# Assimilation of T&S and Altimetry into ocean models with water mass constraints

- Applications of ocean data assimilation
  - Reanalysis, Seasonal forecasting, Operational oceanography
- Key ocean data sets
  - Altimeter, In Situ, SST
- Sequential Assimilation methods
- Constraints from oceanography
  - Vertical projection of altimeter sea level
  - Assimilating Temperature with Salinity corrections
  - Assimilating Salinity as S(T) instead of S(z)
- Detecting and Accounting for Bias Errors
- Conclusions





# **Ocean Reanalysis for Climate Studies**

30 HadCM3 1.25° + 5 yr smooth Heat content anomaly (10<sup>22</sup>J) Sampled 2.5° + 5 yr smooth Sampled 3.75° + 5 yr smooth 20 Sampled 5° + 5 yr smooth Levitus et al. (2000) 10 Ô See Gregory et al. [2004] 1960 1980 2000 2020 Year

Heat content for Anomaly for the upper 3000m

- Global and basin
   Scale Heat Content
- Salinity/Freshwater
   ⇒Hydrological cycle
- changes
- CO2 sequestration
- North Atlantic Deep Water Volume
- Changes in southern ocean T/S properties
- Changes in strait transports eg. Arctic overflows (Dixon et al)





#### **4D Var approach to state estimation**

Main drawback to ocean inverse work is steady state assumption, Ship section data often measured years apart In any case what is exact value of "Mean Circulation"

ECCO group (Stammer and Wunsch) are using Least Squares Cost function approach to model time-evolving circulation over 1992-2002 period, with low resolution model=> 4DVar method

$$J = \frac{1}{2} [(\overline{\zeta} - \overline{\zeta}_{obs})^T \mathbf{W}_{EGM96} (\overline{\zeta} - \overline{\zeta}_{obs}) \\ + (\zeta' - \zeta'_{TP})^T W_{TP} (\zeta' - \zeta'_{TP}) + (\zeta' - \zeta'_{ERS})^T W_{ERS} (\zeta' - \zeta'_{ERS}) \\ + (\tau_{\mathbf{u}} - \tau_{\mathbf{u}obs})^T W_{NSCAT_x} (\tau_{\mathbf{u}} - \tau_{\mathbf{u}obs}) + (\tau_{\mathbf{v}} - \tau_{\mathbf{v}obs})^T W_{NSCAT_y} (\tau_{\mathbf{v}} - \tau_{\mathbf{v}obs}) \\ + (\mathbf{H}_{\mathbf{Q}} - \mathbf{H}_{\mathbf{Q}obs})^T W_{HQ} (\mathbf{H}_{\mathbf{Q}} - \mathbf{H}_{\mathbf{Q}obs}) + (\mathbf{H}_{\mathbf{F}} - \mathbf{H}_{\mathbf{F}obs})^T W_{H_F} (\mathbf{H}_{\mathbf{F}} - \mathbf{H}_{\mathbf{F}obs}) \\ + \sum_{i=1}^{12} (\overline{\mathbf{T}}_{\mathbf{i}} - \overline{\mathbf{T}}_{\mathbf{i}Lev})^T W_T (\overline{\mathbf{T}}_{\mathbf{i}} - \overline{\mathbf{T}}_{\mathbf{i}Lev}) + \sum_{i=1}^{12} (\overline{\mathbf{S}}_{\mathbf{i}} - \overline{\mathbf{S}}_{\mathbf{i}Lev})^T W_S (\overline{\mathbf{S}}_{\mathbf{i}} - \overline{\mathbf{S}}_{\mathbf{i}Lev}) \\ + (\mathbf{T} - \mathbf{SST})^T W_{SST} (\mathbf{T} - \mathbf{SST}).$$

Is this the only way to assimilate ocean data for climate reanalysis studies??





# Historical availability of ocean data



### <u>Upper Ocean T(z) 1993</u>

#### 69733 points

#### Historical T profiles N Atlantic



FIG. 1. Numbers of temperature profiles available from the North Atlantic during 5-yr periods from 1950 to 1994. Dark shaded blocks show the total number, and light-shaded blocks show numbers reaching to 400 m or deeper.

