Observation impact studies: EUMETNET and other

EWGLAM meeting, ECMWF, 2 Oct 2017

Bruce Ingleby and Lars Isaksen

ECMWF

Bruce.Ingleby@ecmwf.int

Outline

- Use of aircraft humidity data
  - Test with data from U.S. aircraft 2014
  - Latest O-B statistics
- Pressure from drifting buoys and ships
- Radiosonde experiments (not EUMETNET)
  - Treatment of radiosonde drift
  - Other: Russian 1 ascent/day, RS41 descent data
- Summary

- Tomas Kral helped, others credited in later slides
- Thanks to EUMETNET for supporting aircraft/buoy work

ECMWF Numerical Weather Prediction (NWP) system

- Background (B) – 12 hour forecast – compared with observations (O), they are combined to make the Analysis – start of next forecast.
- B and O have uncorrelated errors – very useful to look at O-B statistics
- ECMWF produce daily coverage maps and monthly monitoring statistics feedback to data producers – partly via EUMETNET

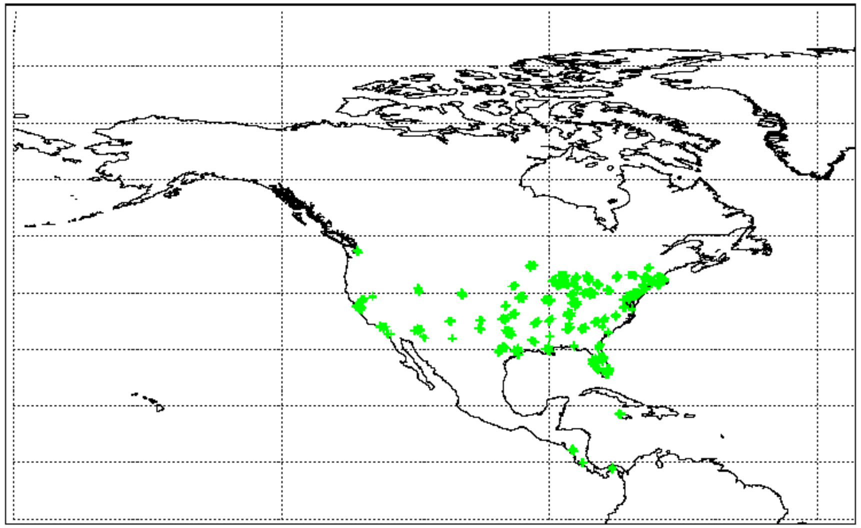
- Assessing usefulness of observations
- Data denial studies (Observing System Experiments or OSEs)
  - Rerun NWP system without certain subsets of observations
- Forecast Sensitivity to Observation Impact (FSOI)
  - Uses adjoint to estimate the contribution of each observation to reducing forecast error 24 hours later (relies on good analysis, linear approximation less good for near-surface variables, doesn't look at cumulative effect)
  - See eg Cardinali (2009, QJ), Lorenc and Marriott (2014,QJ)

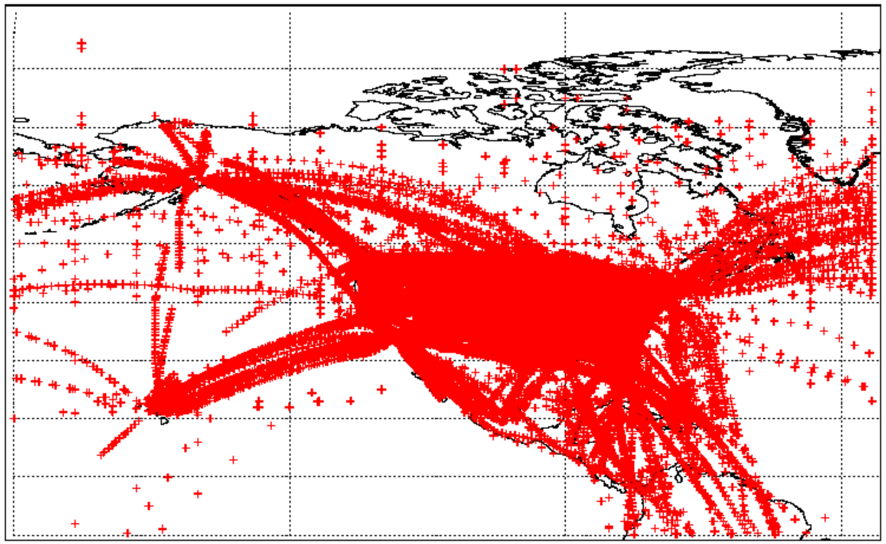
Extra aircraft data – especially humidity

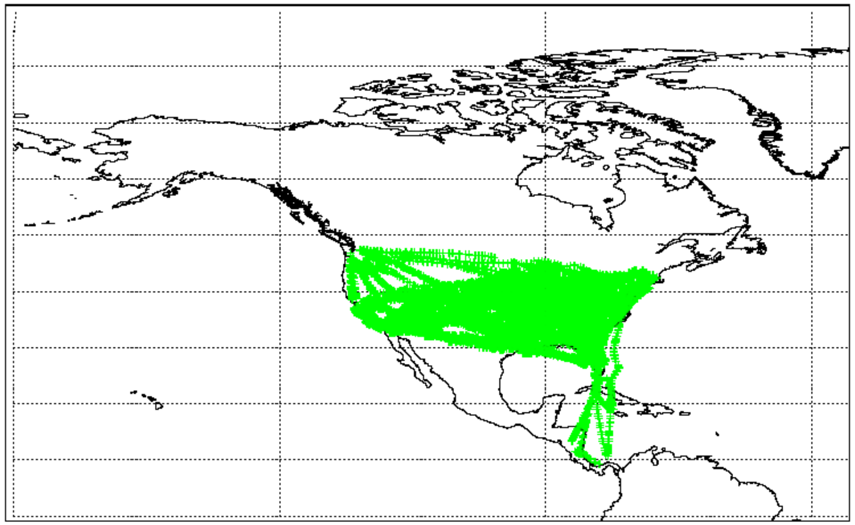
- AMDAR humidity measured using WVSS-II lidar sensor
  - Vance et al (AMT, 2014) – compared with research instruments on FAAM aircraft
  - Petersen et al (BAMS, 2016) – look at short range impact over USA
  - Mainly over North America (>100 US aircraft) also 9 German aircraft
  - Assimilated at ECMWF operationally from March 2016
- TAMDAR – measures wind and temperature too
  - Mainly short-haul flights
  - Capacitive humidity sensor perhaps less good at very low temperatures?
  - Panasonic (was Airdat) commercial data
  - Not used operationally at ECMWF, will concentrate on AMDAR data
- OSEs run for 2011 and 2014 using extra data over North America
- Results shown for 2014 – using ECMWF global NWP system
- A Cress ran experiments using DWD LAM relocated to N America

AMDAR coverage in 2014

- Over airports at low levels – spreads out at upper levels
- Humidity mainly over CONUS
- Temperature more widespread

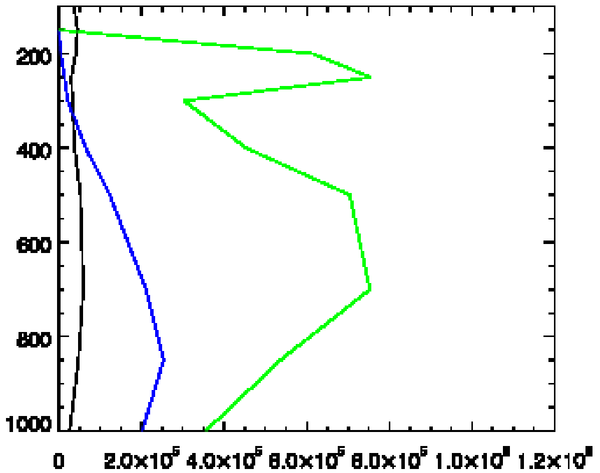

1 July 2014 AMDAR q availability Surf - 700 hPa


1 July 2014 AMDAR T availability 350 - 100 hPa


1 July 2014 AMDAR q availability 350 - 100 hPa

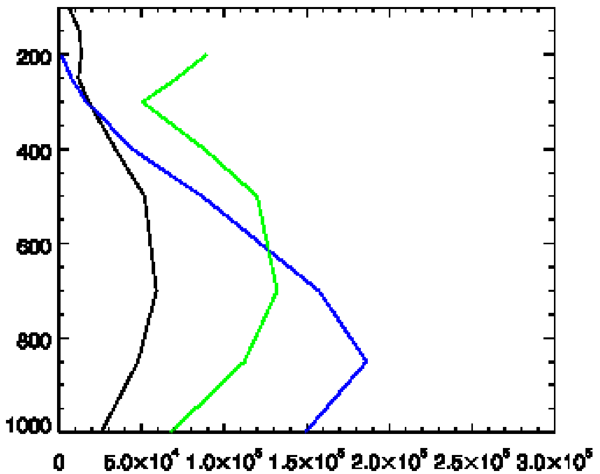
Number of reports (North America) vs pressure

