Predicting the impact of future observing systems

Katie Lean, Mary Borderies, Niels Bormann, Philippe Chambon, Giovanna De Chiara, Sean Healy, Sarah-Jane Lock, Katrin Lonitz, Linus Magnusson

ECMWF Annual Seminar, 16th Sept, Reading, UK

katie.lean@ecmwf.int

Why should we assess future observations in NWP?

Shape the evolution of the global observing system

2,200

lives saved per year in the US from halving forecast errors

Shrader et al., 2023

Justify investment – support space agencies/program approvals

€2.7bn

In asset losses avoided annually by early warning systems in Europe alone

Hallegatte, 2012

Inform satellite design decisions

Advocate for important aspects for NWP

Advanced preparation of NWP system for planned future observations

What do we need to assess future observations for NWP?

In a perfect world we want…

A variety of clear diagnostics that can represent the benefit to global (or regional) NWP

Context with well-established real observations to ground impact estimates

A framework that is:

- Flexible
- Cheap
- Easily maintained with latest NWP advances
- Quick to run

Outline

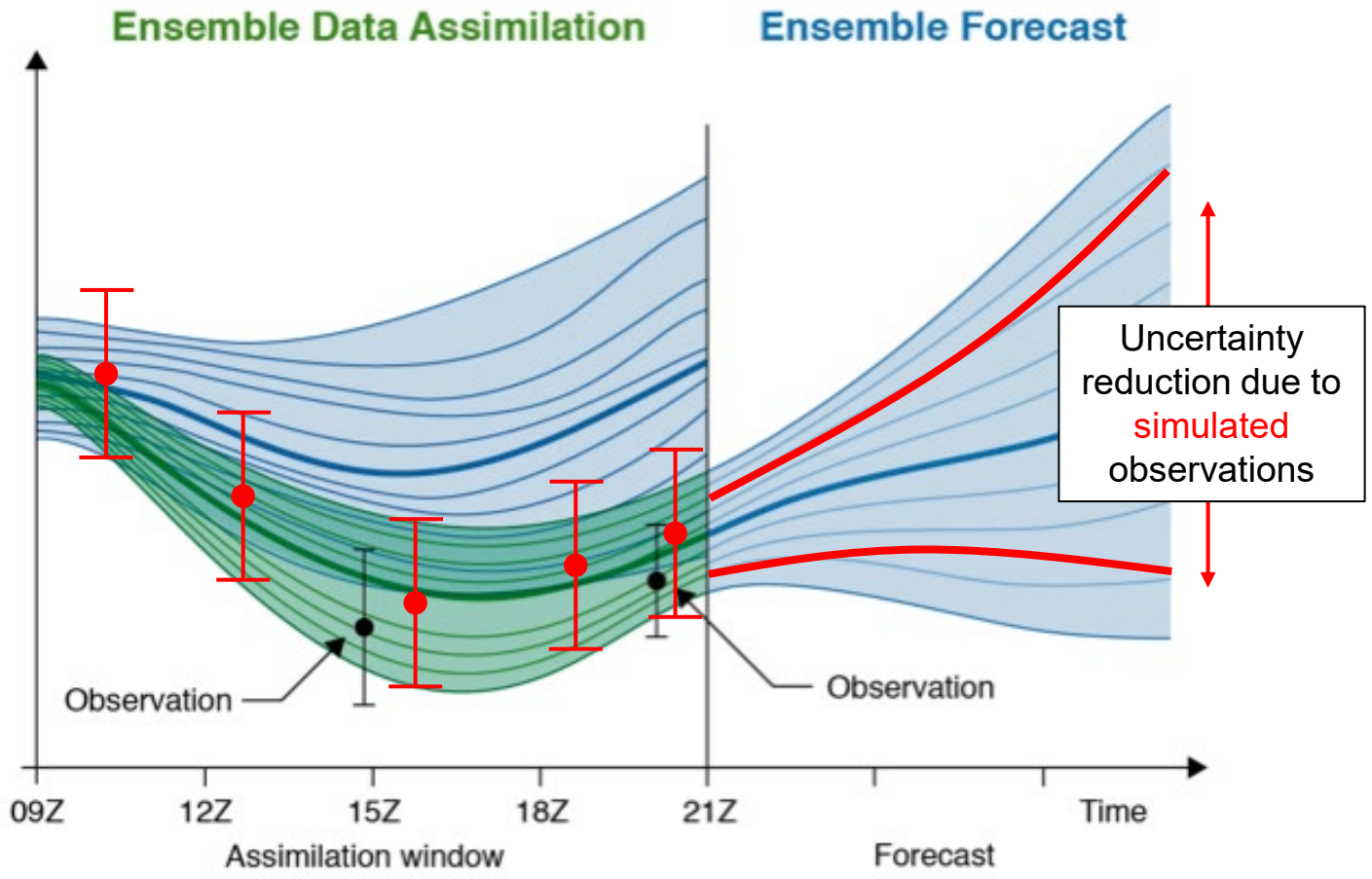
- The Ensemble of Data Assimilations (EDA) method
- Observing System Simulation Experiment (OSSE) method
- EDA and OSSE methods: a complementary pairing

Ensemble of Data Assimilations (EDA) method

The Ensemble of Data Assimilations (EDA) method

- EDA consists of:
  - Finite number of independent cycling assimilation systems
  - Uses real and added simulated observations
  - Observations, forecast model and SSTs perturbed to generate different inputs for each member
- Used operationally to provide
  - Initial conditions for the ensemble forecast
  - Flow-dependent errors for deterministic forecast

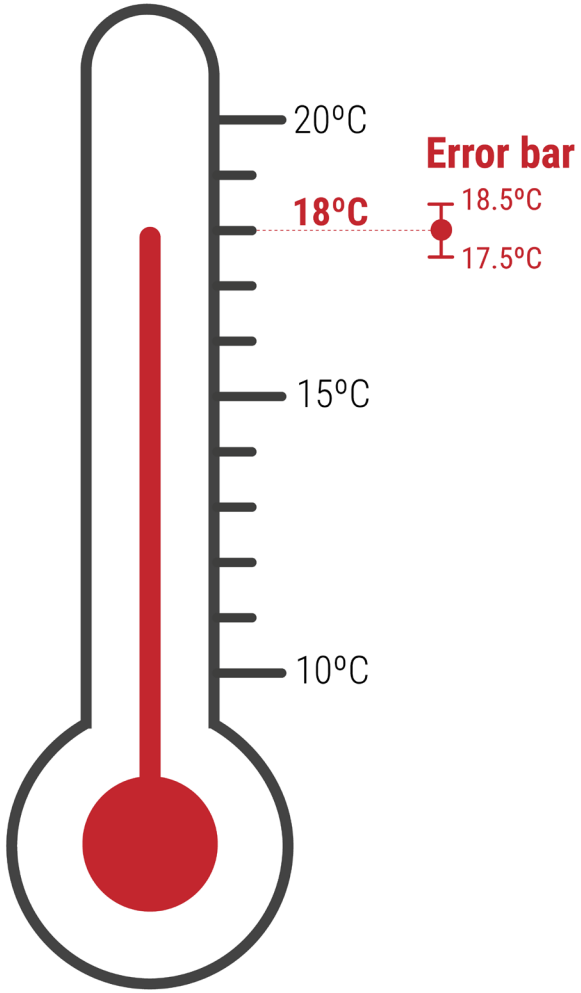
Benefit of additional data measured by reduction in variation across different members – "EDA spread" → reducing forecast/analysis uncertainties



EDA similar to information content study: how does new data improve our knowledge of the atmosphere?

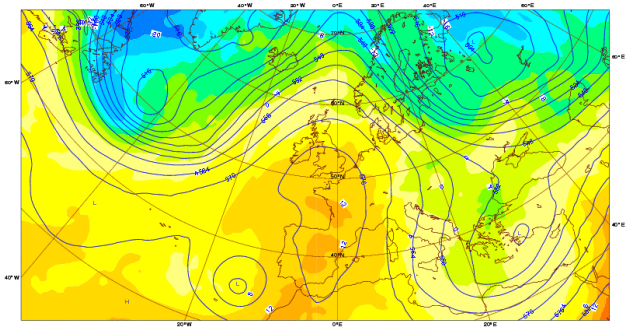
Predicting the value of future observations

- **We can't improve a forecast using simulated data!**
- **BUT we can estimate how new observations will reduce the "error bars" we assign to our forecasts**
- The EDA provides an estimate of the error bar
- Assumes errors of the simulated observations are realistic – these are used to perturb the observations
- Focus on spread changes at 12-hour forecast range – clearest signals from observation changes
- To avoid sampling errors, focus on larger area averages
- Originally developed ~20 years ago for doppler wind impact estimation and subsequently used for radio occultation, microwave sounders, in-situ…



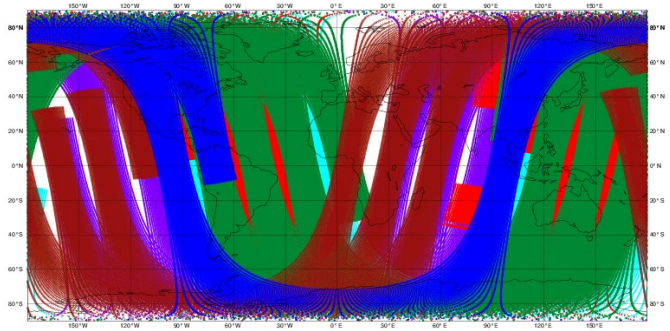
Observation simulation

ECMWF NWP analysis
→ proxy for the 'truth' for the new observations



interpolate

Observation time and location
e.g. sat zenith/azimuth angles etc.



