ACCORD

A Consortium for COnvection-scale modelling Research and Development

~~Potential for~~ use of data assimilation framework for verification

26th EWGLAM 31st SRNWP meeting, 29-28 September 2024, Prague

Eoin Whelan, Magnus Lindskog and James Fannon

ACCORD

A Consortium for COnvection-scale modelling Research and Development

Use of data assimilation framework for verification

26th EWGLAM 31st SRNWP meeting, 29-28 September 2024, Prague

Eoin Whelan, Magnus Lindskog and James Fannon

Outline

- Acknowledgments
- Linking verification with data assimilation (DA)
- Implementation
- Results

ACCORD

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

ACCORD

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, $\mathbf{x}^{\mathbf{a}}$, by **minimising** a penalty function, J:

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

J_b measures the distance of the state to be derived, $\mathbf{x}$, to the background state $\mathbf{x}^{\mathbf{b}}$
J_o measures the distance to the different types of observations.
$\boldsymbol{\delta}_{\mathbf{x}}$ represents the assimilation increments added to the background state, $\mathbf{x}^{\mathbf{b}}$, to form the analysis, $\mathbf{x}^{\mathbf{a}}$.

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:

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

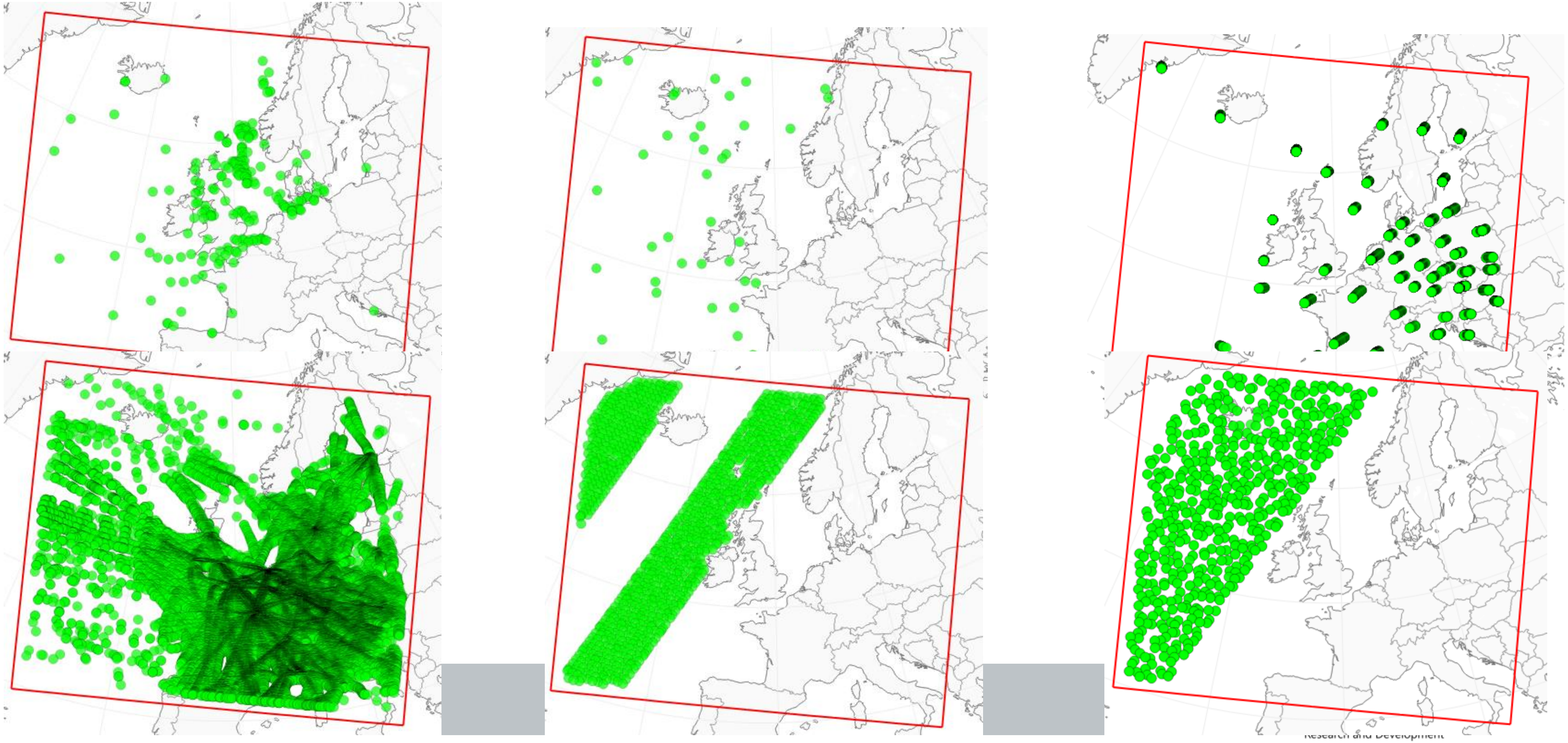
J_b measures the distance of the state to be derived, **x**, to the background state $\mathbf{x^b}$
J_o measures the distance to the different types of observations.
$\boldsymbol{\delta}_\mathbf{x}$ represents the assimilation increments added to the background state, $\mathbf{x^b}$, to form the