Temperature



- TEMP
- AMDAR
- TAMDAR

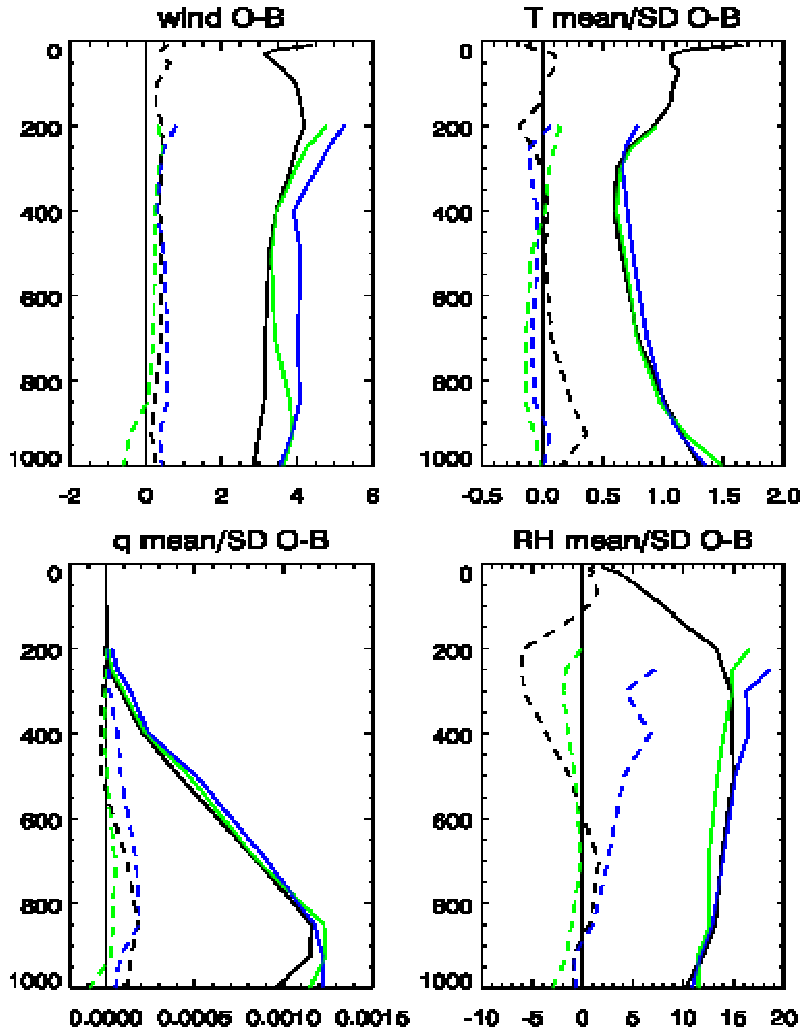
Humidity



~11% of AMDAR (ACARS) reports contained humidity (excluding zero values!)

Fit to forecast (North America): April-June 2014

- Radiosonde
  - Standard levels
- AMDAR (ACARS)
  - Good O-B
  - Esp. humidity
- TAMDAR
  - RH biases at upper levels (small sample)

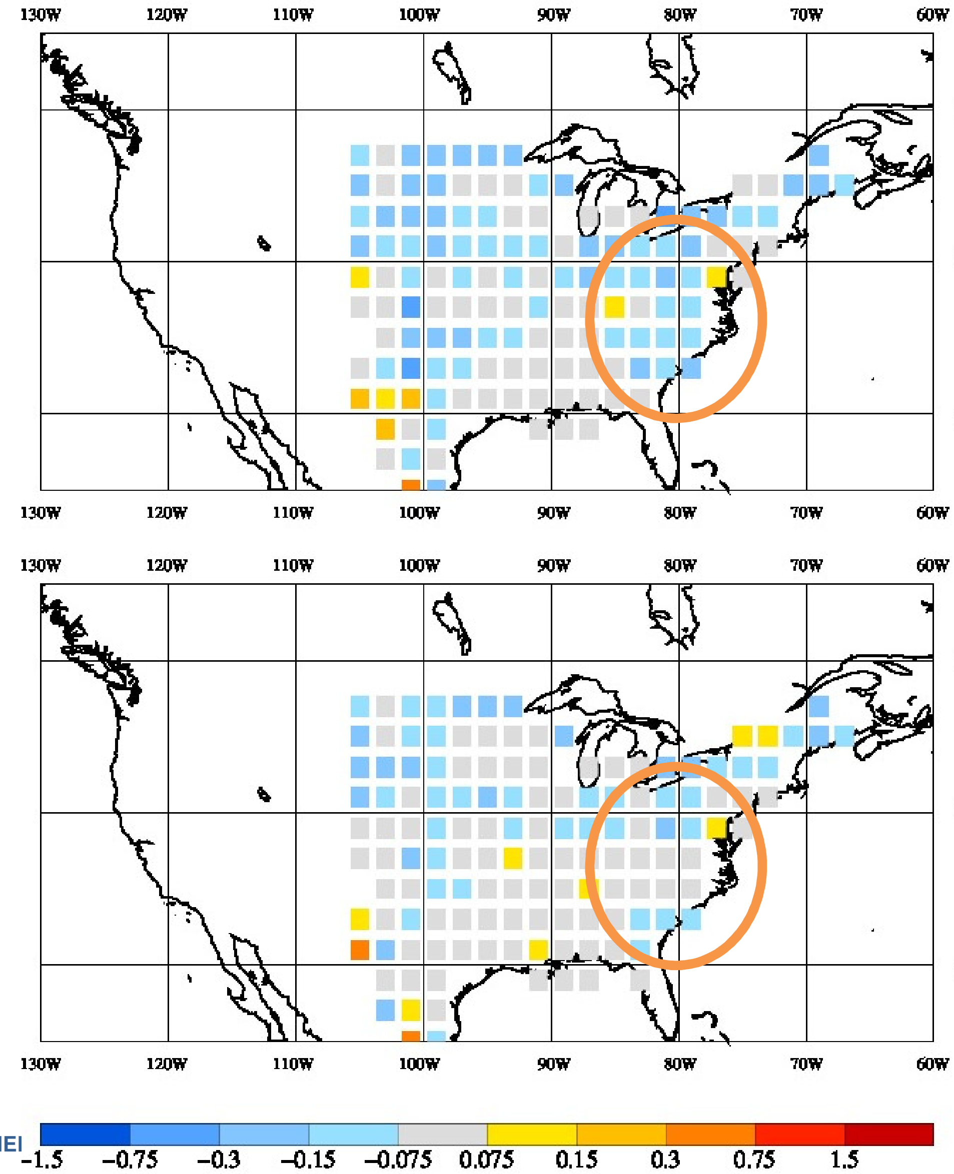


Experiments run

- March – May 2011
- December 2013 – February 2014
- April – June 2014 (comparison with radar precipitation)
- Experiments
  - Control
  - Amdar Humidity
  - TAMDAR*
  - Amdar Humidity (AH) + TAMDAR*
  - Tests with reduced humidity σ_o – period 3
  - **Problem with TAMDAR pressure altitude periods 2/3 rerun*

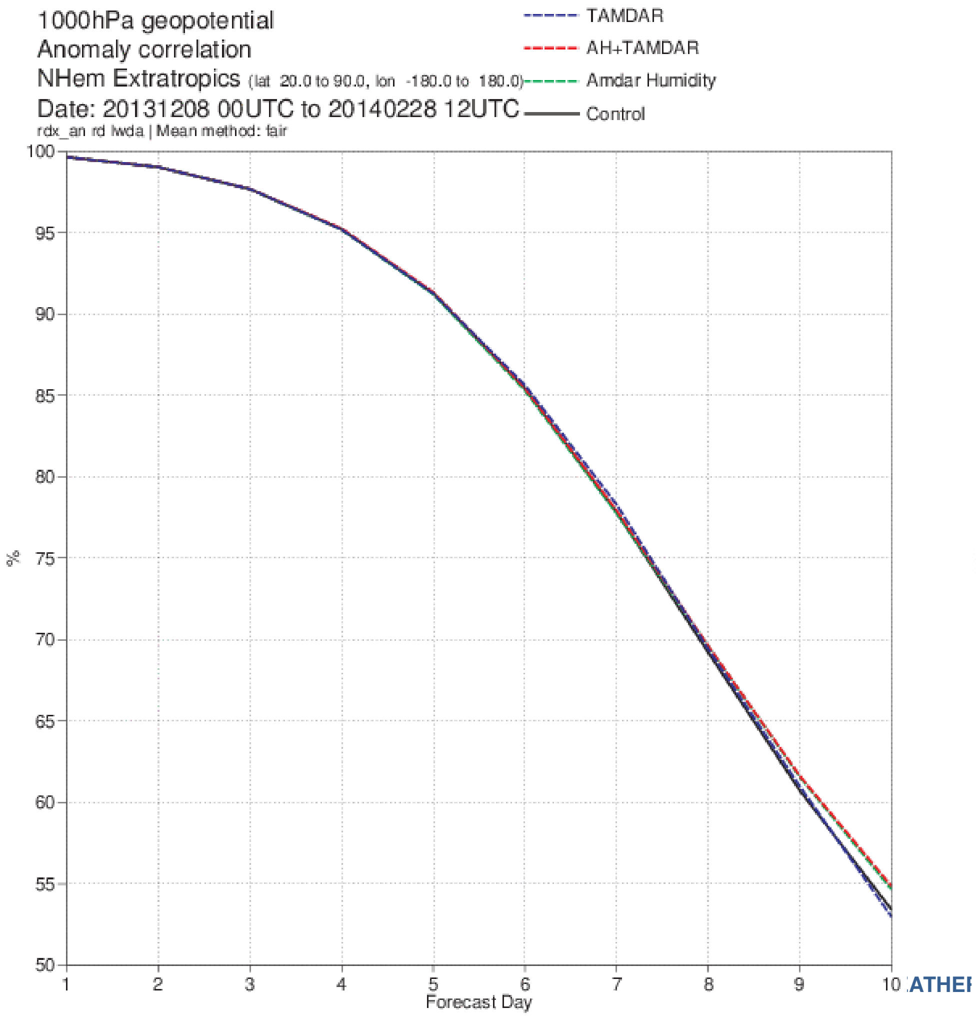
12 hour fit to radar

- Mean difference for control (top)
- AMDAR (below) gives improved mean fit to NEXRAD precipitation (AMJ 2014): -0.020 vs -0.030 ln(rain_rate + 1)
- Also slightly more NEXRAD reports pass quality control
- Plots and statistics: Philippe Lopez (ECMWF)

