Forcings for reanalysis and initialised predictions

Tim Stockdale, ECMWF

Workshop: Pathway to regular and sustained delivery of climate forcing datasets, 28-31 October 2024

ECMWF

Re-analysis and seasonal (and decadal) predictions need forcings

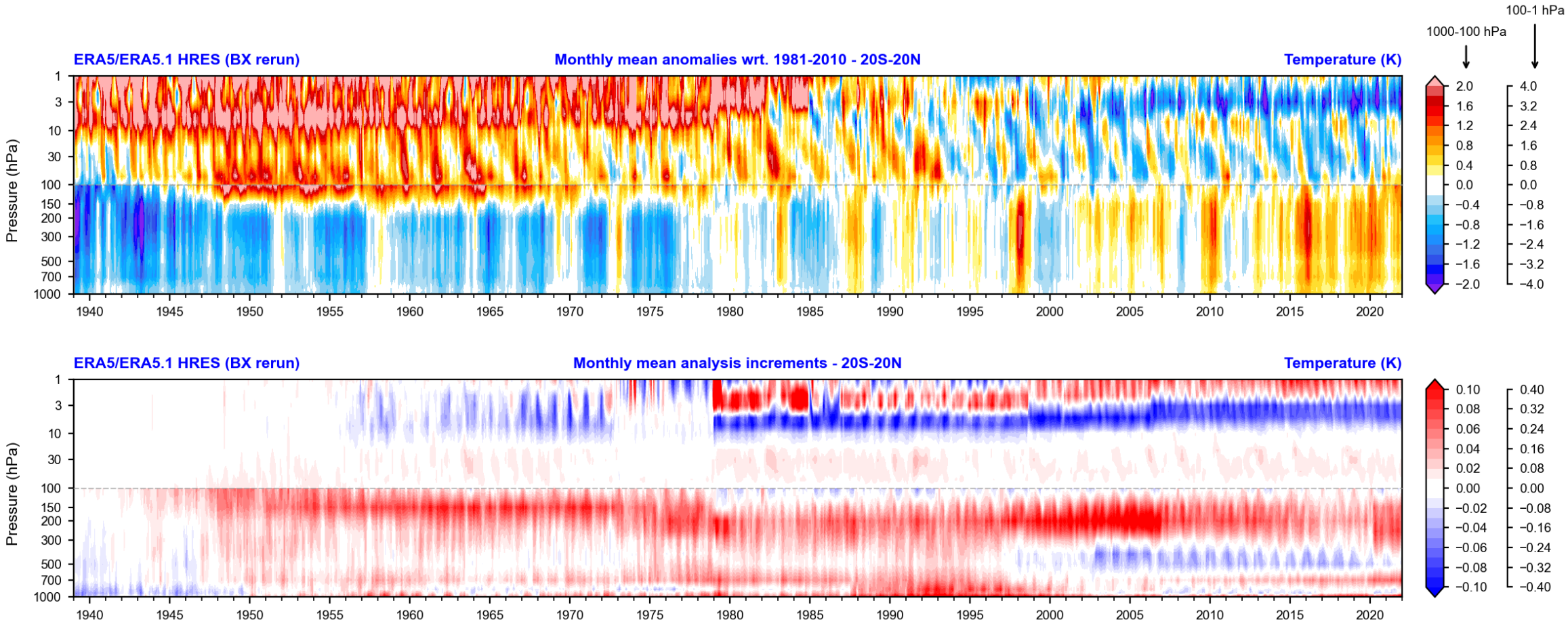
- Reanalysis
  - ERA5 covers 1951 to present, ERA5T runs just a day or so behind real-time
  - Needs to know solar forcing, GHGs, aerosols, SST, sea-ice, land surface etc.
  - In ERA5, these are a mixed bag (some CMIP5, some fixed climatologies).

- SEAS5
  - Re-forecasts from 1981 onwards
  - Real-time forecasts run at about 10am on the 1st of each month, with initial conditions from 0Z.
  - **Operational** (in NWP sense). **Not allowed to use ERA5T**. Products usually generated by 2nd, public release at 12Z on the 5th.
  - Forcing requirements similar to reanalysis, except are needed up to 13 months into the future. (Not SST, obviously, since this is predicted; some CMIP5, some fixed climatologies, volcanic aerosol is nominally a simplified damped persistence).

