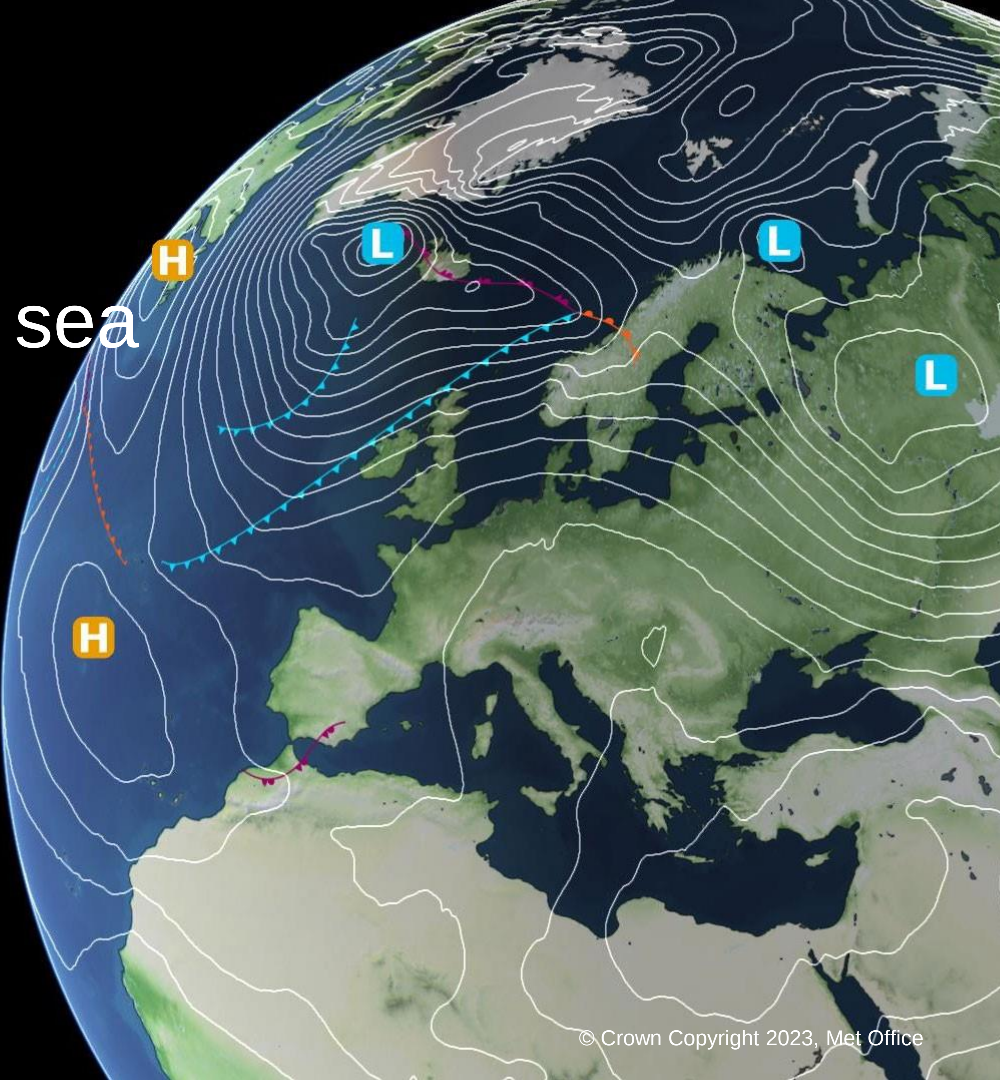
Met Office
Hadley Centre

Reconstruction of sea-surface temperature and sea ice data sets

Nick Rayner
Met Office Hadley Centre

ECMWF Annual Seminar, 6th Sept 2023

With thanks to John Kennedy, Holly Titchner, Owen Embury, Chris Atkinson and Liz Kent for material presented



Overview

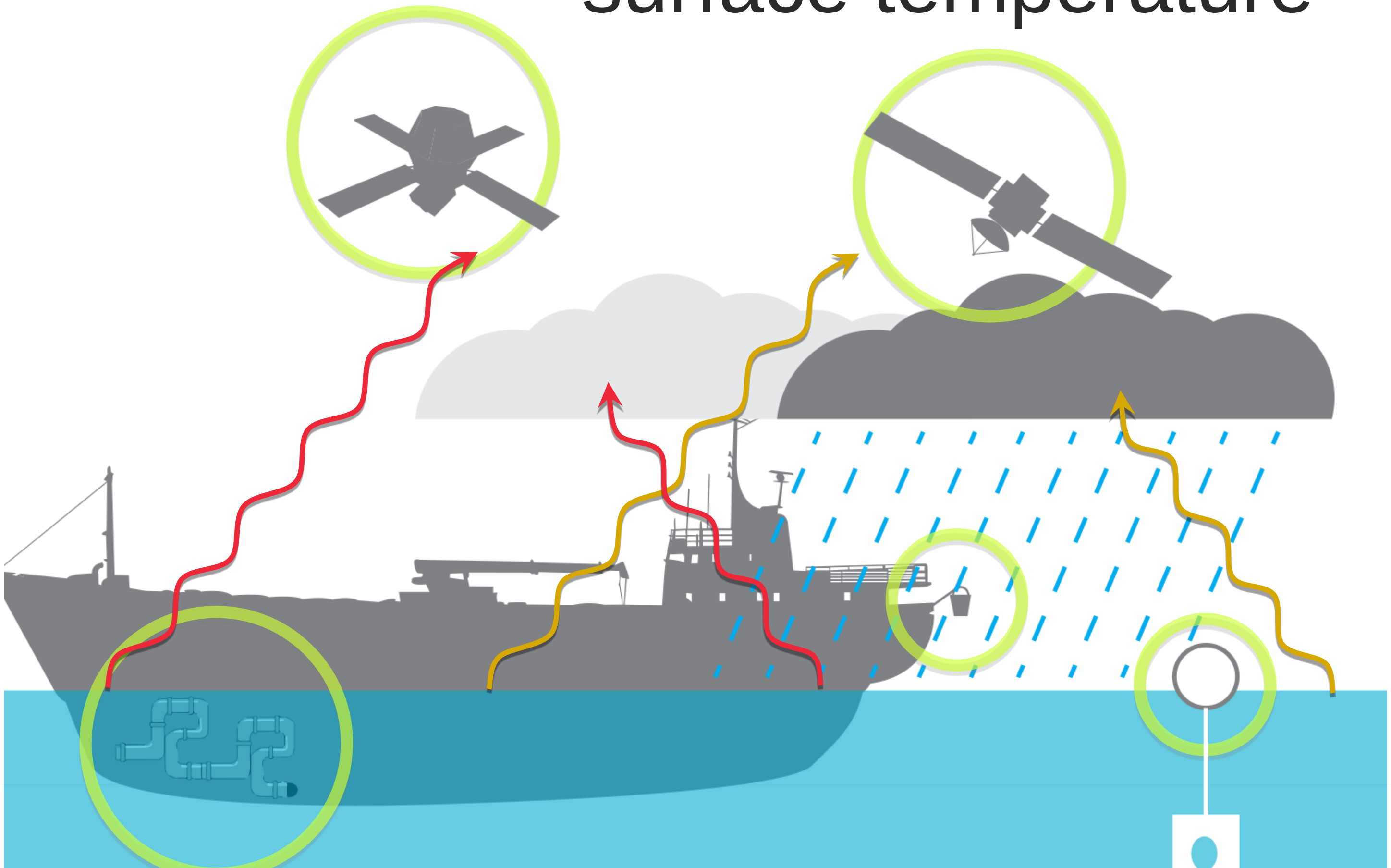
- Key principles of homogeneity, completeness, uncertainty quantification and consistency
- Historical (in situ and satellite) observations and what they represent
- Methods that can be used for reconstruction
- Practical communication of uncertainty information via an ensemble
- Creating usable data
- Outstanding challenges

Key principles

- Homogeneity
  - Avoid confusion, spurious trends and discontinuities
- Completeness
  - The reanalysis needs SST and sea ice fields over the whole ocean
- Practically-conveyed uncertainty estimates
  - That are straightforward to propagate through into the reanalysis output
- Consistency between SST and sea ice
  - To avoid spurious local fluxes or other adverse outcomes
- Usable fields

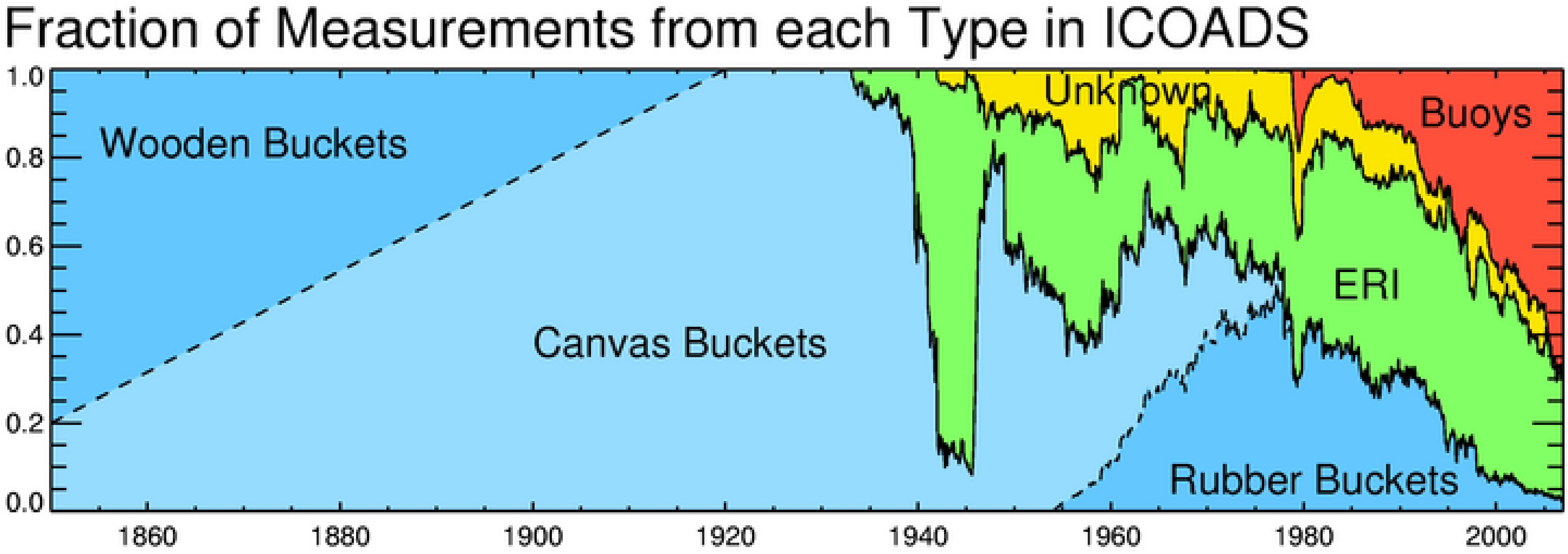
Input observations

Measuring sea surface temperature



Evolution of the *in situ* SST observing system

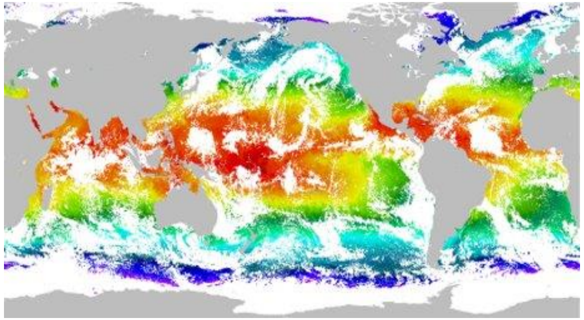
As understood by Kennedy et al 2011

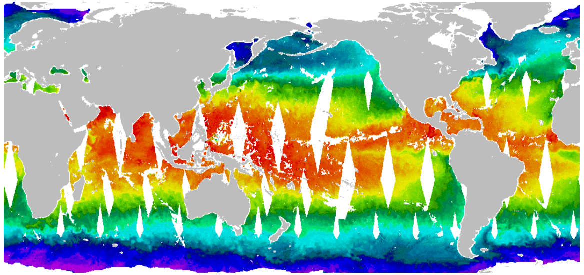


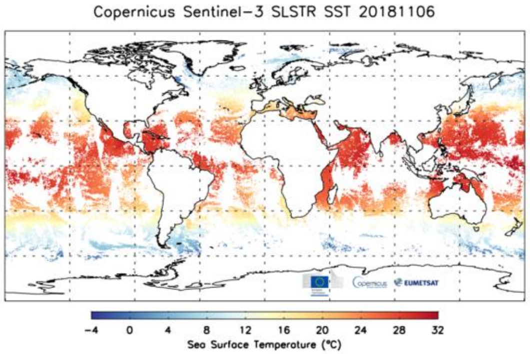
Met Office
Hadley Centre

Types of satellite instruments

- Polar-orbiting infra red AVHRR series has good coverage and spatial resolution and a long, continuous record, but orbits of many of these instruments have drifted over time.
- Polar- and semi equatorial-orbiting microwave instruments have shorter records, but a wide swath and provide information under clouds
- Polar-orbiting infra red ATSR series has relatively poor coverage, but has high spatial resolution, is stable and accurate. Designed for monitoring, has a “dual view” and so can be more robust to atmospheric contamination. Now succeeded by SLSTR.

