Assessment and Validation of the NOCS2.0 Dataset

David I. Berry and Elizabeth C. Kent
National Oceanography Centre
dyb@noc.ac.uk

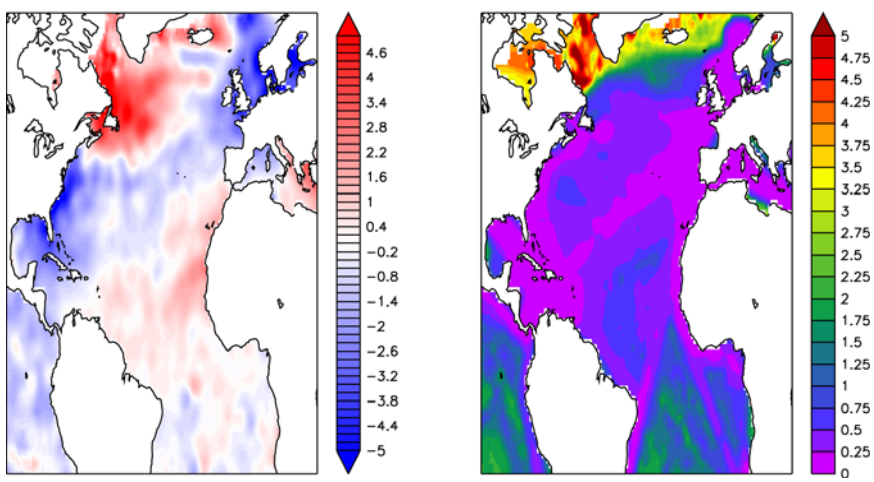
Outline of talk

1) Overview of the NOCS dataset
2) Assessment and validation of the dataset
3) Further work
4) Summary

Overview of the NOCS dataset: what the dataset contains

- Monthly mean 1° estimates of surface fluxes and meteorological parameters over the oceans
  - Surface (10 m) air temperature and humidity
  - Surface (10 m) wind speed
  - Sea surface temperature
  - Sensible and latent (evaporation) heat fluxes
  - Sea level pressure, cloud cover
- "Realistic" sampling / measurement uncertainty estimates for each parameter
- Bias uncertainty estimates made
- Only Voluntary Observing Ships (VOS) observations (ICOADS 2.4 / 2.5) used - independence from other sources maintained (e.g. moored buoys)

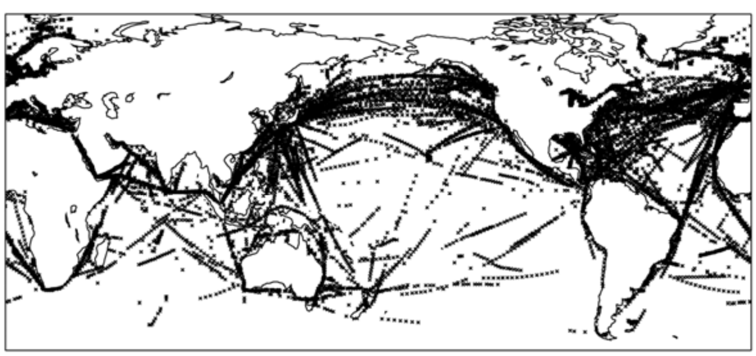
Air temperature anomaly (left) and uncertainty (right) – Dec 2010



Overview of the NOCS dataset: construction method

- Voluntary observing ship observations (VOS) height (and bias) adjusted then averaged onto daily 1° grid
  - Optimal interpolation used following Lorenc (1981) and Reynolds and Smith (1994)
  - Uncertainties for individual observations estimated using semi-variograms following Kent and Berry (2005)
  - Optimal interpolation (OI; e.g. Lorenc 1981, Reynolds and Smith 1994) used to give spatially complete daily fields
  - Uncertainty estimates made as part of OI process
- Daily estimates of the latent and sensible heat fluxes made using bulk formulae (Smith, 1980; Smith 1988)
- Uncertainty in daily fluxes estimated using propagation of errors
- Daily fields averaged taking into account correlation between days

Inhomogeneous sampling



Air Temperature Observations (01–Jul–1994 to 04–Jul–1994)