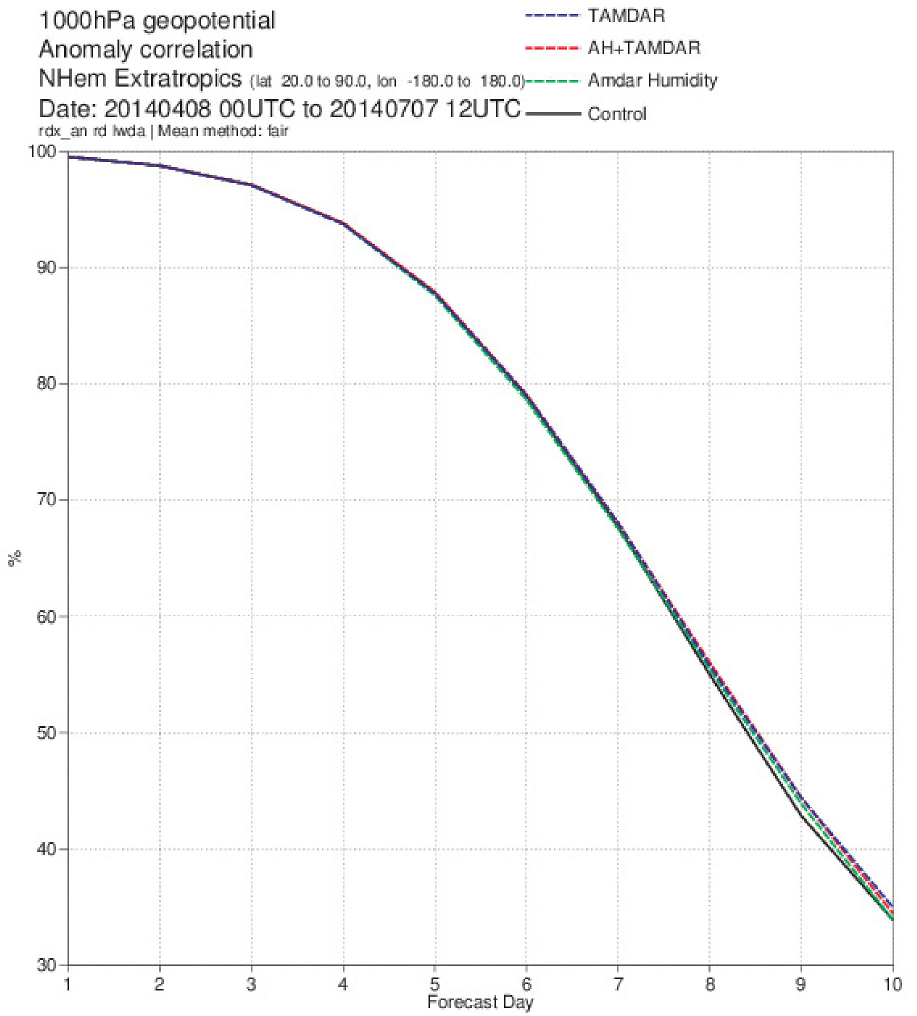


NHem Z1000 vs operational analyses

Dec 2013 – Feb 2014



April – June 2014

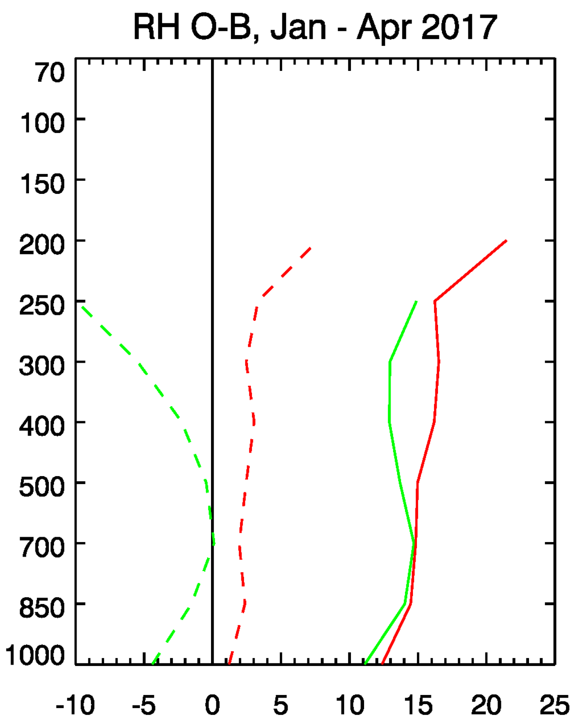


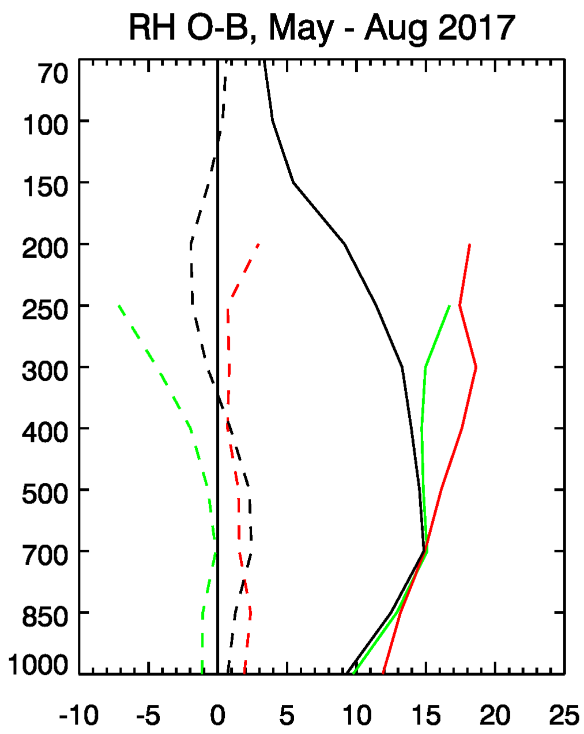
Summary of aircraft studies

- Aircraft data generally good quality (after quality control)
    - AMDAR similar to radiosonde quality, humidity better
    - TAMDAR marginally worse (esp. high levels, small samples)
    - More near-surface, ascent/descent TAMDAR reports
- Generally large scale impact is near neutral Impact vs surface temperature/cloud reports (not shown) noisy
- ECMWF monitoring has helped improve new E-AMDAR humidities
- Both AMDAR and TAMDAR gave improved 12 hour forecasts vs NEXRAD radar/composite 6 hour precipitation fields
- ECMWF used AMDAR humidity operationally from March 2016
- *Petersen et al (2016) very positive: "Automated aircraft water vapor reports, provided by 148 aircraft worldwide through the WMO's AMDAR program, are at least as accurate as rawinsonde observations and have greater influence on 1–2 day NWP forecasts than all other in-situ moisture data over the United States"*

2017 O-B statistics

- RMS OK, but notable biases for E-AMDAR esp. in colder months (250 sample small)
- Comparison with US-AMDAR and E-TEMP (black)





Buoy and ship impact

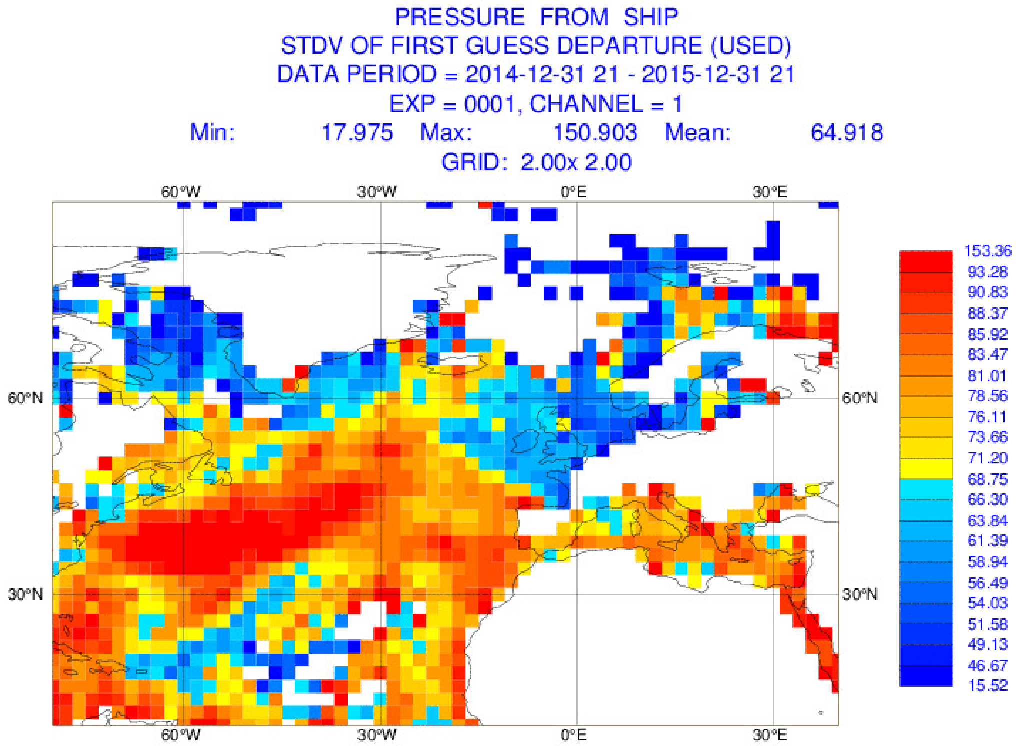
- Diagnostic studies (North Atlantic)
  - FSOI vs baroclinic development areas
- Recent drifter studies for WMO
- OSEs
  - Experiment design
  - Results
- Drifters without pressure