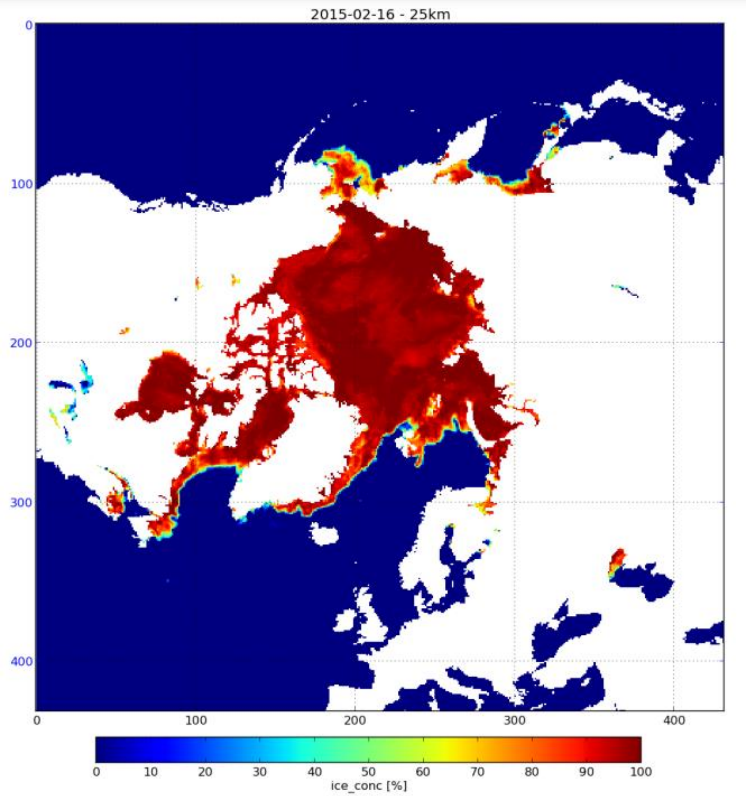






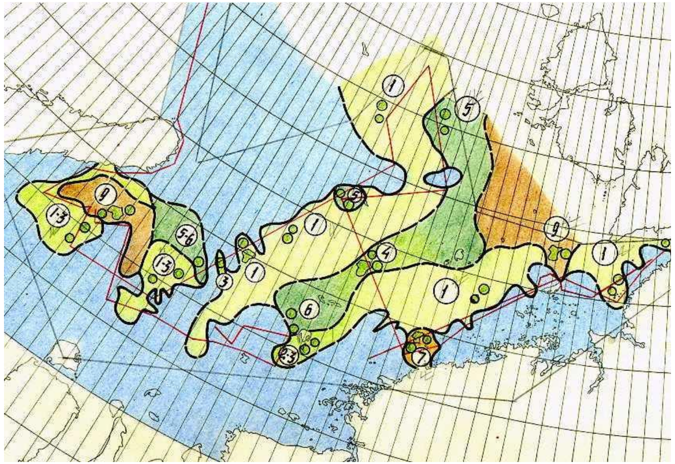
Sea ice information retrieved from satellites

- Example shown is 16th Feb 2015, ESA CCI sea ice 25km
- Concentration is shown 0-100%
- Sea ice concentration is challenging to retrieve in the summer when surface melt-water is present
- The SMMR/SSM/I series provides a 45-year record

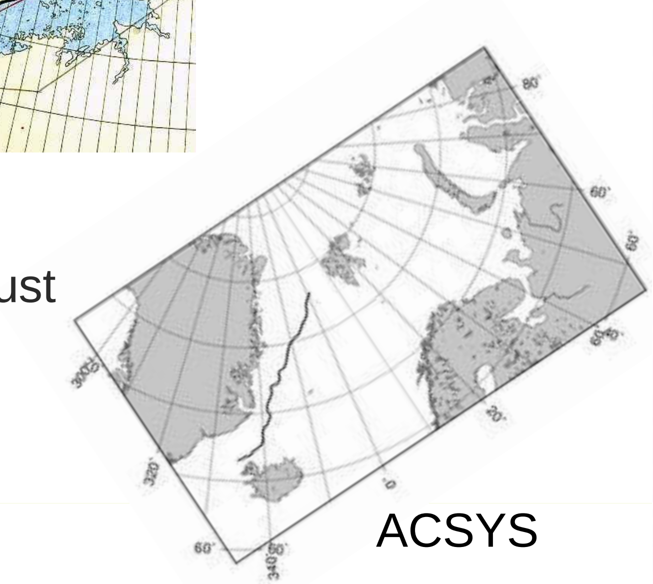


https://dap.ceda.ac.uk/neodc/esacci/sea_ice/docs/SICCI_P2_SIC_PUG_D3.3_Issue_1.1.pdf

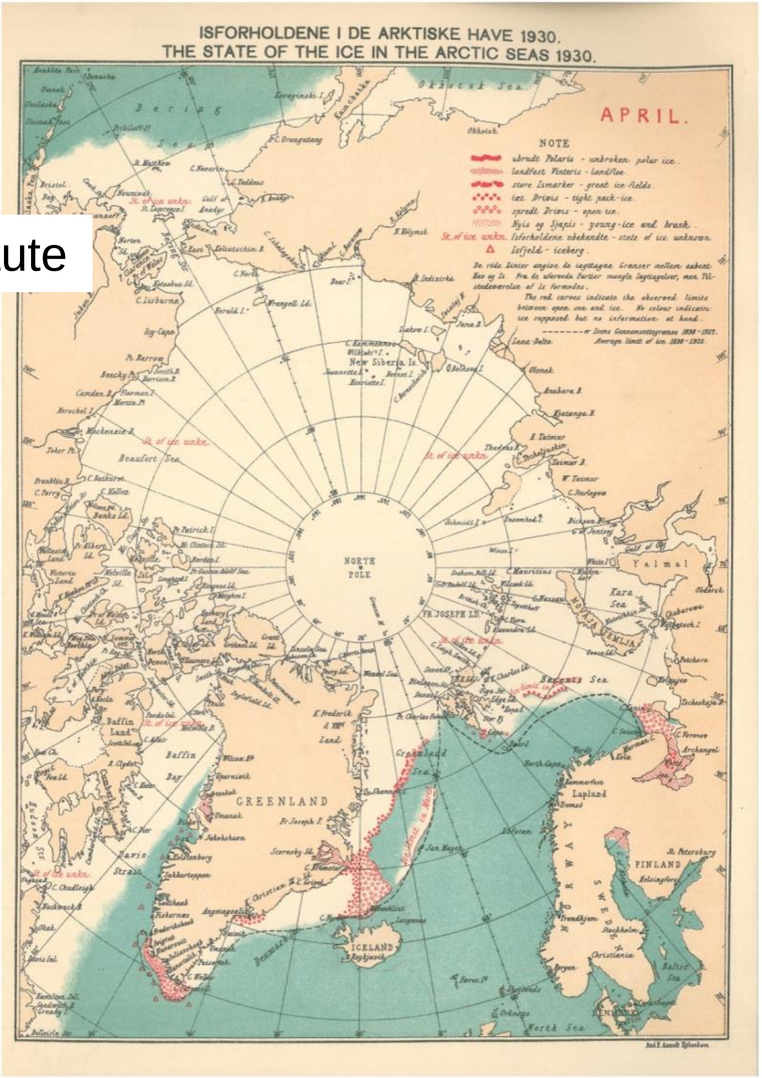
Sea ice charts



AARI (Russian) chart, Kara Sea, August 1933 (figure from NSIDC website)



ACSYS



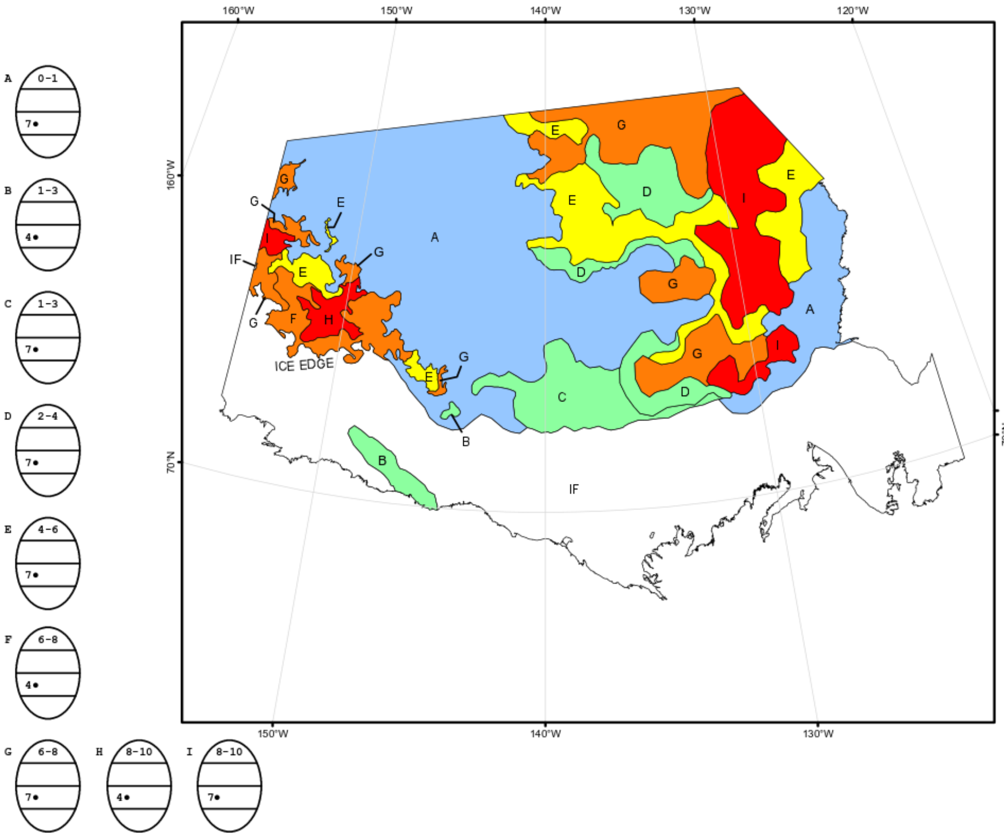
Danish Met Institute

April 1930

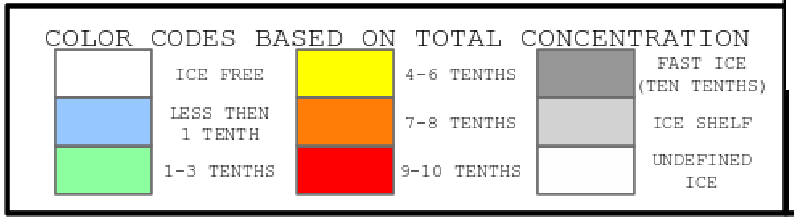
Met Office
Hadley Centre

Modern sea ice charts

National Ice Center chart, Beaufort Sea, September 2015



IF = ICE FREE



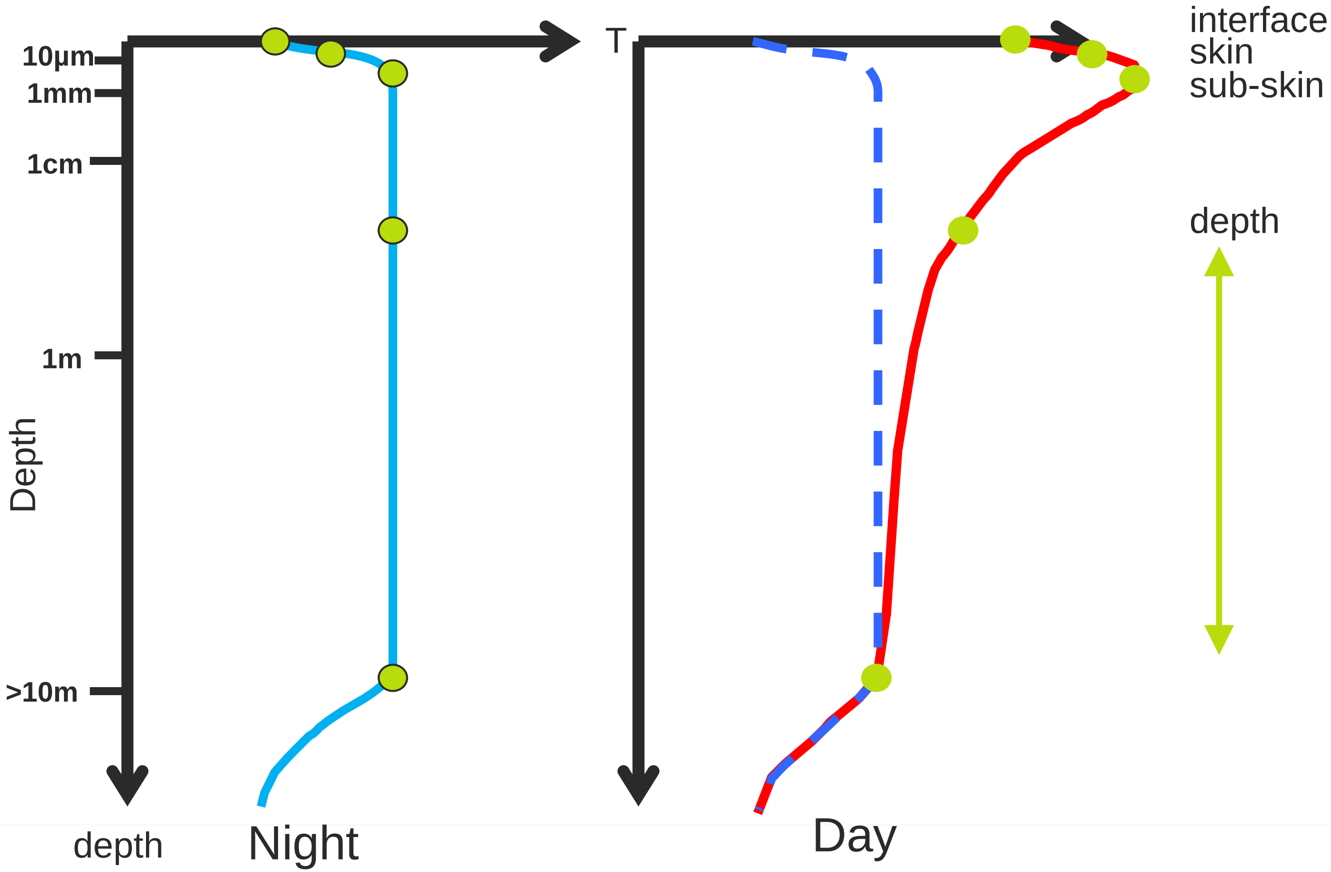
Met Office Hadley Centre

Creating homogeneity

Why homogeneity is important

- Need to account for differences between the measurands of each data source – ensure we know what the data set aims to represent
- Removes the impact of non-climatic discontinuities arising from changes in measurement method/input data source, etc
- Allows the reanalysis to optimally represent actual changes in the climate
- Avoids spurious responses of the reanalysis to non-climatic changes in the data, e.g. large heat fluxes over the polar regions

What is sea-surface temperature?