Overview of the NOCS dataset: construction method

- Voluntary observing ship observations height (and bias) adjusted then averaged onto daily 1° grid
  - Optimal interpolation used following Lorenc (1981) and Reynolds and Smith (1994)
  - Uncertainties for individual observations estimated using semi-variograms following Kent and Berry (2005)
  - Optimal interpolation (OI; e.g. Lorenc 1981, Reynolds and Smith 1994) used to give spatially complete daily fields
  - Uncertainty estimates made as part of OI process
- Daily estimates of the latent and sensible heat fluxes made using bulk formulae (Smith, 1980; Smith 1988)
- Uncertainty in daily fluxes estimated using propagation of errors
- Daily fields averaged taking into account correlation between days

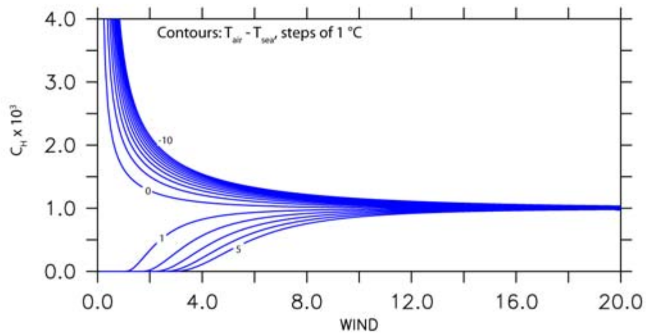
Bulk formulae for turbulent fluxes

$$Q_H = c_p \rho_0 C_H U(T_{air} - T_{sea})$$
$$Q_E = L_v \rho_0 C_E U(q_{air} - q_{sea})$$

c_p = specific heat capacity of air
ρ_0 = density of air
L_v = latent heat of vapourisation
U = wind speed
T_{air} = air temperature
T_{sea} = sea surface temperature
Q_{air} = specific humidity
Q_{sea} = specific humidity at sea surface
C_H = heat transfer coefficient
C_E = moisture transfer coefficient

Transfer coefficients non-linear function of temperature, pressure, humidity and wind speed

Bulk formulae for turbulent fluxes



Dependence of C_H on wind speed and air – sea temperature difference

Overview of the NOCS dataset: construction method

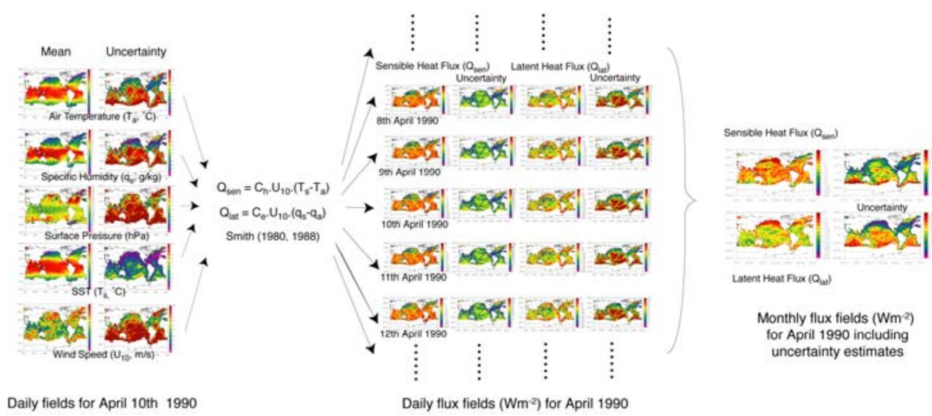
- Voluntary observing ship observations height (and bias) adjusted then averaged onto daily 1° grid
  - Optimal interpolation used following Lorenc (1981) and Reynolds and Smith (1994)
  - Uncertainties for individual observations estimated using semi-variograms following Kent and Berry (2005)
  - Optimal interpolation (OI; e.g. Lorenc 1981, Reynolds and Smith 1994) used to give spatially complete daily fields
  - Uncertainty estimates made as part of OI process
- Daily estimates of the latent and sensible heat fluxes made using bulk formulae (Smith, 1980; Smith 1988)
- Uncertainty in daily fluxes estimated using propagation of errors
- Daily fields averaged taking into account correlation between days