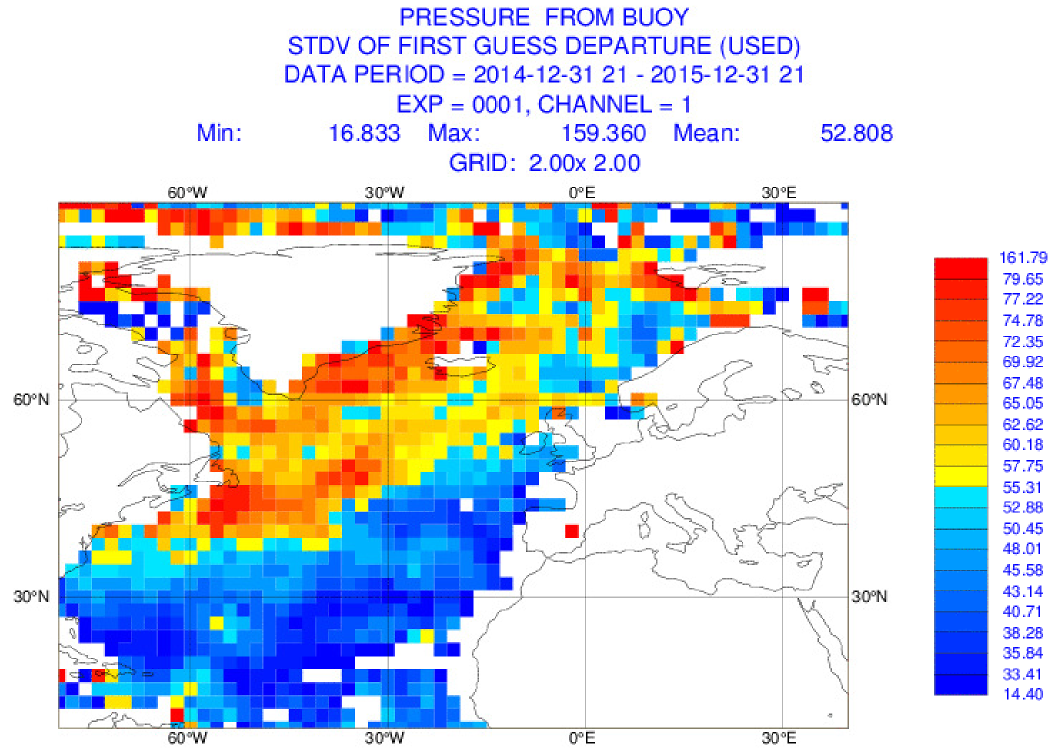
FSOI: Forecast sensitivity to Observation Influence

OSE: Observing System Experiments

Standard deviation of Pmsl O-B 2015

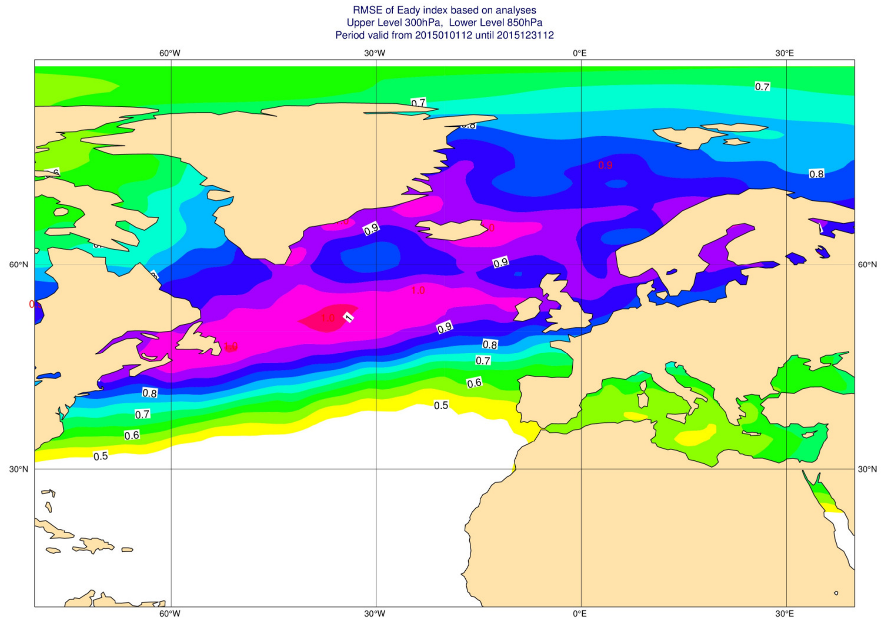
- Highest Standard deviation in ‘baroclinic area’, in Southern part ships worse than buoys
- Some large (ice-related?) Standard deviations in Northern part for buoys

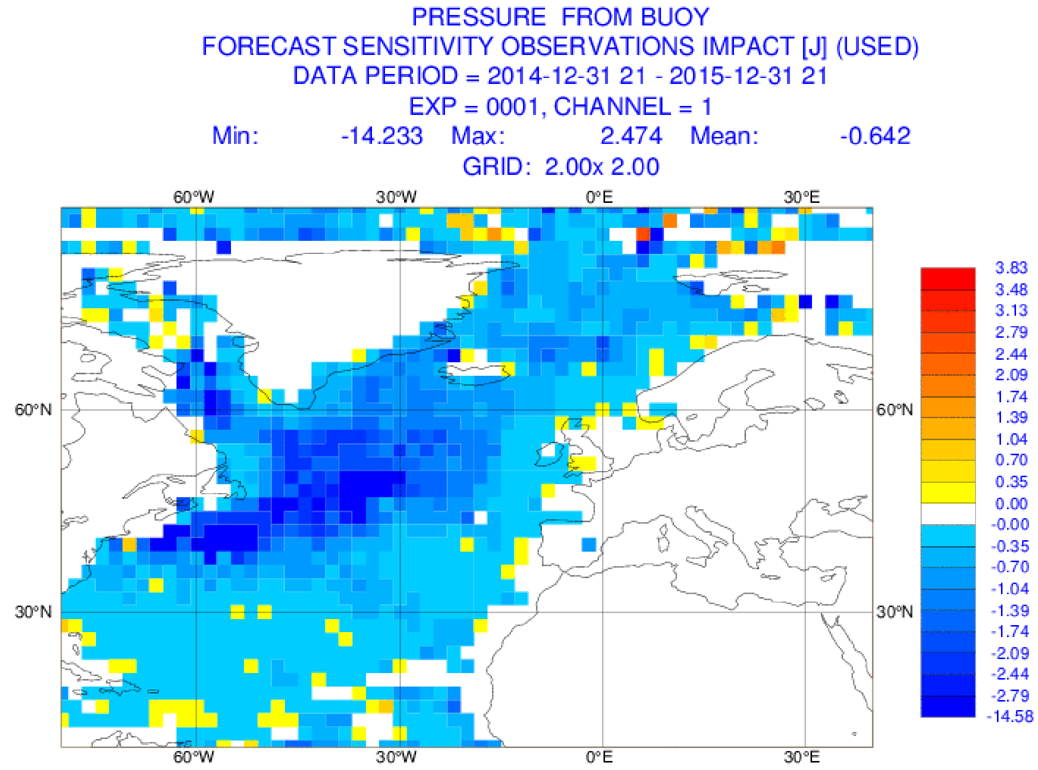




Buoy pressure FSOI vs Eady index, 2015

- Biggest impact over Gulf Stream – linked to SST gradient and development
- Ship FSOI (not shown) similar but noisier
- Eady index calculated between 850 and 300 hPa





Recent ECMWF drifter studies for WMO

- Centurioni et al (2017, BAMS)
- Horányi et al (2017, QJRMS)
- Global data denial studies – found significant impact throughout troposphere
- More impact, even in Northern Extratropics, in Jul-Aug 2012 than in Nov-Dec 2010
- Slightly surprising (expect larger impact in winter?): role of data density (more in N Atlantic, fewer in N Pacific)? or natural variability (yes).
- Nov-Dec 2010 part of high predictability period (Arctic Oscillation –ve)