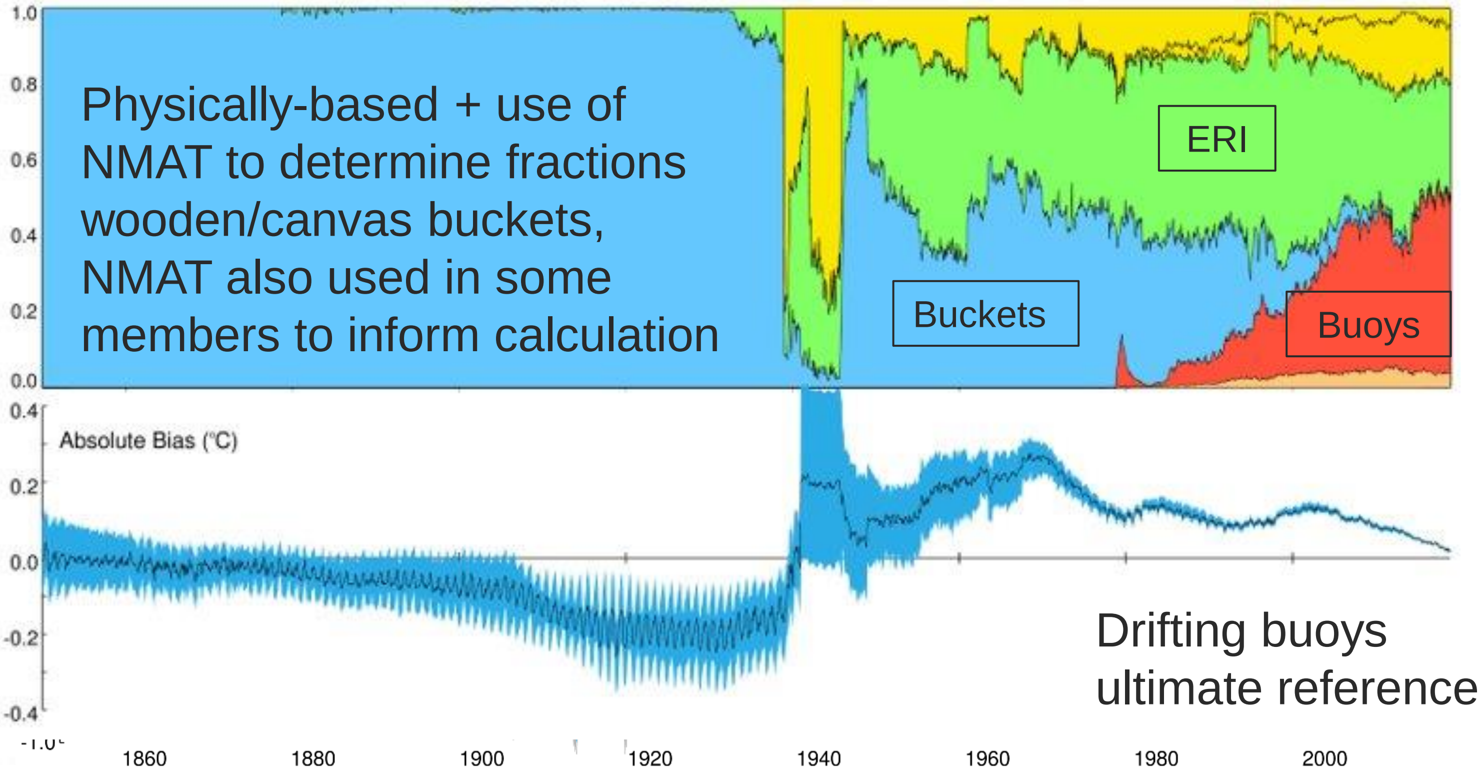
Met Office Hadley Centre

Ways of achieving homogeneity

- Compare everything and develop empirical corrections, relative to a chosen reference
  - Risks picking the wrong reference and biasing the whole system
- Understand each data source physically and correct according to its own biases
  - Then compare to everything else and check consistency
  - But this requires good metadata, which is often lacking
  - However, this allows potential propagation of error structure
- Let the reanalysis handle it – still requires good understanding and metadata
  - Works for assimilated observations but not for a forcing data set?

Met Office Hadley Centre

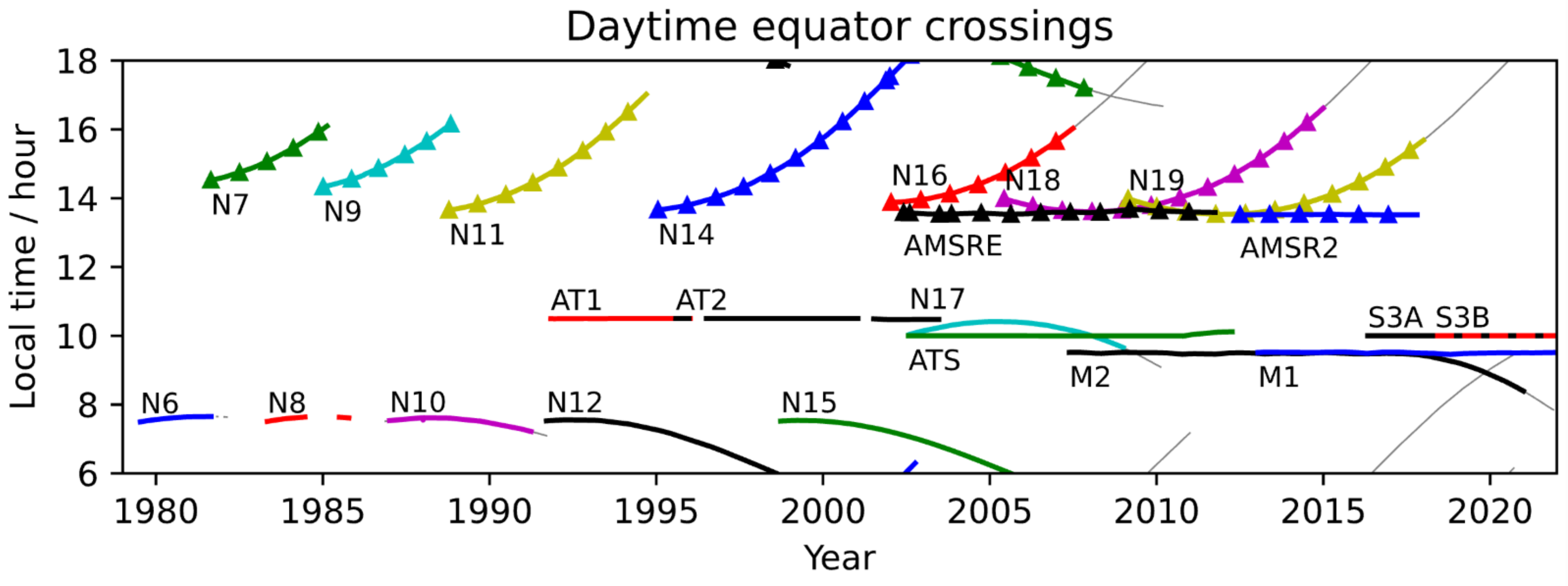
Bias in changing *in situ* SST measurement network, as estimated in the HadSST.4.0.0.0 data set



Oceanographic profiles in upper 10m (not Argo) used to inform some parameter ranges for ERI bias estimate and modern bucket biases

Kennedy et al 2019, https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2018JD029867

SST-measuring satellite orbits



- Correction is made to 1030/2230 local time
- Brightness temperatures harmonised across sensors
- Physically-based bias-aware optimal estimation retrievals

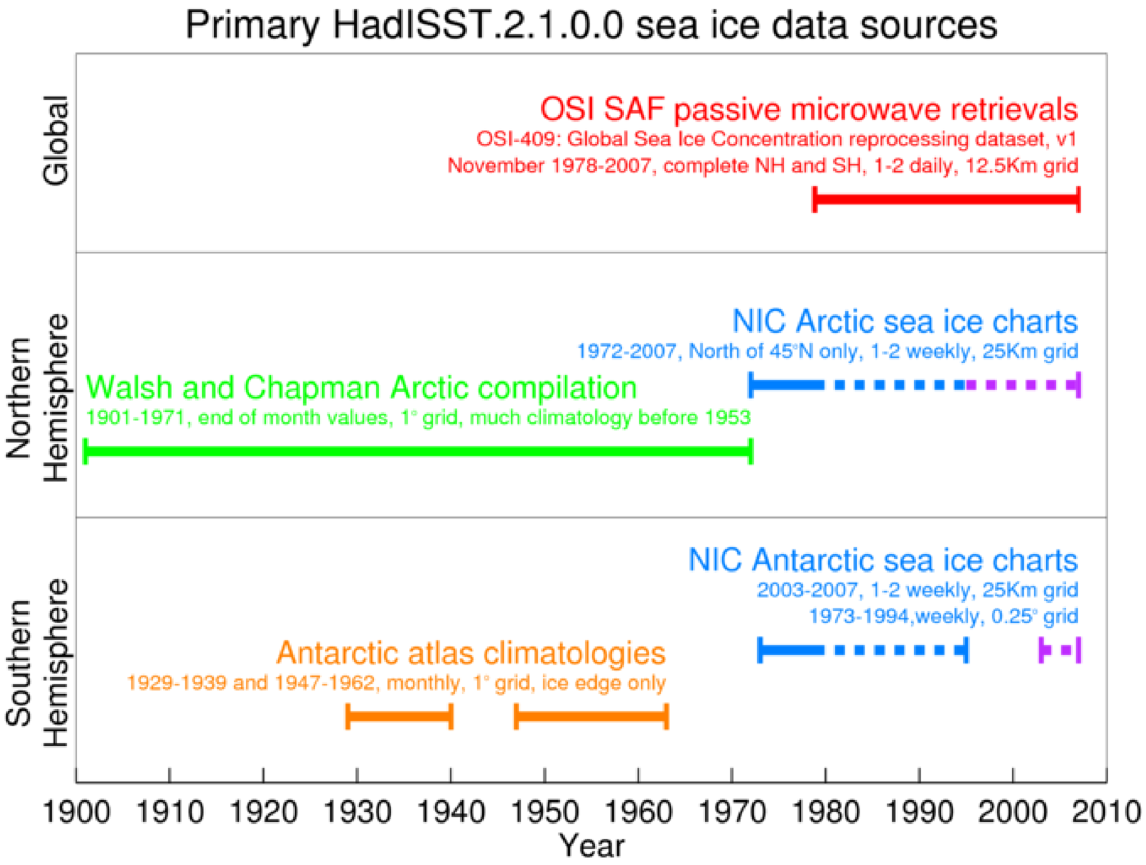
Embury et al 2023, Satellite-based time-series of sea-surface temperature since 1980 for climate applications version 3, *submitted*

Timeline of sea ice data sources

solid = main input data sources

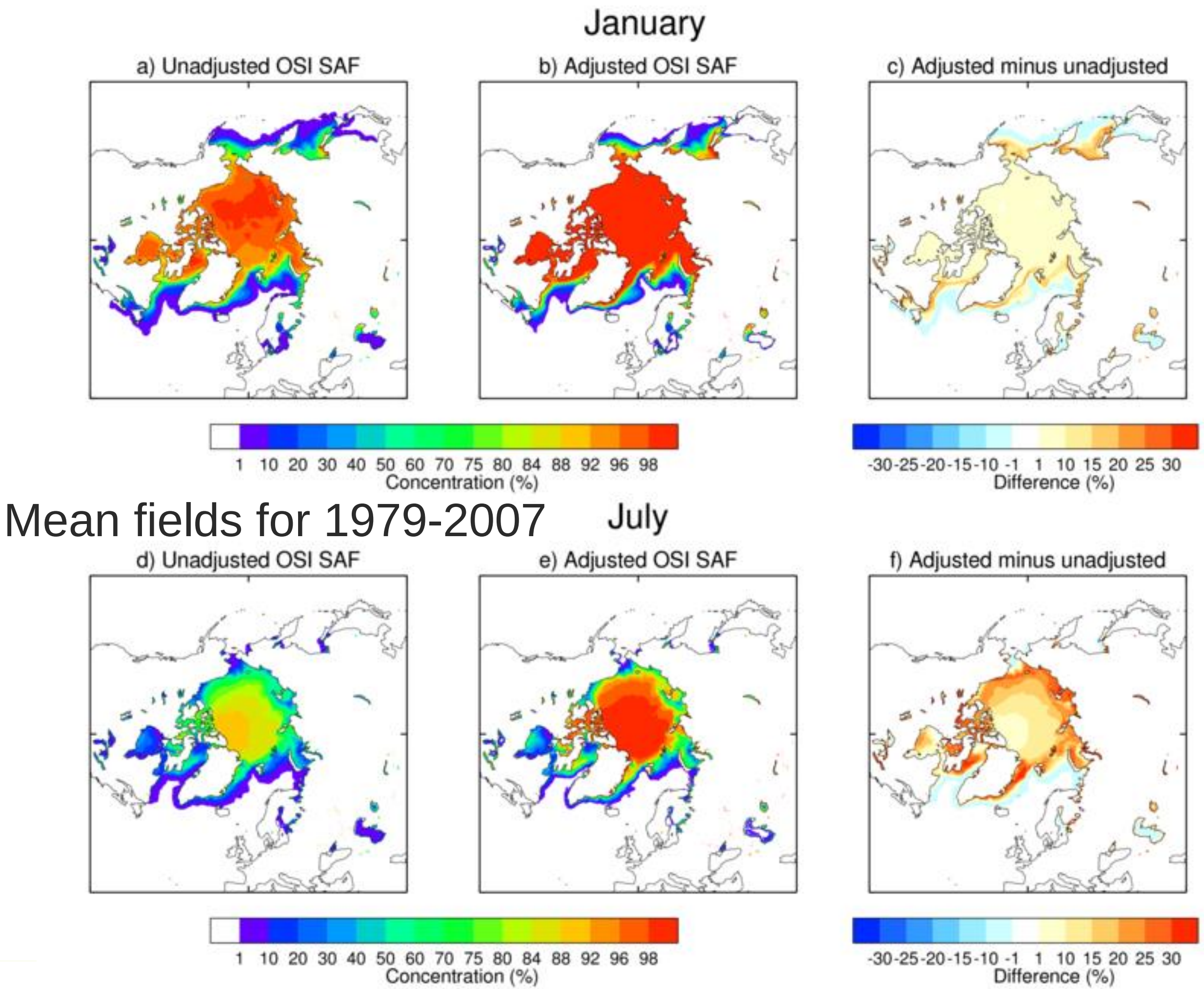
dashed = data used for calculation of bias adjustments only (using overlap periods)

purple = reference against which bias adjustments were calculated



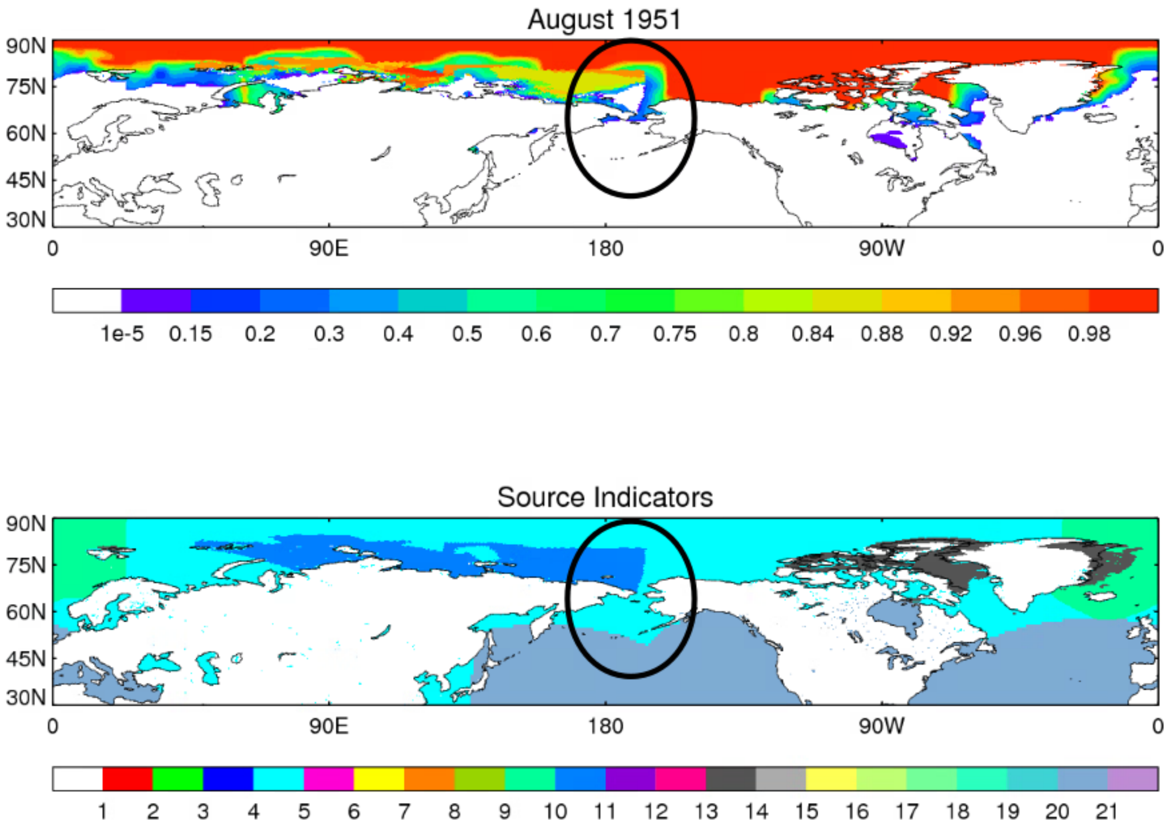
Titchner and Rayner, 2014, https://agupubs.onlinelibrary.wiley.com/doi/full/10.1002/2013JD020316

Adjustments applied to OSI SAF data to bring into line with NIC charts to allow consistency across the historical record.



