

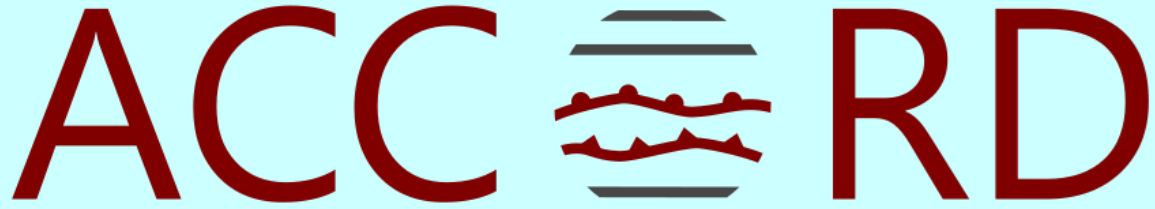
A Consortium for CONvection-scale modelling  
Research and Development

## DA based verification - an update

5th ACCORD All Staff Workshop 31 March - 4 April 2025, Zalakaros and hybrid  
James Fannon and Eoin Whelan

# Outline

- Acknowledgments
- Linking verification with data assimilation
- Implementation
- Early results
- Verification using radiosonde observations



A Consortium for CONvection-scale modelling  
Research and Development

## Acknowledgements

# Acknowledgements

- Based on previous work by Roger Randriamampianina
- Parallel development (and use) by Siebren de Haan
- Independent developments by Météo-France
  - *SCOOPS* (scores using OOPS screening)
- Discussion at EWGLAM 2022
- Presentation at EWGLAM 2024 (very similar to what follows!)

# ACC RD

A Consortium for CONvection-scale modelling  
Research and Development

Links to DA

# Variational DA and the observation operator, $H$

In DA we apply a variational formalism that consists in finding the best possible initial state,  $\mathbf{x}^a$ , by **minimising** a penalty function,  $J$ :

$$J = J_b + J_o = \frac{1}{2} \delta \mathbf{x}^T \mathbf{B}^{-1} \delta \mathbf{x} + \frac{1}{2} (\mathbf{H} \mathbf{x}^b + \mathbf{H} \delta \mathbf{x} - \mathbf{y})^T \mathbf{R}^{-1} (\mathbf{H} \mathbf{x}^b + \mathbf{H} \delta \mathbf{x} - \mathbf{y})$$

$J_b$  measures the distance of the state to be derived,  $\mathbf{x}$ , to the background state  $\mathbf{x}^b$

$J_o$  measures the distance to the different types of observations.

$\delta_{\mathbf{x}}$  represents the assimilation increments added to the background state,  $\mathbf{x}^b$ , to form the analysis,  $\mathbf{x}^a$ .

The observation operator,  $H$ , provides the link that projects the model state to all types of observations used in the data assimilation to enable comparison between the model state (background or analysis) and the observations.

# Variational DA and the observation operator, $H$

In DA we apply a variational formalism that consists in finding the best possible initial state,  $\mathbf{x}^a$ , by **minimising** a penalty function,  $J$ :

$$J = J_b + J_o = \frac{1}{2} \delta \mathbf{x}^T \mathbf{B}^{-1} \delta \mathbf{x} + \frac{1}{2} (\mathbf{H} \mathbf{x}^b + \mathbf{H} \delta \mathbf{x} - \mathbf{y})^T \mathbf{R}^{-1} (\mathbf{H} \mathbf{x}^b + \mathbf{H} \delta \mathbf{x} - \mathbf{y})$$

$J_b$  measures the distance of the state to be derived,  $\mathbf{x}$ , to the background state  $\mathbf{x}^b$

$J_o$  measures the distance to the different types of observations.

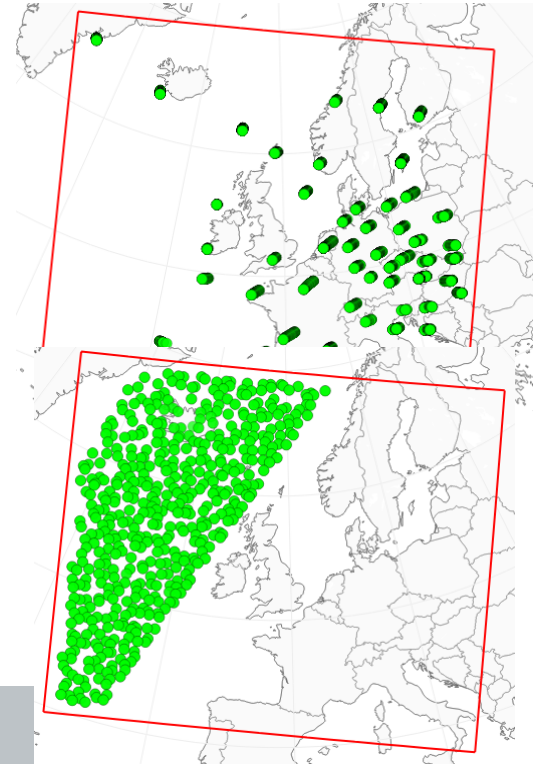
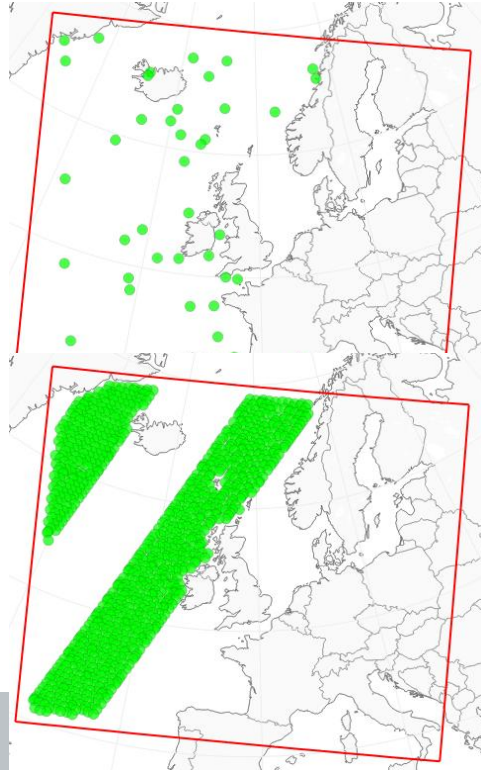
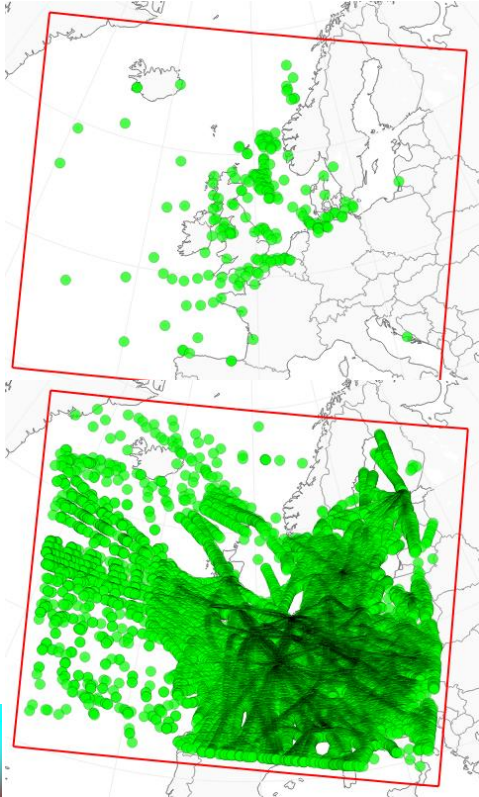
$\delta_{\mathbf{x}}$  represents the assimilation increments added to the background state,  $\mathbf{x}^b$ , to form the analysis,  $\mathbf{x}^a$ .