OSEs for E-SURFMAR study

- Details agreed with E-SURFMAR manager (Paul Poli) in 2016
- Decided to just remove Northern Hemisphere (NH) data – moored buoys not removed
- Decided against "time-thinning" experiments – less realistic

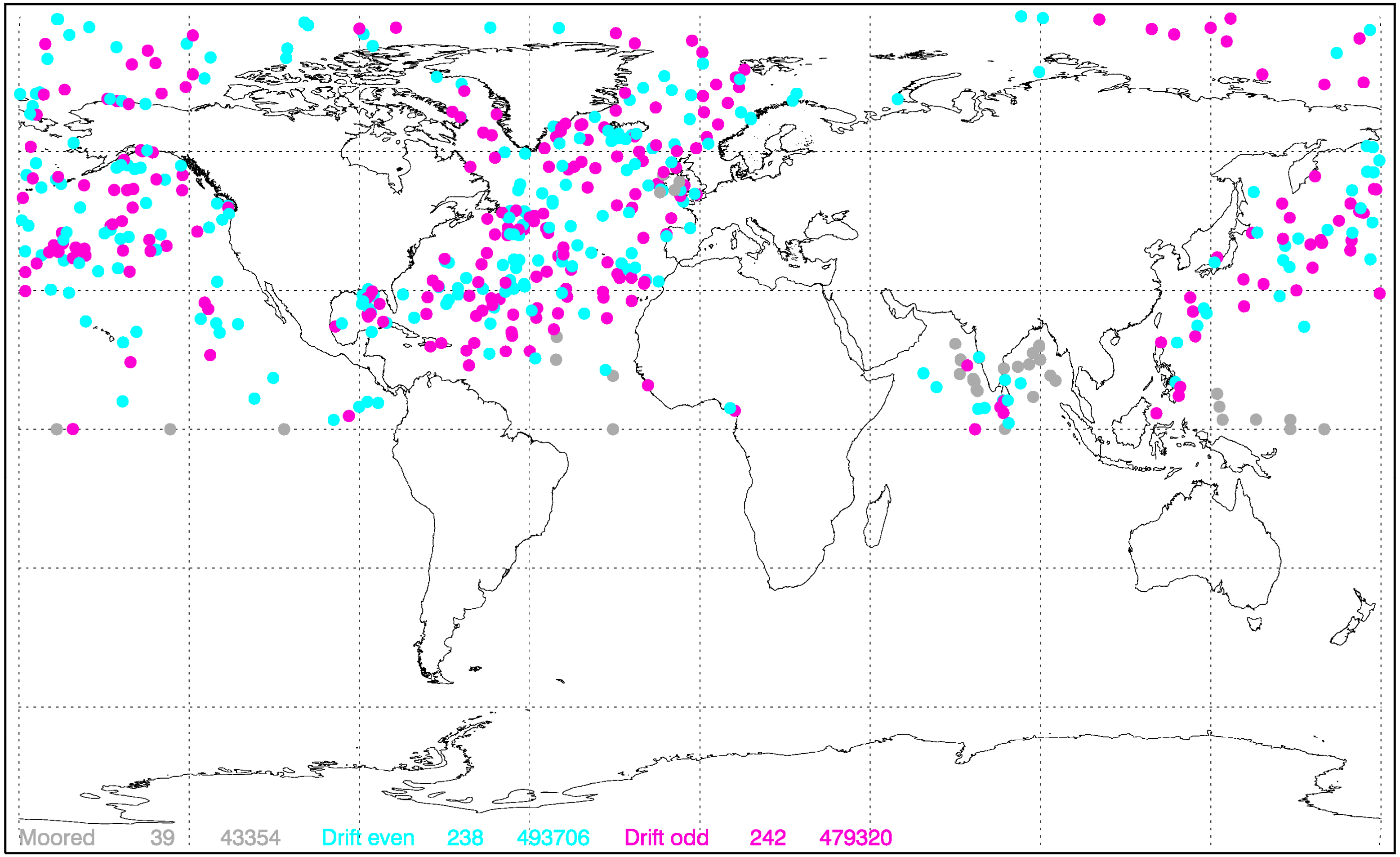
- Control (all data)
- NoNHsh – Remove ship data from NH
- HalfNHdr – Remove ~half drifter data from NH
- NoNHdr – Remove all drifter data from NH
- NoNHdrsh – Remove all drifter and all ship data from NH

- Trials run for November 2015 - February 2016, cy43r1, T399
- Alphanumeric buoy data used (not BUFR)

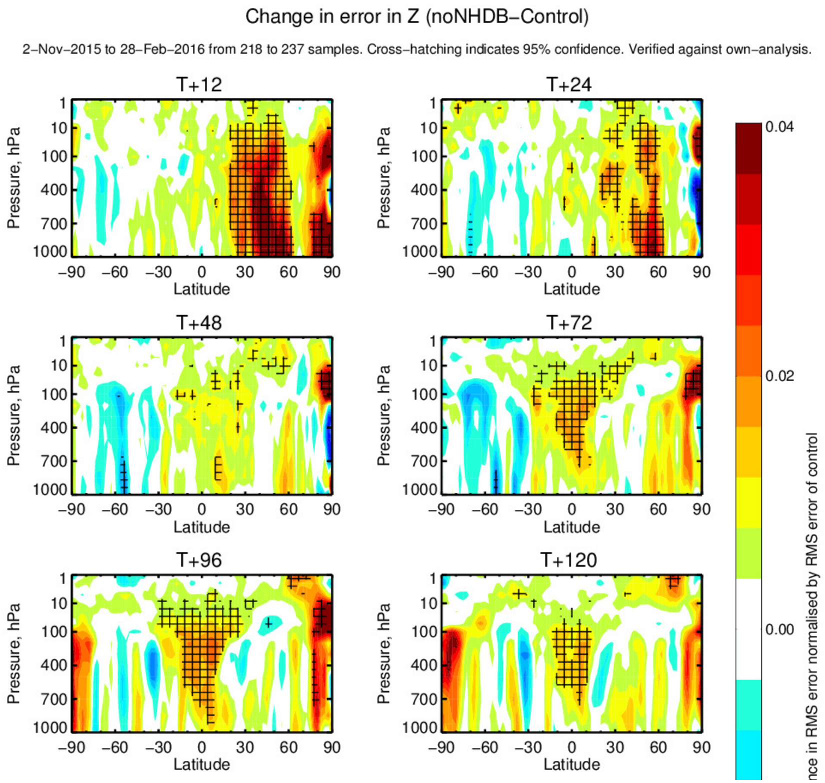
Selecting half the drifters

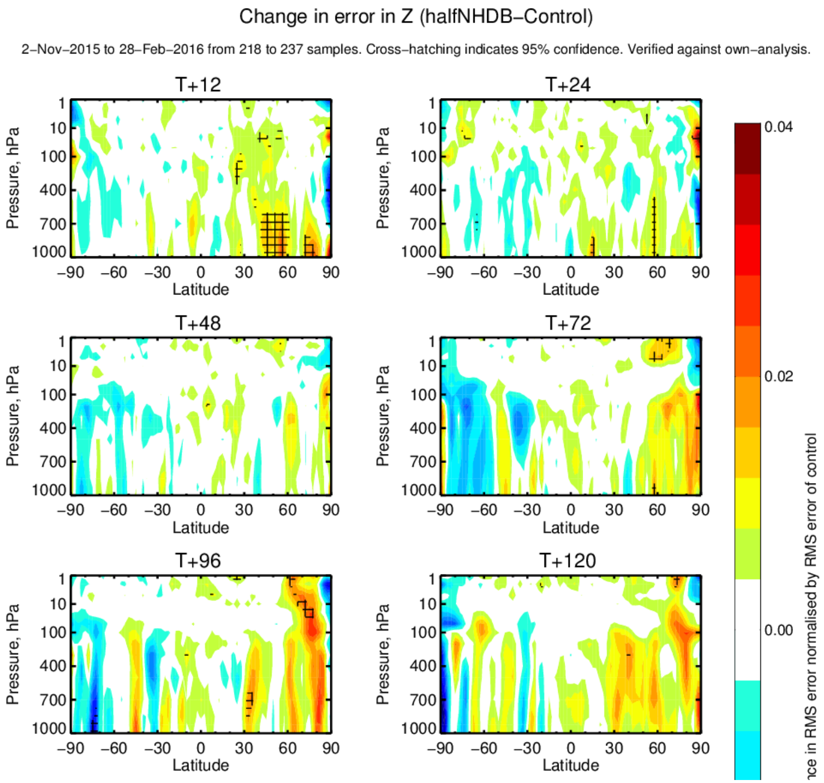
- Drifters with odd identifiers blacklisted (but still monitored)

Nov 2015 - Feb 2016 NH BUOYS with used pressure even/odd (last position)