Titchner and Rayner, 2014, https://agupubs.onlinelibrary.wiley.com/doi/full/10.1002/2013JD020316

Met Office
Hadley Centre



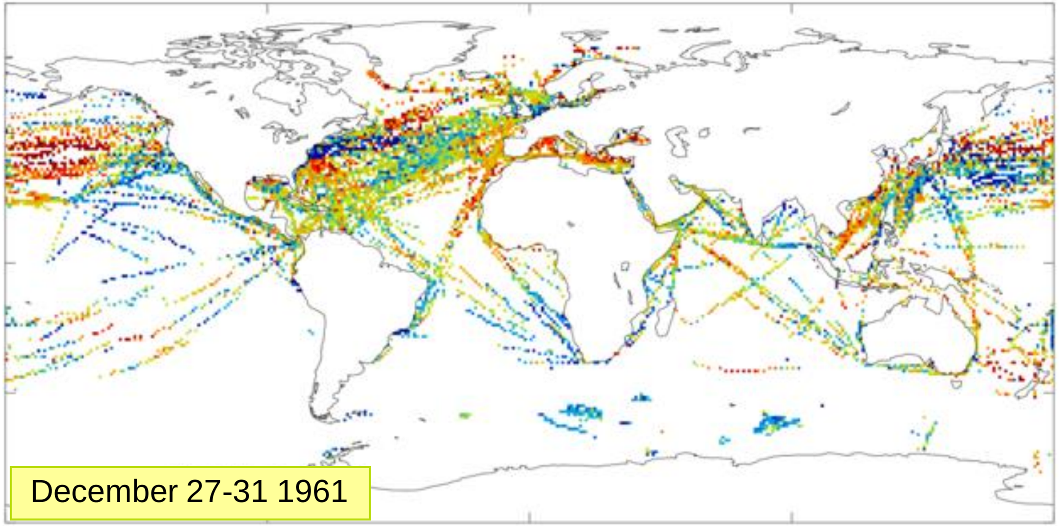

```
01 = "Satellite passive microwave";
02 = "Danish Meteorological Institute";
03 = "Dehn";
04 = "NAVO yearbooks";
05 = "AARI";
06 = "Hill";
07 = "Whaling Records - Complete Sea Ic
08 = "Whaling Records - Partial Sea Ice
09 = "Whaling Records - No Sea Ice";
10 = "DMI yearbook narrative";
11 = "ACSYS";
12 = "Walsh and Johnson";
13 = "JMA charts";
14 = "Kelly ice extent grids";
15 = "Land mask correction fill";
16 = "Change-of-land mask ocean";
17 = "Analog fill - spatial";
18 = "Analog fill - temporal";
```

Data are SIBT1850: Walsh et al 2019, https://nsidc.org/sites/default/files/g10010_v0020_1_0.pdf

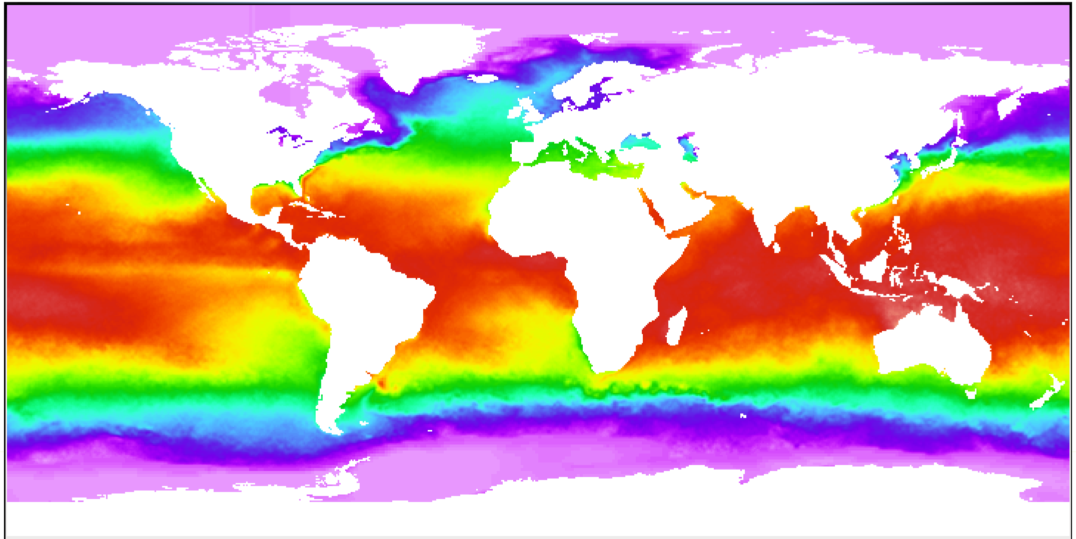
Completeness

Met Office Hadley Centre

Grid of all available SST measurements for December 27th-31st in 1961



What we need to drive a reanalysis



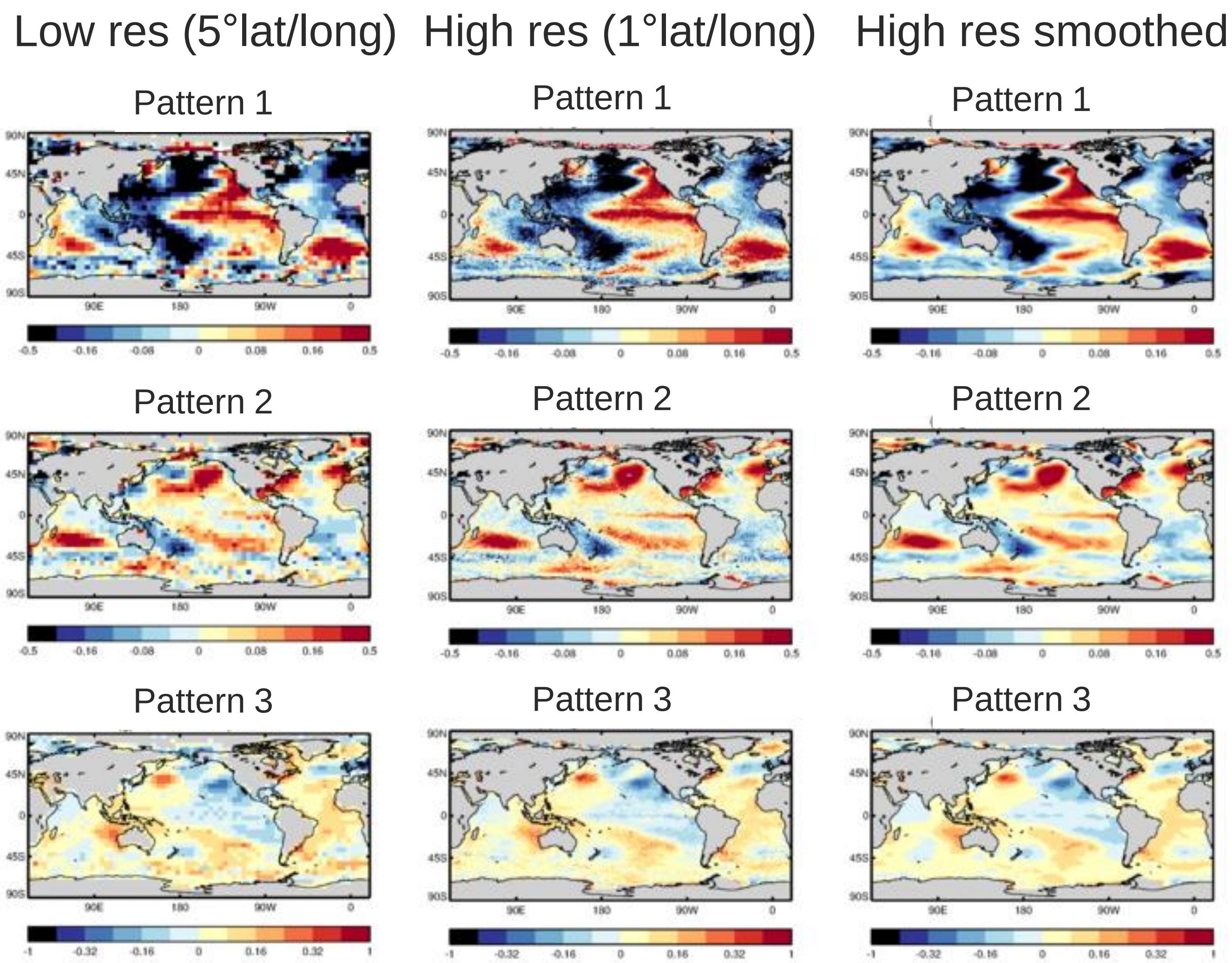
Overview of types of methods used

- Optimal interpolation – OI.v2, Daily OI (1982 onwards)
- 3D-VAR - OSTIA / ESA CCI (1980 onwards)
- Reduced space optimal interpolation – HadISST1
- “Multi-Time-Scale” (MTA) analysis method reconstructs daily SST fields as a sum of a trend, interannual variations and daily changes – COBE-2
- Variational Bayesian Principal Component Analysis + local OI - HadISST2
- Quasi-global and so not used as boundary forcing:
  - Combination of Empirical Orthogonal Teleconnections and a low-frequency smoothing – ERSSTv5 (quasi-global)
  - Kriging – Berkeley Earth (quasi-global)

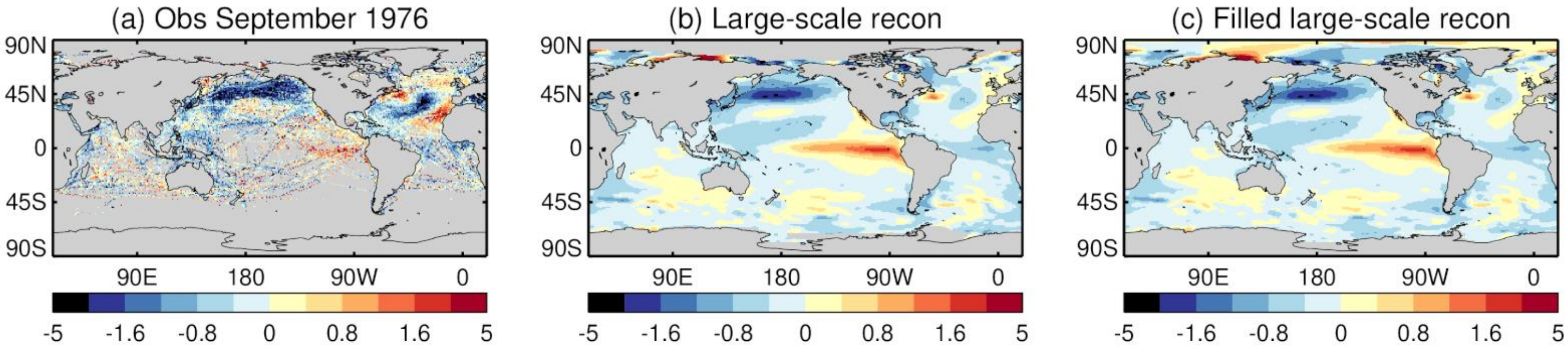
Met Office
Hadley Centre

Covariance patterns used in large-scale reconstruction

Based on all available in situ and satellite data, 1850-2021 HadISST.2.4.0.0, *paper in prep*



Large-scale SST anomaly reconstruction

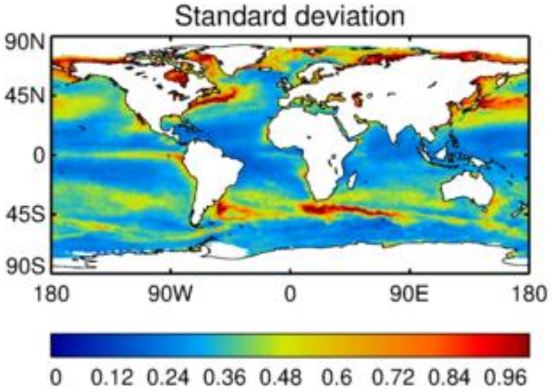


All available in situ data, September 1976

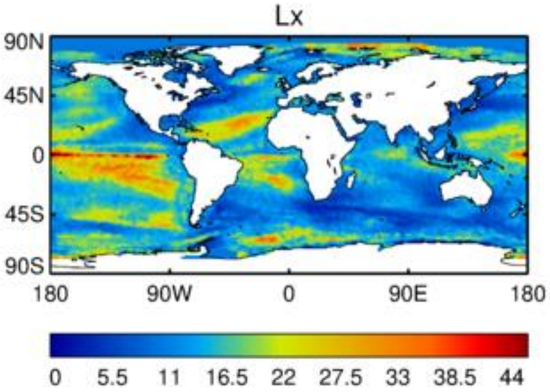
Met Office
Hadley Centre

Parameters of mid-scale reconstruction

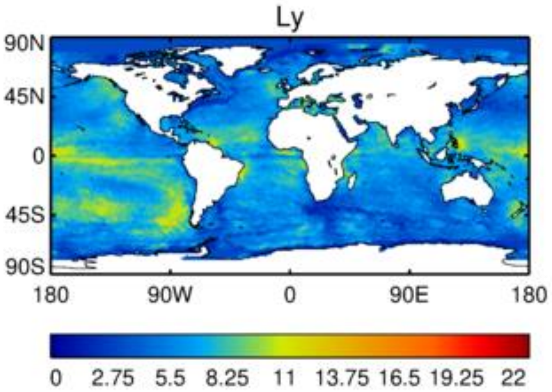
Based on all available in situ and satellite data, 1985-2021 HadISST.2.4.0.0, *paper in prep*



