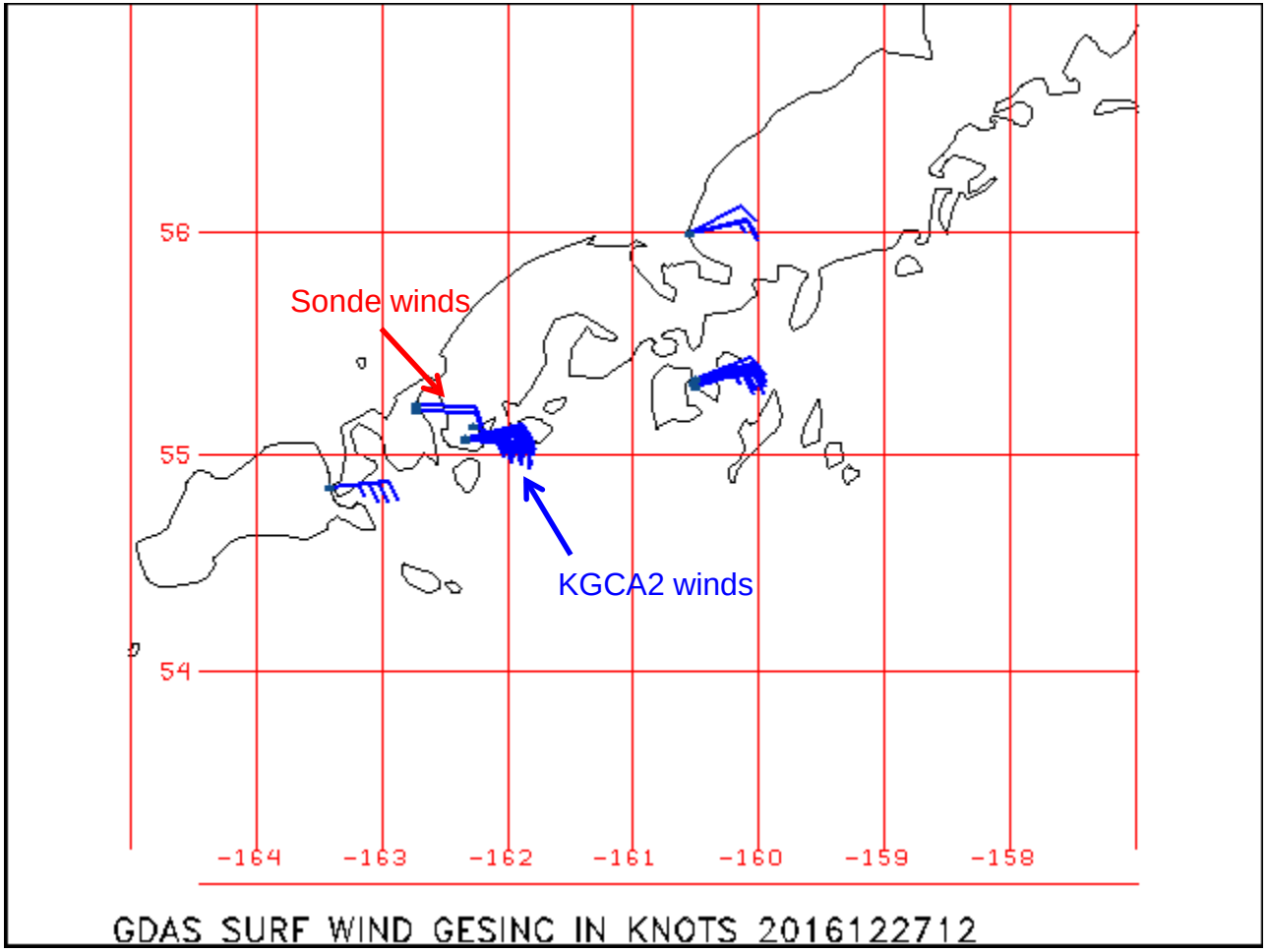
The red dots are deleted reports, while the green passed the new QC

There are many small scale errors getting deleted – should they?