**Observation operator,  $H$ , used in DA can also be used for model verification!**

The observation operator,  $H$ , provides the link that projects the model state to all types of observations used in the data assimilation to enable comparison between the model state (background or analysis) and the observations.

# Screening of observations and $H$ - benefits

Lots of observations!



# Screening of observations and $H$ - benefits

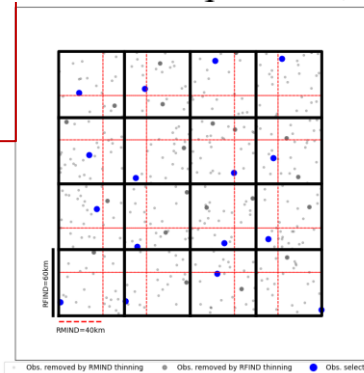
First-guess check to reject observations affected by gross errors:

$$([H(\mathbf{x}_b)]_i - y_i)^2 / \sigma_{b,i}^2 > \text{FgLim} \times \lambda$$

where  $\lambda = 1 + \sigma_{o,i}^2 / \sigma_{b,i}^2$ , FgLim is the rejection limit and  $[H(\mathbf{x}_b)]_i$  denotes the projection of the model state on  $y_i$  observation, where the potential observation bias has been accounted for.  $\sigma_{o,i}$  and  $\sigma_{b,i}$  are the standard deviation of the observation error and background error equivalent, respectively.

Thinning of data to alleviate effects of observation error correlations not represented.

*Illustration of horizontal thinning procedure with finer resolution grid (dashed red) here with grid size 40 km and coarser resolution final grid (full black) here with grid distance 60 km. Here small grey dots represent observations removed by thinning in finer resolution grid, larger grey dots represent the observations removed by final coarser resolution grid and blue dots represent the remaining selected observations.*

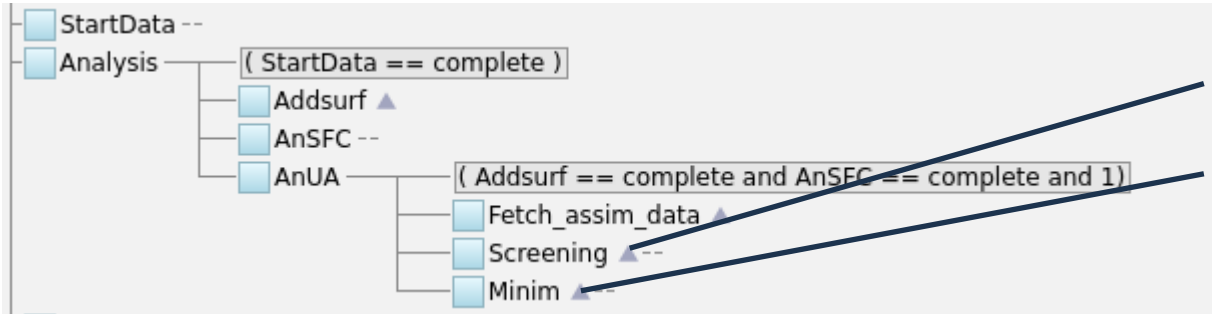


# ACC RD

A Consortium for CONvection-scale modelling  
Research and Development

## Implementation

# Implementation



Screen observations for DA

3D-Var (minimisation of the “J”)



Fetch the forecast files

Use Screening to produce verification statistics

Read ODB and write SQLite

# Data conversion

A few lines of Python (from odbcon <https://github.com/Hirlam/odbcon>)...

```
def convert_db(ifile,opath,obtype):
    # Open and save the filtered odb2 file to a temporary database
    with open(ifile, 'rb') as f:
        df_decoded = odc.read_odb(f, single=True)
    #DBG print(df_decoded)
    sql_tmp = opath + '/sqltmp.db'
    conn = sqlite3.connect(sql_tmp)
    df_decoded.to_sql(obtype, conn, if_exists='replace', index=False)

    # Now read this back in as a standard pandas data frame
    sql_query='select * from '+obtype
    p_df = pd.read_sql(sql_query, conn)
    conn.close()

    #
    #IN lat, lon, level, obs, fcst, sid, fcst_model, fcst_dttm, valid_dttm, lead_time, parameter, units
    #OUT fcst_model, fcst_dttm, valid_dttm, lead_time, sid, parameter, level, fcst, units, lon, lat, obs
    #
    p_df=p_df.loc[:, ['fcst_model', 'fcst_dttm', 'valid_dttm', 'lead_time', 'sid', 'parameter', 'level', 'fcst', 'units', 'lon', 'lat', 'obs'] ]

    # Do some renaming etc
    p_df.rename(columns={'sid': 'SID'}, inplace=True)

    #DBG print(p_df)
    dtg=p_df['fcst_dttm'].iloc[0]
    ll=p_df['lead_time'].iloc[0]
    yyyyymm=str(dtg)[0:6]
    hh=str(dtg)[8:10]

    # Now save this to a permanent sql
    ofile = opath + '/' + 'OFTABLE_' + obtype + '_' + yyyyymm + '_' + hh + '.sqlite'
    connn = sqlite3.connect(ofile)
    p_df.to_sql(obtype, connn, if_exists='append', index=False)
    connn.close()
```