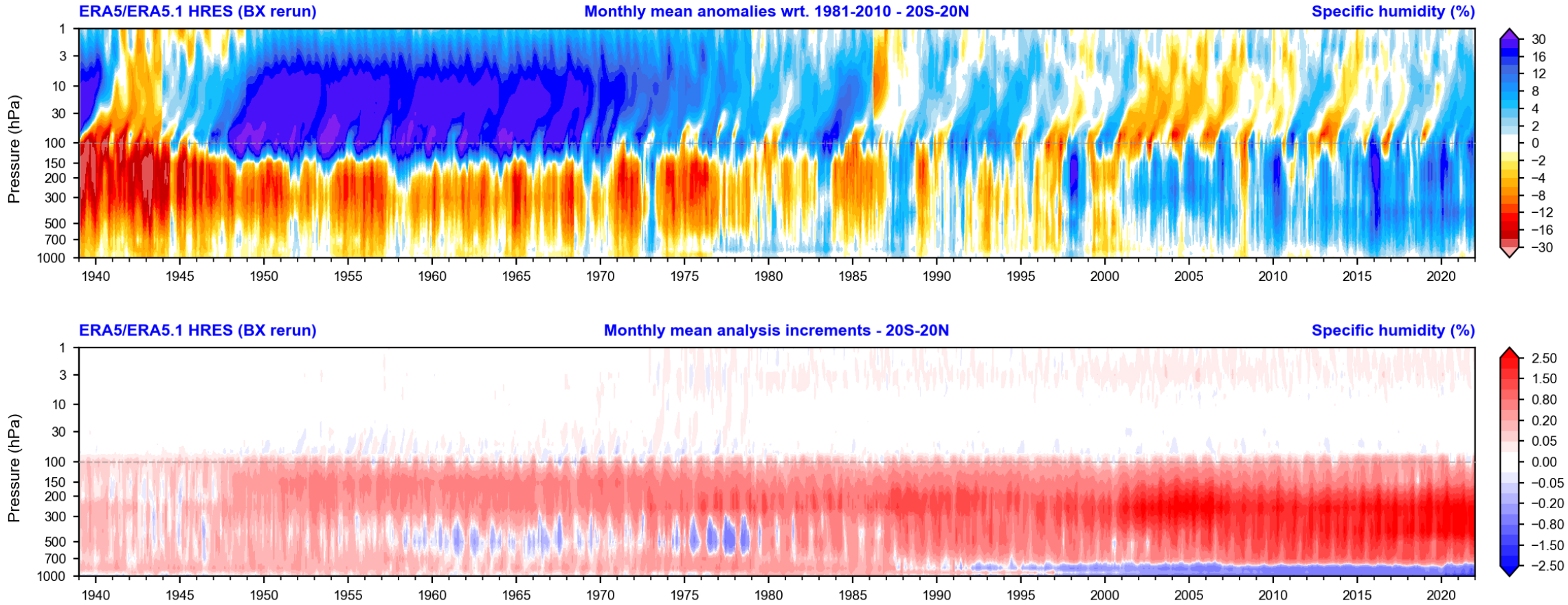
Quality requirements: re-analysis

- High quality and consistency are both *desirable*, to facilitate climate analysis
- However, analysed fields (temperatures, winds etc) are highly constrained by data assimilation
  - (aside: ERA6 uses variational bias correction for most input data, so discontinuities or drifts in model bias characteristics could potentially cause trouble, if anchor observations are insufficient)
  - (aside: some regions e.g. upper stratosphere are less well constrained by data, especially in pre-satellite era, so forcing matters more)
  - (aside: reanalyses have their limits, when data is poor and forcing is uncertain, and that is OK)
- Traceability and credibility are also important – nice to follow community standards, best practice

ERA5 tropical mean temperature analysis, 1939-2021



ERA5 tropical mean specific humidity analysis, 1939-2021



Quality requirements: seasonal forecasting

- Model quality is much more critical
  - Forecasts are for 7-13 months, not 12 hours, plenty of time for things to go wrong

- Real-time forecasts are calibrated against re-forecasts
  - **Critical** that model biases are unchanged
  - Need high degree of consistency for initial conditions and forcings
  - Need to accurately represent both changes in forcing and earth system response
  - Challenging: most datasets are either NRT or ”careful reanalysis” but not both

- The climate change signal has become a key part of a seasonal forecast – not just ENSO

- Not just temperature, other aspects are also challenging (E. Pacific trends, hurricanes etc)
  - Actually, seasonal prediction is an excellent test-bed for climate change / Earth system models

Forcing requirements for ERA6/SEAS6

- Long-lived greenhouse gases. IFS includes seasonally varying geographical structure with long-term time-evolution from CMIP6.

- Ozone. IFS uses a hybrid linear ozone model, captures dependence of ozone on circulation and temperature changes, but needs to know changes in underlying chemistry e.g. from ODS.

- Solar forcing. At moment use only TSI, although in future hope to be able to use time-varying SSI. Will use CMIP7 past values.