#### Now XBTs reach 800m Being superceded by Argo





# **Assimilation for Seasonal Weather Forecasting**

Based on coupled ocean-atmosphere model with ocean data assimilation run in Ensemble mode => 6 months eg. ECMWF









# **Assimilation for Seasonal Weather Forecasting**

- Seasonal forecasting operational at ECMWF, NASA etc.
- TAO buoys provide temperatures and currents to 450m
- Assimilate into coupled ocean-atmosphere global models
- Forecast timescale ~6 months
- Forecasting El Nino onset; Nino 3 surf. Temp. anomalies
- Whole set of climate parameters also predicted, eg. rainfall, surface T anomalies

### ECMWF Forecasts for the 1997 El Nino







# **Operational Ocean Forecasting**

#### Ocean only models forced with winds and fluxes from Met forecasts Assimilating satellite and in situ ocean data (eg. GOOS and GODAE) Products from the EU MERSEA Project North\_Atlantic







#### Key Data for Operational Oceanography









## Key data sets for operational oceanography

#### High resolution SST analyses from Microwave and IR satellites



#### **NCOF OSTIA product (1km resolution)**

http://ghrsst-pp.metoffice.com/pages/latest\_analysis/ostia.html





# **Assimilation for Operational Oceanography**

- Sequential Assimilation schemes, Kalman Filter + variants and simplifications are universally used
- Although Operational NWP benefited from 4DVar eg. ECMWF, 4DVar is expensive in oceans, and models too BIASED?
- Key information is the Background Error Covariance
  - Needed to link Observation, SSH or T(z), S(z) profile with changes in whole model state vector
  - Error covariances poorly known from observations
  - Can use physically based covariance information relating
    - Altimetric SSH with T,S profiles in the water column
    - Temperature profiles T(z) with Salinity profiles S(z)
    - Horizontal correlations of S(T) compared with S(z)
- Consider some idealised assimilation experiments in order to understand ocean assimilation constraints (Altimeter data)





## **Quasi-Geostrophic Box model of the Subtropical** and Subpolar ocean gyres



**Vertical correlations of \psi, q completely different => useful ENVIRONMENT** RESEARCH COUNCIL ENVIRONMENTAL SYSTEMS SCIENCE CENTRE

 $Dq_{2.3}/Dt = 0$ 

Haines (1991) University of Reading

### **Twin experiment assimilation of** $\psi_1$ **every 40 days**

 $\Psi_1 - \Psi_4$ 

 $q_1 - q_4$ 





Note that  $q_2 - q_4$  still converge



#### Twin experiment in OCCAM 36 level model assimilating Sea surface height



#### Note that subsurface T,S still converge



ESA Summer School Frascati August 2006 The University of Reading

Fox et al 2001a

## **Relationship between Altimeter and In Situ Assimilation**







## **Relationship between Altimeter and In Situ Assimilation**







# $\frac{Altimeter \ assimilation \ by \ thermocline}{displacement \ \Delta h}$

Cooper and Haines (1996) extended idea to Primitive Equation models

$$\mathbf{Dq}(\boldsymbol{\rho})/\mathbf{Dt} = \mathbf{0} \qquad \mathbf{q} = \frac{\mathbf{f}}{\rho_0} \frac{\partial \rho}{\partial \mathbf{z}}.$$

Model **q(p)** is preserved by Assimilation provided;

$$\Delta \rho = \frac{\partial \rho}{\partial z} \Delta \mathbf{h}, = \text{Isopycnal displacement}$$
  
Solve for  $\Delta \mathbf{h}$  by assuming deep

pressure unchanged

$$\Delta \mathbf{p}(\mathbf{0}) = \mathbf{g} \int_{-\mathbf{H}}^{\mathbf{0}} \Delta \rho \mathbf{dz},$$



(Different closure to Haines 1991;  $\psi_1$  observed,  $\mathbf{q}_2, \mathbf{q}_3$  and  $\psi_4$  from model)





## **Assimilation of Satellite Altimeter**







## Nice features of altimeter assimilation with conserved water masses

- Simple to apply (don't need pre-calculated covariances)
- Can derive implied vertical covariances analytically so can incorporate into standard assimilation methods
- Vertical covariances are automatically time and flow dependent
- Conservation of water properties allows other assimilation data to determine water mass properties and volumes. Particularly important for Reanalysis and Climate work => gives method similar properties to 4DVar
- Has been used at;
  - UK Met Office, ECMWF, Mercator (in SEEK filter), HYCOM
- Other Assimilation methods sharing water conservation property
  - Oschlies and Willebrand 96: Velocity covariances
  - Gavart and De Mey 97: Potential density depth covariances
  - Greatbatch et al 2001:- Semi-prognostic method





## Assimilation of T profile data

- Historically T(z) profiles make up vast bulk of in situ ocean measurements
- MBTs down to 400m before 1970's
- XBTs down to 800m after 1970's
- Voluntary Observing Ship program
- Highly non-isotropic coverage
- Basis of Levitus estimates of climatic warming of oceans
- El Nino TAO array also mainly T(z) to 400m
- Being superceded by Argo profiling floats
  - T and S down to 2000m with near isotropic coverage