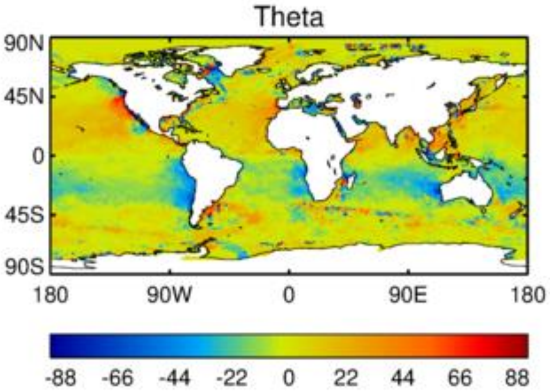
Zonal length scale



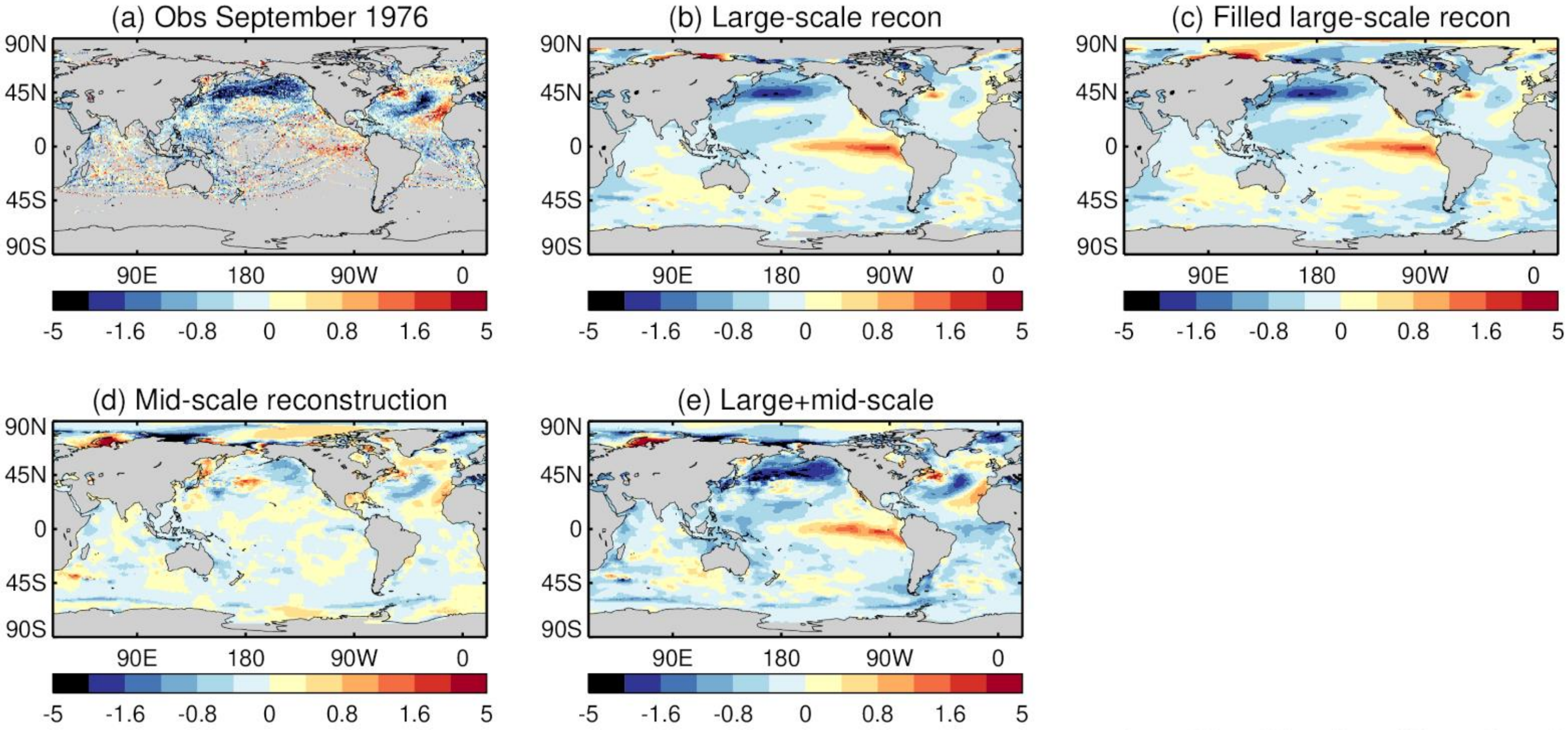
Meridional length scale



Angle of local axis rotation



Mid-scale SST anomaly reconstruction



Conveying uncertainty

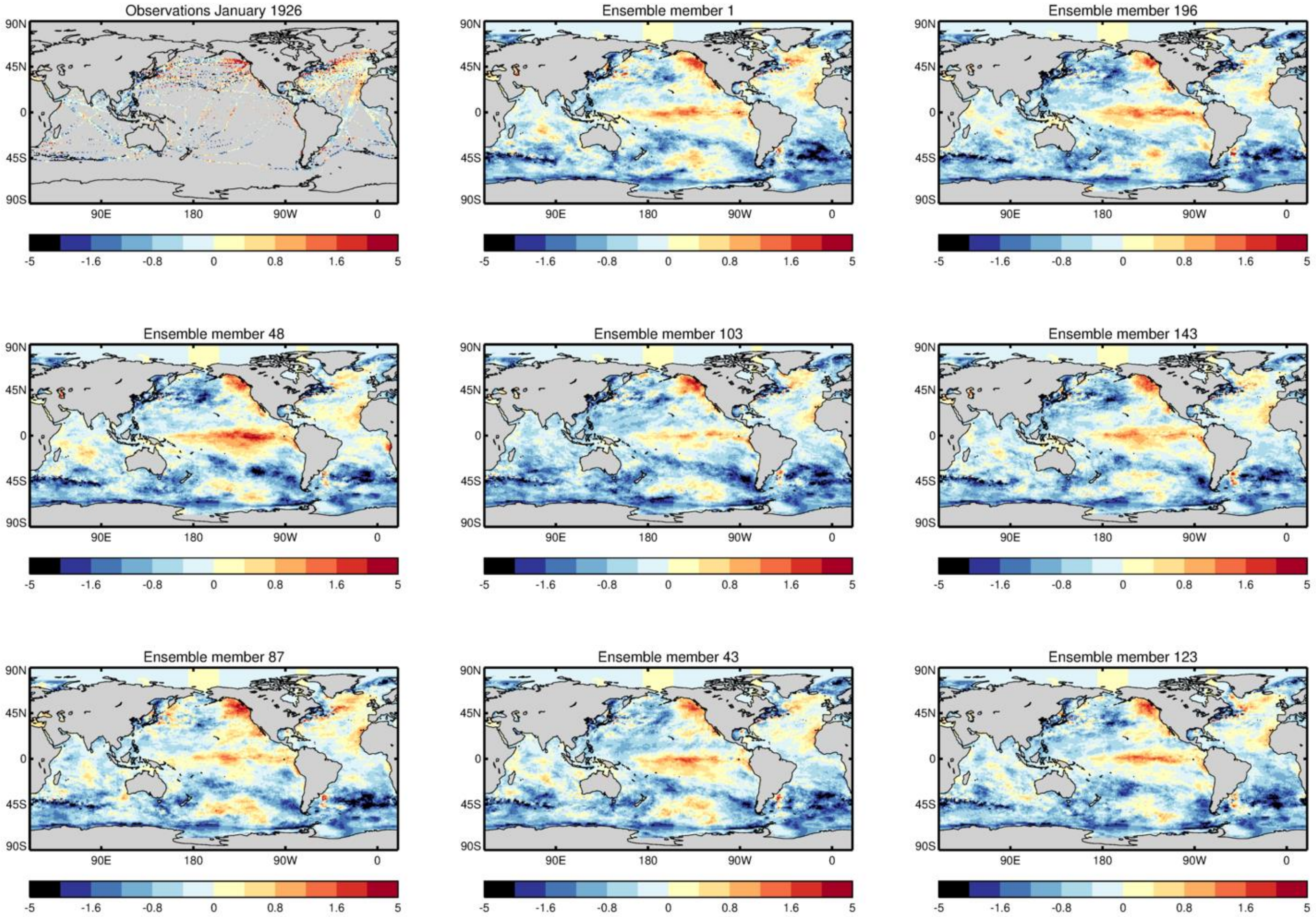
Met Office Hadley Centre

Generating an ensemble (HadISST.2.4.0.0)

- Generate 20 ensemble members from 200 in situ ensemble members
- Each weighting series for the 1-degree covariance patterns used in the large-scale reconstruction has its own time-varying uncertainty estimate
- Fitting a Gaussian process to these weights allows a sample to be drawn from the posterior – this results in a sample reconstruction, one for each of the 20 ensemble members

Met Office Hadley Centre

Large-scale analysis samples January 1926

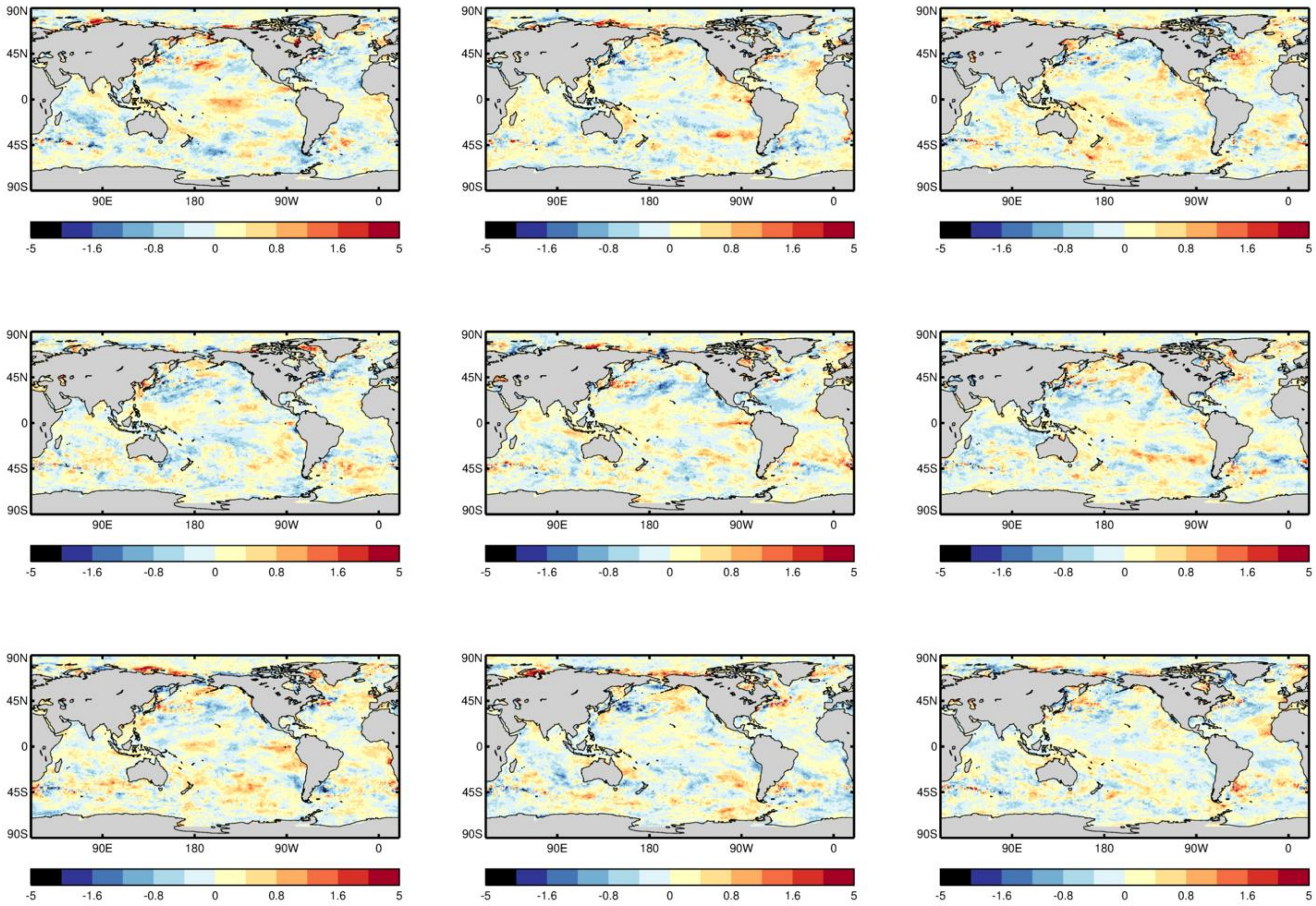


Generating an ensemble (HadISST.2.4.0.0)

- Generate 20 ensemble members from 200 in situ ensemble members
- Each weighting series for the 1-degree covariance patterns used in the large-scale reconstruction has its own time-varying uncertainty estimate
- Fitting a Gaussian process to these weights allows a sample to be drawn from the posterior – this results in a sample reconstruction, one for each of the 20 ensemble members
- 10,000 samples are drawn from the mid-scale covariance and samples at successive times are forced to correlate with each other as determined by satellite measurements
- Samples of the observation error are added and the covariance samples and uncertainty samples used to form a mid-scale analysis sample