Overview of the NOCS dataset: construction method

- Voluntary observing ship observations height (and bias) adjusted then averaged onto daily 1° grid
  - Optimal interpolation used following Lorenc (1981) and Reynolds and Smith (1994)
  - Uncertainties for individual observations estimated using semi-variograms following Kent and Berry (2005)
  - Optimal interpolation (OI; e.g. Lorenc 1981, Reynolds and Smith 1994) used to give spatially complete daily fields
  - Uncertainty estimates made as part of OI process
- Daily estimates of the latent and sensible heat fluxes made using bulk formulae (Smith, 1980; Smith 1988)
- Uncertainty in daily fluxes estimated using propagation of errors
- Daily fields averaged taking into account correlation between days

See Berry and Kent (2009, 2011) for further details

Flux calculation strategy



Assessment and validation

- Cross validation experiments
  - Observations excluded from OI and compared to output fields
- Comparison to independent buoy observations
  - Observations from research moorings provide independent validation datasets
  - But limited number of locations / deployment lengths
- Comparison to input data
  - Not independent but provides sanity check to make sure we get out what we put in

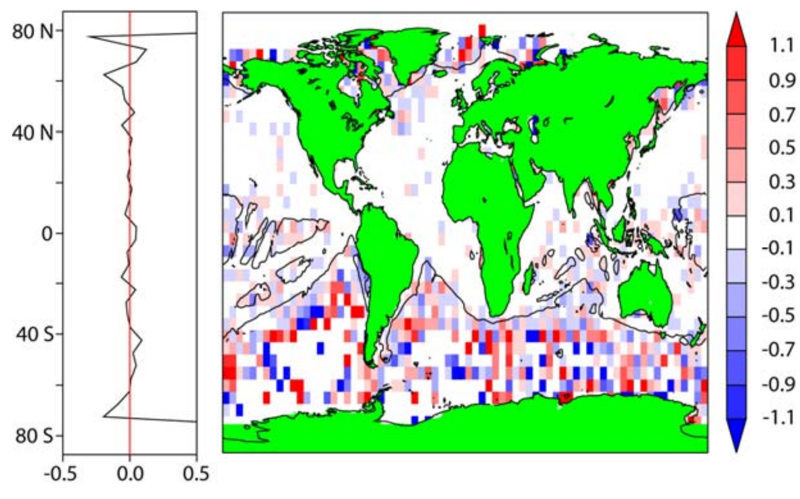
Cross validation

- Optimal interpolation only optimal when error co-variances known
- Method should produce unbiased fields even when error co-variances poorly defined
- Uncertainty estimates may be biased
- In NOCS2.0 a basic error covariance model is used (Gaussian, fixed isotropic length scales)
- Method tested using two different cross-validation experiments, examining the (potential) bias in the mean fields and bias in the uncertainty estimates

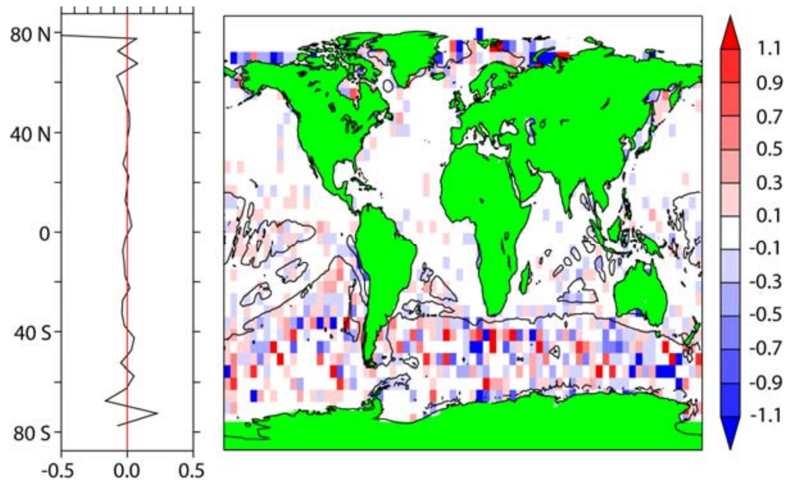
Cross validation – bias in the mean fields due to OI

- Optimal interpolation run 10 times excluding 10 % of observations randomly each run
- For excluded observations, OI – observation differences calculated
- Differences averaged across ensemble of runs to give an estimate of the bias introduced by the method
- Similar results seen using larger ensembles