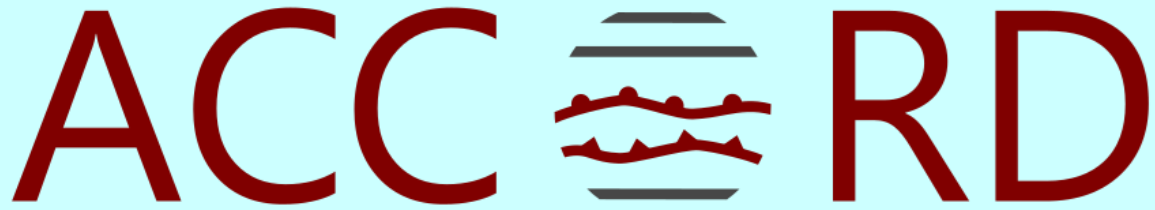
# Data conversion

A few lines of Python opens many avenues for data formats ...

```
#!/bin/env python3
import pandas as pd
import sys
import os
import argparse
#import cdc as odc
import pydc as odc
import sqlite3

def convert_db(ifile,opath,obtype):
    # Open and save the filtered odb2 file to a temporary database
```

Instead of sqlite3, we can as easily write to other formats that have a Python interface



A Consortium for CONvection-scale modelling  
Research and Development

## Results – James

# Results – Really nice interface!

**Experiment**  
Allobsver Sample

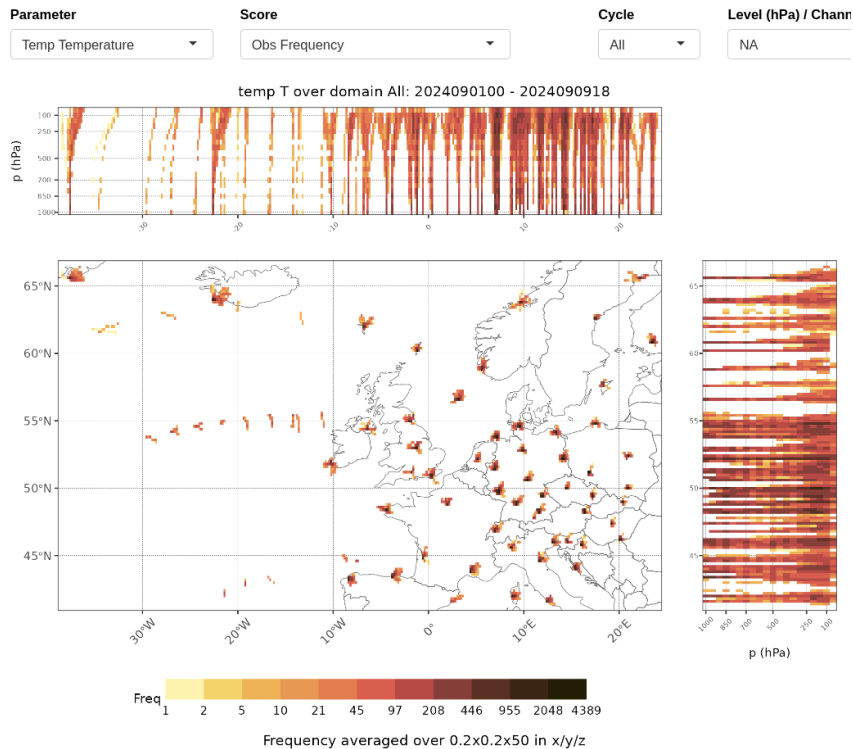
**Year**  
2024

**Date**  
2024/09/01 00Z - 2024/09/07 18Z

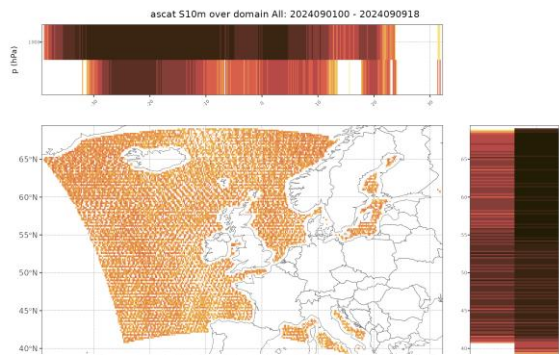
Surface **Temp** Scorecards

Summary Prof Signif

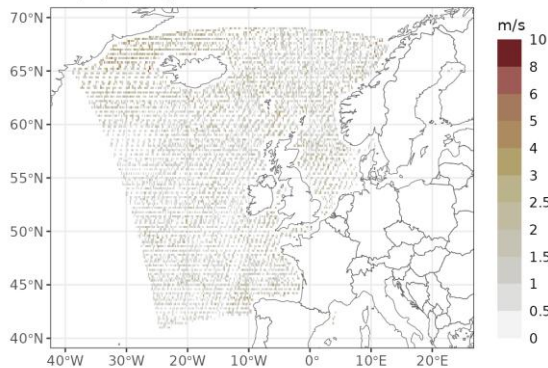
**Station selection**  
 All stations  
 Ireland+UK