Buoy KGCA2 Near Land was Having Negative Impact on the Analyses

- **In December 2016 buoy KGCA2 was over water near complex orography near Cold Bay Alaska resulting in cases with large impact on the analyses**
- **Note that some global analysis systems still do not use land surface winds as they can often have small scale features the models can't resolve**
- **Some marine reports near land can have similar problems**
- **The KGCA2 winds were probably correct, but when the wind direction was in a certain range, the model background was poor**
- **The analysis would then make fairly large changes that are probably wrong due to small scale errors in the model**
- **Should such buoys be on our reject-list, or use some smarter rules for when the data are useful to the large scale analysis?**
- **The next two graphics show large differences in nearby OBs as well as in the OB-BG for one case around this buoy**
- **This was a strong case, but there are many other similar but weaker cases where there may be negative impact from good marine reports close to land**

Select surface wind (observations – background) in knots
12Z 27 December 2016 Case

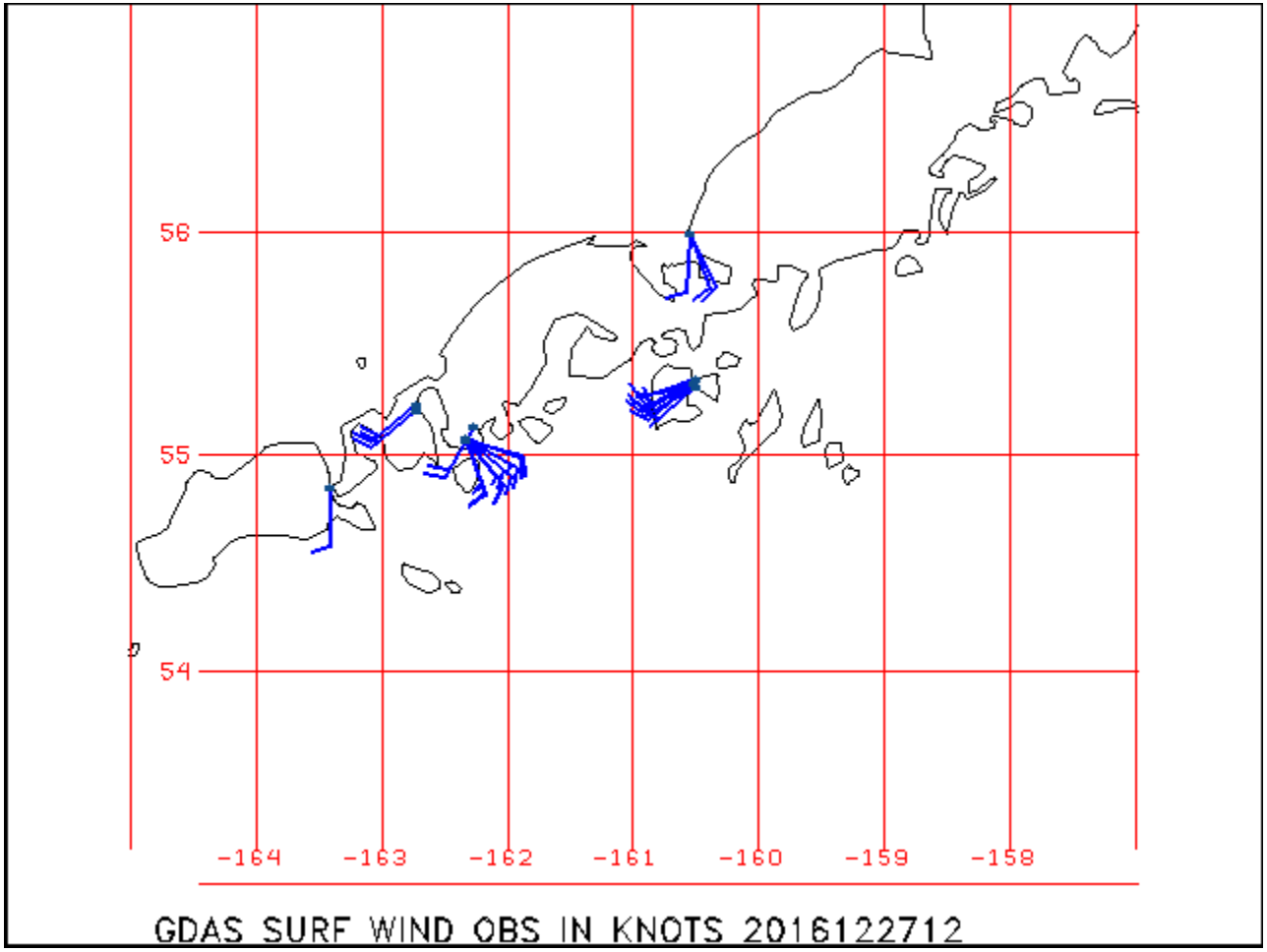


Note OB-BG varies much over small distances

KGCA2 looks over land in gempak graphic

Some of the Cold Bay sonde winds were deleted by the VARQC even though they agreed more with the background

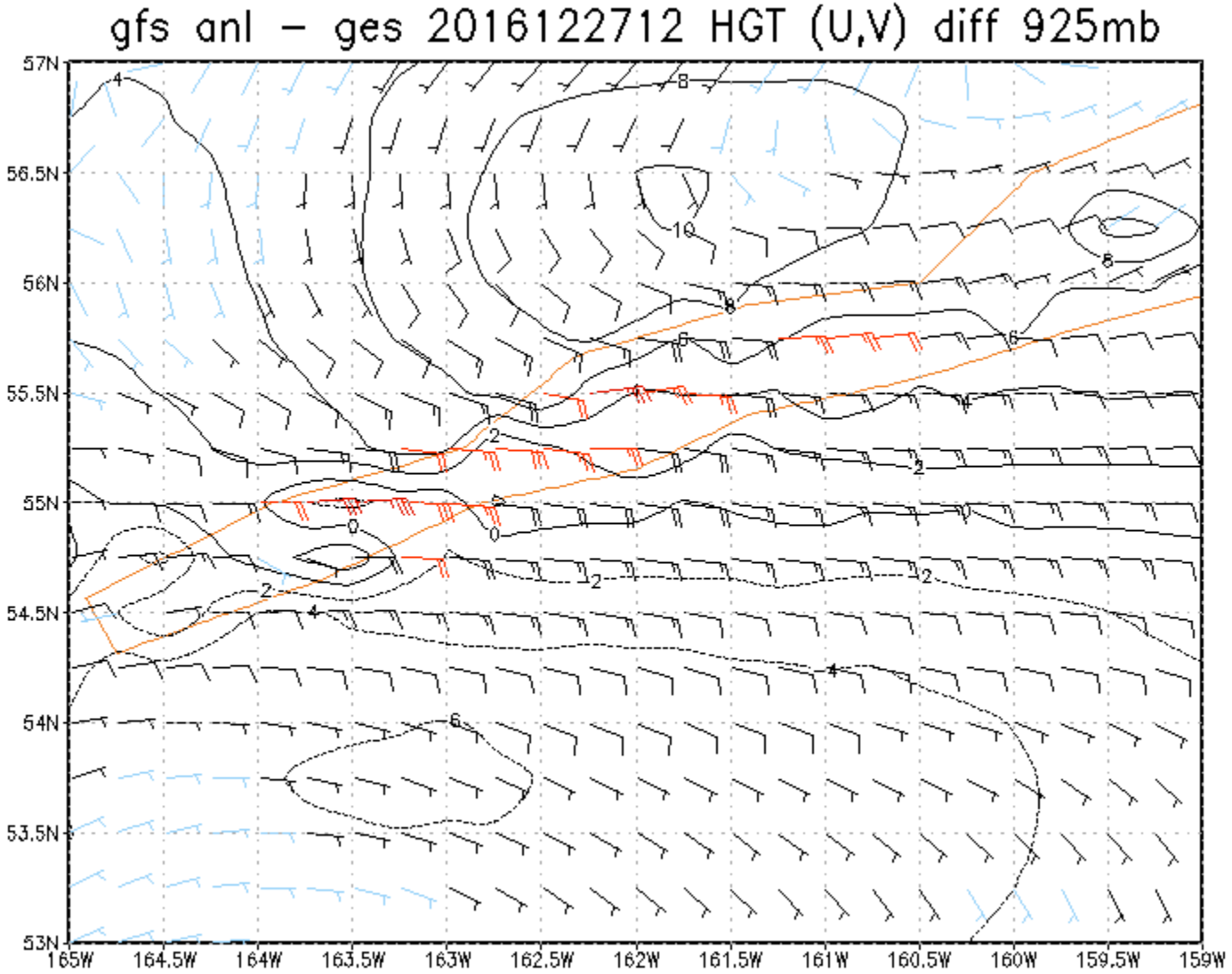
Select surface wind observations in knots
12Z 27 December 2016 Case