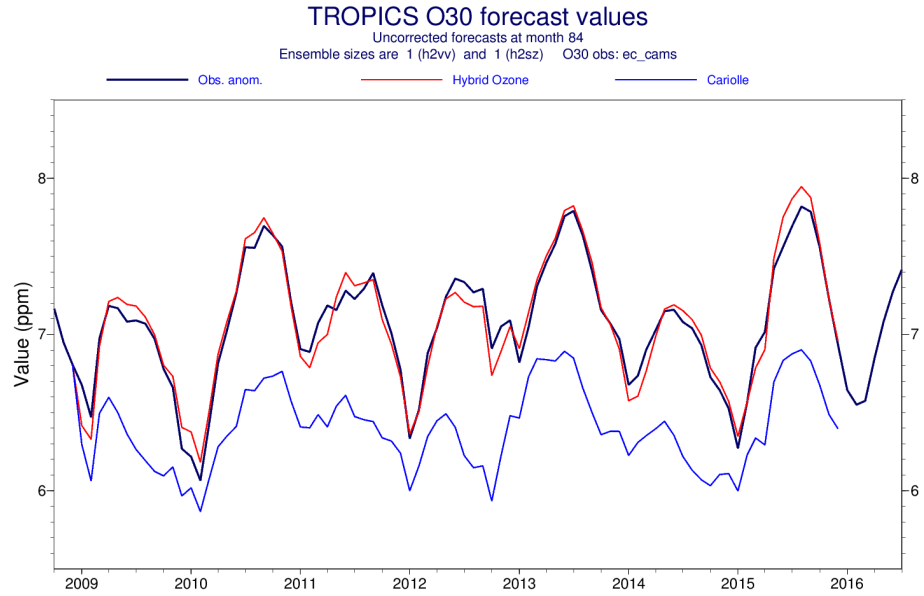
- Tropospheric aerosols. Will switch from fixed monthly to decadally-varying monthly climatology. Derived from specified emissions using IFS-COMPO configuration developed by CAMS.

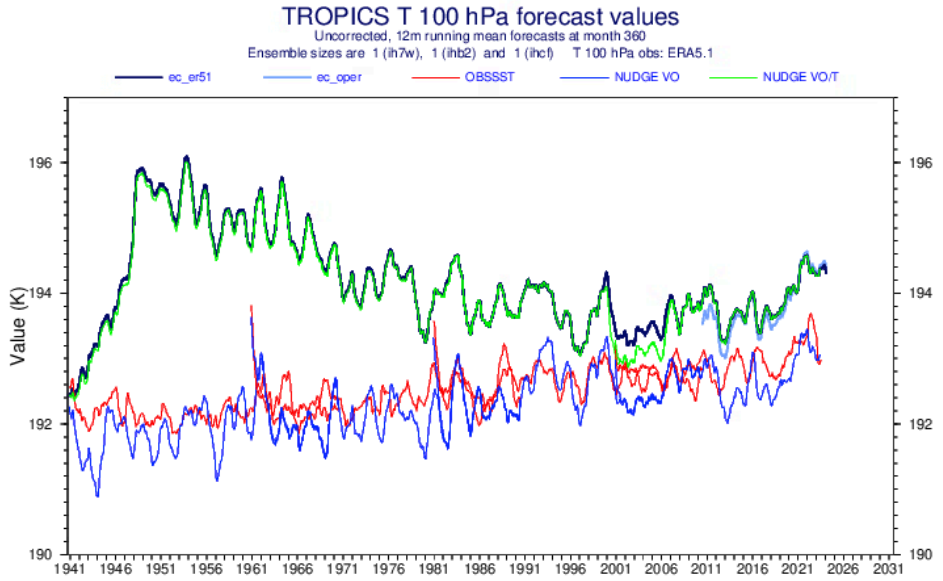
- Volcanic aerosols. Will use combination of GloSSAC2.2 and EVA_H with CMIP7-specified emissions. Need method for updating MSVOL04 in NRT, plans at the moment are very ad-hoc.

- Lake, vegetation and land use changes. Not in ERA6, but in ERA6-Land (CERISE project)