# Results - ASCAT

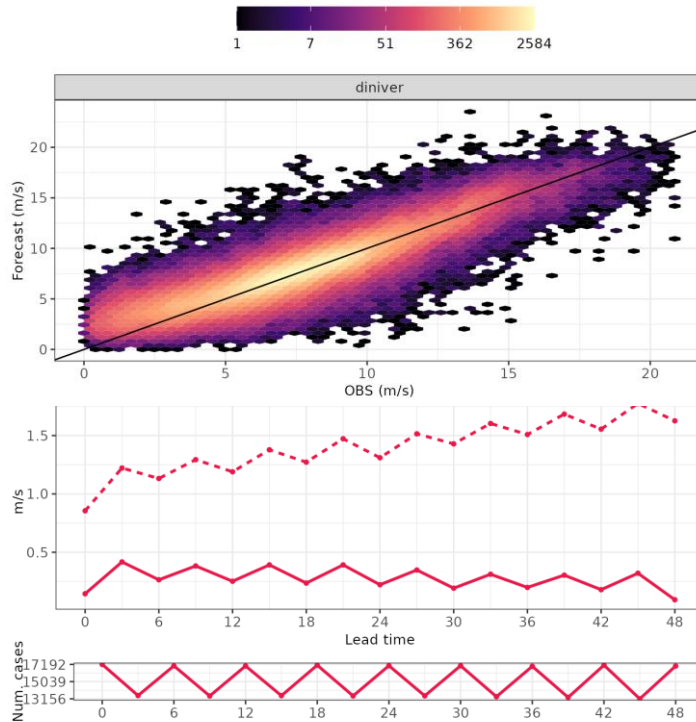


Rmse : Asc\_at\_s10m : 2024-09-01-00 - 2024-09-07-18 (28 cycles)  
All data : 00,06,12,18Z cycles used: Valid hour = 12Z



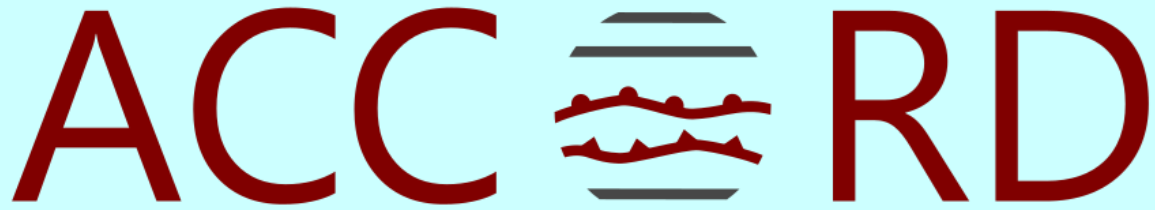
Obs/forecasts rounded to closest 0.2x0.2 grid

Scatterplot : Asc\_at\_s10m : 2024-09-01-00 - 2024-09-07-18 (28 cycles)  
Used (00,06,12,18) +3, 6, ..., 48 : All stations



# Results

- UWC West:
  - Screening by verification processing d-suite data in real-time
  - Comparison with o-suite data pending
- MetCoOp:
  - Used for Cycle 46 experiments
  - Used for a pre-op MEPS suite (without radar)
- HIRLAM/ACCORD:
  - Has been used for research/development Cycle 46 experiments
- Still need a common visualisation for the statistics



A Consortium for CONvection-scale modelling  
Research and Development

## Radiosonde observations – James

# Data

- Period: 2024/09 – 2025/01.
- “Verification by Screening” data from DINI-EPS mbr000 d-suite
- “dini” vobs (created from operational BUFR):
- “dini” forecasts from DINI-EPS mbr000 vfld

# Methodology

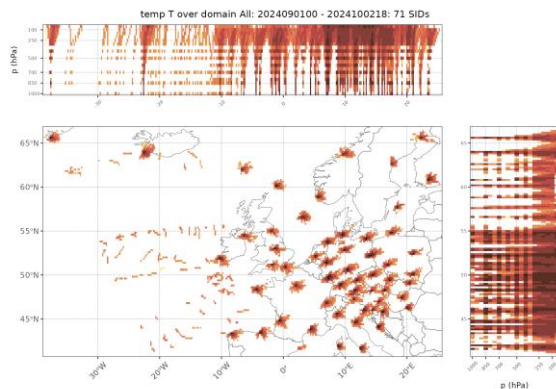
- Read data on a monthly basis (mimic operations).
- Use forecast cycles 00-21Z every 3h (maximise sample size).
- Using lead times 0-48 every 3h (maximise sample size).
- Only reading whole pressure every 10hPa.
- Profiles at valid hours 00Z and 12Z.
- Temperature (T), specific humidity (Q), and wind speed (S).
- Focus:
  - Is verification by screening doing something sensible?
  - Does verification by screening give significantly different results to the traditional methodology?

# Conclusions 1/3 (based on plots below)

- Station coverage:
  - Need to look at adding 3023, 3808, 13275 to OFCTABLE.
  - Some rogue ships in Germany?
- Comparing obs in OFCTABLE vs vobs:
  - Generally similar, some large outliers however (not clear why).
  - For Q the differences are more systematic, where the OFCTABLE observations are drier than the vobs at mid-levels (500-850hPa).
- Comparing forecasts in OFCTABLE vs vfld:
  - Generally reasonable, diffs increase at upper-levels (as expected), particularly for wind speed.
  - Some surprisingly large diffs at 1000hPa – why?

# T: station coverage (2024/09)

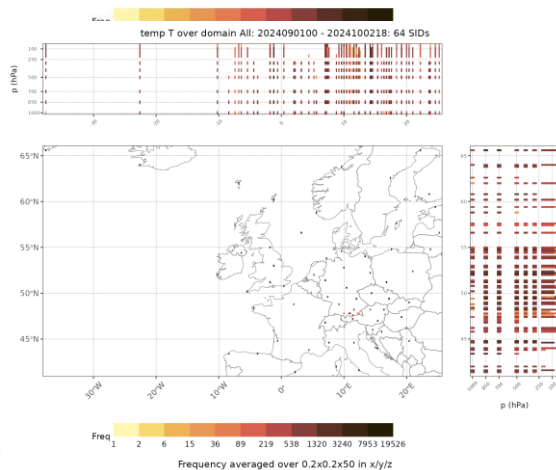
OFCTABLE



Extra stations in OFCTABLE:

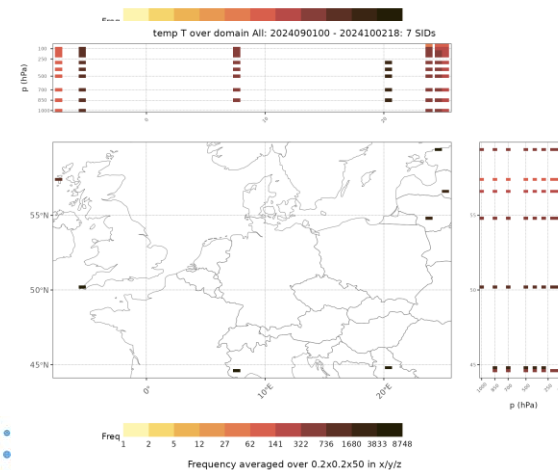
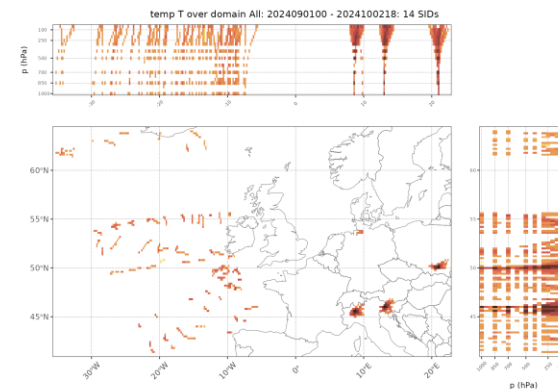
- Mostly ships.
- 12575, 16045, and 16064 (Poland, Italy).
- Note erroneous ship data (ASDE09) over Germany.

VOBS



Extra stations in VOBS:

- Stations near border excluded due to REDZONE
- 3023, 3808 missing from OFCTABLEs.



# Conclusions 2/3 (based on plots below)

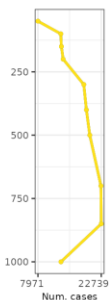
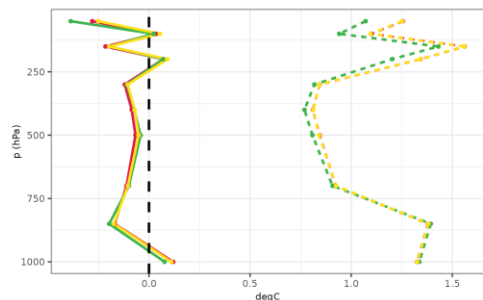
- Temperature profiles:
  - Reduction in STDE at higher levels (300hPa+) of  $\sim 0.1\text{K}$  when using the drift-corrected method. Impact on bias more mixed, limited to  $0.05\text{K}$ .
- Humidity profiles:
  - Impact on STDE is relatively small, limited to  $0.01\text{g/Kg}$ .
  - Impact on bias limited to  $0.01\text{g/Kg}$  when looking at drift-impact only (i.e. using OFCTABLE obs and comparing OFCTABLE vs vfld forecasts).
  - Impact on bias more significant when comparing “Verif by Screening” and “vfld/vobs” methods i.e. increased bias of  $0.025 - 0.1\text{g/kg}$  in mid-levels. This is due to drier observations in OFCTABLE!
- Wind speed profiles:
  - Reduction in STDE at all levels, up to  $0.2-0.3\text{m/s}$  at upper-levels (300hPa+).
  - Bias more mixed, but generally reduced winds ( $0.1\text{m/s}$ ) at 1000hPa.

# T: Profile comparison (2024/11)

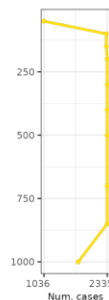
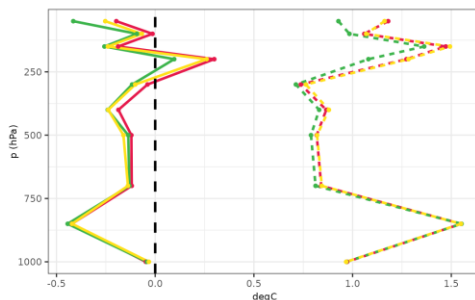
Left-right: All, Ireland+UK, Netherlands (operational) stations. Top/bottom: Valid at 00/12Z.

Vfld vs vobs (traditional approach), Vfld vs OFCTABLE obs, Drift corrected vs OFCTABLE obs

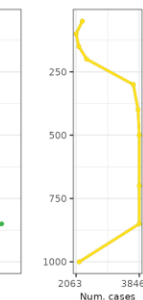
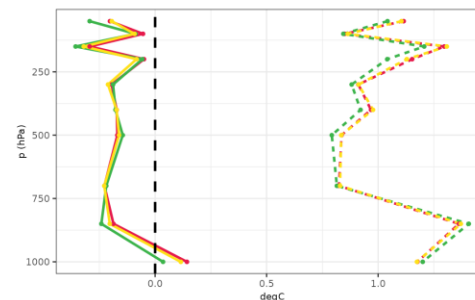
Bias, Stde : T : 2024-11-01-00 - 2024-11-30-21 (240 cycles)  
 All stations (58) : 00,03,06,09,12,15,18,21Z cycles: + 0, 3, ..., 45, 48 : Valid hour = 00Z



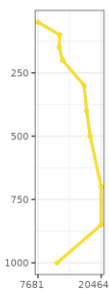
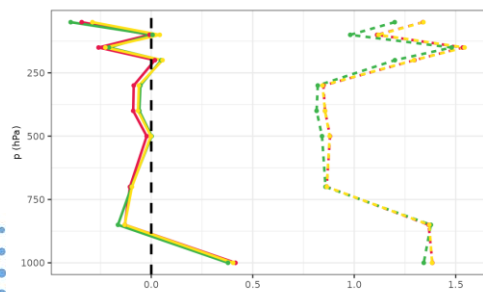
Bias, Stde : T : 2024-11-01-00 - 2024-11-30-21 (240 cycles)  
 IE\_EN stations (6) : 00,03,06,09,12,15,18,21Z cycles: + 0, 3, ..., 45, 48 : Valid hour = 00Z