Observation operator to convert model variables to observed quantities e.g. RTTOV

Add perturbations e.g. reflecting random instrument noise to create realistic characteristics

Simulated observations ready for assimilation

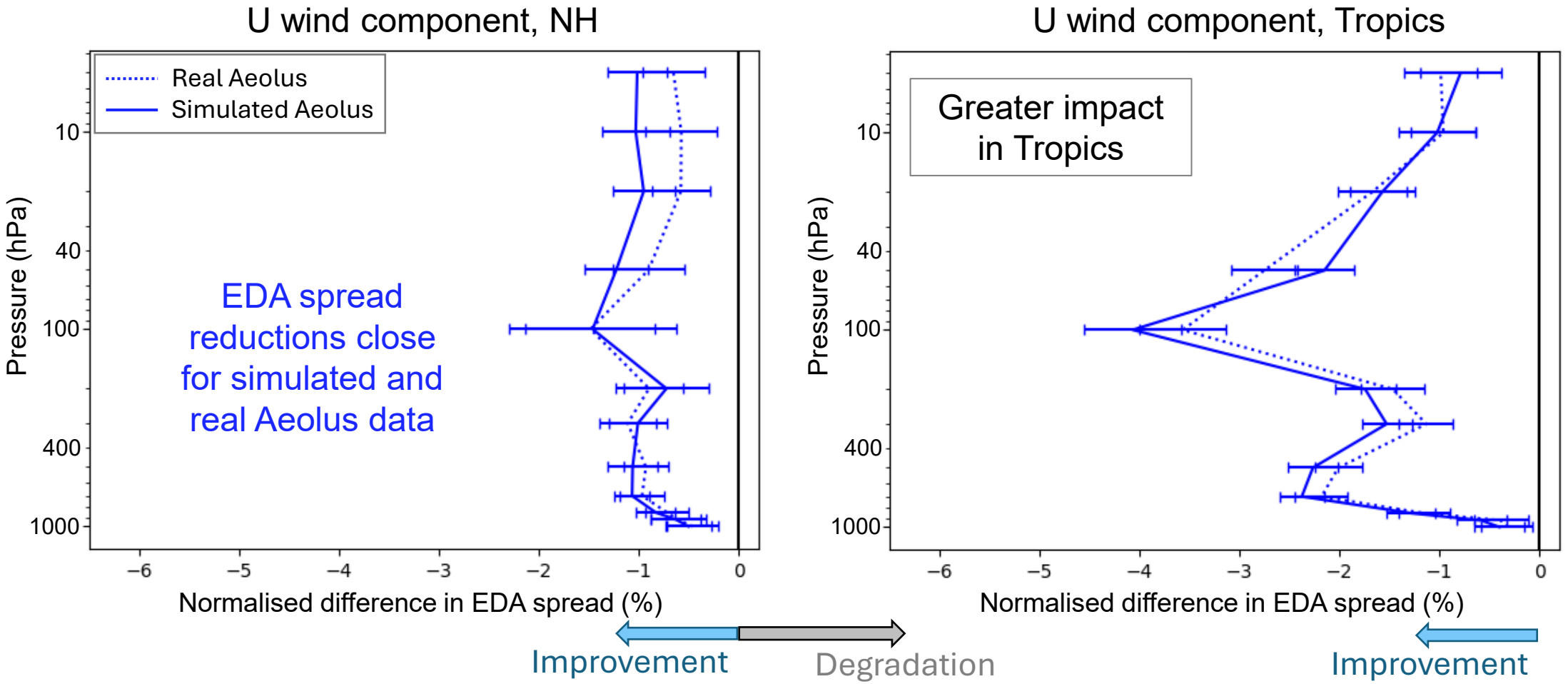
Accompanying observation errors

Start to finish: steps for assessment with the EDA

Preparation	Simulate future observations Modify up-to-date EDA configuration to ingest and assimilate new observations
Development	Define assimilation strategy and accompanying observation errors for new data Calibrate/context with changes to real observations
Experiment	Run separate EDA experiments (typically one-month period) that use: - baseline of real observations - scenarios that change real/simulated observation use
Analysis	Choose suitable set of diagnostics for assessing EDA spread reduction

Verifying agreement between real and simulated Aeolus impact

Study to support a follow-on mission to Aeolus (Doppler Wind Lidar)



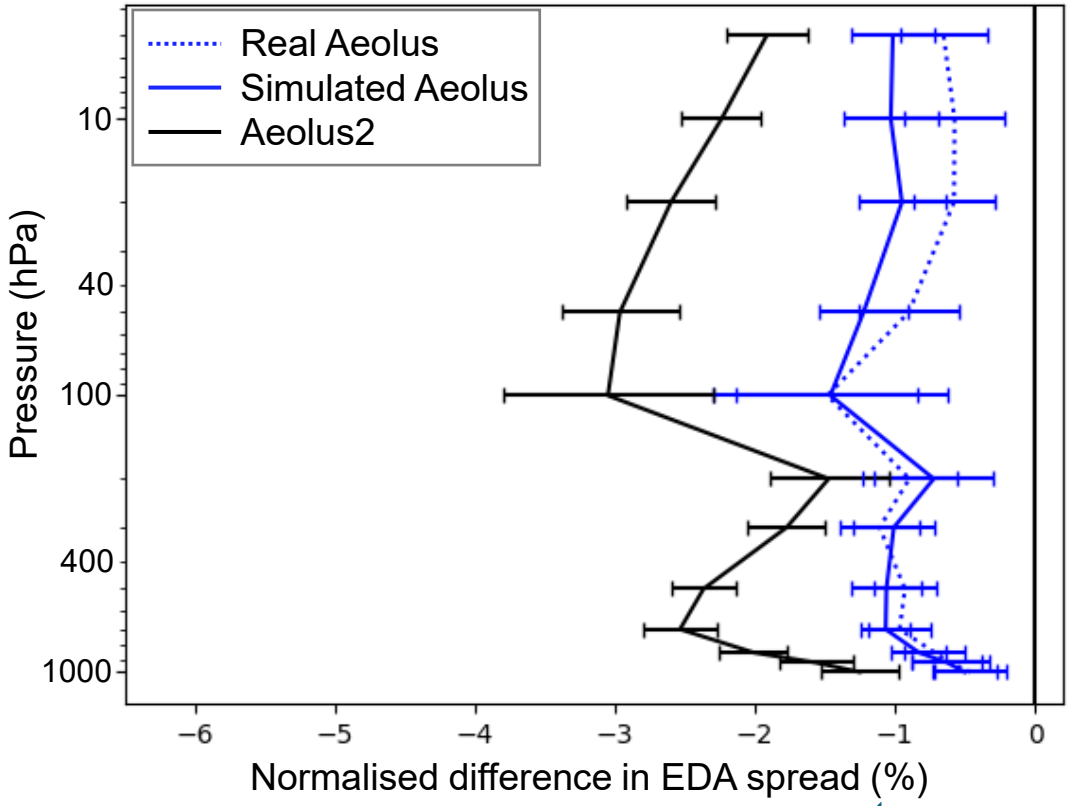
Changes relative to baseline of real observing system

EUMETSAT

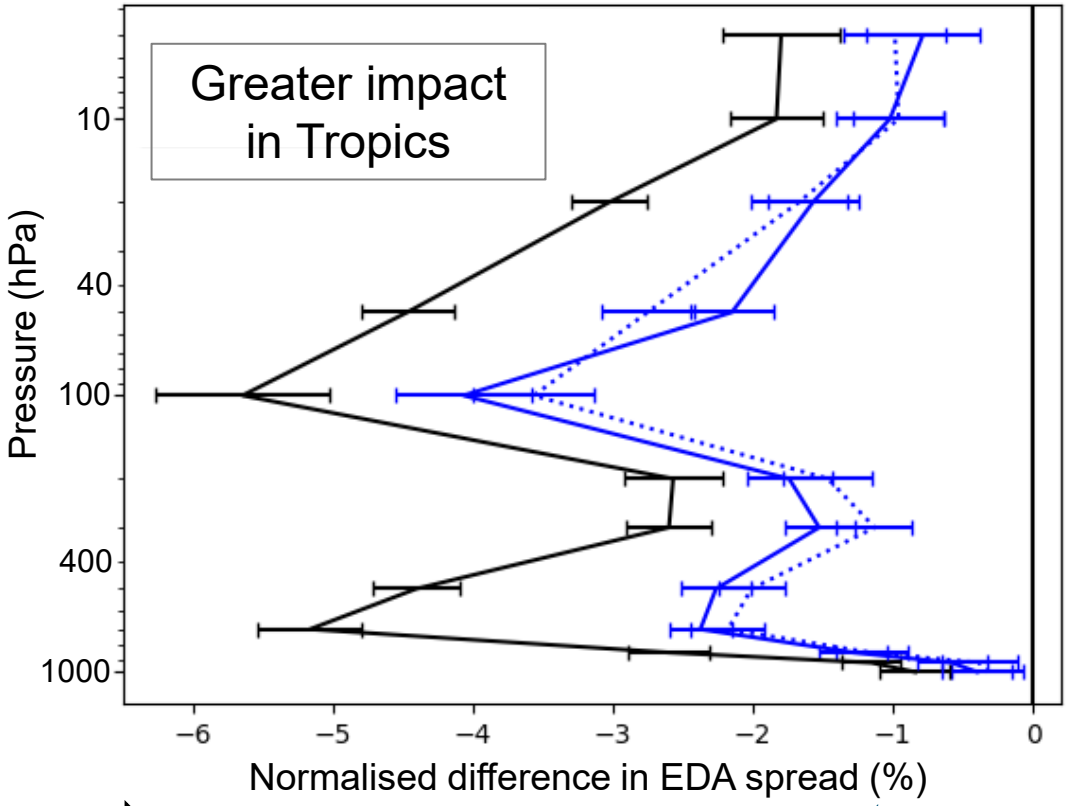
Significant additional benefit for potential Aeolus follow-on

Study to support a follow-on mission to Aeolus (Doppler Wind Lidar)

U wind component, NH



U wind component, Tropics



Improvement ⟵ ⟶ Degradation

Improvement ⟵

Improvements in design result in ~2-3 x impact of first Aeolus

EUMETSAT

Informing design aspects of constellations of small satellites

- Recent significant interest in technology on small satellites
- Aim to probe design aspects of constellations of small satellites carrying microwave (MW) sounders

No. of satellites: how does the benefit scale?

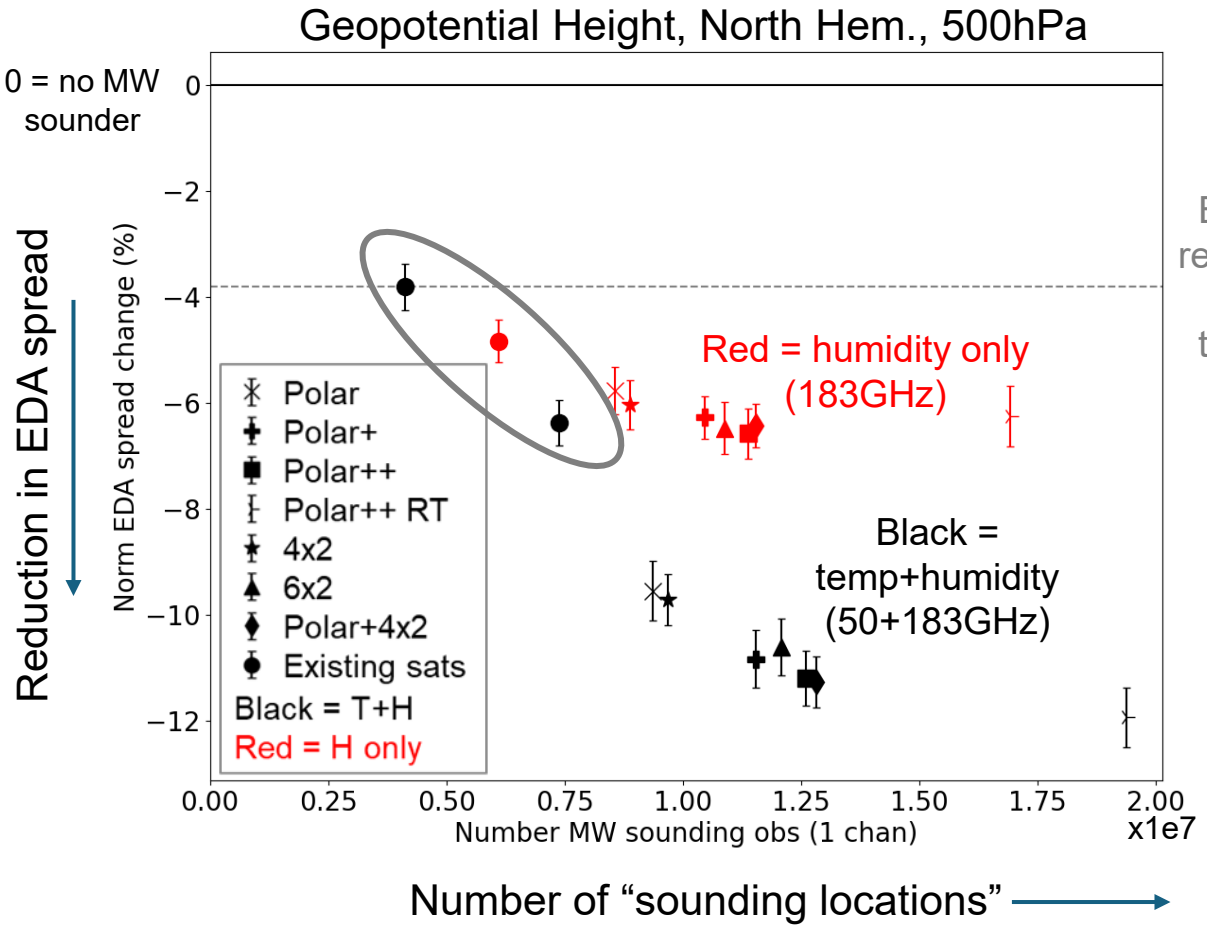
Frequency accommodation: Humidity sounding only (183GHz) or temperature + humidity (50 + 183GHz) sounding?