Ozone

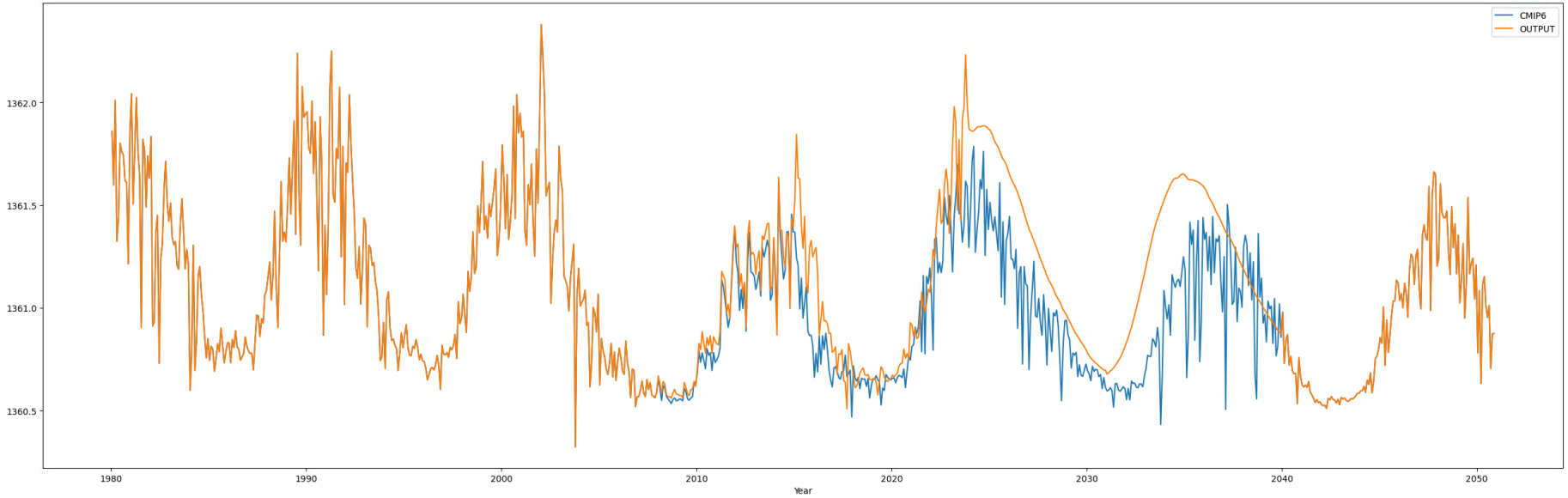
- Hybrid Linear Ozone model
  - Accurate for recent decades, although does not capture all polar variability.
  - Scheme coefficients depend on stratospheric chemistry, but are empirically derived from ozone analyses.
  - Need to define suitable coefficients for past decades (pre-1993). **Using CMIP6 ozone dataset** to help with this.
  - Possible lack of consistency between tropical tropopause ozone estimates and early radiosonde temperature data.
- Longer term, could make use of full chemistry model, consistency of chemistry inputs would be important.





Solar forcing

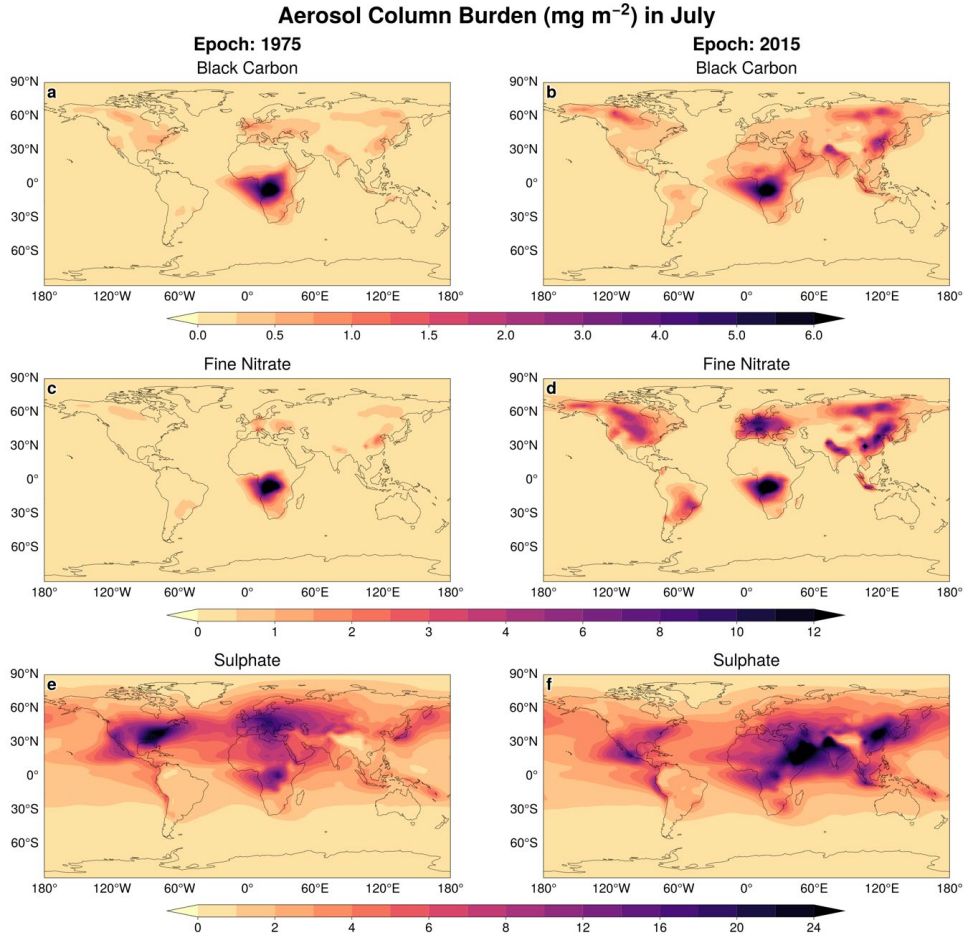
- IFS has been using CMIP6 TSI annual values.
- For ERA6/SEAS6 we are updating this to CMIP7 (for past data)
  - Still need to make the latest adjustment to absolute scaling in HEPPA-SOLARIS 4.4
  - We also need future values. Have done a home-made forecast using latest NASA Marshall sunspot forecast and certain scaling assumptions. At present are not planning to update this - might this change?