Mean difference OI – excluded observations: Air temperature, 1974



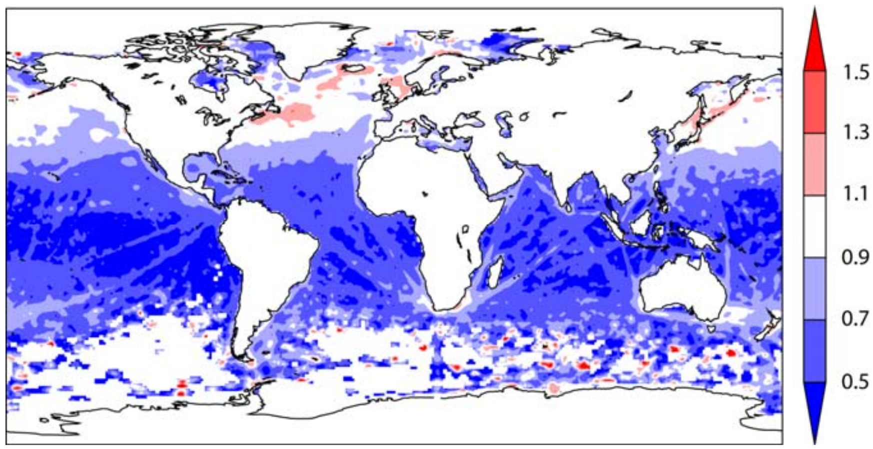
Mean difference OI – excluded observations: Sea surface temperature, 1974



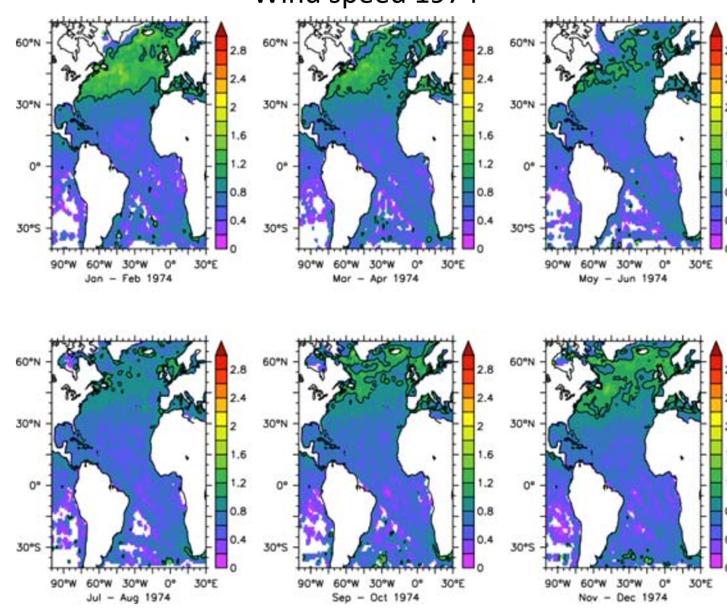
Cross validation – bias in the uncertainty estimates due to OI

- Optimal interpolation run 10 times excluding 50 % of observations randomly each run
- Standard deviation of daily fields across ensemble calculated (i.e. standard deviation of the mean)
- Ensemble standard deviations of the daily values compared to the mean error estimates
- Ratio of standard deviation to uncertainty estimate gives estimate of whether we are over- or underestimating the uncertainty
- Results unreliable in poorly sampled areas

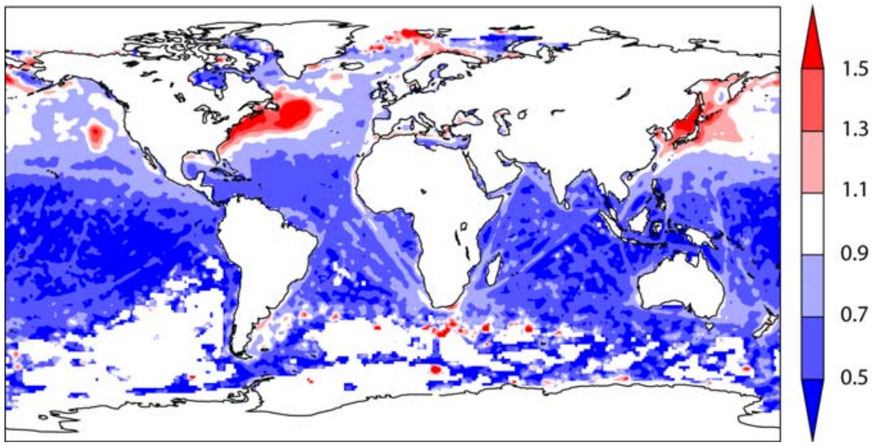
Annual mean ratio of standard deviation to error estimates
Wind speed 1974