OB winds are changing much over small scales

Notice for small changes in location, the winds change considerably

This plot is from the Master grib ANL Minus 6-hour Forecast from the 12Z GDAS on 27 December 2016 – Winds are in knots

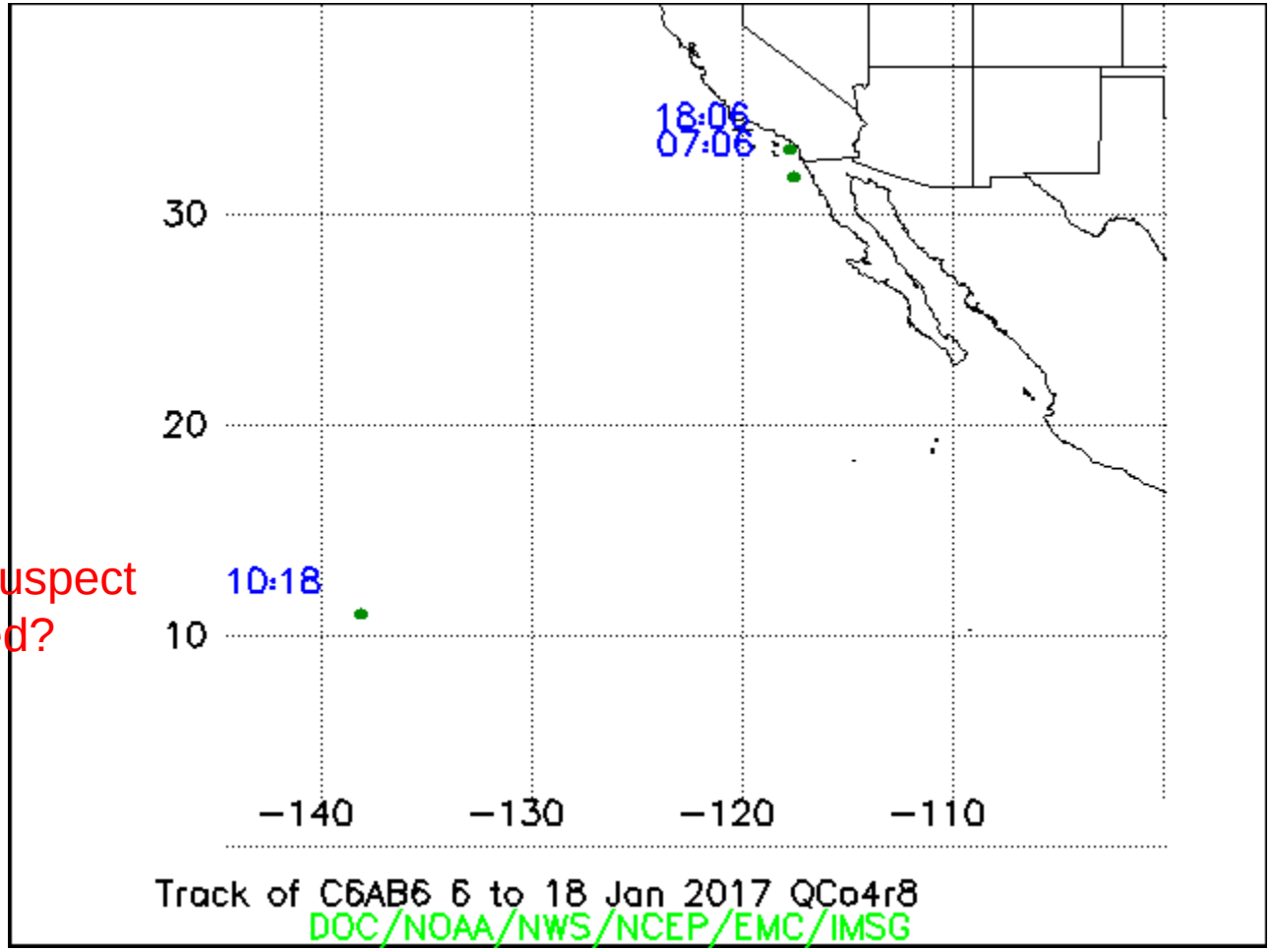


These large analysis – background changes are probably hurting the forecast and are due to mostly to small scale data features

An Example of a Ship with Passing Location Errors?