Geopotential height (Z) verification vs own analyses

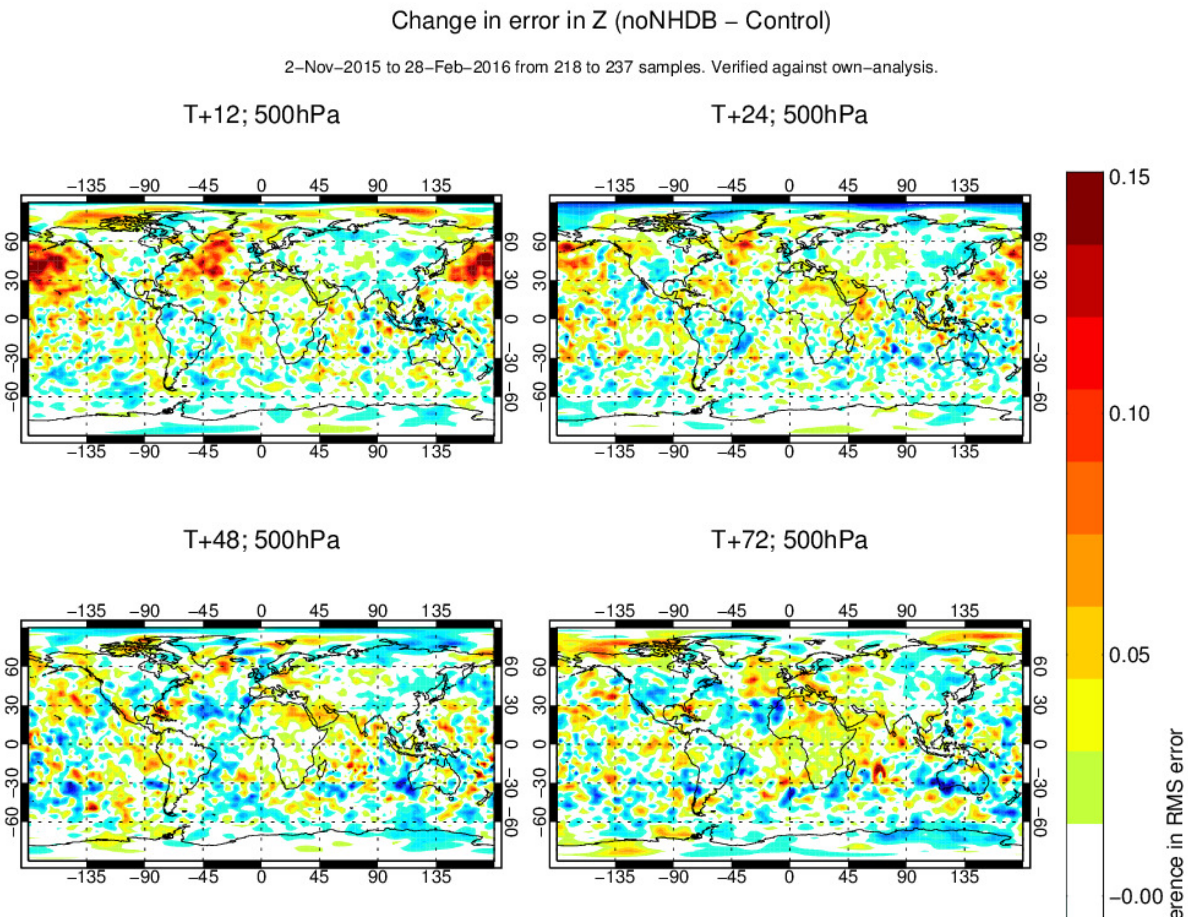




No drifters (left) has much larger impact than removing half of drifters (right).
Red colours – experiment worse, hatching statistically significant.

Spatial distribution

- At short range the impact is mainly over western North Atlantic/Pacific – then rapidly becomes less localised.



Results for Drifter rms(O-B), hPa ~T+12 verification

	20°-50°N	50°-90°N	EUCOS	N. Pacific
Number	540724	450719	469553	495993
Control	0.719	0.978	0.853	0.844
NoNHsh	0.721	0.978	0.854	0.845
HalfNHdr	0.737	1.002	0.868	0.870
NoNHdr	0.773	1.079	0.901	0.949
NoNHdrsh	0.777	1.080	0.904	0.952

EUCOS: 10°-90°N, 70°W-40°E; N. Pacific: 10°-90°N, 120°E-100°W
Larger errors at higher latitudes (50°-90°N) and data has more impact there.
Control: N Atlantic ~ N Pacific data, but data has more impact in N Pacific.

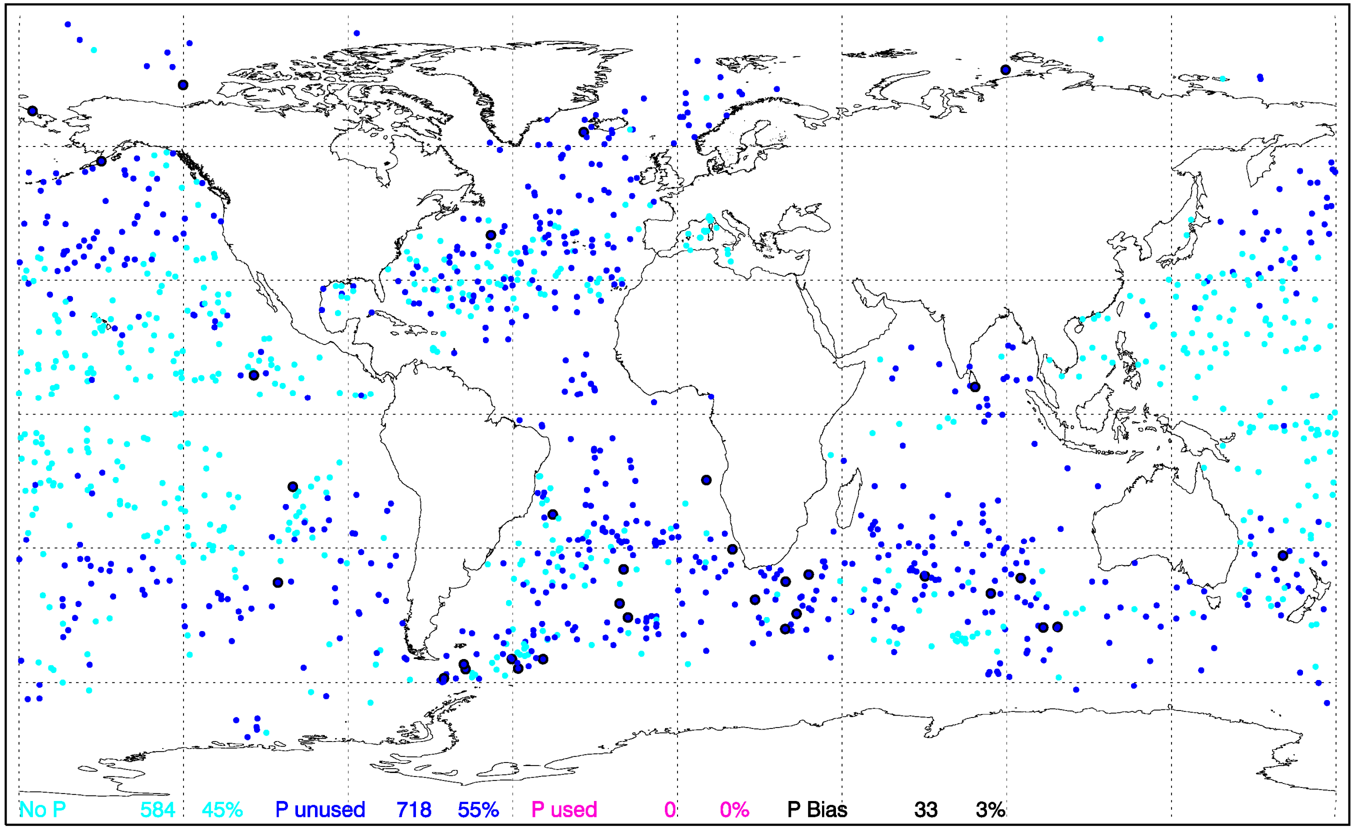
Results for Drifter rms(O-B), hPa ~T+12 verification

	20°-50°N	50°-90°N	EUCOS	N. Pacific	EUCOS%
Number	540724	450719	469553	495993	
Control	0.719	0.978	0.853	0.844	100.0
NoNHsh	0.721	0.978	0.854	0.845	99.4
HalfNHdr	0.737	1.002	0.868	0.870	69.9
NoNHdr	0.773	1.079	0.901	0.949	4.5
NoNHdrsh	0.777	1.080	0.904	0.952	0.0
All/None	*0.925*	*0.906*	*0.944*	*0.887*	

EUCOS: 10°-90°N, 70°W-40°E; N. Pacific: 10°-90°N, 120°E-100°W
Larger errors at higher latitudes (50°-90°N) and data has more impact there.
Control: N Atlantic ~ N Pacific data, but data has more impact in N Pacific.
Impact of ship data small (mostly low lats), larger without drifter data. Sampling bias?
Removing half of drifters has 25-30% of effect of removing all drifters+ships.
Low latitude bias small; 50°-90°N 0.12 hPa for Control -> 0.20 for NoNHdr