Met Office
Hadley Centre

Mid-scale analysis samples

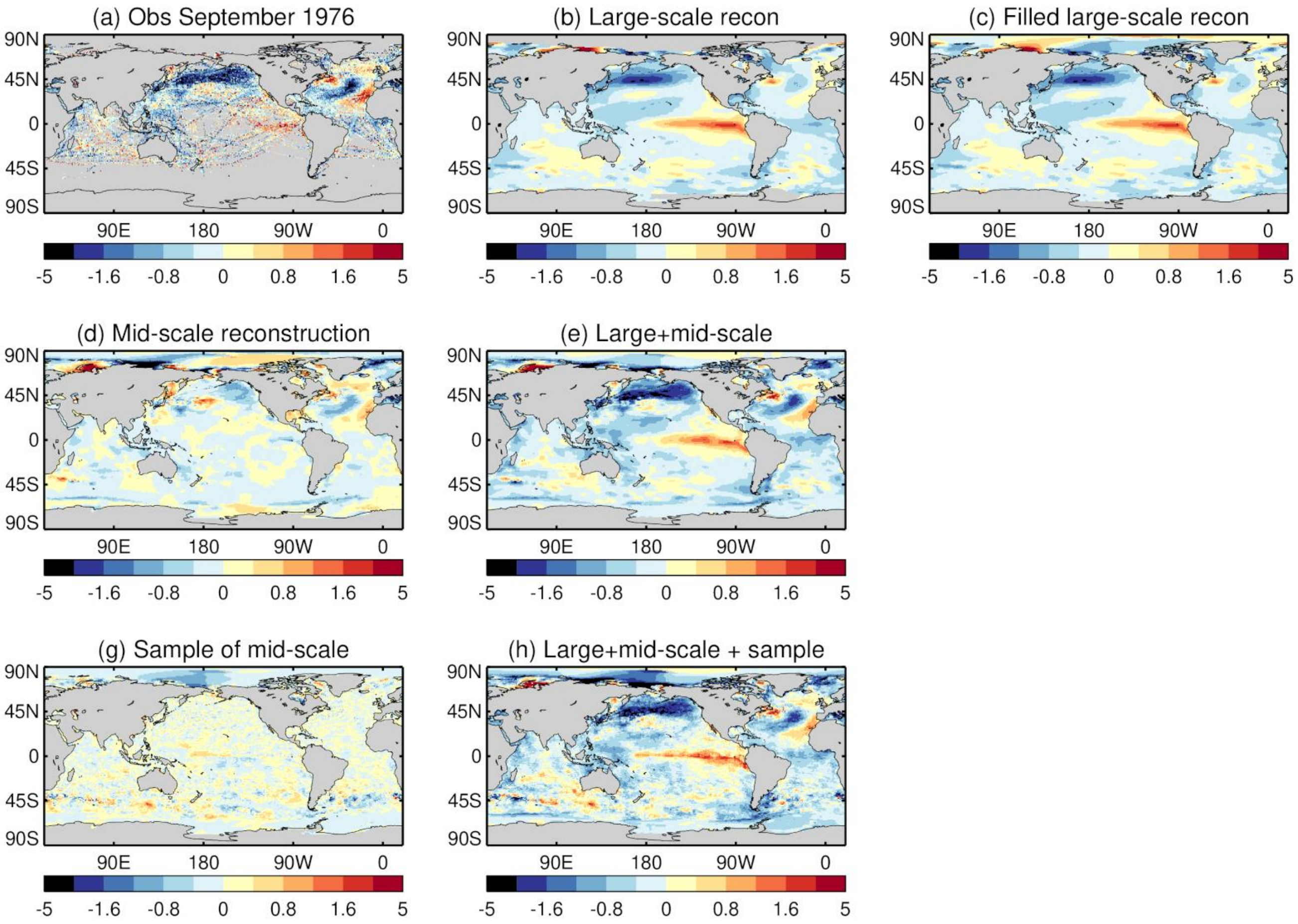


Generating an ensemble (HadISST.2.4.0.0)

- Generate 20 ensemble members from 200 in situ ensemble members
- Each weighting series for the 1-degree covariance patterns used in the large-scale reconstruction has its own time-varying uncertainty estimate
- Fitting a Gaussian process to these weights allows a sample to be drawn from the posterior – this results in a sample reconstruction, one for each of the 20 ensemble members
- 10,000 samples are drawn from the mid-scale covariance and samples at successive times are forced to correlate with each other as determined by satellite measurements
- Samples of the observation error are added and the covariance samples and uncertainty samples used to form a mid-scale analysis sample
- Each ensemble member = large-scale reconstruction sample + mid-scale analysis + mid-scale analysis sample

Constructing an SST anomaly ensemble member

September 1976



Consistency between SST and sea ice

What this means in practice

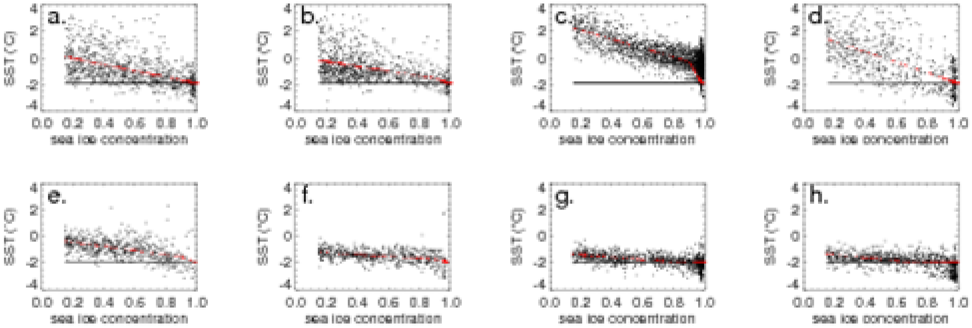
- Few measurements available from which to form an analysis in ice-covered regions
- Need to make informed assumptions about how SST varies in ice-covered grid boxes
- Sea ice concentration usually used to determine SST
- These proxy observations then used in the analysis

How to make SST and sea ice consistent – different methods

- COBE-2: SST estimated using quadratic functions of SIC that reflect the empirical relationship between SIC and observed SST at the same positions on the same day; freezing point depends on salinity
- OSTIA/SST CCI: For regions with greater than 50% concentration the background is relaxed towards −1.8C in the ocean
- HadISST/OI.v2: quadratic relationships between SST and sea ice concentration
- HadISST.2.4.0.0: linear relationships between SST and sea ice concentration
- Nielsen-Englyst et al (2023): combined SST and IST analysis on 0.05 degree daily grid

HadISST.2.4.0.0 method

- Linear relationships between SST and sea ice concentration
- Estimated separately for each hemisphere, each calendar month.
- Vary with longitude in overlapping 21-degree longitude bands.
- Separate relationships are developed for outlying regions



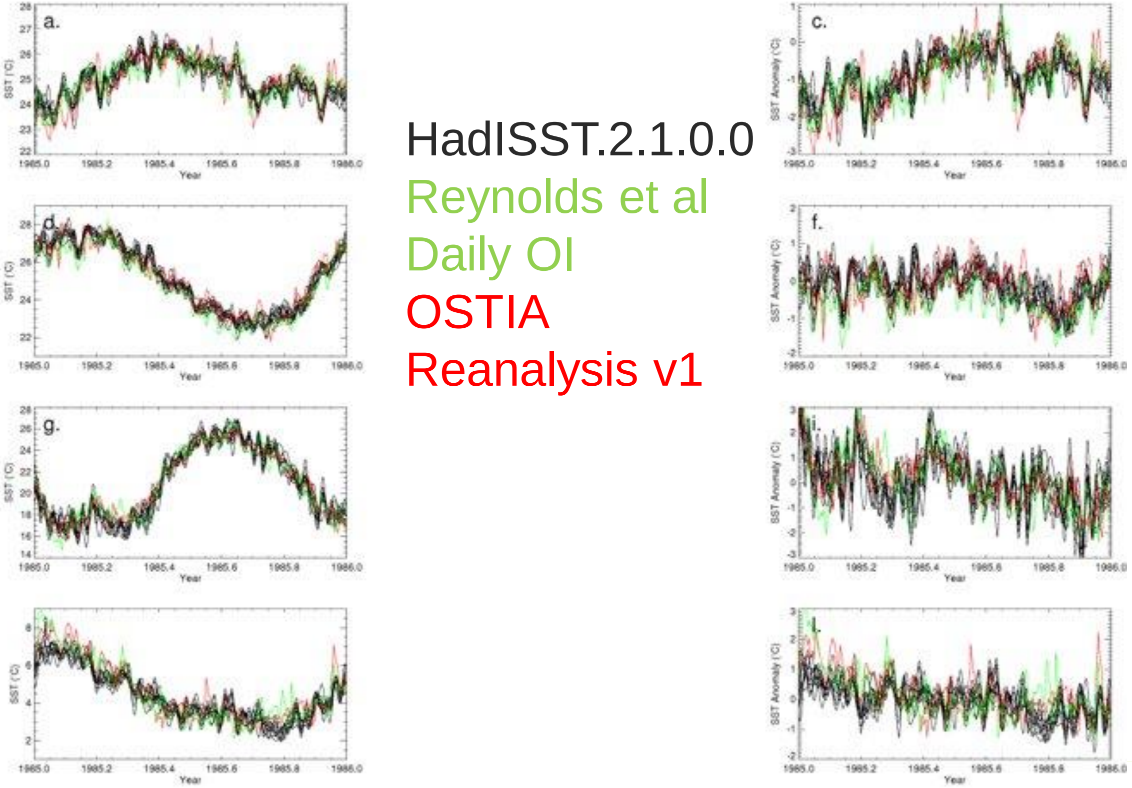
Sea surface temperature (°C) versus sea ice concentration, 1982-2007: a)-d) Arctic; e)-h) Antarctic; a) and e) January; b) and f) April; c) and g) July and d) and h) October.

Further considerations for making usable fields

Met Office
Hadley Centre

Creating dailies from monthly/pentad fields (HadISST.2.1.0.0)

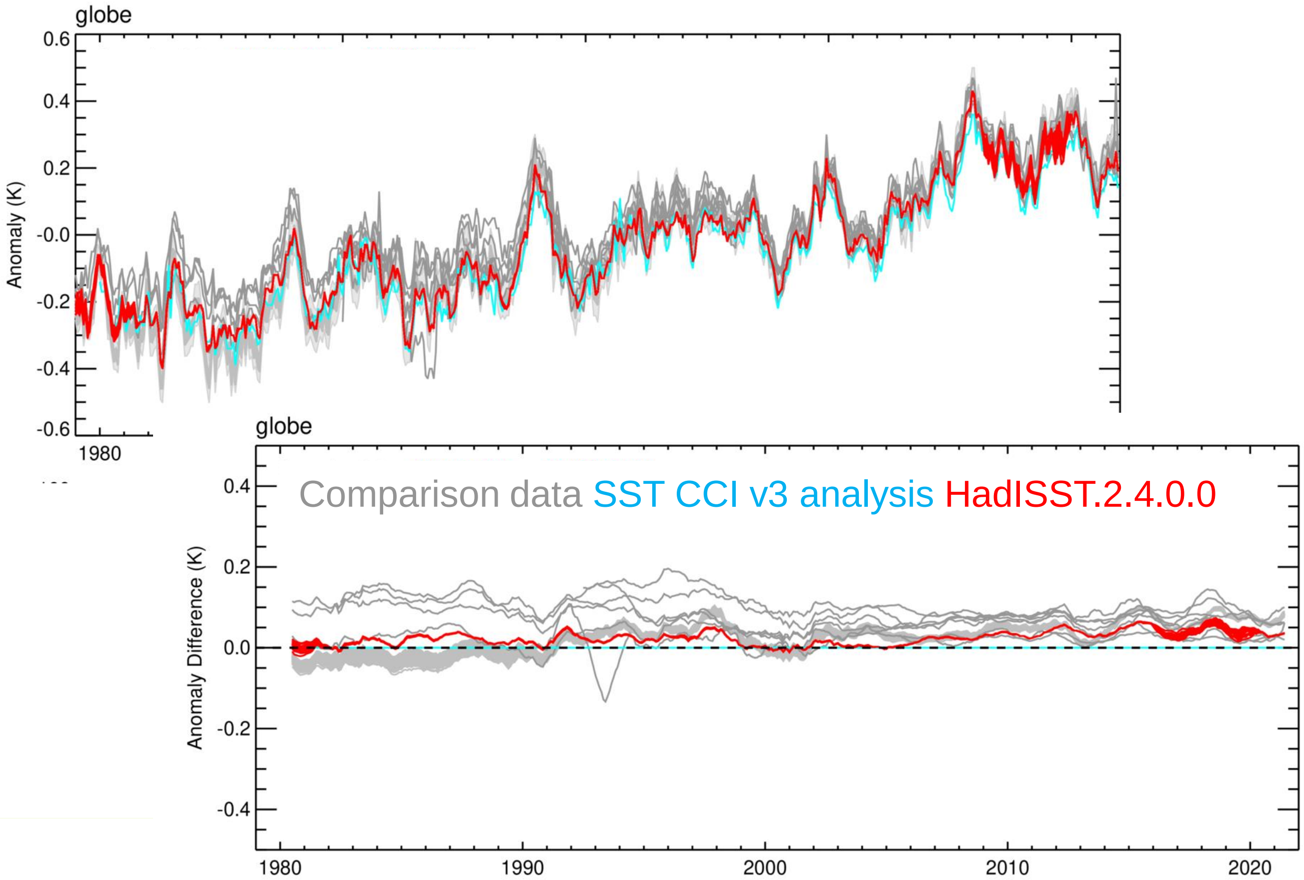
- 1 degree anomaly analysis on monthly or pentad resolution combined with 0.25 degree lat/long daily climatology to provide daily actual SST fields
- Fit cubic spline to monthly/pentad anomalies to interpolate to dailies