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:

$$([H(\boldsymbol{x}_{\mathrm{b}})]_i - y_i)^2/\sigma^2_{\mathrm{b},i} > \mathrm{FgLim} \times \lambda$$

where $\lambda = 1 + \sigma^2_{\mathrm{o},i}/\sigma^2_{\mathrm{b},i}$, FgLim is the rejection limit and $[H(\boldsymbol{x}_{\mathrm{b}})]_i$ denotes the projection of the model state on y_i observation, where the potential observation bias has been accounted for. $\sigma_{\mathrm{o},i}$ and $\sigma_{\mathrm{b},i}$ 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.

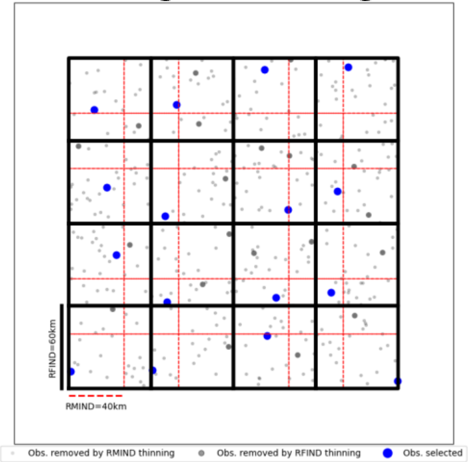
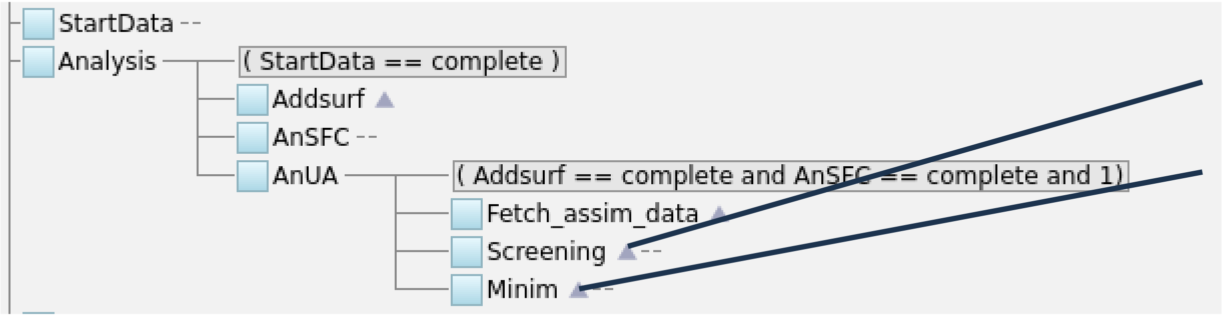


Illustration of horizontal thinning procedure with finer resolution grid (dashed red) here with grid size 40 km and coarser resolution final grid (full black) here with grid distance 60 km. Here small grey dots represent observations removed by thinning in finer resolution grid, larger grey dots represent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

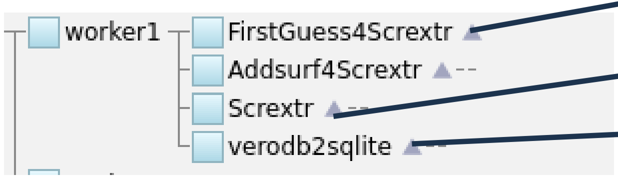
For tuning of screening, see poster by Sanchez Arriola et al.