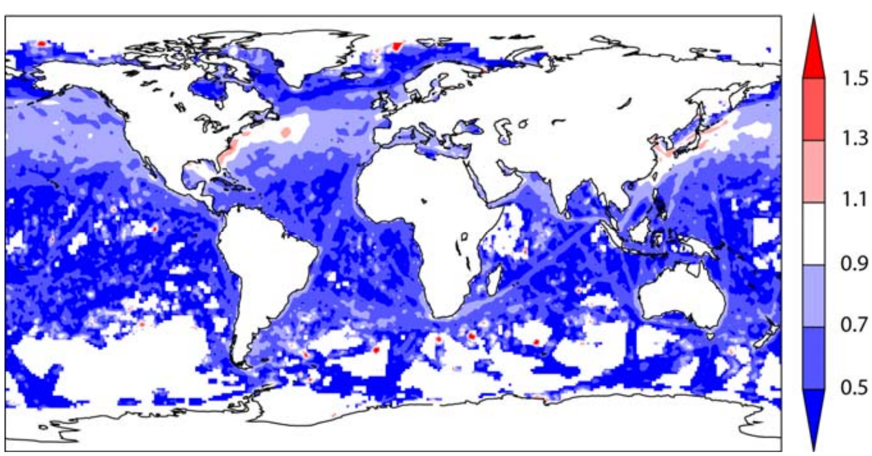
Bi-monthly mean ratio of standard deviation to error estimates
Wind speed 1974



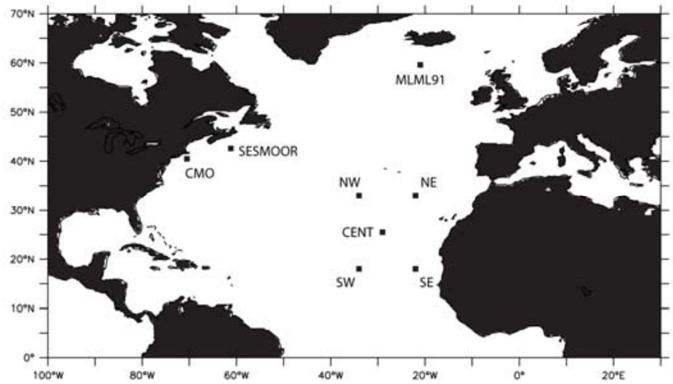
Annual mean ratio of standard deviation to error estimates
Air temperature 1974



Annual mean ratio of standard deviation to error estimates
Air temperature 2010



Comparison with moored buoy data



Data from Woods Hole Upper Ocean Mooring Data Archive at http://uop.whoi.edu/uopdata/

Comparison to WHOI UOP Moorings – Latent Heat Flux

Mooring	NOCS2.0	OAFlux	ERA40	NCEP1	NCEP2	NOCS v1.1
SW	-14.2 ± 47.1 (17.6)	20.6 ± 20.1	*5.6 ± 21.1*	-6.4 ± 29.1	-12.4 ±	0
SE	*1.3 ± 43.1 (17.1)*	11.1 ± 10.4	7.9 ± 20.7	-6.4 ± 25.4	-14.6 ±	-5
NW	-1.2 ± 34.3 (16.9)	8.6 ± 22.9	*-0.4 ±*	9.2 ± 25.8	1.5 ± 30.0	-14
NE	13.7 ± 38.6 (16.9)	13.1 ± 22.3	*7.5 ± 19.0*	8.7 ± 24.4	29.0 ±	4
CENT	10.2 ± 45.2 (16.9)	12.98 ±	*1.2 ± 20.8*	7.2 ± 28.2	2.3 ± 31.5	-9
SESMOOR	*-1.9 ± 73.1 (18.0)*	-2.4 ± 76.8	-18.3 ±	-31.3 ±	-52 ± 99.0	N/A
CMO	*-10.2 ± 36.2*	-14.5 ±	-35.0 ±	-37.1 ±	-50.1 ±	N/A
MLML91	*1.4 ± 22.5 (17.1)*	5.5 ± 13.2	2.3 ± 11.8	1.9 ± 13.8	-7.9 ± 19.7	N/A