(Shown here with CMIP6 scaling for comparison purposes, not final values)

Tropospheric aerosols

- Developed a time-varying tropospheric aerosol climatology
  - Using IFS-COMPO (CAMS), sets of 24h forecasts from 1951 to 2019
  - Based on Cy48r1 chemistry and aerosol optical properties
  - Emissions come from CEDS v_2021_04_21, CMIP6 fire emissions (GFED v4s).
  - Funded by CONFESS project as a one-off
- Future possibilities (funding permitting)
  - extend time-varying climatology to NRT and keep up to date.
  - Repeat climatology with new latest model version
  - Move to interactive aerosols, at least for some species. Still need emissions, though.

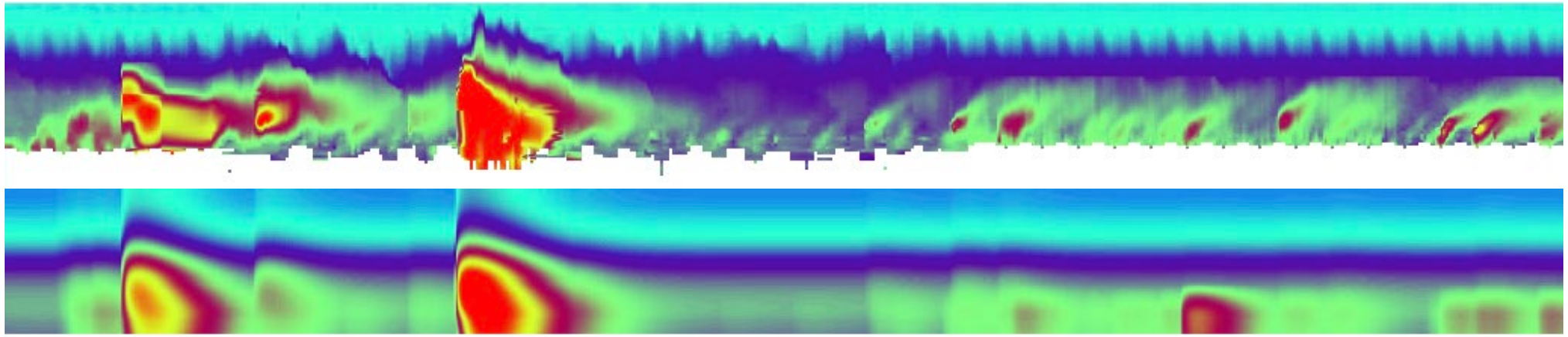


The CONFESS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004156.

Volcanic aerosols

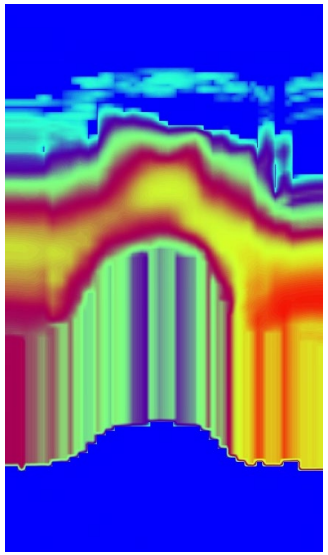
- IFS representation has previously been very crude
  - Missing entirely from early re-analyses
  - Vertically integrated GISS data, applied with fixed vertical functions
  - SEAS4, SEAS5 – damped persistence based on 3 numbers in a namelist

- ERA6 and SEAS6 have available GloSSACv2.2 and EVA_H with specified eruptions
  - Prior to 1979, entirely dependent on Thomas Aubrey's CMIP7 eruption dataset
  - Can then use MSVOL04, plan to then add important eruptions by hand. **Can we do better for NRT?**