Drifting buoys without pressure sensors

45% (Cyan coloured) don't report pressure, mainly Pacific but also Atlantic



Summary of buoy/ship studies

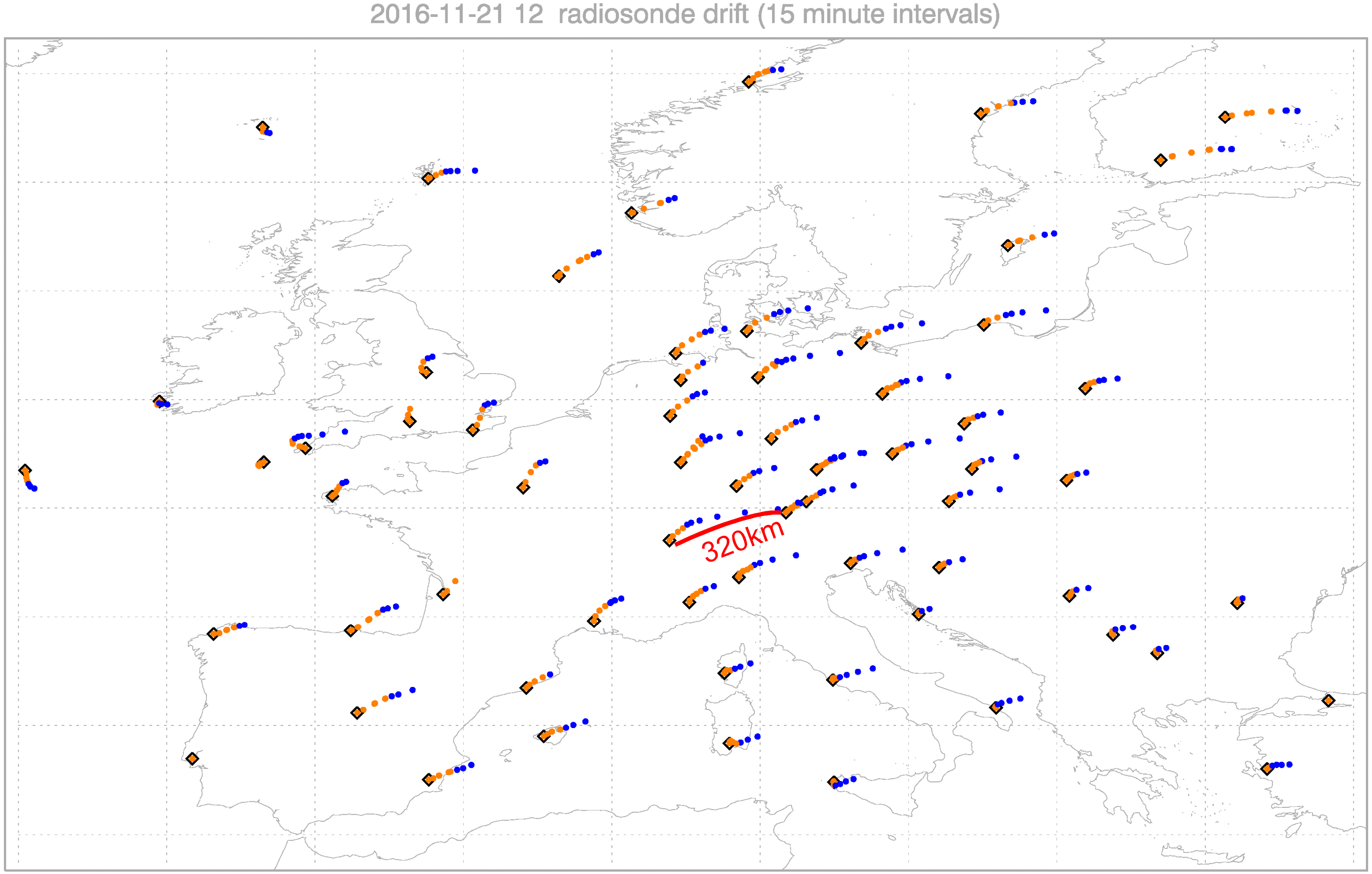
- Areas of large FSOI link to baroclinic development – also seen on global scale (and data sparse Arctic areas) – priority for pressure measurements
- Buoy pressure better quality than ship, ship quality variable (similar to results shown by Ingleby, 2010, JTech)
    - Caveat: drifting buoys don't provide flux estimates
- OSEs performed for Nov 2015 – Feb 2016:
- Northern Hemisphere drifter pressure has significant impact (ship much less especially with drifters present)
- Removing ~half the drifters has 25-30% of short range impact of removing all drifters+ships
- At short range biggest geopotential height changes in west of ocean basins – links to FSOI results and baroclinic development areas (above)
- More impact at high latitudes (fewer pressure drifters in (sub)tropics) and in the Pacific

Radiosonde drift processing

- A radiosonde ascent takes about two hours and during that time it can drift over 200 km horizontally (~4000 points with 2-second reporting)
- Currently the profile is treated as vertical and instantaneous!
- In (new) drift processing we split the ascent into 15 minute chunks (could use smaller intervals)
- Minor changes to vertical thinning were made – use up to 3% more levels
- Implemented within 45r1to go operational February 2018
- Only affects native BUFR reports (that include displacement) – mainly from Europe and Australia/NZ for the trial period, ~17% of stations

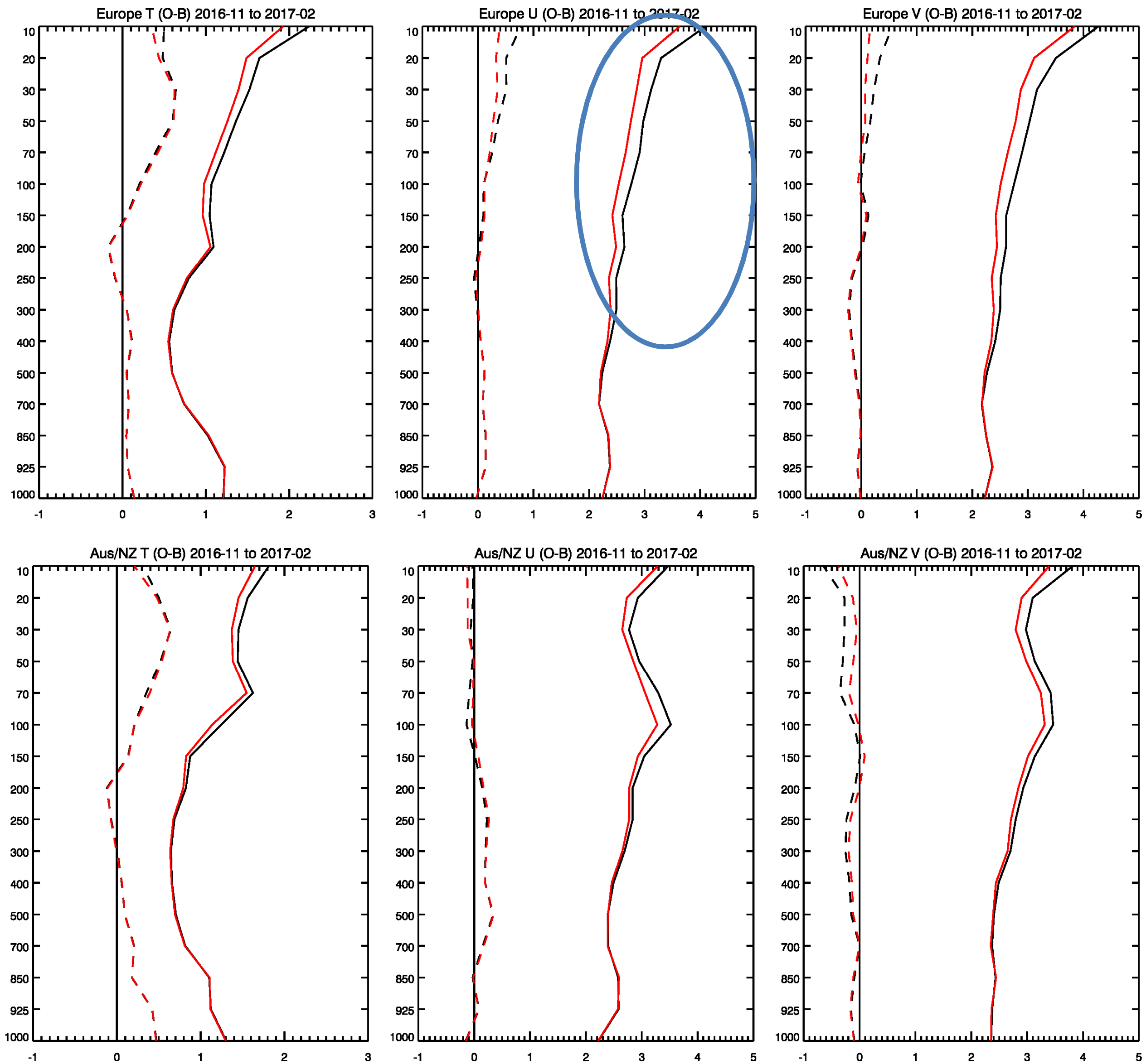
Example of large drift – 12 UTC 21 November 2016

- Black diamonds – launch, levels to 100 hPa, levels above 100 hPa
- French (LoRes) and Polish (just upgraded) BUFR not used at the time