Types of orbit: polar only or combined with low inclination?

- Use small satellite based on real instrument with well-established set of channels
- Design set of scenarios that vary:
  - Number of satellites (from 8 to 20)
  - Test humidity only and full channel sets separately
  - Use a selection of polar only and polar/inclined orbit options

Larger spread reduction with more MW observations

8-28 June 2018



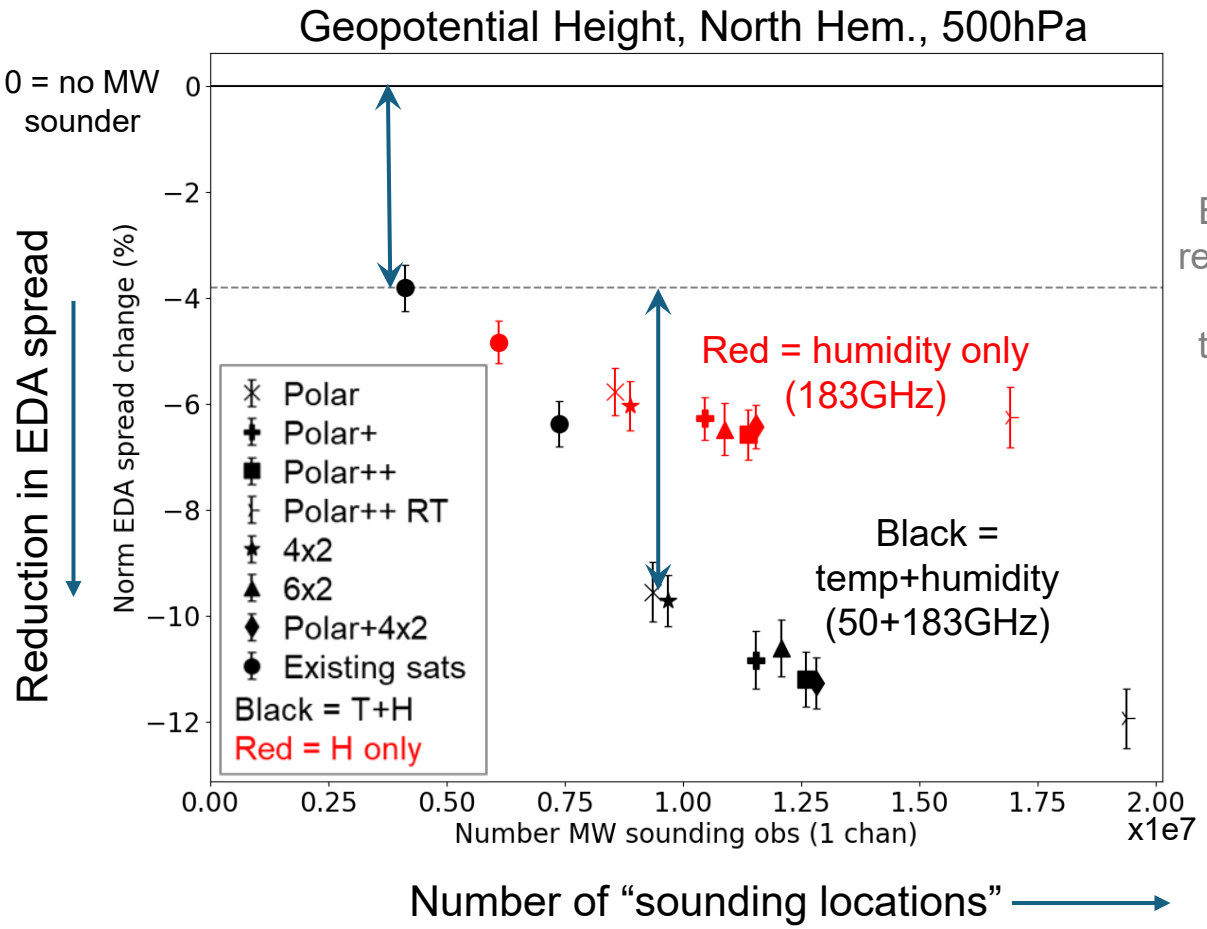
Baseline using real observations + 4 x real temp+humidity sounders

Clear additional benefit of temp sounding channels…
…But benefit from humidity possibly underestimated e.g. from lower resolution EDA spread calculation
Worth accommodating 50GHz temp channels where possible

Expts with existing data fit well with simulated data results

Larger spread reduction with more MW observations

8-28 June 2018



Baseline using real observations + 4 x real temp+humidity sounders

Significant impact of adding 8 sats - larger than that of 4 existing T+H sounders

Rate of spread reduction slowing for more obs
Benefit of more satellites gradually diminishes and may be less cost-effective

Clear additional benefit of temp sounding channels…
…But benefit from humidity possibly underestimated e.g. from lower resolution EDA spread calculation
Worth accommodating 50GHz temp channels where possible

Expts with existing data fit well with simulated data results

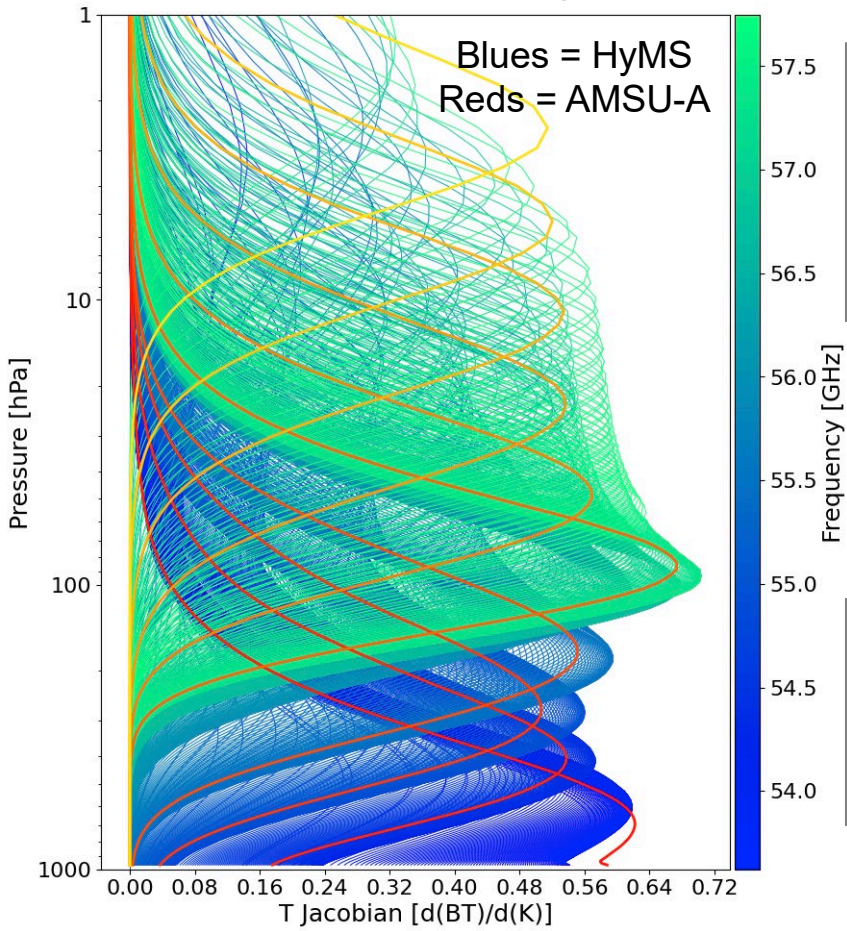
Application to new technology: hyperspectral MW sounders

- New technology provides significant departure from heritage MW sounders
- First study of global NWP impact
- Potential instrument based on real Hyperspectral Microwave Sounder (HyMS) but with more idealised configuration

Challenges:

- Development of new assimilation strategy – use of ~1000 channels vs ~17 for current sounders
- Channel usage – high-spectral resolution could provide highest information content but data volumes not practical
- Observation error modelling especially handling correlations between channels

Vertical sensitivity of 845 assimilated temp-sounding channels



Finer spectral sampling captures small absorption features

More diverse vertical sampling of atmosphere

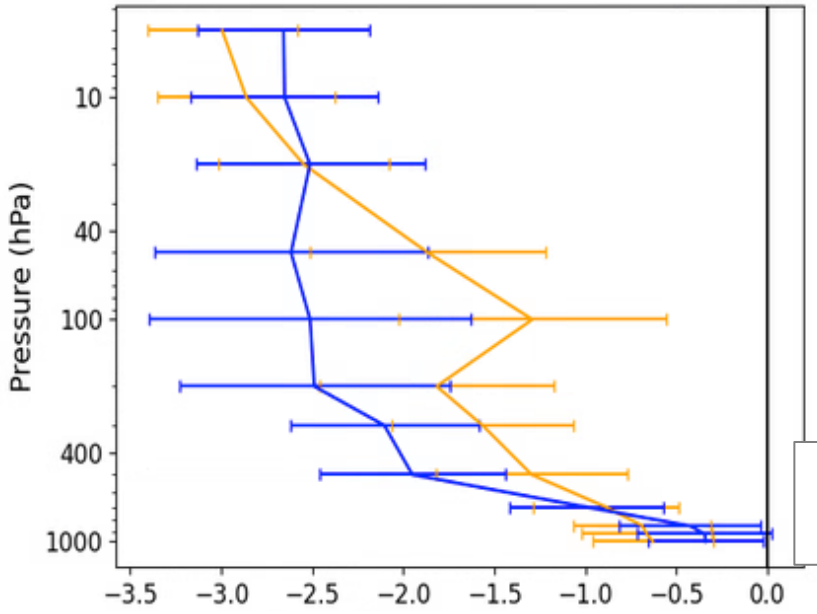
Covers new parts of frequency spectrum not yet exploited by current MW