Implementation



Screen observations for DA

3D-Var (minimisation of the “J”)



Fetch the forecast files

Use Screening to produce verification statistics

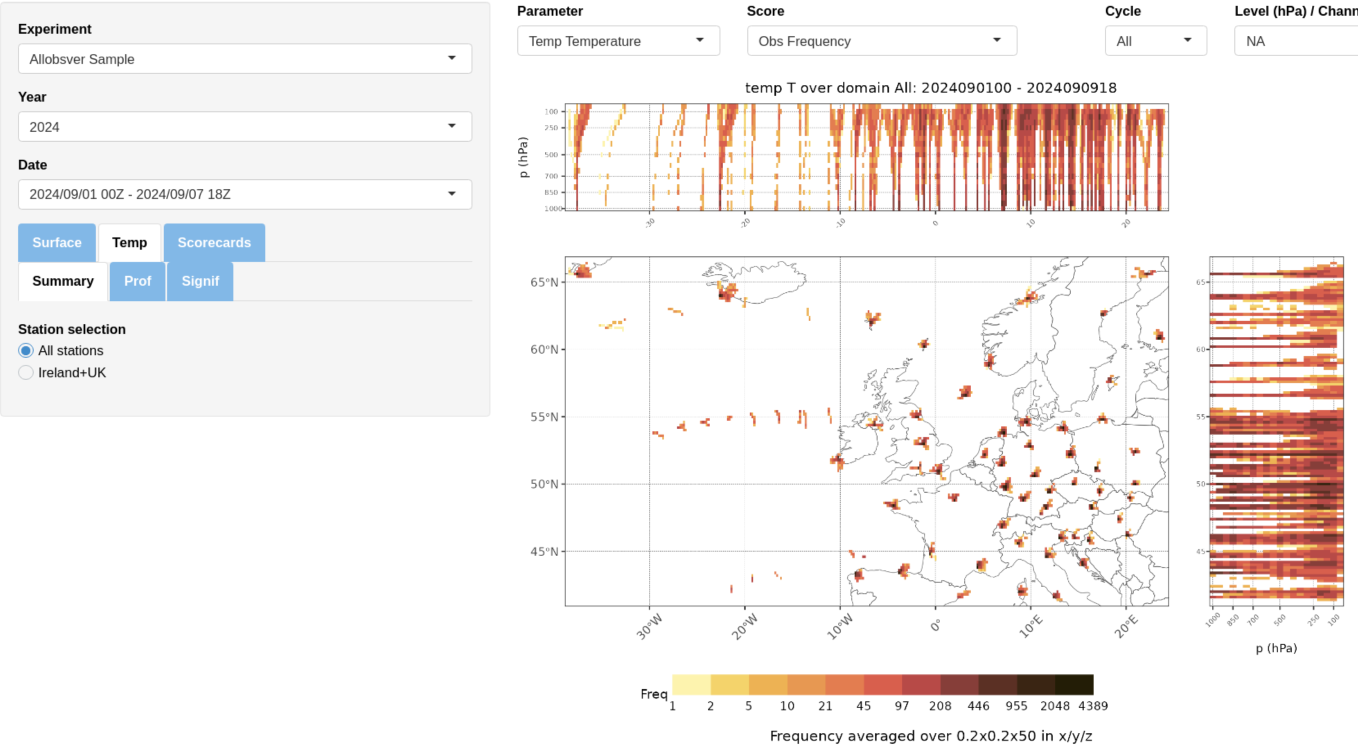
Read ODB and write SQLite

Data conversion

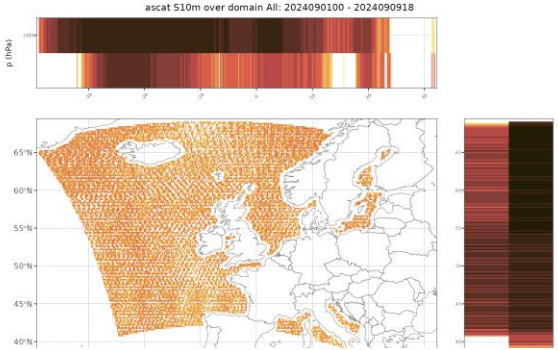
A few lines of Python …

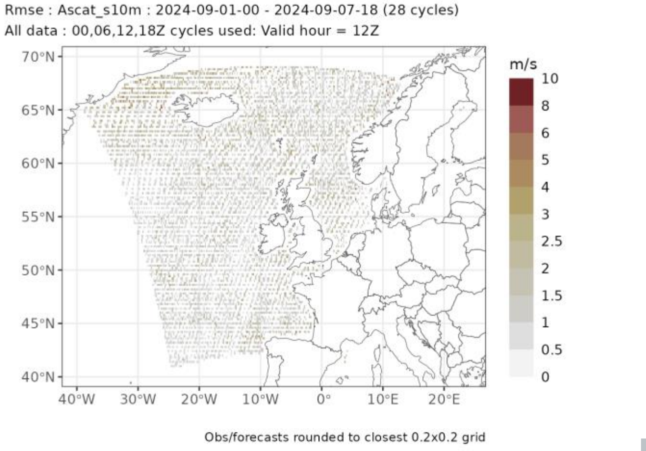
```
def convert_db(ifile,opath,obtype):
  # Open and save the filtered odb2 file to a temporary database
  with open(ifile, 'rb') as f:
      df_decoded = odc.read_odb(f, single=True)
  #DBG print(df_decoded)
  sql_tmp = opath + '/sqltmp.db'
  conn = sqlite3.connect(sql_tmp)
  df_decoded.to_sql(obtype, 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]
  yyyymm=str(dtg)[0:6]
  hh=str(dtg)[8:10]

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

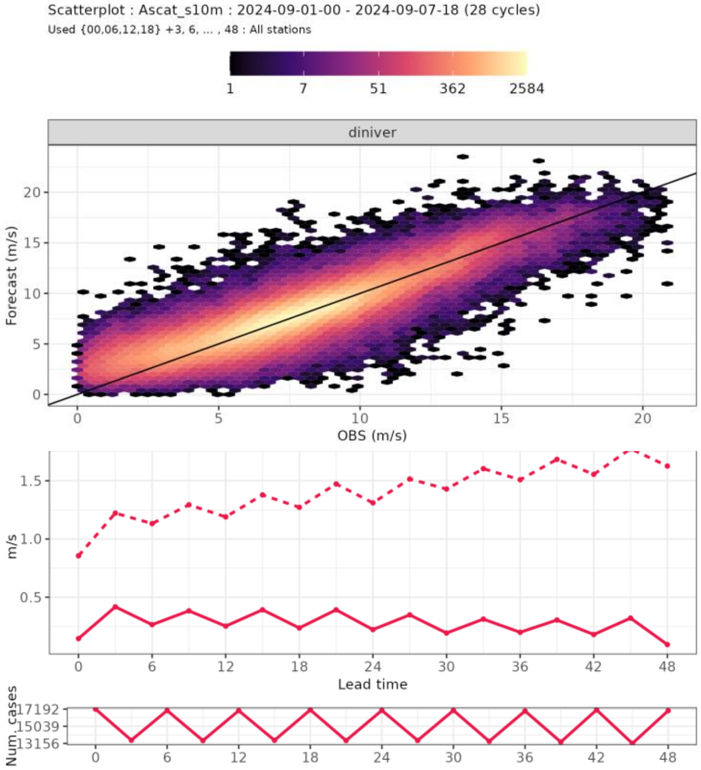
Results – Really nice interface!



Results - ASCAT







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