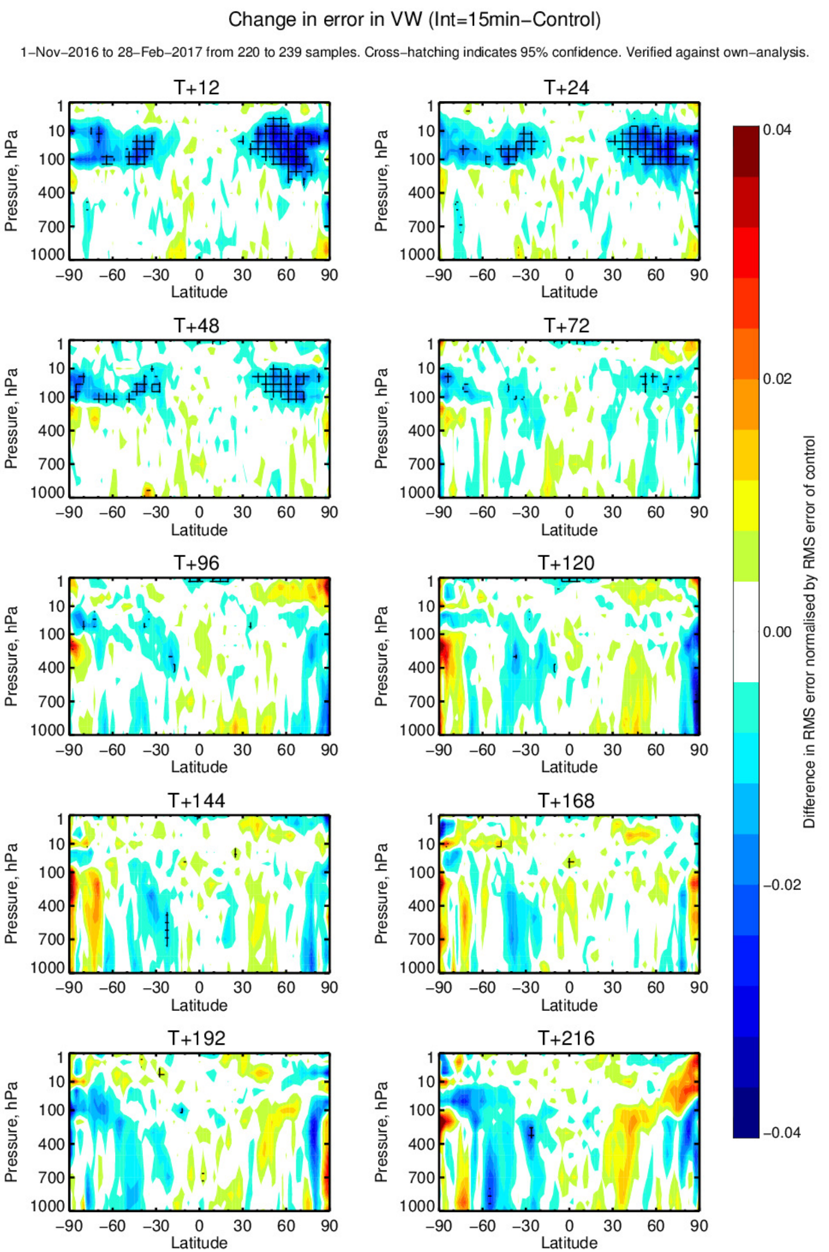


Mean and rms O-B statistics

- Data for Nov 2016 – Feb 2017
- Europe (top), Australia/NZ (bottom)
- Assimilated BUFR TEMP standard levels only (to get clean comparison)
- Good improvements at 200 hPa and above – including wind biases
- More improvement for wind than temperature (less competition)
- Improvement comes mainly from better H operator (rather than B)
- More drift in the winter hemisphere



ECMWF EUROPEAN CENTRE



Impact of drift processing

- Ascents split into 15 minute chunks versus operation-like control
- Vector wind RMS differences verified against own analysis, Nov 2016 – Feb 2017 (sample size >220)
- Blue means improvement
- Short range – much smaller rms in ET stratosphere, mainly due to smaller increments
- Less impact on temperature – satellite data
- Medium range some improvements in troposphere, especially Southern Extratropics, also seen for Z
- This is using drift from ~15% of radiosonde stations, bigger impact with more (including tropics)

Radiosonde drift processing and LAMs

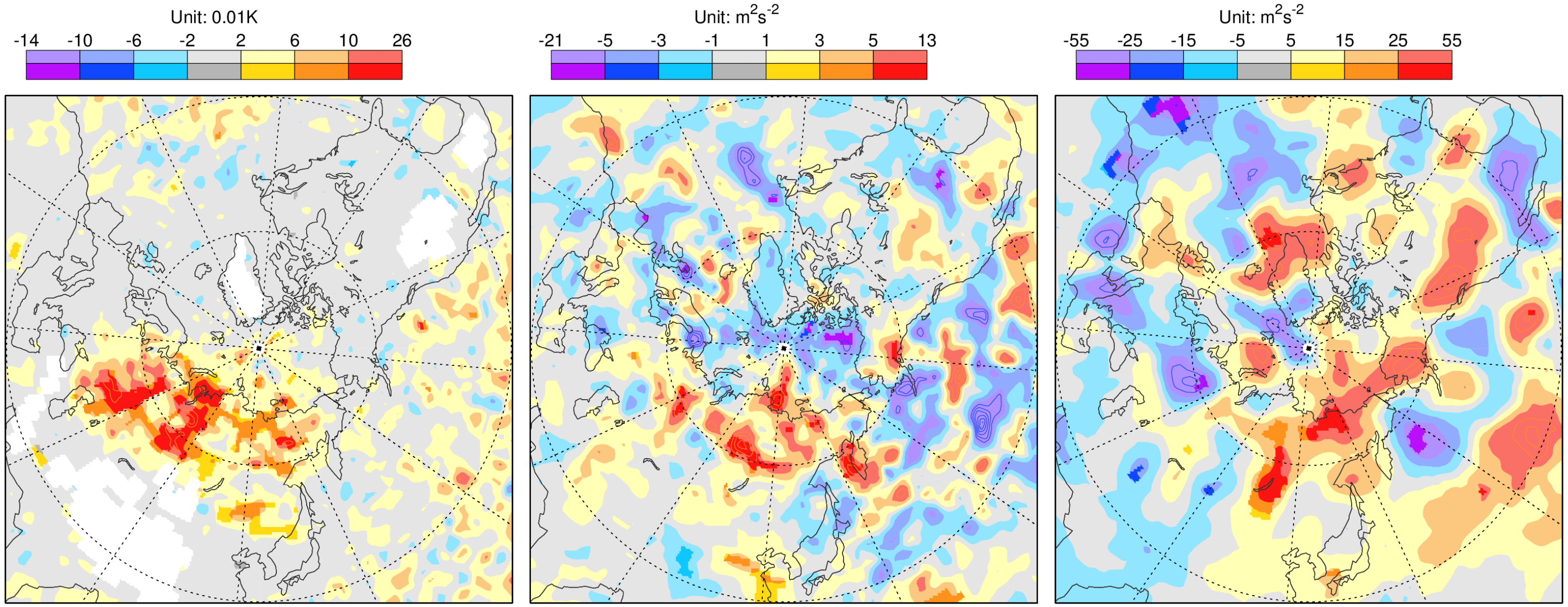
- Laroche and Sarrazin (2013) found good impact in Candadian global NWP
- Ingleby and Edwards (2015, ASL) found improved upper level O-B in UKV
  - Best to use both correct time and lat/long
  - Dow and Macpherson ran UKV trial – neutral impact (implemented)
  - Because windy cases dominated by boundary conditions?
- Particularly if 1 hour analysis window used each radiosonde ascent will span two or three windows – various practical issues
- Need for earlier (500 hPa) radiosonde report?
- Use in verification?? Verification doesn't time interpolate fields.

Impact of 1 vs 2 Russian sondes per day (EC newsletter 149)

- Short range impact mainly at low levels over Russia (plot from Mark Rodwell)
- Spreads downstream and upwards (~6 months NWP improvements to NH at D+5)
- This is in Winter (no tropospheric sat. radiances over snow) less impact in Summer
- No wind profilers, very few aircraft ascents/descents over Russia
- Russian radiosonde quality worse than average in UTLS (less difference in LT)

Mean change in RMSE (Test minus Control) for the period 20131201-20140228. Saturated colours indicate significance at the 5% level

(a) T850 D+1 (b) Z500 D+2 (c) Z200 D+5



Radiosonde notes

- We can run data denial tests using large subsets (110 Russian sondes)
- We can't sensibly run tests for 1 or 2 radiosondes, S:N too small
- We can look at FSOI statistics – more impact for remote sites

- Ongoing migration to (HiRes) BUFR radiosonde data – Europe largely complete, but only ~35% of stations worldwide provide good BUFR
- "An assessment of different radiosonde types", ECMWF TM 807, 2017
- Work on updating/reducing radiosonde bias correction (can't continue to use RS92 as a reference as it is phased out)
- Germany, UK, Finland are looking at descent profiles from RS41 – after balloon burst (quality varies, some usable), ECMWF will process them soon

Summary

- AMDAR humidity generally good quality and useful
- Biases in E-AMDAR need a closer look

- BUOY pressure very important, especially in development areas (shown by FSOI) and data sparse regions
- Increased data density likely to improve forecasts (but diminishing returns – general feature)

- Radiosonde drift treatment improves O-B statistics and global forecasts – less impact on LAM forecasts?
- Radiosondes still important as a reference

Migration to (High-Resolution) BUFR reports

- ~800 stations globally
- 180 RS92, 87 RS41
- 27% send HiRes BUFR ☺
- 8% send LoRes BUFR ☺
- 38% reformatted TEMP ☹
- 27% send no BUFR ☹

- Migration was meant to be complete Nov 2014!
- US roll-out, Japan/China working on HiRes
- Two NMSs use RS41 but don't send good BUFR ☹

August 2017: Radiosonde BUFR availability/type