Positive impact from hyperspectral MW sounders

Early results show hyperspectral MW achieves similar or even greater positive impact than heritage sensor

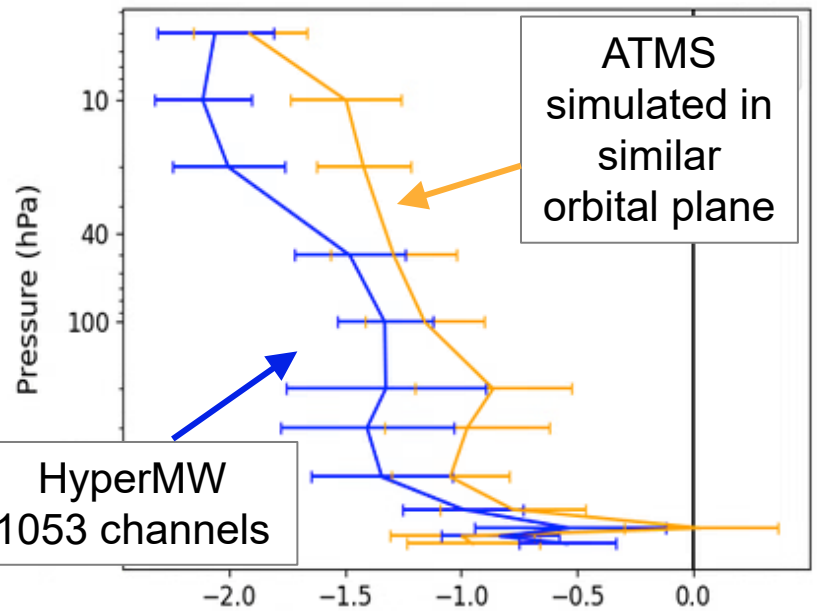
Quality control applied conservatively guided by current restrictions on current frequencies

Geopotential height, NH



Normalised difference in EDA spread (%)

Temperature, SH



Normalised difference in EDA spread (%)

Improvement

Investigation into impact sensitivity to observation error setting resulted in indirectly accounting for correlated observation errors (inflating by 100%)

Ongoing work to determine relative impact from better spectral coverage vs. higher spectral resolution

Extending the ensemble method to longer lead times

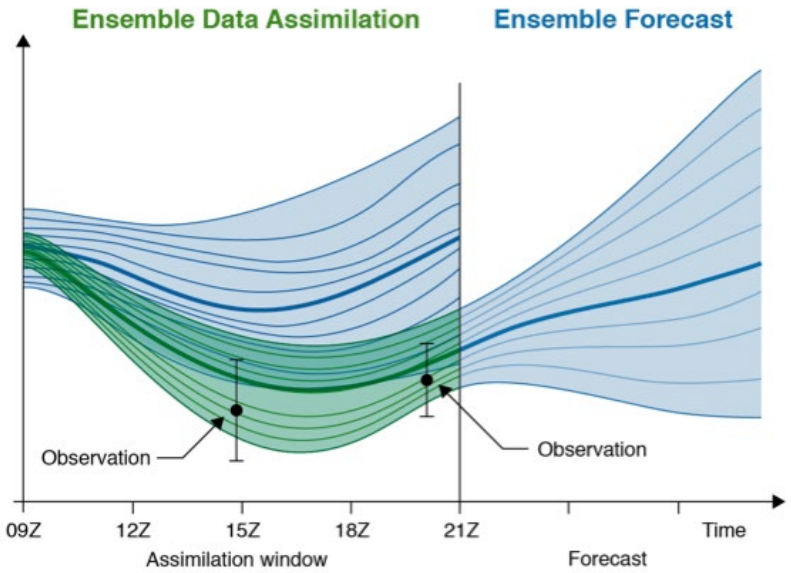
- Possible to run longer forecast ranges from EDA for spread calculations

Or…

- Ensemble forecasting system (ENS) with the complete initial conditions perturbations, including:
  - re-centring of the EDA perturbations (scale-dependent re-centring on the deterministic analysis)
  - singular vectors (initial condition perturbations in the directions of rapid growth in uncertainty)
- Combine EDA and ENS systems to potentially produce more reliable error statistics at medium-range

In both cases:

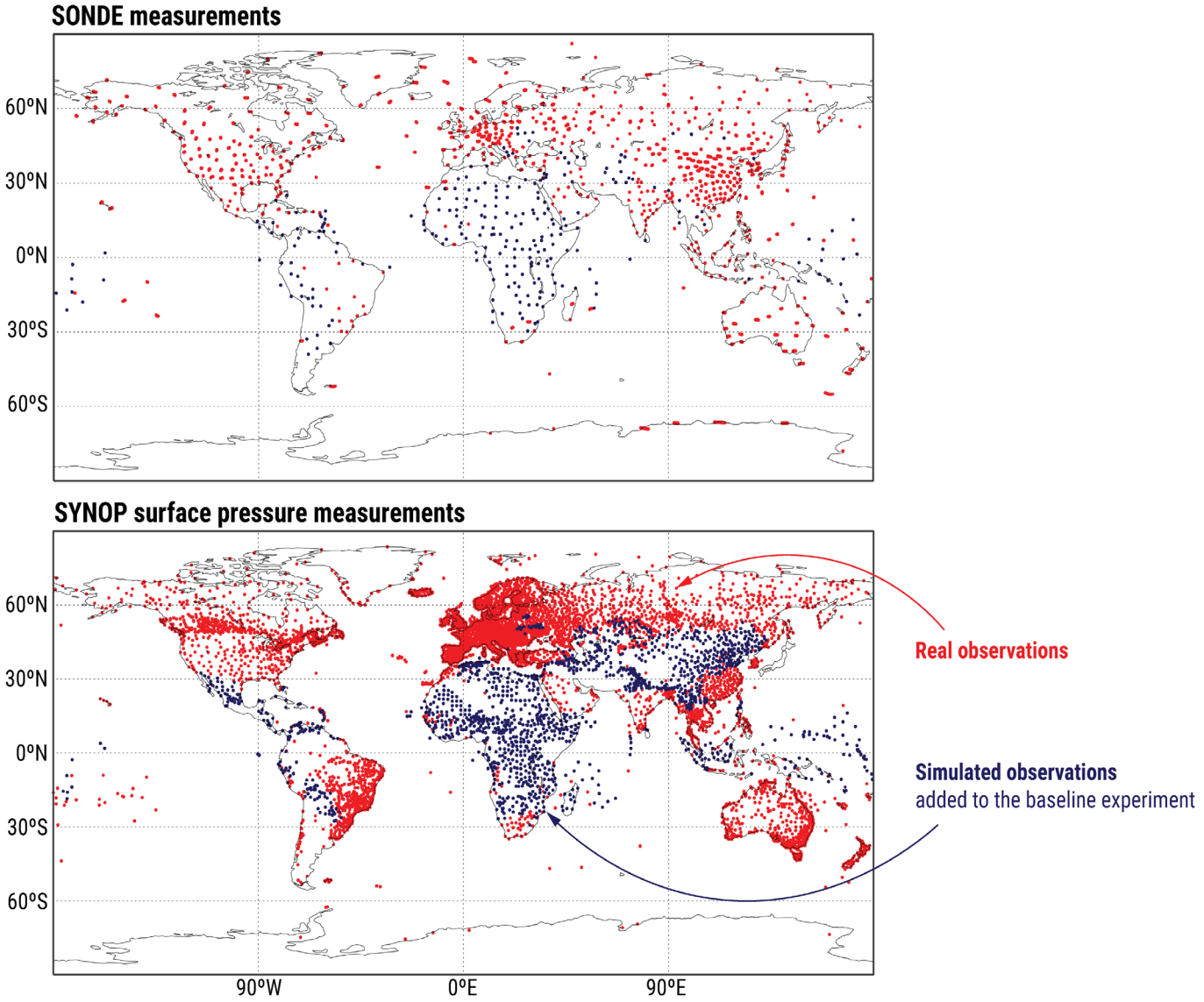
- Run for longer experiment period to get more robust statistics
- Lower sampling errors for investigating smaller geographical areas e.g. for tropical cyclone development



2 different systems:
EDA members provide input to ensemble (ENS) forecasting system

Expanding coverage of surface, upper-air and marine stations

- Aim to advise on targeted expansion of surface land stations/radiosondes/buoys (Global Basic Observing Network [GBON] network)
- Short-range EDA spread reduction showed that observations that filled “gaps” in coverage were most important
- Short-range results identify 2 most promising candidates
- Recent work to extend to evaluate impact at medium-range lead times (3-10 days)