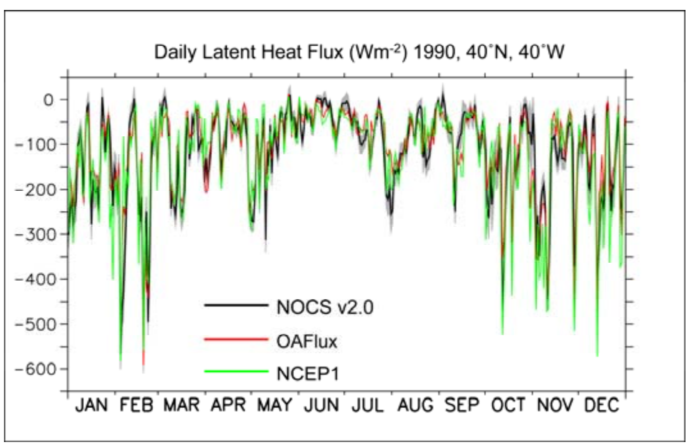
Mean daily latent heat flux difference (W m^{-2}) (product - buoy) over period of buoy deployment. Mean ± sdv (uncertainty). Values for NOCS v1.1 are for monthly mean fluxes (Josey, 2001).

Comparison to WHOI UOP Moorings – Sensible Heat Flux

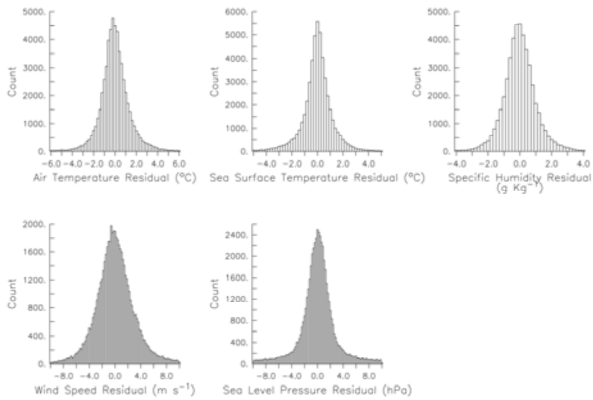
Mooring	NOCS2.0	OAFlux	ERA40	NCEP1	NCEP2	NOCS v1.1
SW	-7.9 ± 8.9 (4.0)	-2.3 ± 4.8	-5.9 ± 5.5	-11.2 ± 10.4	*-1.6 ± 7.2*	-4
SE	*-0.2 ± 8.0 (3.6)*	4.7 ± 5.5	-0.3 ± 4.9	-1.7 ± 7.8	5.5 ± 5.7	1
NW	*-0.1 ± 7.2 (3.4)*	0.7 ± 6.8	-3.8 ± 7.6	-1.9 ± 9.9	0.8 ± 8.8	0
NE	2.7 ± 7.6 (3.4)	*0.2 ± 5.1*	-2.5 ± 5.6	-4.1 ± 8.4	1.0 ± 6.6	3
CENT	-0.5 ± 8.1 (3.3)	-1.3 ± 4.3	-4.5 ± 5.0	-4.4 ± 7.5	*0.5 ± 6.2*	-1
SESMOOR	*-2.2 ± 52.3 (8.1)*	-10.1 ± 56.6	-13.6 ±72.2	-31.9 ± 78.6	-37.1 ± 86.0	N/A
CMO	*-4.1 ± 26.3 (5.3)*	-5.3 ± 15.7	-13.2 ± 26.5	-16.4 ± 28.5	-18.9 ± 47.0	N/A
MLML91	4.0 ± 9.7 (3.6)	2.4 ± 5.6	2.3 ± 7.3	*0.1 ± 8.4*	4.5 ± 7.9	N/A

Mean daily sensible heat flux difference (W m^{-2}) (product - buoy) over period of buoy deployment. Mean ± sdv (uncertainty). Values for NOCS v1.1 are for monthly mean fluxes (Josey, 2001).