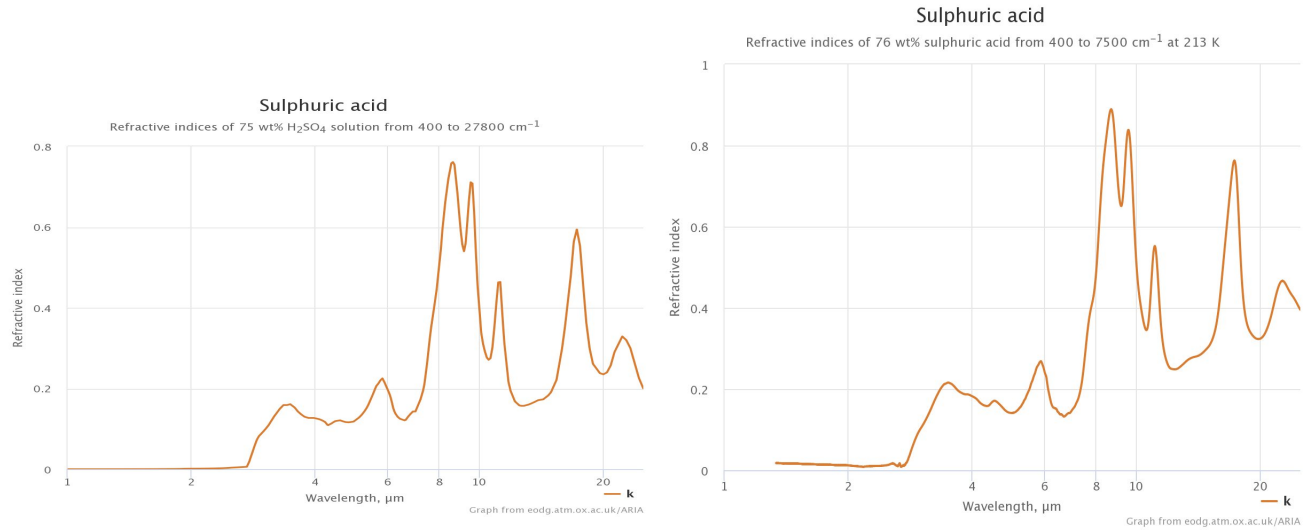


Volcanic aerosol – an aside on optical properties

- IFS optical code needs aerosol mass density (kg/kg) as input, then does its own radiative calculations based on pre-computed tables using a fixed particle size distribution.
- Optical properties a function of particle size distribution and complex refractive indices.
- IFS values for both were quite inaccurate, needed upgrading.
- Used ecaeropt code for calculations (Mie theory)



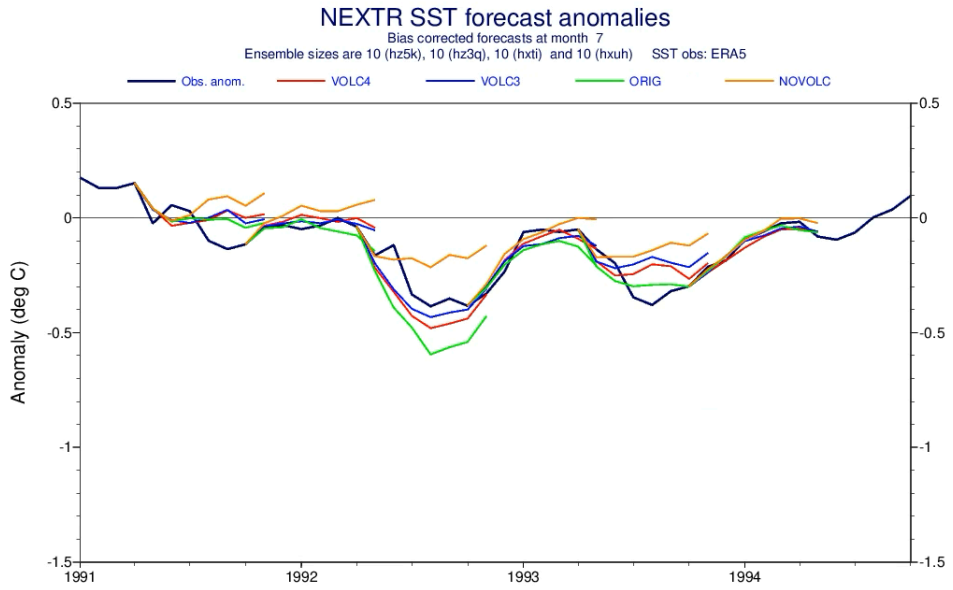
GLOSSACv2 data: post-Pinatubo snapshot showing vertical and latitudinal structure



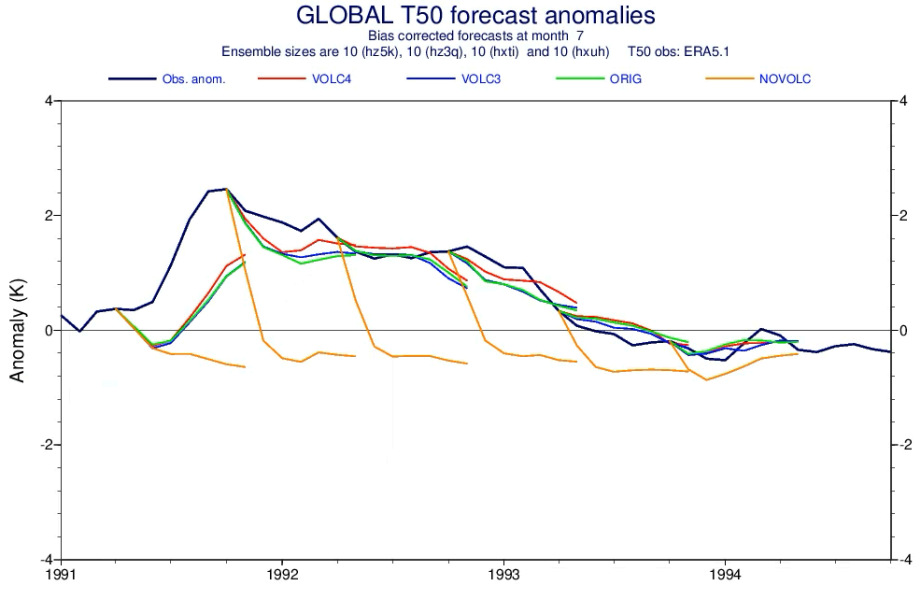
Sulphuric acid refractive index imaginary component from (Palmer and Williams (1975) (left) and Lund Myrhe et al (2003) (right)