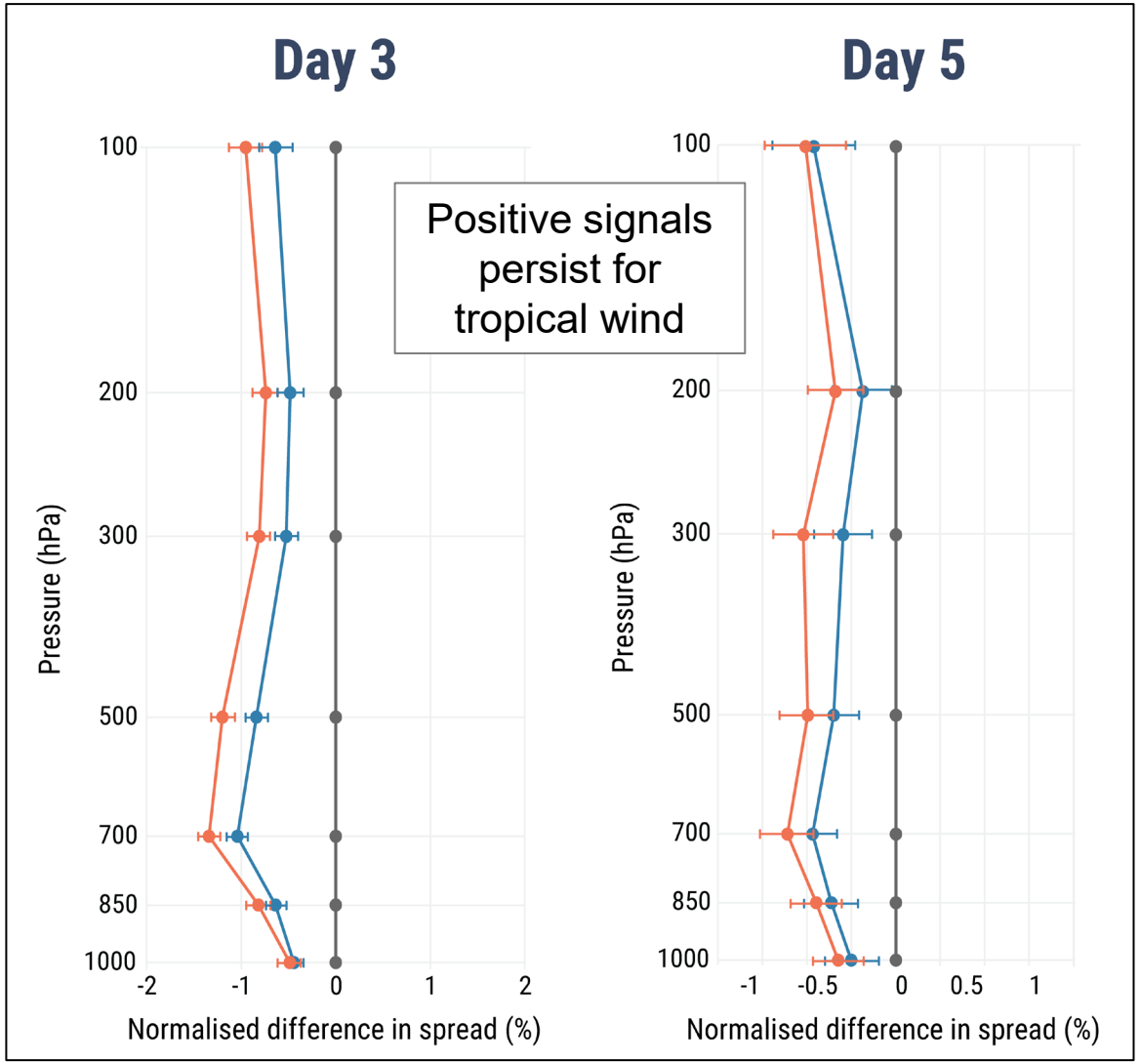


SOFF
Systematic Observations Financing Facility

Clear medium-range impact in tropical wind

- Forecast accuracy improvements are expected to extend into the medium-range in the Tropics to around day 7
- Similar, statistically significant improvements found for both surface and upper-air parameters
- Using EDA alone or EDA+ENS configuration produced similar relative reductions

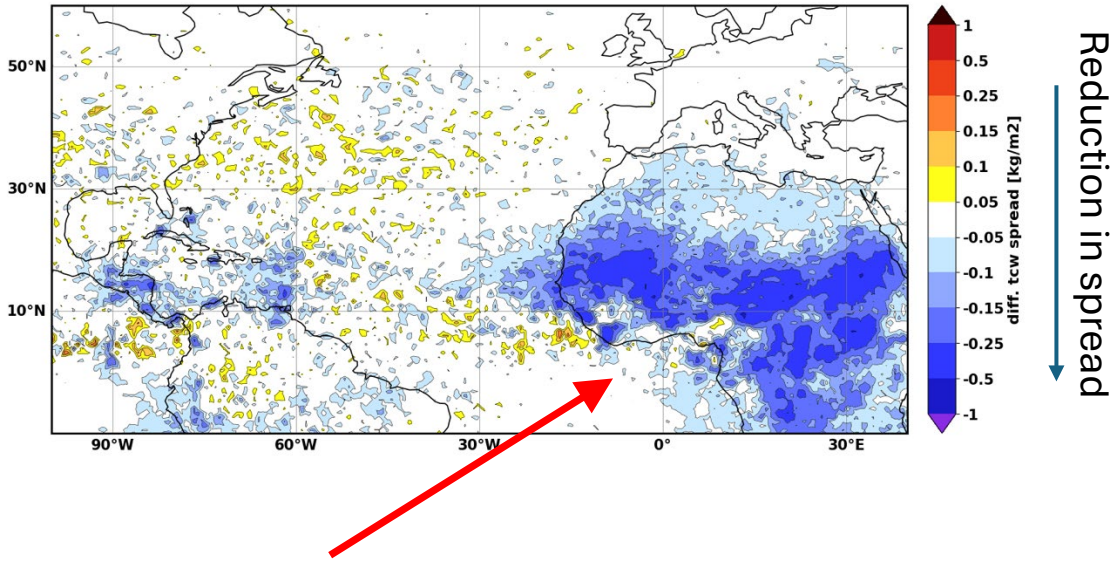
Spread reduction for zonal wind



Control
Scenario 1 (less dense)
Scenario 3 (more dense)

Westward propagation of information from Africa into the Atlantic

Total Column Water at **0h**



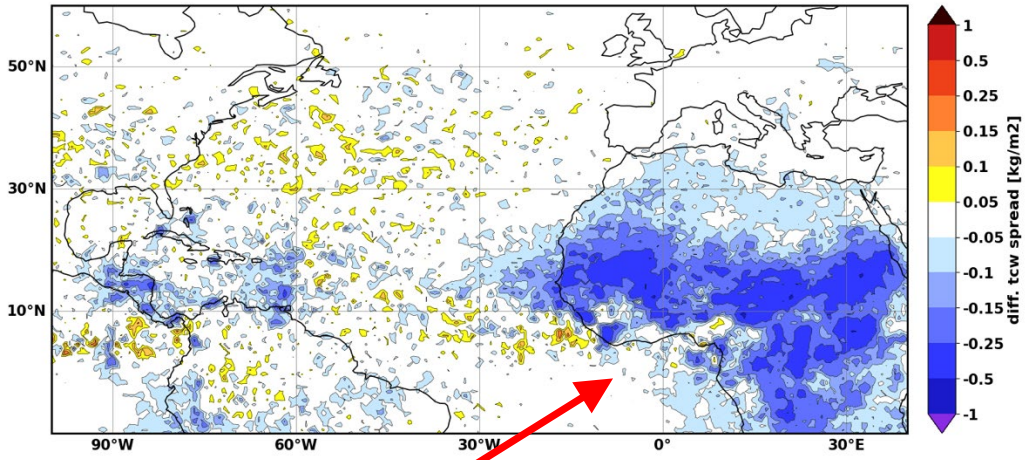
Improvements in Africa…

TCW can be used as a proxy for precipitation

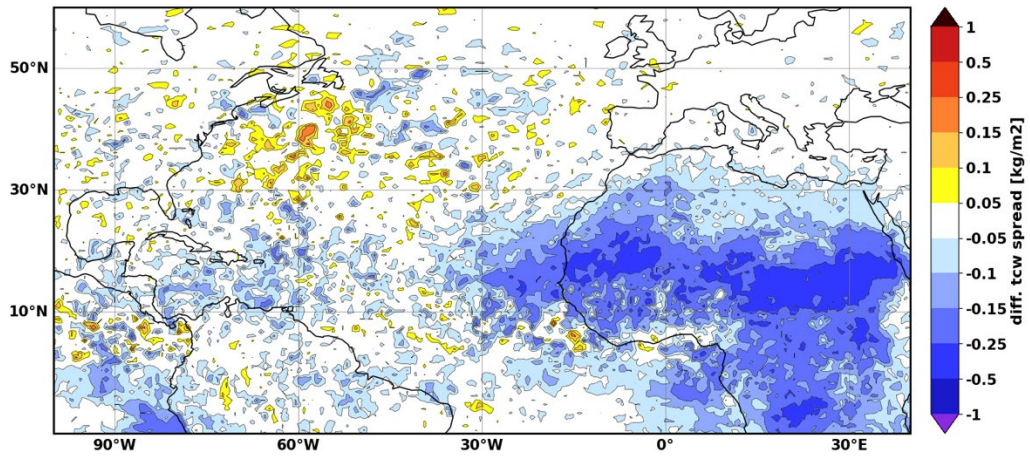
Important factor in African Easterly Waves which can feature organised convective precipitation

Westward propagation of information from Africa into the Atlantic

Total Column Water at **0h**



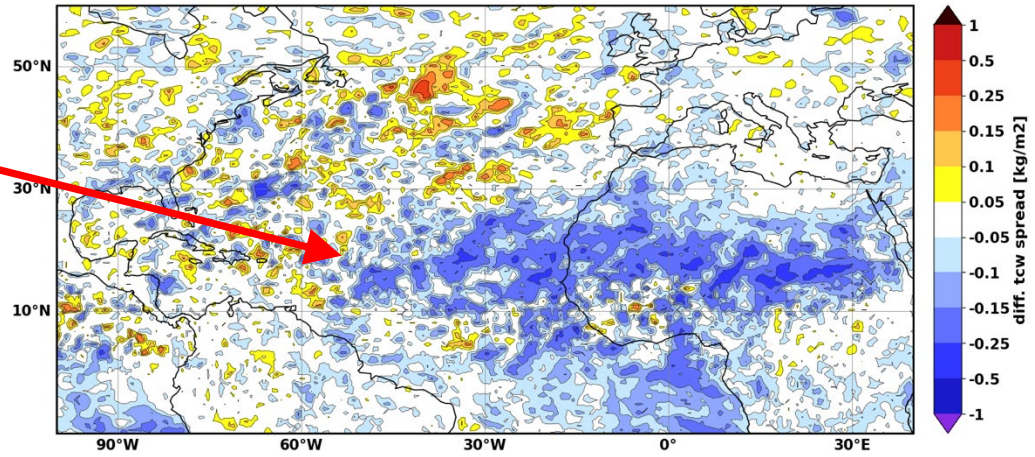
Total Column Water at **+24h**



Reduction in spread

Improvements in Africa…
…lead to improvements in the Atlantic as the forecast range increases