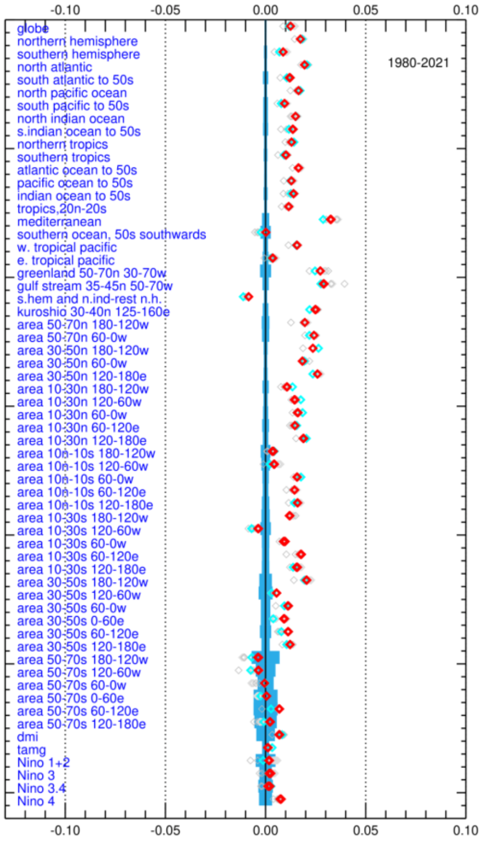


SST

SST anoms

Employ families of related data sets



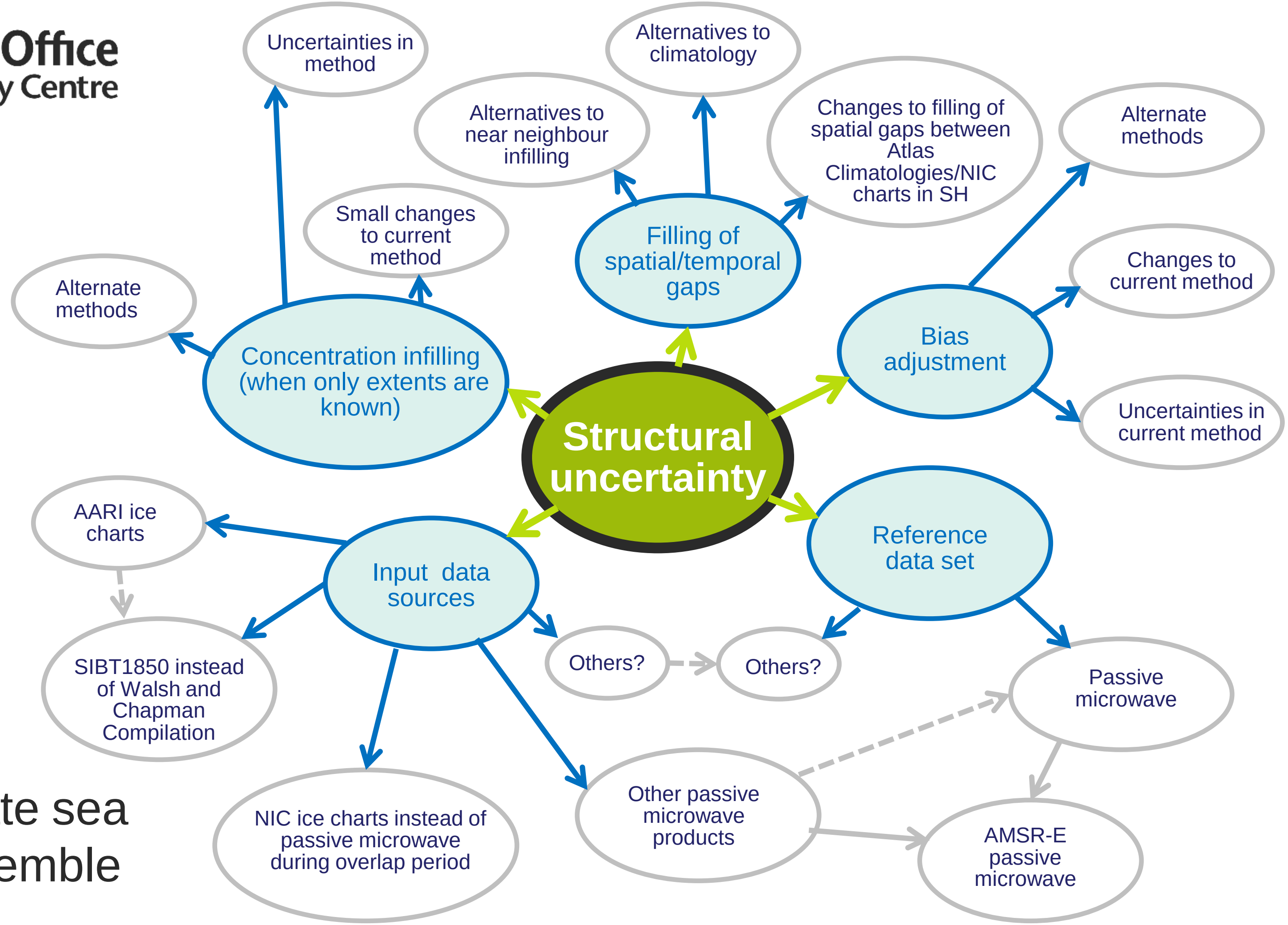


Outstanding challenges

Outstanding historical SST work

- Use feedback from reanalysis to improve bias corrections?
- Create an efficient pipeline for incorporating newly digitised historical SST observations and metadata into global data bases – new ML data rescue methods will increase the need for this
- Pull through recent new research and understanding into improved forcing data sets more quickly
- Understand in situ SST measurements better from a metrological point of view to improve uncertainty estimates
- Do we need to understand how to represent diurnal variability?

Met Office
Hadley Centre



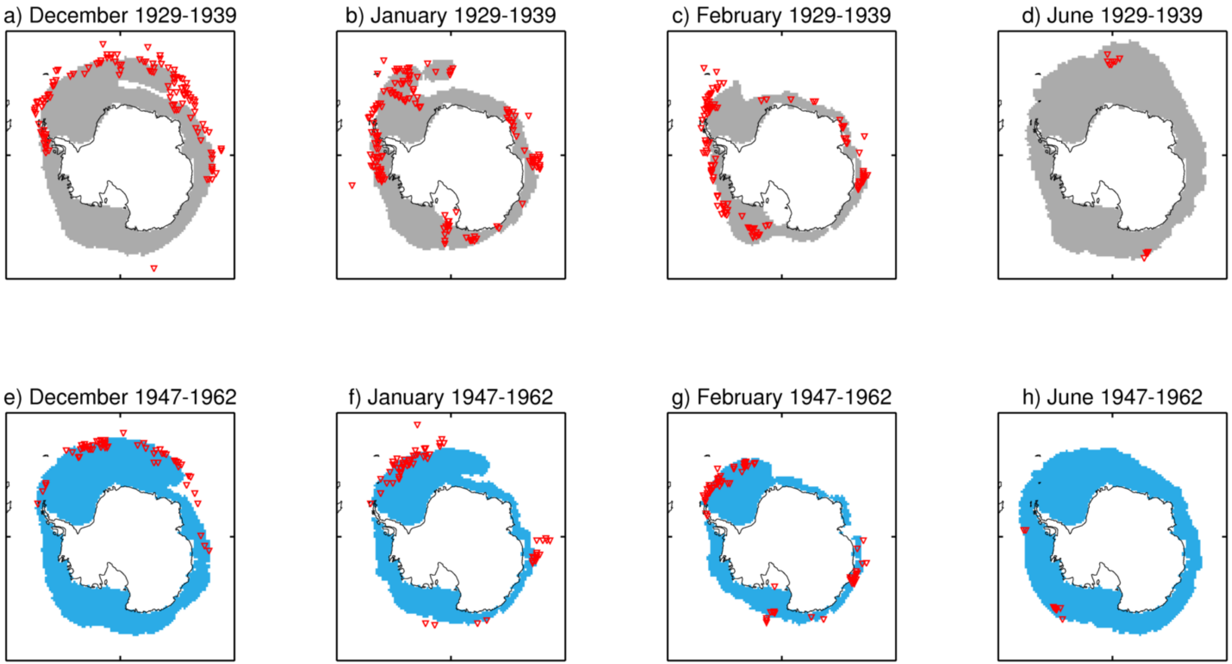
Generate sea ice ensemble

Need more historical Antarctic sea ice observations

German climatology, 1929-1939

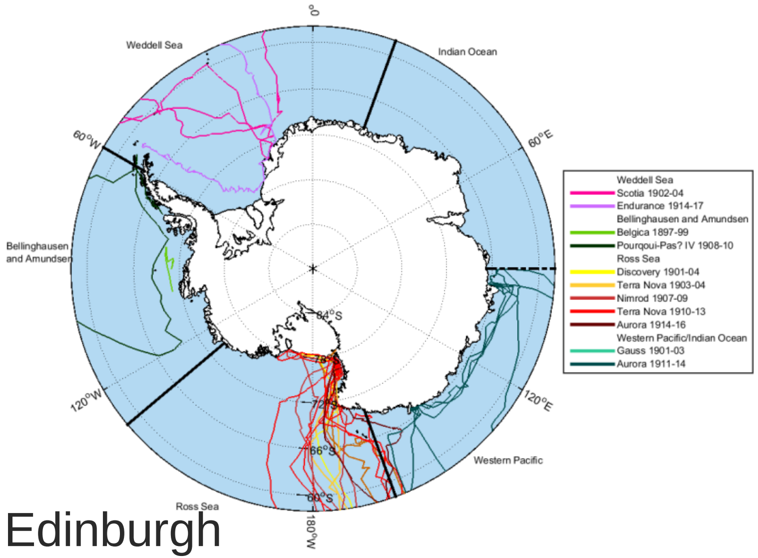
Russian climatology, 1947-1962

Southern Ocean Ice Reports



Titchner and Rayner, 2014, https://agupubs.onlinelibrary.wiley.com/doi/full/10.1002/2013JD020316

Recovery of historical Antarctic sea ice



Edinburgh & Day (2016) https://tc.copernicus.org/preprints/tc-2016-90/tc-2016-90.pdf

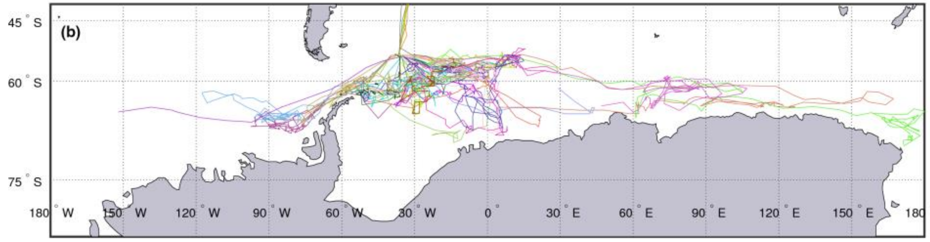
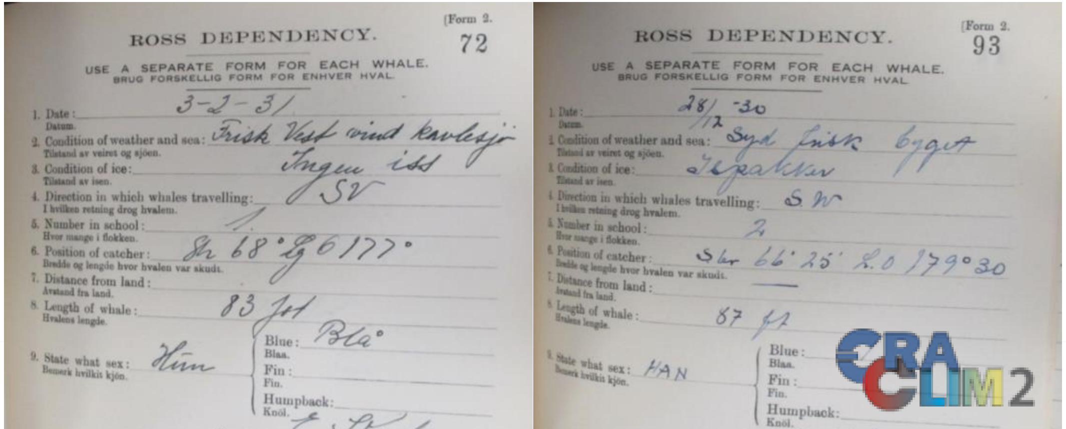


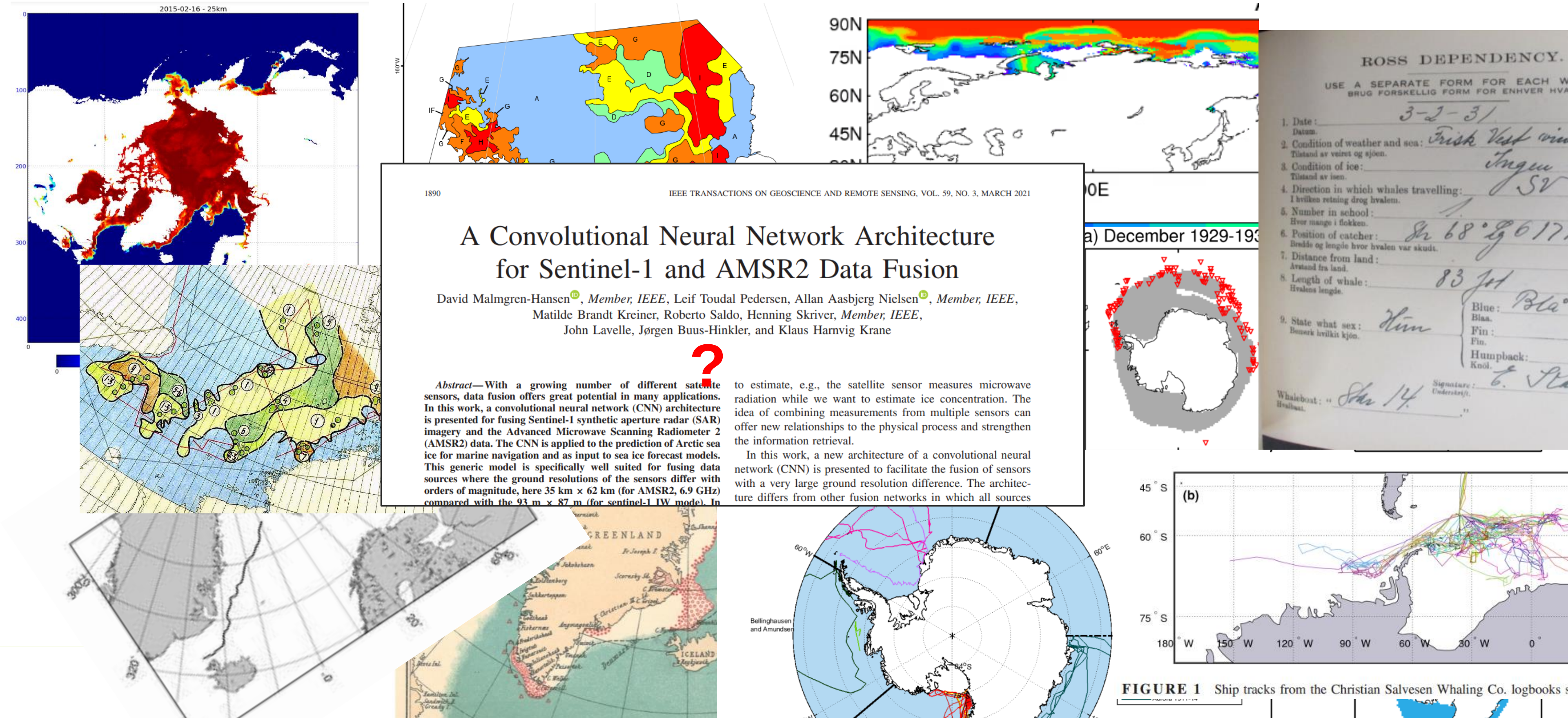
FIGURE 1 Ship tracks from the Christian Salvesen Whaling Co. logbooks separated into 'Ship ID' groups and denoted by different colours;

Teleti et al (2018), https://rmets.onlinelibrary.wiley.com/doi/pdf/10.1002/gdj3.65



Right: Whaling Records – Archive of Sea Mammal Research Unit, University of St. Andrews, Clive Wilkinson

Combination of sea ice data types using machine learning



1890 IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, VOL. 59, NO. 3, MARCH 2021

A Convolutional Neural Network Architecture for Sentinel-1 and AMSR2 Data Fusion

David Malmgren-Hansen, *Member, IEEE*, Leif Toudal Pedersen, Allan Aasbjerg Nielsen, *Member, IEEE*, Matilde Brandt Kreiner, Roberto Saldo, Henning Skriver, *Member, IEEE*, John Lavelle, Jørgen Buus-Hinkler, and Klaus Harnvig Krane

***Abstract*—With a growing number of different satellite sensors, data fusion offers great potential in many applications. In this work, a convolutional neural network (CNN) architecture is presented for fusing Sentinel-1 synthetic aperture radar (SAR) imagery and the Advanced Microwave Scanning Radiometer 2 (AMSR2) data. The CNN is applied to the prediction of Arctic sea ice for marine navigation and as input to sea ice forecast models. This generic model is specifically well suited for fusing data sources where the ground resolutions of the sensors differ with orders of magnitude, here 35 km × 62 km (for AMSR2, 6.9 GHz) compared with the 93 m × 87 m (for sentinel-1 IW mode). In**

to estimate, e.g., the satellite sensor measures microwave radiation while we want to estimate ice concentration. The idea of combining measurements from multiple sensors can offer new relationships to the physical process and strengthen the information retrieval.