## Assimilation of T profile data at ECMWF (pre 2000)

- Assimilation of TAO T profiles from tropical Pacific has been main focus of all seasonal forecasting projects
- ECMWF use OI assimilation, Smith et al (1991)
- 10 days of data assimilated together
- T(z) profiles vertically interpolated to model levels
- Separate horizontal OI on each model level 1500km zonal, 200km meridional scales at Equator
- Observations and model T data given equal weight
- TAO T profiles only reach 450m
- Salinity not updated







## Mean meridional sections 30W from 1990-98 **ECMWF** assimilation experiments





sections

sections



#### **T-profile assimilation with T/S conservation**



ENVIRONMENTAL SYSTEMS SCIENCE CENTRE

Frascati August 2006

University of Reading

# ECMWF Forecasts of Nino3 temperatures

5 member ensembleforecasts started every3 months from 1993-1997= 100 forecasts



Segschneider et al 2001

C-OI = Original T assimilation C-T+A= + Altimeter data C-TA+TS = + T/S conservation scheme





### **<u>Complementarity between Altimeter</u>** <u>and T profile assimilation</u>

- Altimeter = Vertical thermocline displacement
- T-profile = S(T) preserved + displacement
- Both preserve S(T) which neither observe
- Both preserve <u>volume(ρ)</u> or volume(T) except
- T-profile => changes in volume(ρ) in upper water column where observations made





## Assimilation of S profile data

- Ocean salinity difficult to measure (Conductivity corrected for T)
- CTD measured from research ships: eg. Section data such as WOCE or specific local research programs
- Very little historical data
- Important climatic signals in Salinity variability
  - Great Salinity Anomaly; Dixon et al. (1996)
  - Changes in polar-equator salinity gradient => changes in hydrological cycle (evaporation-transport-precipitation) Curry et al (2003)
  - Controls density structure in polar oceans
- Much Salinity variability is highly correlated with T variability
  - How to take advantage of this during assimilation of S data?
  - Otherwise S(z) gives very little additional information over T(z)





# **Salinity variance in z and T space**

104 CTD profilesover 10 days inW. Equatorial Pacific

Reduced variance in S(T) suggests value of T/S preservation during assimilation

Model Representivity of S(T) probably better than for S(z) or T(z)





Troccoli and Haines (1999)



## **Ratio of S variance on depth surface and Isotherms**



Ratio (S(z) variance / S(T) variance) in 1x1° bins for 40 years of data. The 300m depth and the mean isotherm at that depth define salinities. Bins with ratio > 1 black; ratio < 1 dark grey.

Haines et al (2006)





### **Bermuda Salinity timeseries**



S(z) has more Variability at High Frequencies Dynamical Origin

Remaining variability in S(T) Lower Frequency Thermodynamic Origin

Lower Representivity error

Different Spatial Scales too



#### **One point S correlation maps HadCEM 1/3 model**



ENVIRONMENTAL 3131 EM3 SCIENCE CENTRE



Correlation of salinity on 12C temp surface

DATA SET: hadcem\_s(t)\_15\_60\_-22



Expect error Covariances of S(T) to be larger Scale than S(z) => Useful in assimilation of Salinity data, especially for Reanalysis



## **Assimilation of Salinity at ECMWF**

Two stage salinity assimilation process (Implemented by Arthur Vidard)

- 1) TH99: S(T) unaltered by T assimilation.
  - $\Delta S(T) = 0$ ;  $\Delta S_T(z) \neq 0$
- 2) Salinity assimilation:  $\Delta S(T) \neq 0$ ;  $\Delta S_{S}(z) \neq 0$

Idea: perform a second OI using T+S data to correct the T/S relationship (Haines et al 2006: Mon. Weath. Rev.)



$$S_{a}(\mathcal{F}_{a}) = S_{a}'(\mathcal{F}_{a}) + K'((S_{bo}(\mathcal{F}_{a})) - HS_{b}(\mathcal{F}_{a})))$$







Also, second salinity increment is independent of the 1st!





## Recovering THC strength from an ECMWF Ocean Reanalysis

# Thermohaline overturning circulation in the North Atlantic





ESA Summer School Frascati August 2006



Heat flux

# **Conservation properties in sequential assimilation**

#### **Altimeter Assimilation**

**Displacement**  $\Delta h \Rightarrow$  **Gross Isopycnal geometry** 

+ Currents (geostrophy)

- Volume and T/S properties preserved on isopycnals
- Adiabatic (Thermodynamically Reversible)

#### **T** Profile Assimilation

#### T(z) => Isothermal Water Volumes

- S(T) properties preserved (since salinity is not observed)
- Volumes and T/S preserved below deepest observation

#### S(T) Assimilation

- S(T) => Isopycnal Water Properties
- Large scale, slow variations associated with ventilation and climatic change