Total Column Water at **+72h**

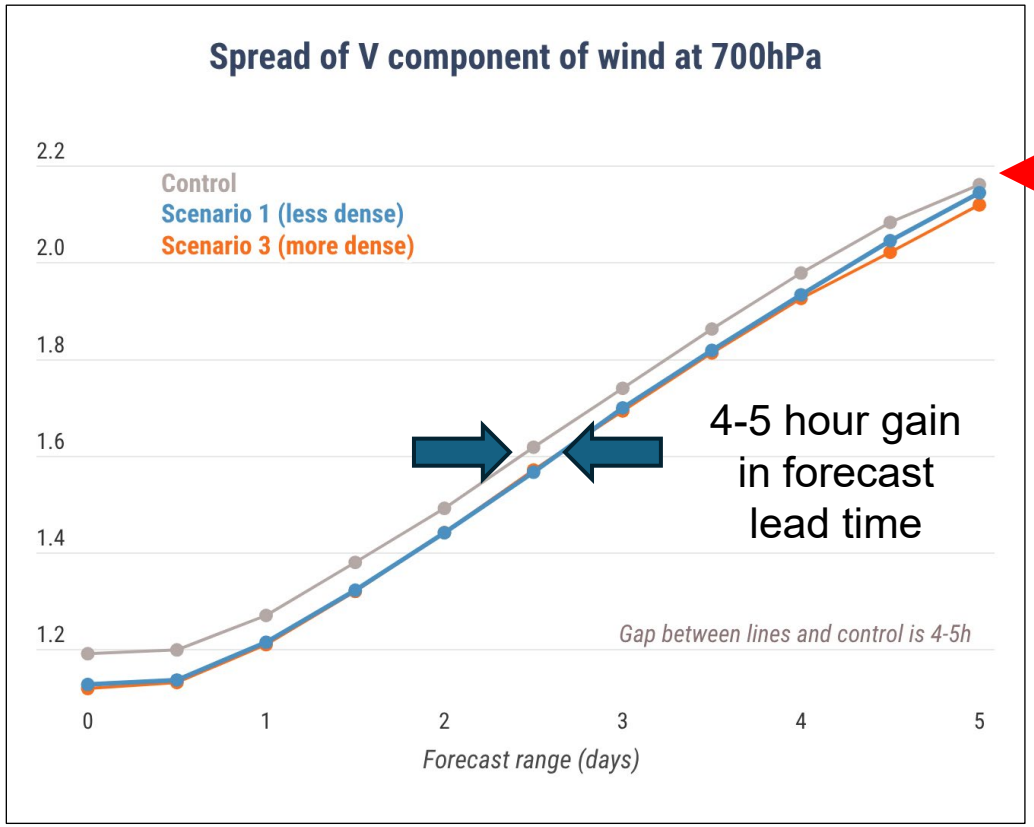


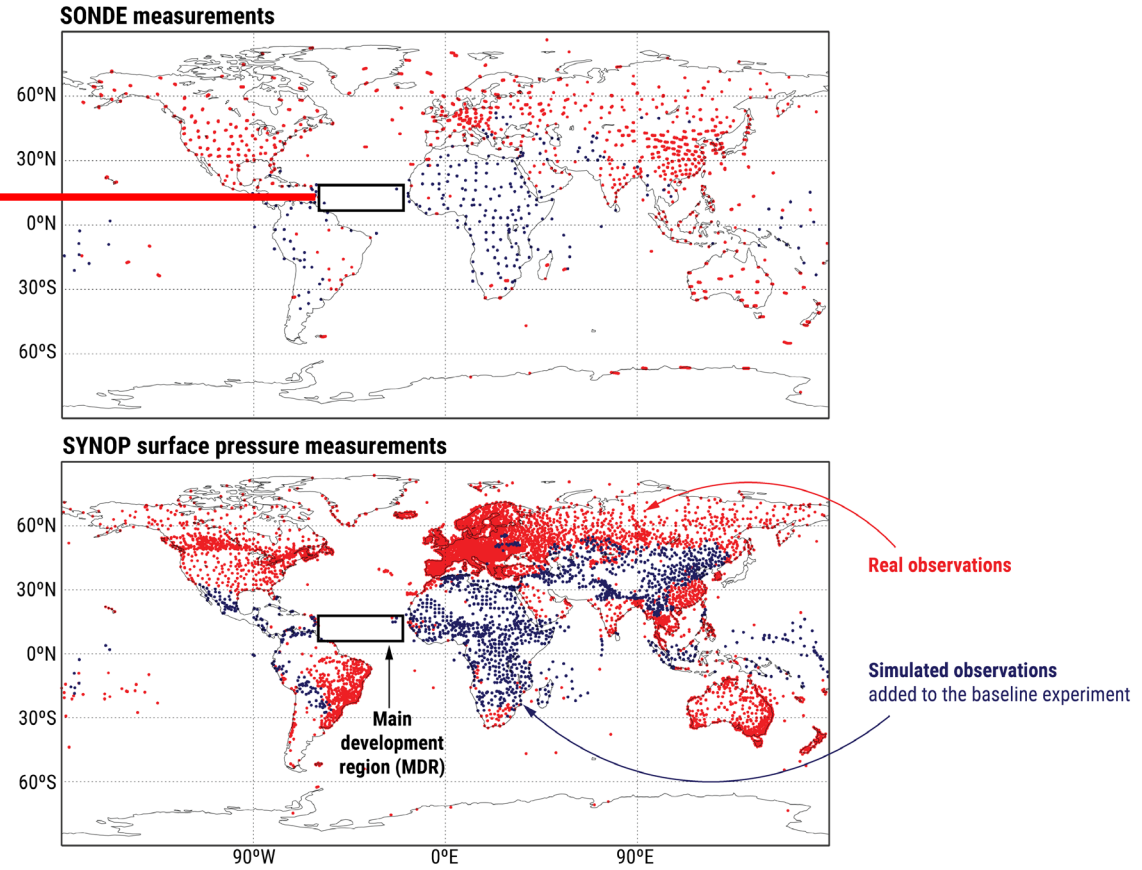
TCW can be used as a proxy for precipitation

Important factor in African Easterly Waves which can feature organised convective precipitation

Improvements in the Atlantic “main development region”

Reduction in spread from adding surface/upper air sites





Around 60% of Atlantic tropical cyclones develop from African Easterly Waves
Meridional wind at 700hPa useful AEW diagnostic
Around **4-5 hour gain in forecast lead time** for parameters that often “seed” tropical cyclone development

Translating to forecast impact

- Good evidence of relationship between EDA spread changes and forecast error reduction…
- …But no neat one-to-one relation to OSE style metrics
- Impacts usually in relative terms to real observations
- Established that real and equivalent simulated data can have similar EDA impacts
- Where technology is similar, possible to draw parallels between spread reduction and forecast impact

Expt comparison	Z 500hPa EDA spread reduction	Z 500hPa T+48h OSE impact
3 x real temp+humidity sounder	~4%	~4%
6 x small satellites	~4%	~4%?

- Ongoing work to investigate medium-range impacts

EDA advantages

Operationally maintained system

Simulation of future data only

Array of key impact variables available, tailored to observation type



EDA limits

Complicated connection to more tangible forecast impact measures

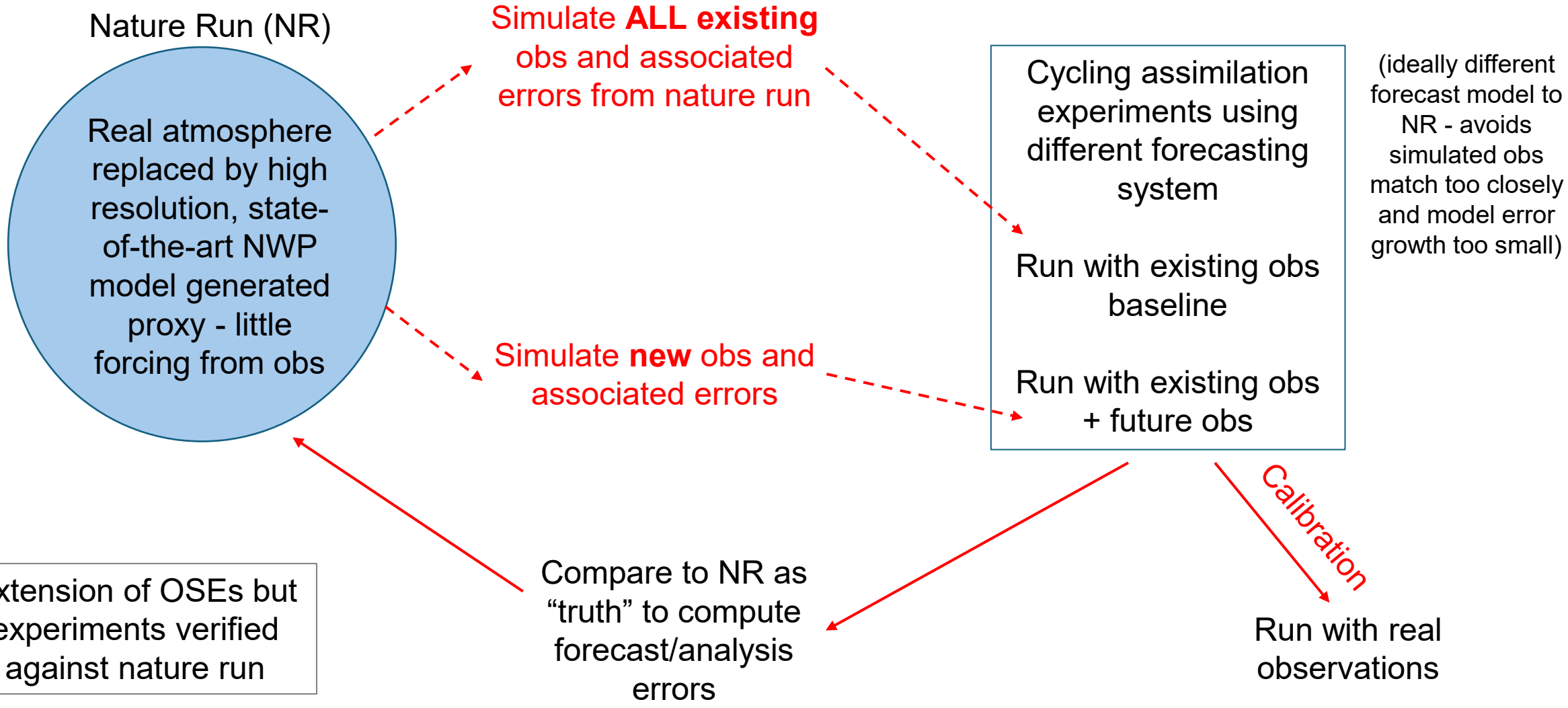
Ensemble sampling noise issues prevent analysis of fine scale features

Relies on accurate estimates of observation error for simulated data

Short-range impact cheaper and faster but for medium-range process more expensive and slower

Observing System Simulation Experiments (OSSEs)

OSSE set up


Nature Run (NR)
Real atmosphere replaced by high resolution, state-of-the-art NWP model generated proxy - little forcing from obs
Simulate **ALL existing** obs and associated errors from nature run
Simulate **new** obs and associated errors
Cycling assimilation experiments using different forecasting system
Run with existing obs baseline
Run with existing obs + future obs
(ideally different forecast model to NR - avoids simulated obs match too closely and model error growth too small)
Calibration
Run with real observations
Compare to NR as "truth" to compute forecast/analysis errors
Extension of OSEs but experiments verified against nature run

Start to finish: steps for running an OSSE

Preparation/ Development	Produce Nature Run (NR) – “true state” reference **All** observations (real and future) simulated from NR Develop assimilation strategy and accompanying obs errors for future data
Calibration	Add simulated measurement errors to all observations Calibrate for realistic impacts
Experiment	Run assimilation experiments with: - Different observation scenarios of all simulated obs with and without new data
Analysis	Diagnostics can be similar to OSEs using NR as reference (-> see previous talk by Fabien Carminati)