- **The next slide shows select reports from ship C6AB6 that are likely in a wrong locations, but had no high speed jump**
  - **The ship had been moving very slowly near the southern California coast, and then in 84 hours it jumped to around 11N and 138W, which was a distance of 1937 miles at a speed of 23.1 mph**
  - **There were 2 reports in this area, with bad fits to the model BG**
  - **These appear to be double typos in location such as 11 versus 31 for the latitude as well as a longitude typo**
  - **After the 2 bad reports, the locations jumped back to the southern California coast taking 171 hours and a speed of 11.6 mph**
- **When there are large differences in time between reports, it is difficult to be sure if locations are correct and we also then get few reports**
- **If we had automated reporting, there would not be big time gaps**

Due to large time differences, the OB to the southwest
Is possible, but is it correct and should we give it less weight?



Should this suspect OB be deleted?

Here 10:18 stands for 18Z on the 10th, but this style of plot can make other plots too busy

New QC Changes for Aircraft Data

- **The ground speeds of aircraft are QCed for both excessive speed and too slow speeds as a function of pressure**
- **Since many aircraft do not report time in seconds even though they can have multiple reports with the same time, the times are simply adjusted equally over a minute divided by the number with the same time – this needs improvement**
- **QC can be done to eliminate high vertical velocities (VV)**
- **The inter observational travel speed checks used in QC logic are better using ground speeds minus the most probable expected speed for that pressure, and need to include using air speeds**
- **There are useful diagnostic prints for each aircraft report before and after the NRLACQC and the new QC**
- **These include: difference in time and distance, ground speed, VV, heading and various symbols to show problems like excess ground speeds, which OBs failed QC etc.**
- **This makes it easy to check on cases where the QC did not succeed**
- **Additional new track-checking stats are made for OB to next OB tests to compare with the multi-inter observational QC tests**

Some Stats on MDCRS Track-checking Experiments 1-5 November 2017

New Variational QC experiments: H high speed checks, L low speed checks,
V vertical velocity (VV) checks, D minimum distance check 100 miles, d 0.1 mile
The high VV limits were a very crude 40 m/sec, and new limits will be studied

Symbols:
HS num of high speed, LS num low speed, HV num high VV, VD a large velocity vector dif, P prepbufr, N NRLACQC, Q new QC
For the NRLACQC, an OB is here considered rejected if TQM and WQM=13
So HLd means the new QC checked for High and Low ground speeds with small minimum distances between OBs
HSP num of High Speed in prepbufr – HVN num of High VV after NRLACQC, etc
There were 2,182,101 OBs total, 181,437 UPS

EXP	HSP	HSN	HSQ	LSP	LSN	LSQ	HVP	HVN	HVQ	VDP	VDN	VDQ
HD	159	10	5	39257	26406	39059	2705	1162	2570	5020	982	4300
HLD	159	10	5	39257	26406	5	2705	1162	2504	5020	982	3769
Hd	9099	6062	11	39257	26406	37008	2705	1162	780	5020	982	249
HLd	9099	6062	10	39257	26406	5	2705	1162	739	5020	982	238
V	9099	6062	6973	39257	26406	38891	2705	1162	64	5020	982	1772