Volcanic aerosol – impact

- Impact on stratospheric temperature modelled quite well, except in first 6 months after eruption
- Impact on SST seems credible, but true impact is somewhat uncertain due to poor El Nino forecast
- Would need time-varying particle size distribution to get initial response right if using observed 525nm extinction



NHEM SST evolution using new and original optical properties compared to an experiment with no volcanic eruption



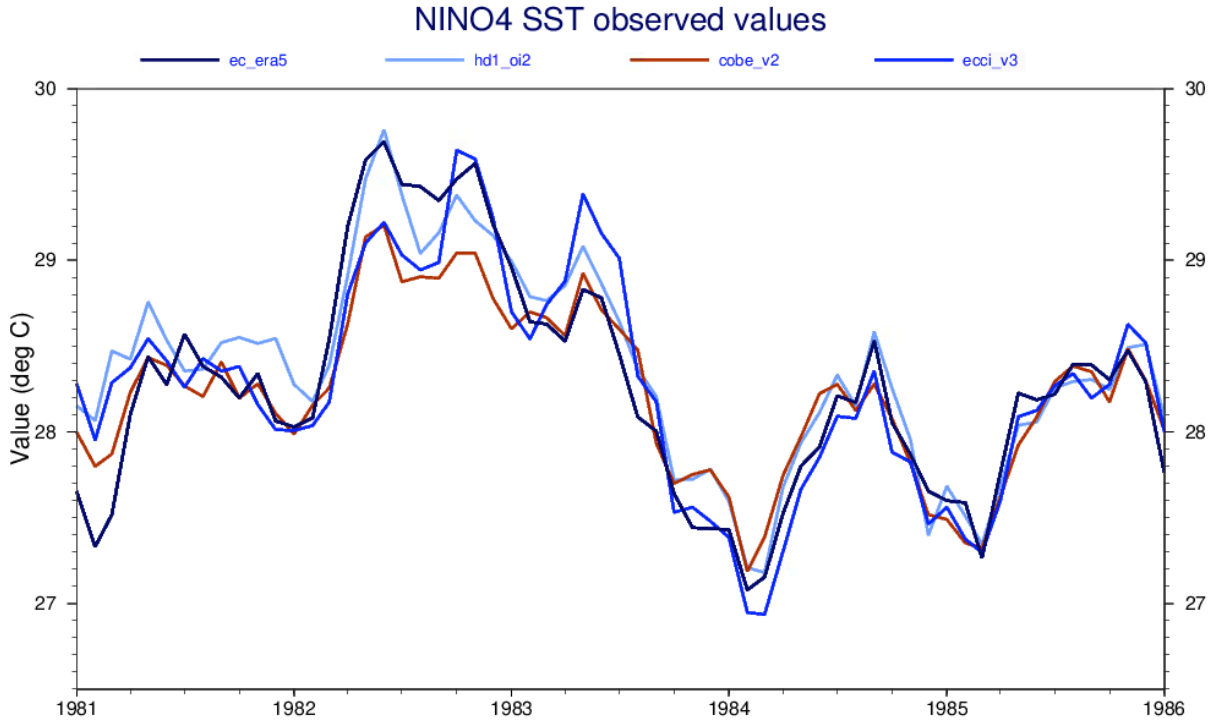
Global 50 hPa T evolution using new and original optical properties compared to an experiment with no volcanic eruption

Lakes, vegetation, land use

- ERA6 is about to start, cannot include any changes in these
- ERA6-land will be a high-resolution land-surface product, run later
  - Will be able to pick up in-progress development on time-varying lakes, and maybe other aspects of time-varying land surface, part of CERISE project
  - Turns out to be quite hard to make consistent time-varying vegetation datasets, even in the satellite era.
  - Aiming to use LUH2, will they be able to switch to LUH3?