Calibration to ensure realistic impacts

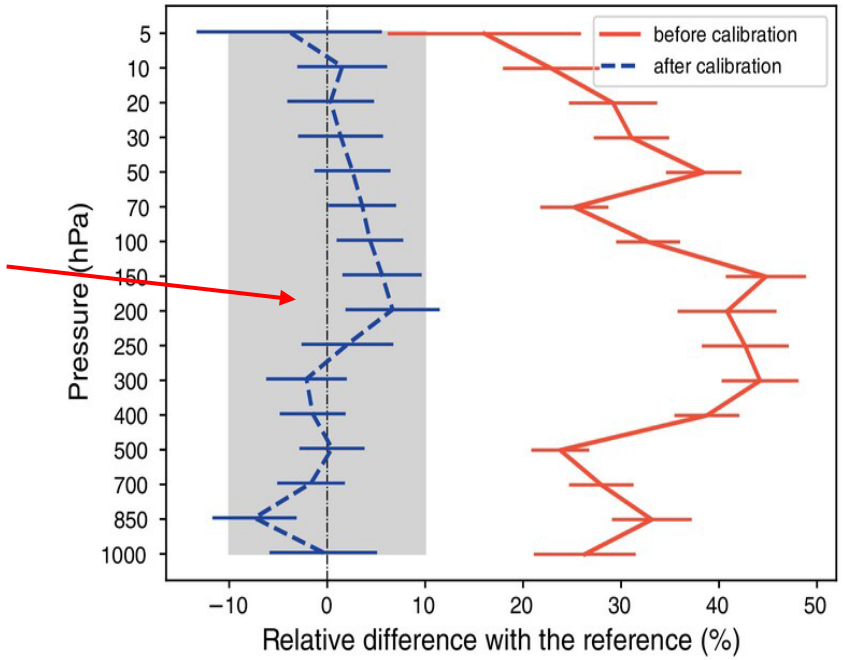
- Aim to achieve same weighting in assimilation for simulated obs as real counterparts

Possible options:

- Tune perturbations to get good agreement of first-guess departure statistics between real and simulated cases

Zero reference = first guess departure std. dev. for real obs

Simulated radiosonde observation first guess statistics agree well with real obs after calibration



(b) radiosondes, temperature

Figure from Revoire et al. 2024
https://doi.org/10.1002/qj.4747

Calibration to ensure realistic impacts

- Aim to achieve same weighting in assimilation for simulated obs as real counterparts

Possible options:

- Tune perturbations to get good agreement of first-guess departure statistics between real and simulated cases
- Obtain realistic forecast impacts assimilating real or simulated dataset
- Compare forecast-error growth OSSE vs. nature run, OSE vs. reference
- Horizontal correlations of departure statistics
- Number of observations assimilated
- Contributions to the cost function

Can be a long process but once well calibrated, system can be reused

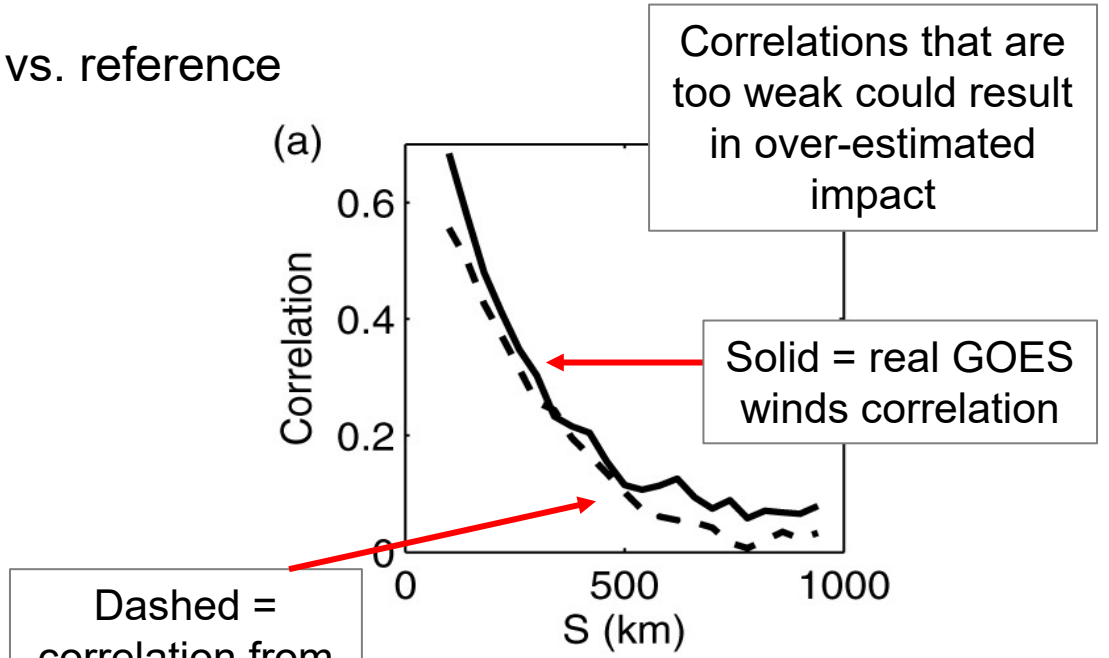


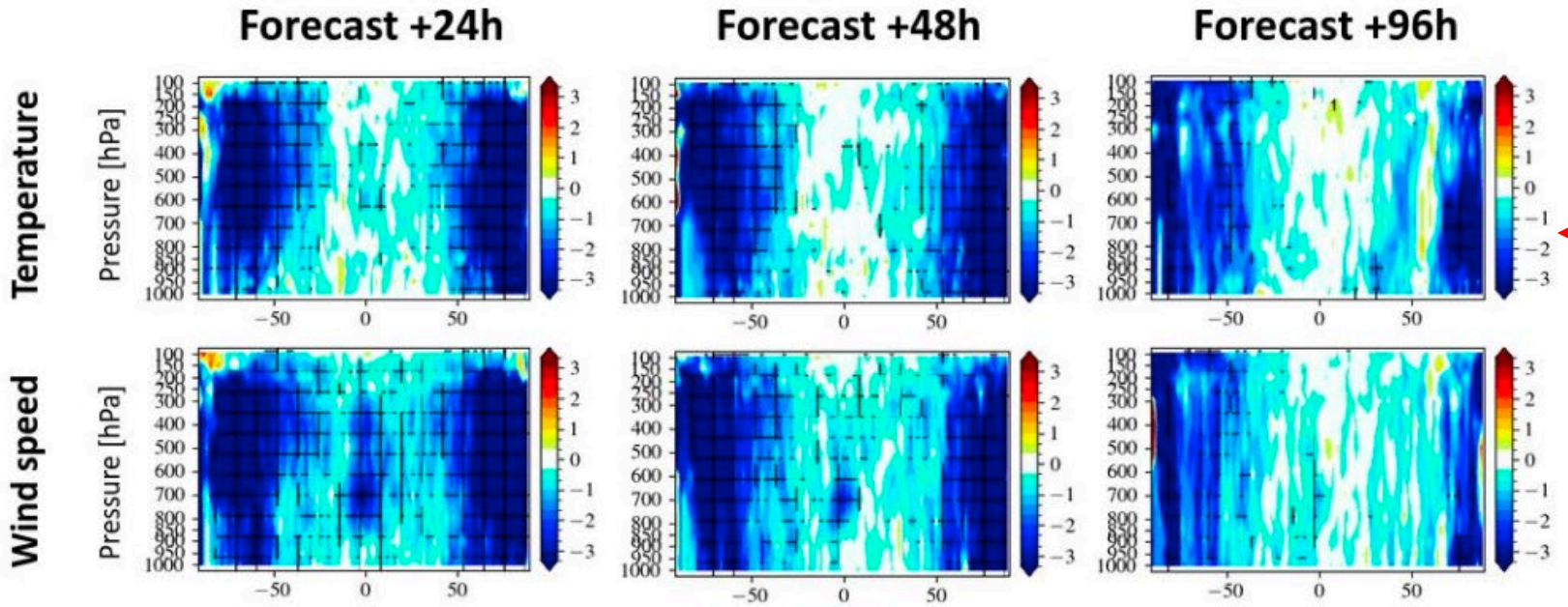
Figure from Enrico et al. 2013
https://doi.org/10.1002/qj.2027

Significant impact from WIVERN - Doppler Cloud Radar

WIVERN produces Horizontal Line Of Sight (HLOS) winds from Doppler Radar operating at 94GHz

Assessment to support selection as Earth Explorer 11 mission

Difference of standard deviation of forecast errors (2 x 3 months)



Positive impact up to 4 days ahead

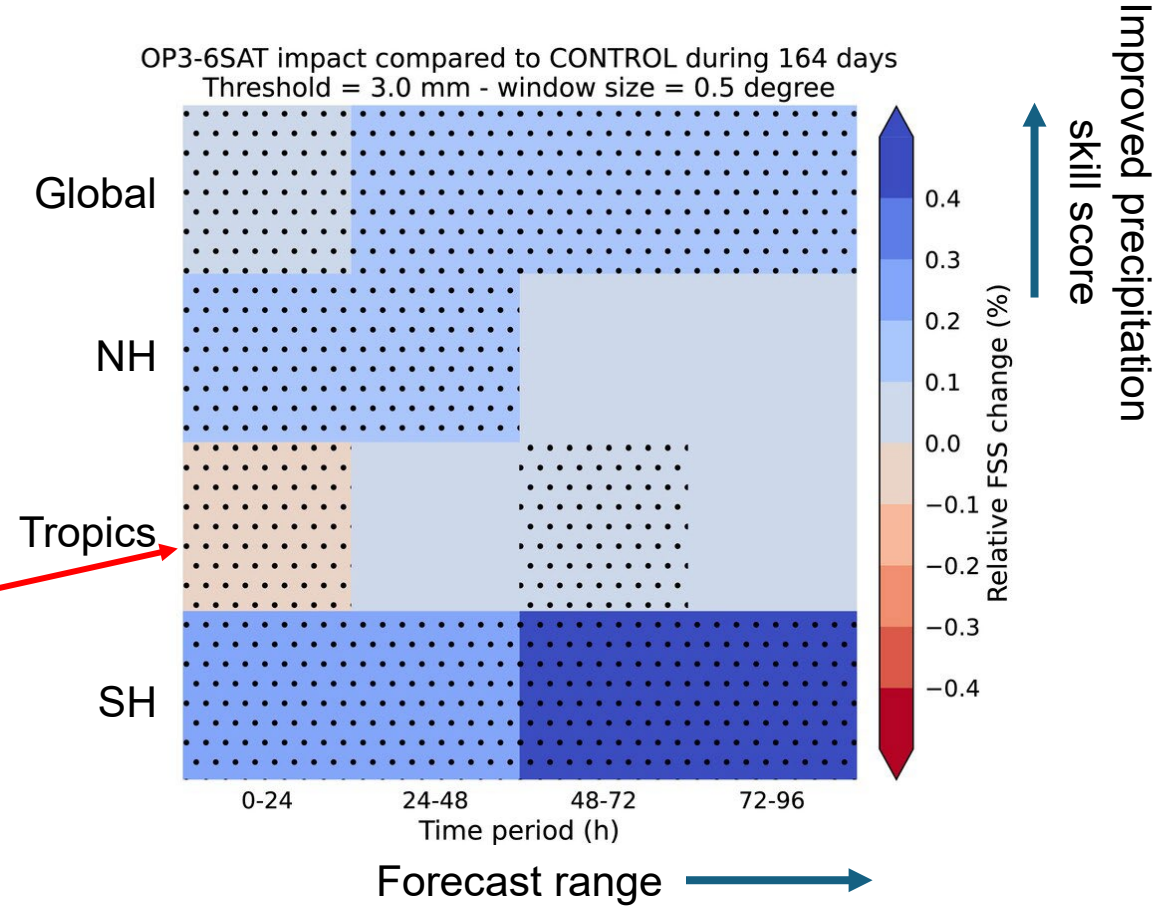
Lower impact in tropics for mass variables