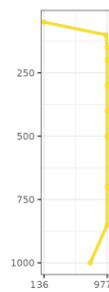
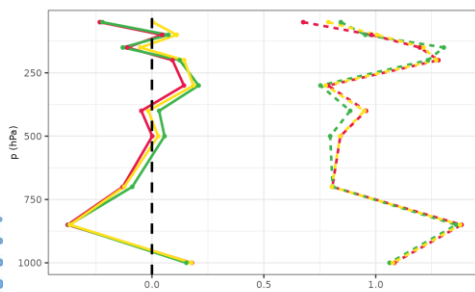
Bias, Stde : T : 2024-11-01-00 - 2024-11-30-21 (240 cycles)  
 NL\_OP stations (10) : 00,03,06,09,12,15,18,21Z cycles: + 0, 3, ..., 45, 48 : Valid hour = 00Z



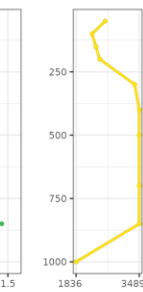
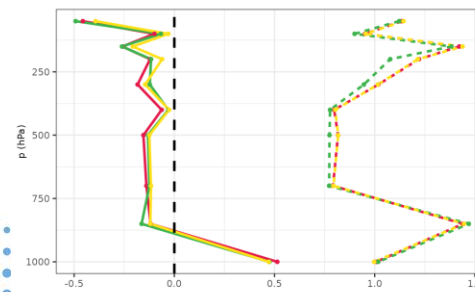
Bias, Stde : T : 2024-11-01-00 - 2024-11-30-21 (240 cycles)  
 All stations (58) : 00,03,06,09,12,15,18,21Z cycles: + 0, 3, ..., 45, 48 : Valid hour = 12Z



Bias, Stde : T : 2024-11-01-00 - 2024-11-30-21 (240 cycles)  
 IE\_EN stations (6) : 00,03,06,09,12,15,18,21Z cycles: + 0, 3, ..., 45, 48 : Valid hour = 12Z



Bias, Stde : T : 2024-11-01-00 - 2024-11-30-21 (240 cycles)  
 NL\_OP stations (10) : 00,03,06,09,12,15,18,21Z cycles: + 0, 3, ..., 45, 48 : Valid hour = 12Z



# Conclusions 2/3 (based on plots below)

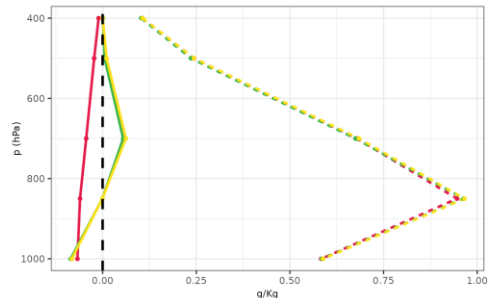
- Temperature profiles:
  - Reduction in STDE at higher levels (300hPa+) of  $\sim 0.1\text{K}$  when using the drift-corrected method. Impact on bias more mixed, limited to  $0.05\text{K}$ .
- Humidity profiles:
  - Impact on STDE is relatively small, limited to  $0.01\text{g/Kg}$ .
  - Impact on bias limited to  $0.01\text{g/Kg}$  when looking at drift-impact only (i.e. using OFCTABLE obs and comparing OFCTABLE vs vfld forecasts).
  - Impact on bias more significant when comparing “Verif by Screening” and “vfld/vobs” methods i.e. increased bias of  $0.025 - 0.1\text{g/kg}$  in mid-levels. This is due to drier observations in OFCTABLE!
- Wind speed profiles:
  - Reduction in STDE at all levels, up to  $0.2-0.3\text{m/s}$  at upper-levels (300hPa+).
  - Bias more mixed, but generally reduced winds ( $0.1\text{m/s}$ ) at 1000hPa.

# Q: Profile comparison (2024/11)

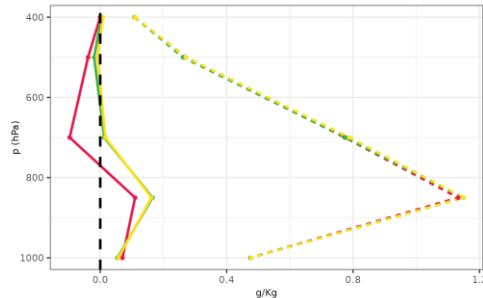
Left-right: All, Ireland+UK, Netherlands (operational) stations. Top/bottom: Valid at 00/12Z.

Vfld vs vobs (traditional approach), Vfld vs OFCTABLE obs, Drift corrected vs OFCTABLE obs

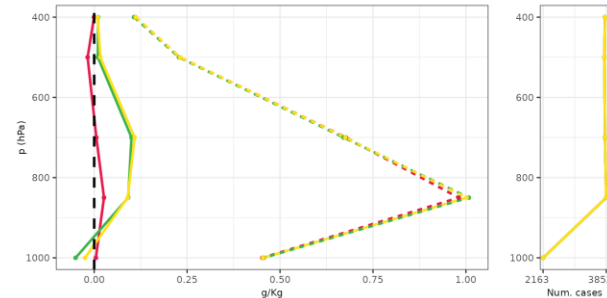
Bias, Stde : Q : 2024-11-01-00 - 2024-11-30-21 (240 cycles)  
 All stations (59) : 00,03,06,09,12,15,18,21Z cycles: + 0, 3, ..., 45, 48 : Valid hour = 00Z



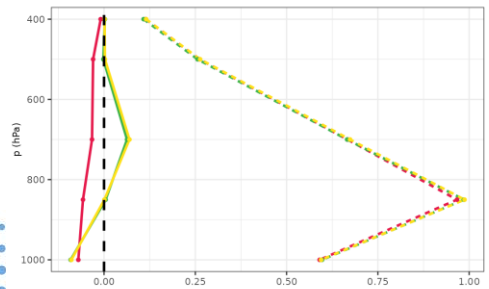
Bias, Stde : Q : 2024-11-01-00 - 2024-11-30-21 (240 cycles)  
 IE\_EN stations (7) : 00,03,06,09,12,15,18,21Z cycles: + 0, 3, ..., 45, 48 : Valid hour = 00Z