Is it a forcing? – but it is important: SST

- SST is a key field for re-analysis, either of the atmosphere or the ocean.
- Might think this was well observed for the satellite era, but think again ...
  - Substantial differences between what the 1982-1983 El Nino looks like in different SST datasets
  - Has been a lot of work on which datasets are most plausible for use in ERA6 and OCEAN6



Thoughts on Sustained mode: Extension vs update

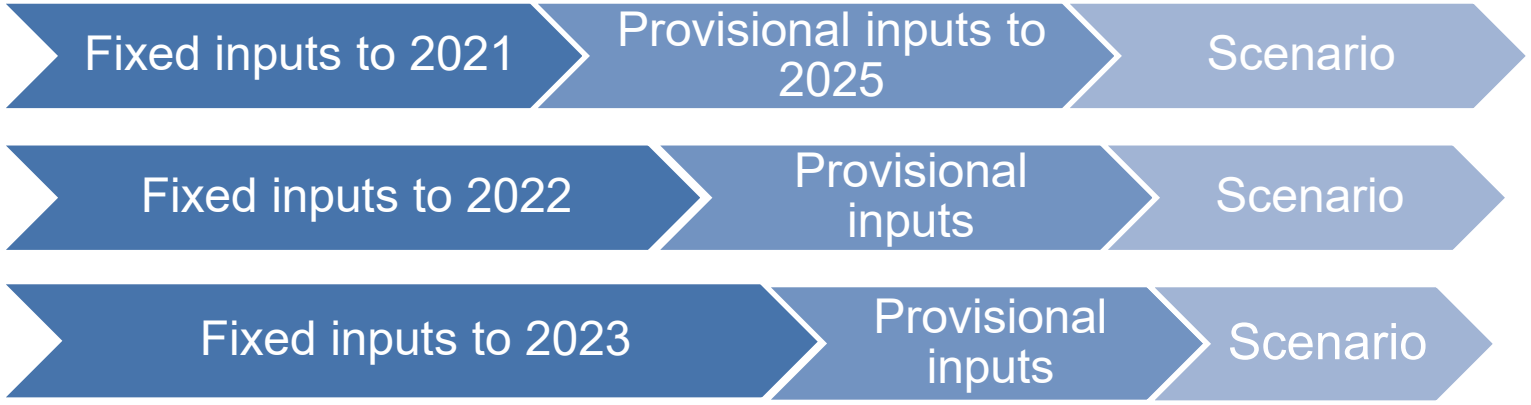
- **Extension**: Use same methodology, same/similar data sources, to extend time-series of data; values for earlier dates do not change. **Needed** for applications with heavy investments and infrequent updates (re-analyses, decadal forecasts, seasonal forecasts, MIP-era expts)

- **Update**: Revision of methodology and/or data sources, whole time-series changes. Only when debugging early release data, or creating a new generation of re-analyses/forecasts. Datasets are never perfect, once initial debugging done we fix the data/system for a given generation, all new bug-fixes go into a “latest” version but the fixed version is not changed.

- For sustained science and applications, we need both of these

But how to manage extensions, if the input data is not NRT, and/or subject to frequent revision of most recent values? (eg CEDS).

Extension: two paradigms for dealing with dataset extensions



Maybe OK for decadal, scenario updates.

Maybe less good for reanalysis, seasonal prediction, which will always be using provisional data.

NRT data streams must be identified (or possibly developed), with a high degree of compatibility with earlier data.

Gives access to up-to-date forcing, where this can be done reasonably accurately.

Still need predicted values for forecasts, but not reanalyses

Two paradigms for extensions: possible examples

- Annual increments to data, 6-18+ months behind, including prospective forcings for near future
  - Greenhouse gas concentrations (long-lived tropospheric) CO2, CH4, NOx, CFCs etc.
  - Anthropogenic industrial, transport and heating emissions
  - Land use changes

- Daily NRT updates, plus updated projected values for use by forecast systems
  - Biomass burning emissions
  - Volcanic emission data (volcanic aerosol loading calculated assuming no further eruptions)
  - Possibly solar data

What specific issues might benefit from better coordination/funding?

- (Leaving the question of **how** this might be done to others)
- Emissions data for chemistry/aerosol models?
  - Fire emissions, anthropogenic emissions, natural emissions
  - What about consistent long-reach reanalyses and real-time analyses, are they needed or not?
- NRT updates of volcanic eruption datasets?
  - At the moment we plan to rely on the VOLRES email group plus CAMS analyses of SO2 to help guess the numbers to use. **Can we be more organized?**
  - Is there value in pursuing NRT updates of observational datasets such as GloSSAC? Would they always be too slow to be useful for reanalysis and forecast systems? Are there other users for NRT?
- Solar forcing
  - how NRT are operational TSI/SSI data, e.g. from NNL
  - Can someone operationalize **forecasts of TSI/SSI?** How useful would this be?
- Can we produce consistent historical and NRT ozone datasets?
  - How important is this, or do we say that groups need to model this themselves if they want updated values?