Most beneficial in polar regions/extra-tropics

Time to reach threshold on forecast error indicates gain in lead time of ~40mins in NH, 1h20mins in SH - important for decision making in extreme events

Precipitation improvement from microwave sounders

OSSE to support the funding case for EPS-Sterna: constellation of small satellites carrying microwave sounders



Short-range degradation in tropics linked with humidity field degradation

Likely due to issue in representation of convection in NWP model

Global positive impacts on precipitation from MW constellation

Similar results for constellations with 3-8 satellites – precipitation harder to improve with more data

Figure from Revoire et al. 2024
https://doi.org/10.1002/qj.4747

OSSE advantages

Diagnostics intuitive with full array of variables similar to OSE

Provides long-range lead times (need long experiment for statistical significance)

Possibility to investigate case studies and high-impact events



OSSE limits

Simulation of all observations and accurate estimates of errors required

Initial calibration is difficult and needs care

NR and separate assimilation system takes a lot of resource and needs expert knowledge in data assimilation/modelling

Harder to maintain and keep up-to-date with latest NWP

EDA and OSSE methods: a complementary pairing





Individually powerful and complementary together

- EDA and OSSE methods both provide a wealth of information with different perspectives

EDA considers the uncertainty in the forecast – the change in the error bar on the forecast variable

OSSE evaluates a specific realisation of forecast error (compared to “truth” from nature run) – the change in the absolute error on the forecast variable

- Individually can provide important insights on future observations
- Together provide opportunity to probe different aspects and corroborate conclusions with very different systems

A case for EPS-Sterna: Assessing MW sounders with EDA and OSSEs

FSOI-like metric: Impact of small satellites against loss of Metop-B

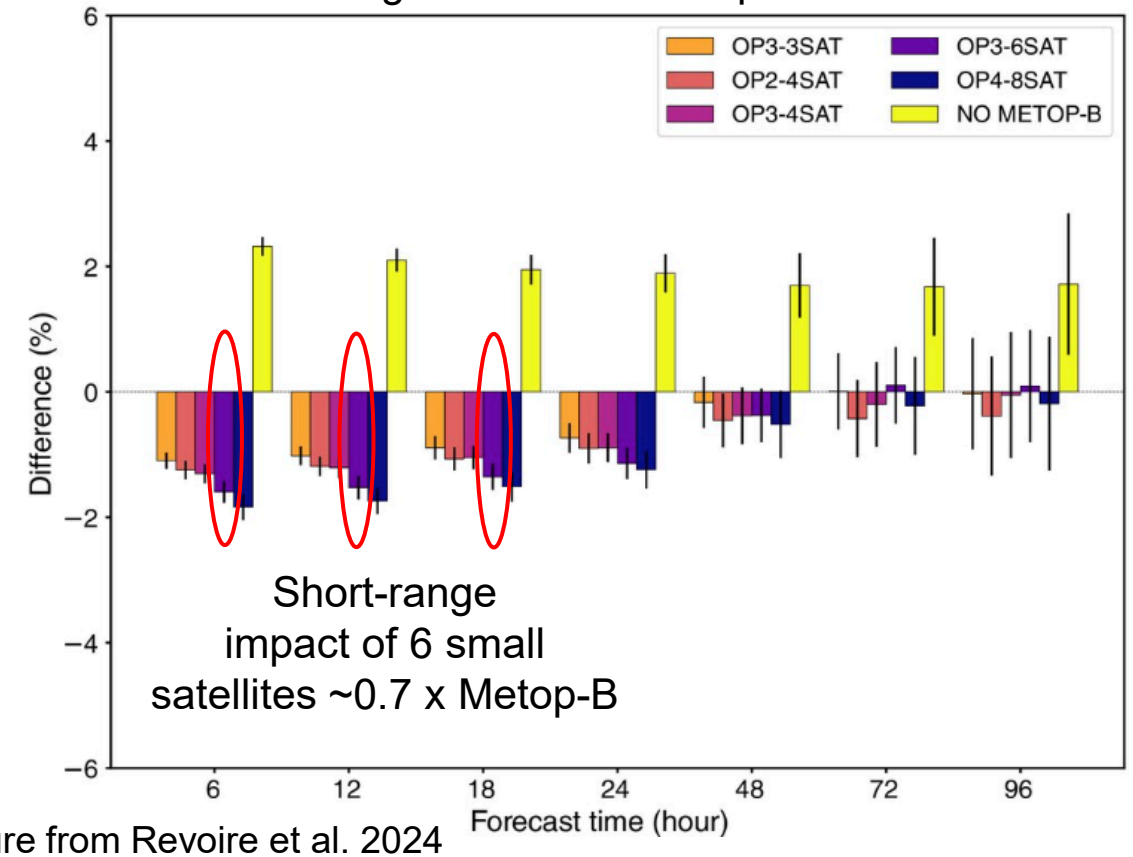
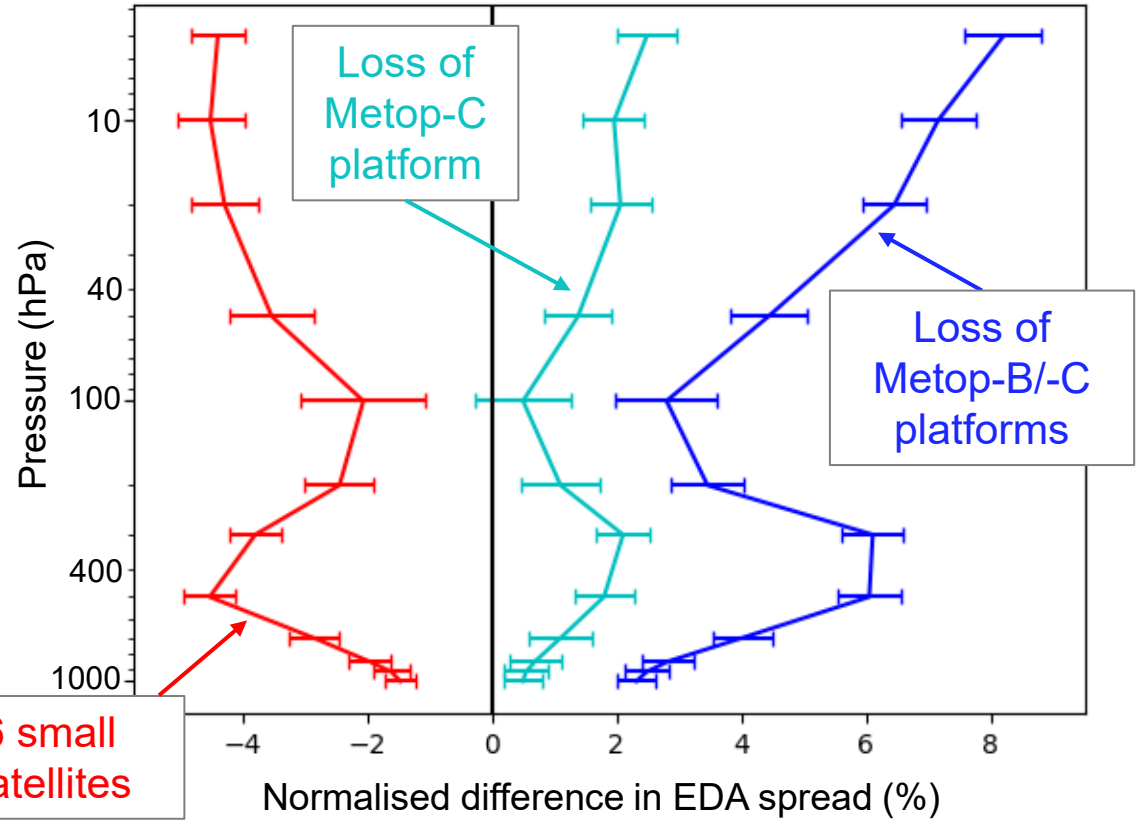


Figure from Revoire et al. 2024
https://doi.org/10.1002/qj.4747

Geopotential height, NH



Both studies conclude that relative to a full Metop platform, a constellation of small satellites carrying MW sounders provides significant impact

Funding for EPS-Sterna recently approved!

Common challenges

- Calibration to real observation impacts is difficult
- Careful interpretation is needed, keeping in mind the method limitations
- Deciding on the selection of relevant diagnostics to focus on – traditional key NWP variables and tailoring to specific target areas of new data
- Choosing suitable experiment periods and baseline of observations
- **Limited by the sophistication of the current data assimilation system**

The future of looking at the future

- Forecasting systems evolve (incl. physics-based NWP systems), and any such studies only reflect the current status
- Physics-based arguments identifying observations gaps for totally novel observations are important
- AI has the potential to change the impact of (future) observations
- AIFS and DOP systems highlight the value of current high-quality observations
- Could eventually open different avenues to doing these kinds of studies
- In the future we will need a good mix of observations to make good forecasts - physics-based modelling such as in the EDA/OSSE is currently still the best way to guide such developments

- "Prediction is very difficult, especially if it's about the future."
  — *Niels Bohr (Nobel Laureate in Physics)*

References

- Errico, R.M., Yang, R., Privé, N.C., Tai, K.-S., Todling, R., Sienkiewicz, M.E. and Guo, J. (2013), Development and validation of observing-system simulation experiments at NASA's Global Modeling and Assimilation Office. Q.J.R. Meteorol. Soc., 139: 1162-1178. https://doi.org/10.1002/qj.2027
- Rivoire, L., Marty, R., Carrel-Billiard, T., Chambon, P., Fourrié, N., Audouin, O., et al. (2024) A global observing-system simulation experiment for the EPS–Sterna microwave constellation. *Quarterly Journal of the Royal Meteorological Society*, 150(762), 2991–3012. Available from: https://doi.org/10.1002/qj.4747
- Rivoire, L., Marty, R., Carrel-Billiard, T., Chambon, P., Fourrié, N., Audouin, O., et al. (2024) A global observing-system simulation experiment for the EPS–Sterna microwave constellation. *Quarterly Journal of the Royal Meteorological Society*, 150(762), 2991–3012. Available from: https://doi.org/10.1002/qj.4747
- Shrader, Jeffrey & Bakkensen, Laura & Lemoine, Derek. (2023). Fatal Errors: The Mortality Value of Accurate Weather Forecasts. SSRN Electronic Journal. 10.2139/ssrn.4491219.

Thank you!



www.ecmwf.int