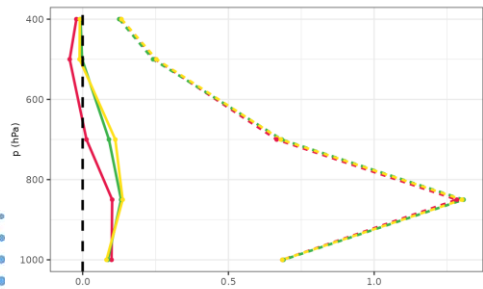
Bias, Stde : Q : 2024-11-01-00 - 2024-11-30-21 (240 cycles)  
 NL\_OP stations (10) : 00,03,06,09,12,15,18,21Z cycles: + 0, 3, ..., 45, 48 : Valid hour = 00Z



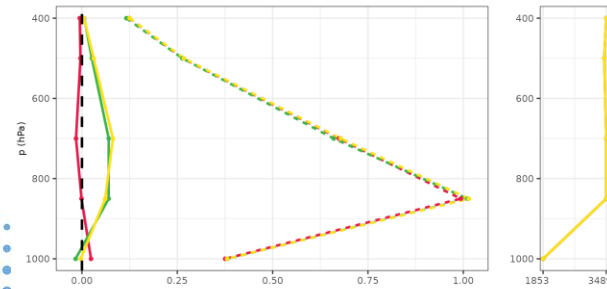
Bias, Stde : Q : 2024-11-01-00 - 2024-11-30-21 (240 cycles)  
 All stations (59) : 00,03,06,09,12,15,18,21Z cycles: + 0, 3, ..., 45, 48 : Valid hour = 12Z



Bias, Stde : Q : 2024-11-01-00 - 2024-11-30-21 (240 cycles)  
 IE\_EN stations (7) : 00,03,06,09,12,15,18,21Z cycles: + 0, 3, ..., 45, 48 : Valid hour = 12Z



Bias, Stde : Q : 2024-11-01-00 - 2024-11-30-21 (240 cycles)  
 NL\_OP stations (10) : 00,03,06,09,12,15,18,21Z cycles: + 0, 3, ..., 45, 48 : Valid hour = 12Z



# Conclusions 2/3 (based on plots below)

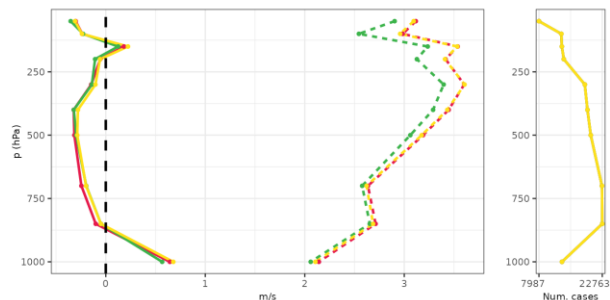
- Temperature profiles:
  - Reduction in STDE at higher levels (300hPa+) of  $\sim 0.1\text{K}$  when using the drift-corrected method. Impact on bias more mixed, limited to  $0.05\text{K}$ .
- Humidity profiles:
  - Impact on STDE is relatively small, limited to  $0.01\text{g/Kg}$ .
  - Impact on bias limited to  $0.01\text{g/Kg}$  when looking at drift-impact only (i.e. using OFCTABLE obs and comparing OFCTABLE vs vfld forecasts).
  - Impact on bias more significant when comparing “Verif by Screening” and “vfld/vobs” methods i.e. increased bias of  $0.025 - 0.1\text{g/kg}$  in mid-levels. This is due to drier observations in OFCTABLE!
- Wind speed profiles:
  - Reduction in STDE at all levels, up to  $0.2-0.3\text{m/s}$  at upper-levels (300hPa+).
  - Bias more mixed, but generally reduced winds ( $0.1\text{m/s}$ ) at 1000hPa.

# S: Profile comparison (2024/11)

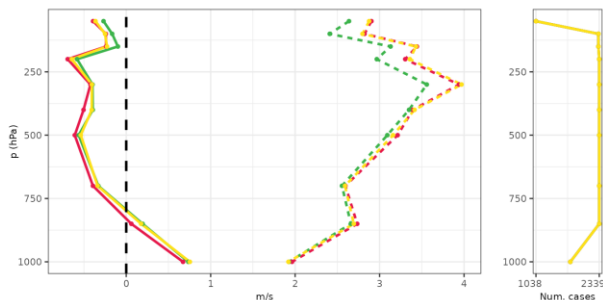
Left-right: All, Ireland+UK, Netherlands (operational) stations. Top/bottom: Valid at 00/12Z.

Vfld vs vobs (traditional approach), Vfld vs OFCTABLE obs, Drift corrected vs OFCTABLE obs

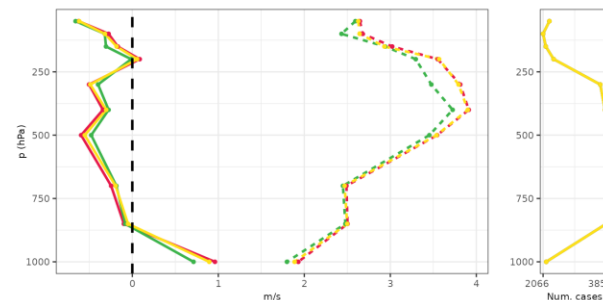
Bias, Stde : S : 2024-11-01-00 - 2024-11-30-21 (240 cycles)  
 All stations (59) : 00.03.06.09.12.15.18.21Z cycles: + 0, 3, ..., 45, 48 : Valid hour = 00Z



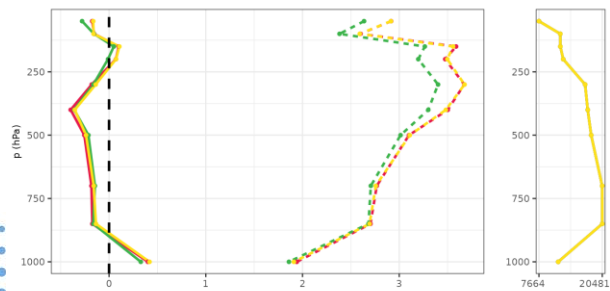
Bias, Stde : S : 2024-11-01-00 - 2024-11-30-21 (240 cycles)  
 IE\_EN stations (7) : 00.03.06.09.12.15.18.21Z cycles: + 0, 3, ..., 45, 48 : Valid hour = 00Z