In this work, a new architecture of a convolutional neural network (CNN) is presented to facilitate the fusion of sensors with a very large ground resolution difference. The architecture differs from other fusion networks in which all sources

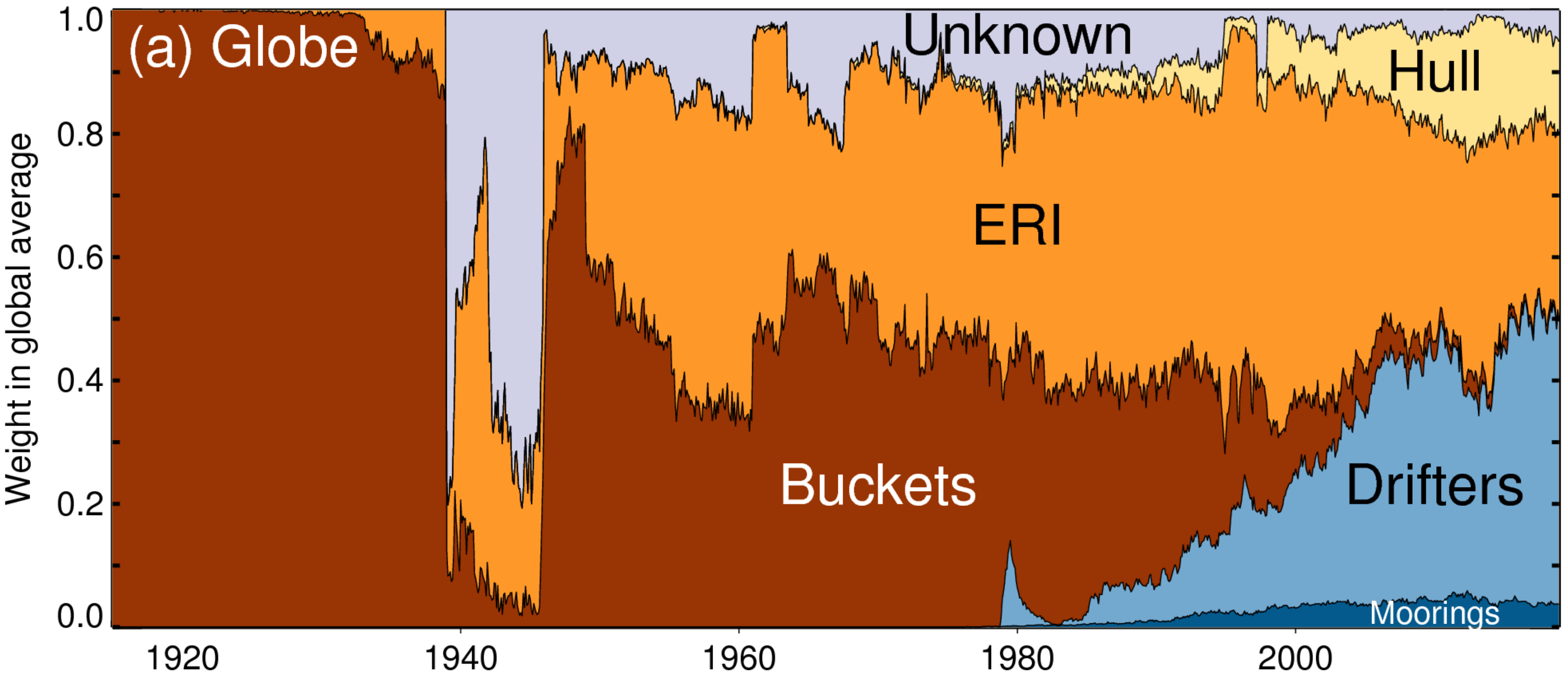
FIGURE 1 Ship tracks from the Christian Salvesen Whaling Co. logbooks

Summary

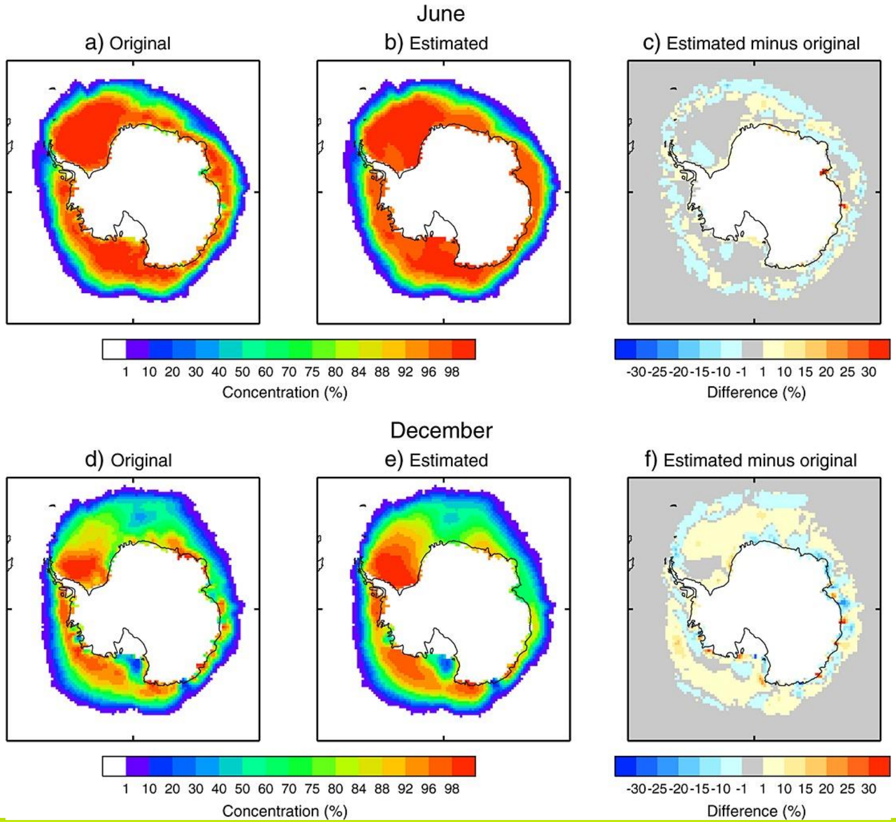
- Reconstructing SST and sea ice data sets is a research-intensive, multi-step process that encompasses homogenisation, infilling and uncertainty quantification, inter alia
- The science behind SST data set reconstruction is mature, but more can be done to enable the pull through of newly digitised data and metadata and understanding to improve data sets further
- Satellite-derived records of SST are now so well-understood that families of high- and low-resolution data sets can be created that are more consistent with one another
- Much work is still needed on reconstruction of sea ice data, including data rescue, development of novel data combination techniques and ensemble generation

Extra slides

Composition of ICOADS as understood by Kennedy et al 2019



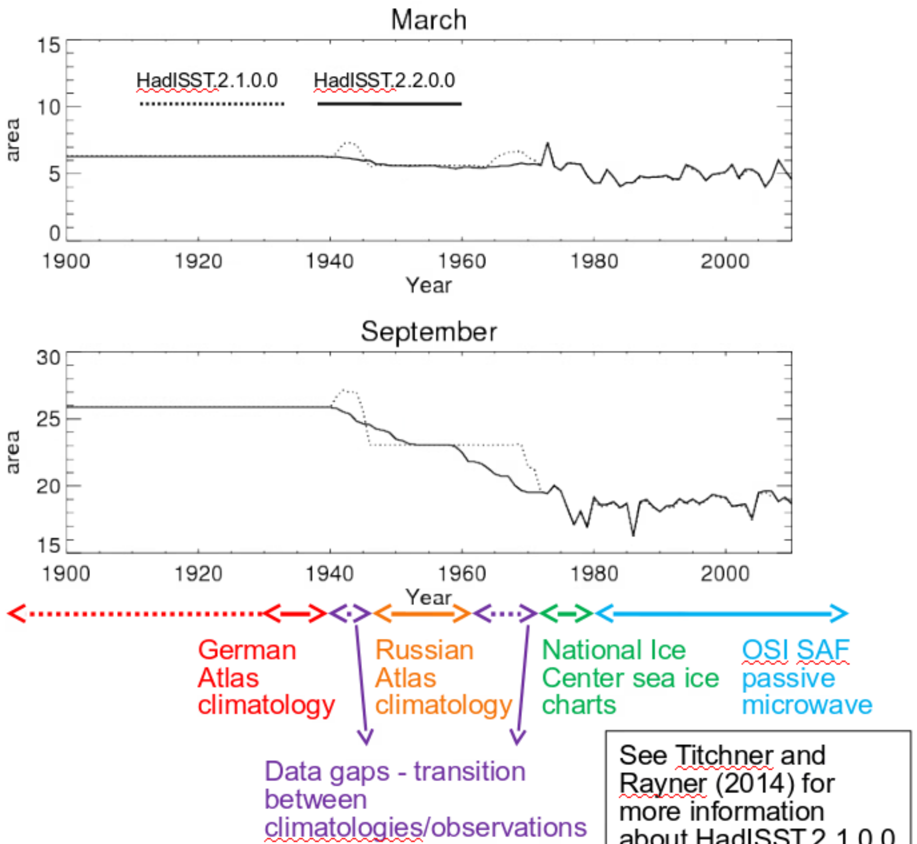
Infilling concentrations



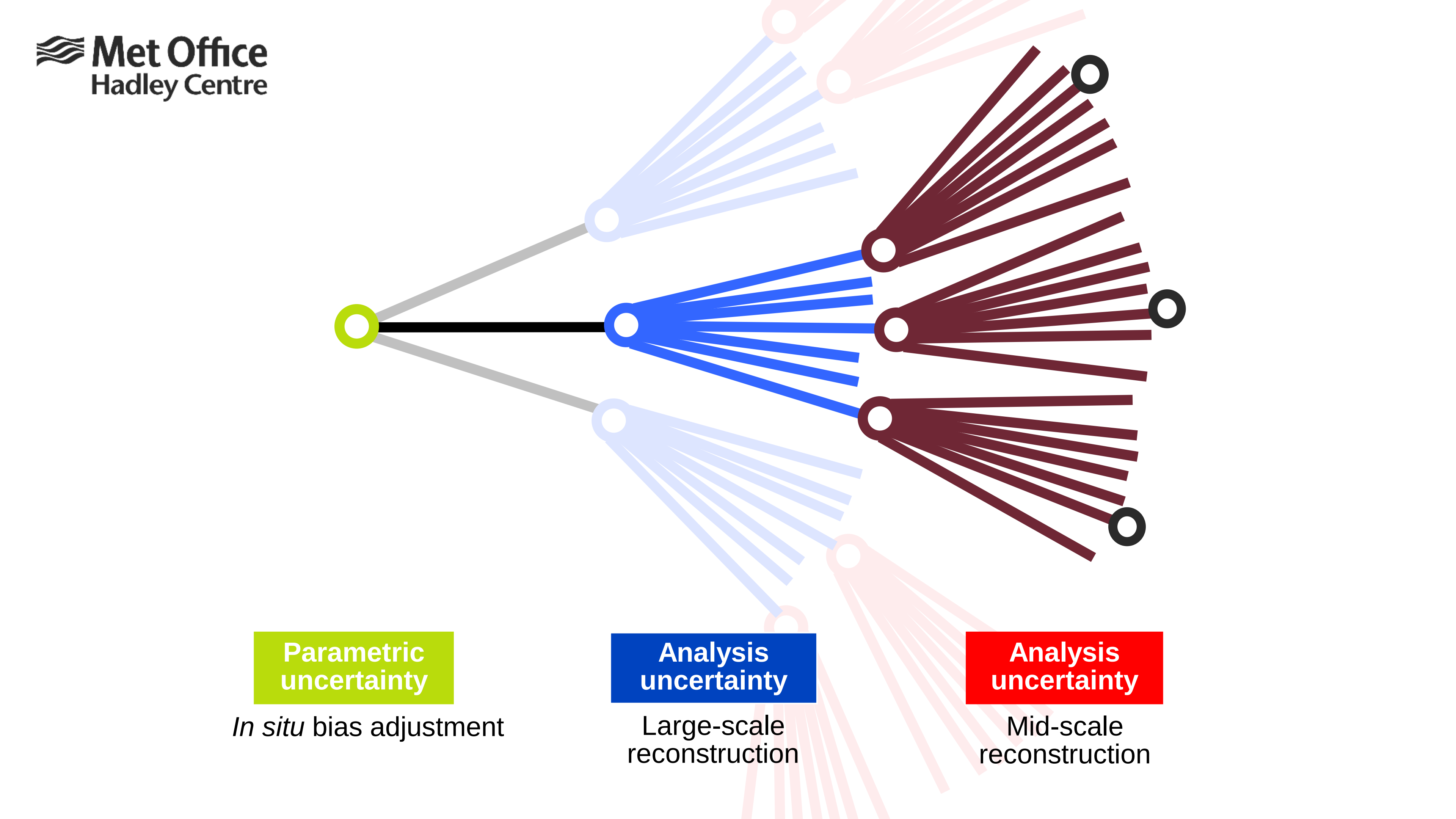
Titchner and Rayner, 2014, https://agupubs.onlinelibrary.wiley.com/doi/full/10.1002/2013JD020316

In HadISST.2.1.0.0 sea ice concentrations were linearly interpolated – creates unrealistic concentration fields and sudden changes in extents

In HadISST.2.2.0.0 sea ice the location of the ice edge is linearly interpolated at each 1 degree longitude



See Titchner and Rayner (2014) for more information about HadISST.2.1.0.0


Met Office
Hadley Centre
Parametric uncertainty
In situ bias adjustment
Analysis uncertainty
Large-scale reconstruction
Analysis uncertainty
Mid-scale reconstruction