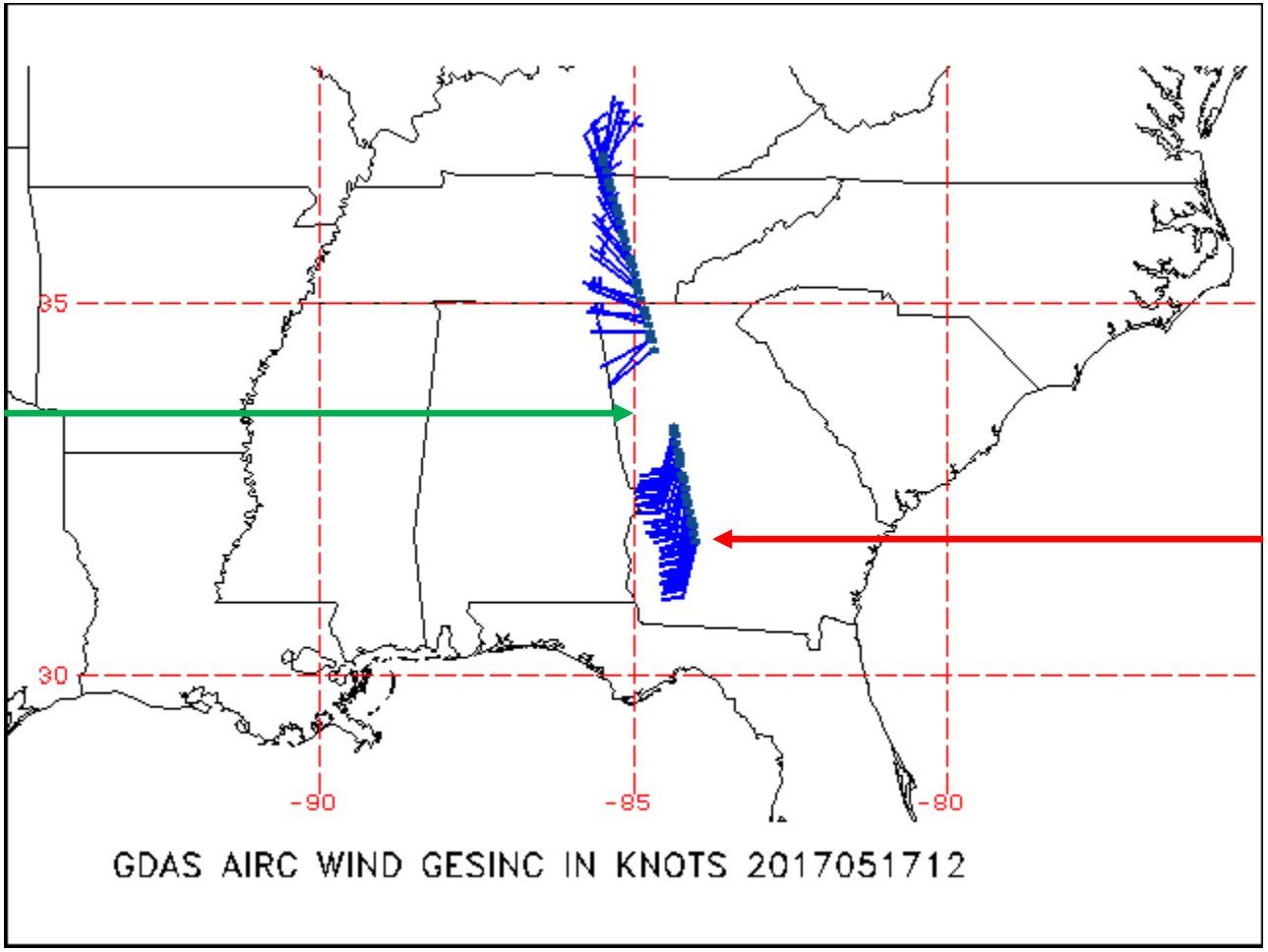
UPS Data Only

EXP	HSP	HSN	HSQ	LSP	LSN	LSQ	HVP	HVN	HVQ	VDP	VDN	VDQ
HLD	0	0	0	3154	1623	1	1006	54	1006	2107	207	2102
HLd	2243	1045	0	3154	1623	1	1006	54	9	2107	207	2

For large distance high speed errors the NRLACQC and new code are comparable

This example also found by QC differences to the NRLACQC has a group of Obs at consistent but wrong locations that mostly passed QC