North Atlantic Latent Heat Flux (Wm^{-2})

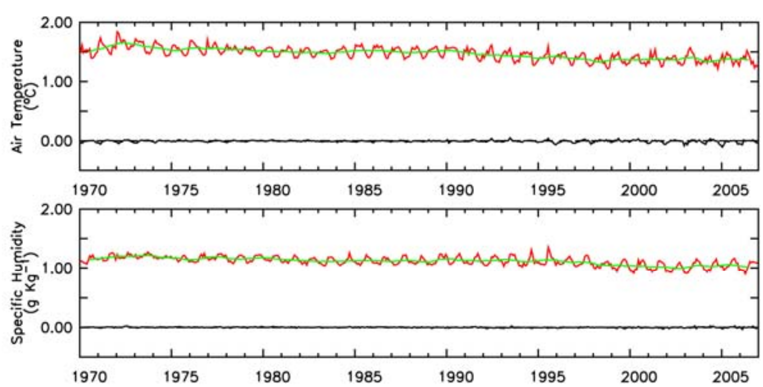


Comparison to VOS observations (January 1993)



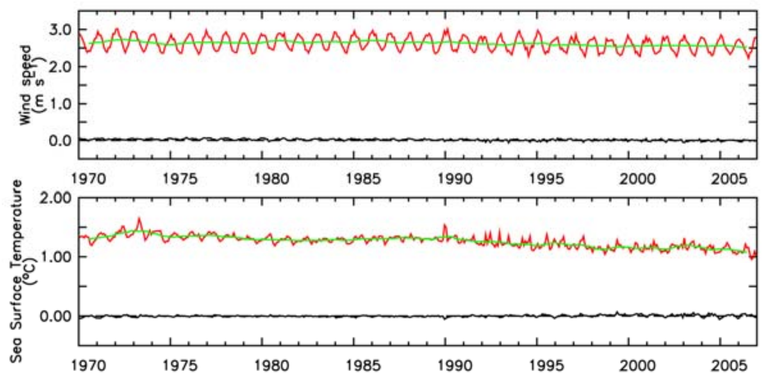
Differences between observations and interpolated values in NOCS2.0
(not independent)

Comparison to VOS observations



RMS differences between OI fields and input data

Comparison to VOS observations



RMS differences between OI fields and input data

Summary

- NOCS Dataset contains estimates of all variables required to estimate the surface sensible and latent heat fluxes + the fluxes
- Uncertainty estimates provided alongside the data
- NOCS Dataset shown to be unbiased relative to input data (not independent)
- Cross validation of mean fields also shows output from OI to be unbiased
- NOCS dataset compares favorably to other datasets when compared to independent observations from research moorings
- Uncertainty estimates right order of magnitude – but improvements can be made through improved error covariance estimates

Future Work

- Improvement of length scales
  - Currently fixed length scales used
  - Preliminary work begun using variable length scales
- Extension backwards in time
  - Initially back to 1954 when metadata begins to be available
  - Examine possibility of extending back in time further (problems with data coverage, availability of certain variables and metadata)
- Expansion to other variables
  - Version 1 of the dataset included wind stress and precipitation
  - We aim to include these in version 2 as well but further work needed
- Hi-resolution version
  - Sub daily and < 1° spatial resolution
  - Use of other data sources (e.g. satellite fields, drifting buoys) but loss of independence
  - Equal area grid

References / Further Information

- Berry, D. I., and E. C. Kent, 2009: A New Air - Sea Interaction Gridded Data Set from ICOADS with Uncertainty Estimates. *Bulletin Of The American Meteorological Society, 90, 645 - 656, DOI: 10.1175/2008BAMS2639.1.*

- Berry, D. I., and E. C. Kent, 2011: Air-Sea Fluxes from ICOADS: The Construction of a New Gridded Dataset with Uncertainty Estimates. *International Journal of Climatology, In press, DOI: 10.1002/joc.2059*

- *http://www.noc.soton.ac.uk/noc_flux/* (links to dataset, above papers and background research)

Questions?