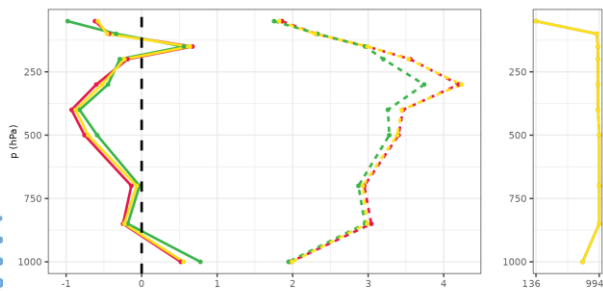
Bias, Stde : S : 2024-11-01-00 - 2024-11-30-21 (240 cycles)  
 NL\_OP stations (10) : 00.03.06.09.12.15.18.21Z cycles: + 0, 3, ..., 45, 48 : Valid hour = 00Z



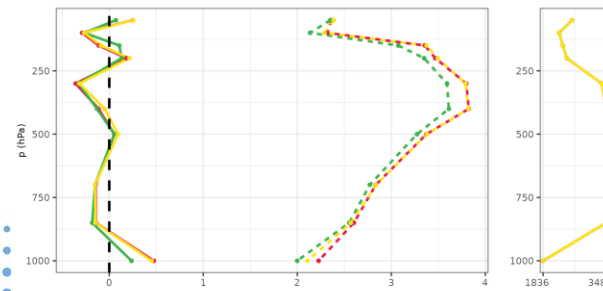
Bias, Stde : S : 2024-11-01-00 - 2024-11-30-21 (240 cycles)  
 All stations (59) : 00.03.06.09.12.15.18.21Z cycles: + 0, 3, ..., 45, 48 : Valid hour = 12Z



Bias, Stde : S : 2024-11-01-00 - 2024-11-30-21 (240 cycles)  
 IE\_EN stations (7) : 00.03.06.09.12.15.18.21Z cycles: + 0, 3, ..., 45, 48 : Valid hour = 12Z



Bias, Stde : S : 2024-11-01-00 - 2024-11-30-21 (240 cycles)  
 NL\_OP stations (10) : 00.03.06.09.12.15.18.21Z cycles: + 0, 3, ..., 45, 48 : Valid hour = 12Z

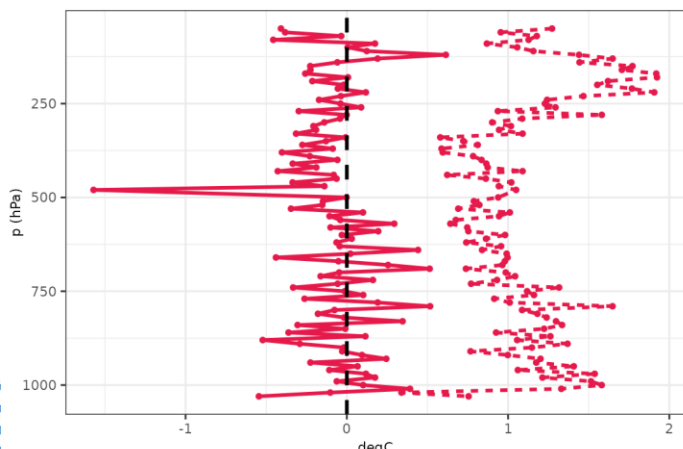


# Conclusions 3/3

- Need to be careful with “useable” obs from “Verif by Screening”.
  - Should obs on nearby pressure levels be combined?
  - If reading whole hours only, only have ~14 “useable” levels (increase of ~4 levels compared to standard approach).
  - Left: read all whole levels in OFCTABLE. Right: after num. cases filtering.

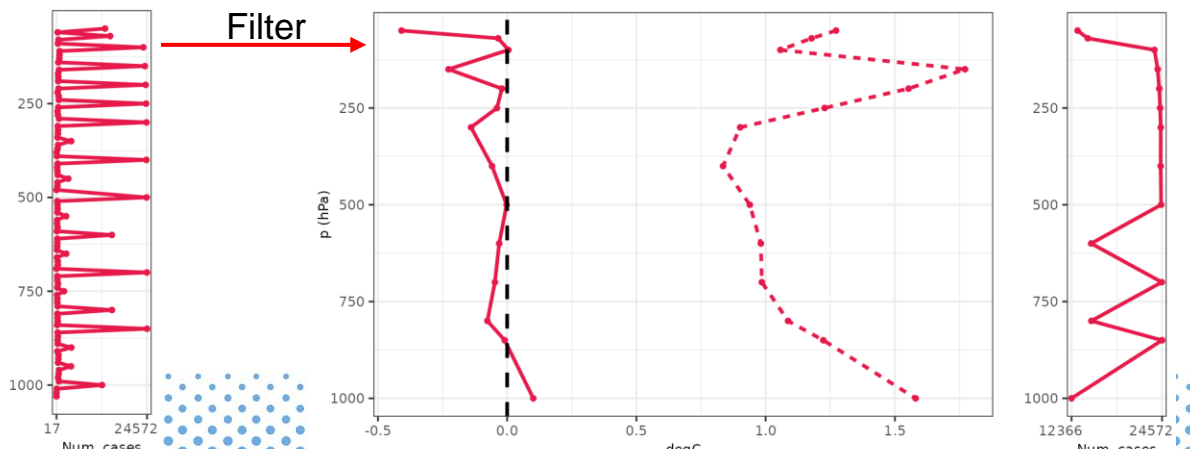
Bias, Stde : T : 2025-01-02-00 - 2025-01-31-21 (240 cycles)  
 All stations (70) : 00,03,06,09,12,15,18,21Z cycles: + 0, 3, ..., 45, 48 : Valid hour = 00Z

— Bias    - - - Stde    -•-•- diniver



Bias, Stde : T : 2025-01-02-00 - 2025-01-31-21 (240 cycles)  
 All stations (70) : 00,03,06,09,12,15,18,21Z cycles: + 0, 3, ..., 45, 48 : Valid hour = 00Z

— Bias    - - - Stde    -•-•- diniver



# Summary Conclusions

- Infrastructure developed to use DA screening to produce verification statistics
- Data readable by harp - <https://github.com/harphub>
- Visualisation needs to be further developed
- Further exploit DA observations
- Implement 4D-Var screening