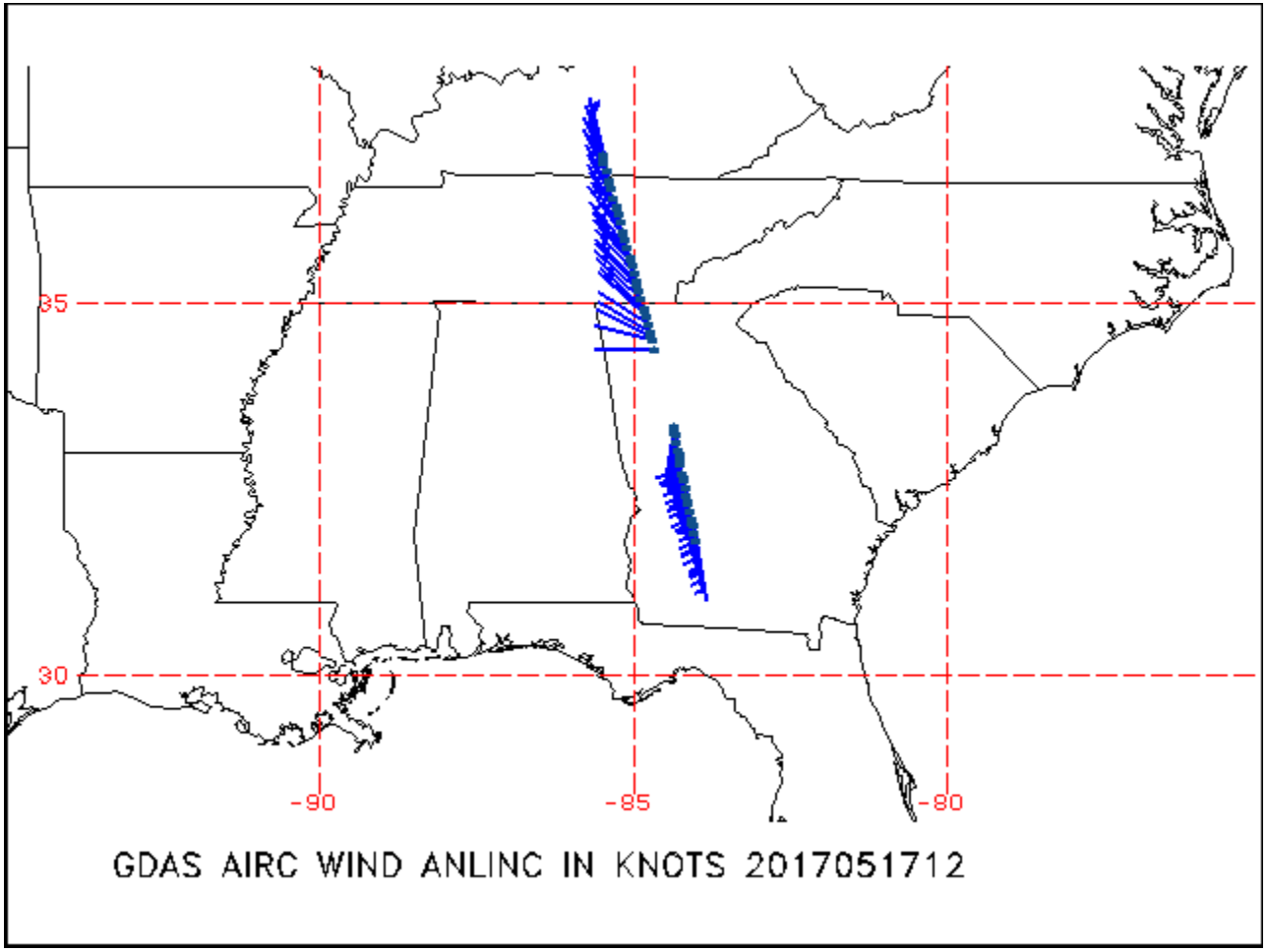
The gap in time was too big for the jump in location



This group of reports appears to be at wrong locations and have larger OB-BG, which is a difficult problem

This is from aircraft V3HUOFRA 12Z 17 May 2017

This slide shows some slightly bigger AN-BG in the wrong data in the more south group



This is also from aircraft V3HUOFRA 12Z 17 May 2017