# **WOCE Atlantic** Section A16

Water property distributions give qualitative information on circulation pathways

Note: Water mass origins AIW, \_\_\_\_\_ NADW, \_\_\_\_ ABW \_\_\_\_\_



A16 Silicate umol/kg

Lagrangian conservation of water properties important in assimilation



300 1000 1300 3000 2300 3000 3300 4000 4300 3000 3300 60000 -35 -35 -15 45

The University of Reading

175

115

105

95

23

73

63

35

45

35

23 13

Frascati August 2006

## Improving T through S assimilation

- Temperature assimilation can improve salinity directly since S(T) conserved
- Salinity assimilation can also improve Temperature, but only indirectly through improved advection
- Obs Background errors Preliminary results from ENACT project reanalysis 1993-2001



NATURAL ENVIRONMENT RESEARCH COUNCIL ENVIRONMENTAL SYSTEMS SCIENCE CENTRE



Mean Observation Minus Background in selected regions for temperature and salinity

#### Twin experiment in OCCAM 36 level model assimilating Sea surface height



#### Note that subsurface T,S still converge



Fox et al 2001a



## **Bias and diagnostics of bias in ocean models**

- Much of optimal assimilation theory assumes that the models and observations are unbiased. This is definitely not the case for ocean models
- Detection of bias is easy: if the innovations do not average to zero then the model (or data) is biased
- In this case one of the main effects of data assimilation is to counteract the bias eg. model drift
- Methods used to correct for bias in ocean models
  - 'Pressure Correction' Bell et al (2000)
  - Semi-prognostic method Greatbatch....
- Having detected bias it should be accounted for in assimilation error analysis or else the weighting of new observations will be poorly handled
- Need to have a bias model





### **Accounting for Bias in Data Assimilation**

- Dee (2006) Review in QJRMS
- 3D Variational formulation easiest to understand (derivable from Bayesian analysis; Drecourt et al; 2006)

```
2J(x,b,c) = (y-b-x)^{T}R^{-1}(y-b-x) + 
(x-x^{f}+c)^{T}B^{-1}(x-x^{f}+c) + 
(b-b^{f})^{T}O^{-1}(b-b^{f}) + 
(c-c^{f})^{T}P^{-1}(c-c^{f})
```

Minimise J wrt x,b,c

y =observation

x =model state

**b** = **observation bias** 

c =model forecast bias

Superscript f are forecast values

R =observation error covariance B =model background error covariance

- O =observation bias error covariance
- P =model forecast bias error covariance

Observation operators have been omitted





#### **Accounting for Bias in Data Assimilation**

• Solution (Analysed variables <sup>a</sup>)  $\mathbf{x}^{a} = (\mathbf{x}^{f} - \mathbf{c}^{f}) + \mathbf{K} \{(\mathbf{y} - \mathbf{b}^{f}) - (\mathbf{x}^{f} - \mathbf{c}^{f})\}$   $\mathbf{b}^{a} = \mathbf{b}^{f} + \mathbf{F} \{(\mathbf{y} - \mathbf{b}^{f}) - (\mathbf{x}^{f} - \mathbf{c}^{f})\}$   $\mathbf{c}^{a} = \mathbf{c}^{f} + \mathbf{G} \{(\mathbf{y} - \mathbf{b}^{f}) - (\mathbf{x}^{f} - \mathbf{c}^{f})\}$ 

- $K = (B+P) [B+P+O+R]^{-1}$   $F = O [B+P+O+R]^{-1}$  $G = P [B+P+O+R]^{-1}$
- or  $x^a = (x^f c^a) + K_1 \{ (y b^a) (x^f c^a) \}$   $K_1 = B [B + R]^{-1}$
- y =observation x =model state b =observation bias c =model forecast bias

- R =observation error covariance
- B =model background error covariance
- O =observation bias error covariance
- P =model forecast bias error covariance

Usual problems are: (i) Knowing the Covariance errors (ii) Sequential 3DVar requires bias models for  $b^{f}(t+1) = M_{b}[b^{a}(t)];$   $c^{f}(t+1) = M_{c}[c^{a}(t)];$ 





## **Comments on Bias Modelling**

- Known Biases { $b^{f}(t)$ ;  $c^{f}(t)$  known a priori eg. previous runs}
  - $x^{a} = (x^{f}-c^{f}) + K \{(y-b^{f}) (x^{f}-c^{f})\} \qquad K = (B+P)[B+P+O+R]^{-1}$
  - $b^{f}(t) = 0$ ;  $c^{f}(t) = 0$  is particular case
  - (B+P) total model err cov.; (O+R) total obs. err.
- Persistent Biases { $b^{f}(t+1) = b^{a}(t)$ ;  $c^{f}(t+1) = c^{a}(t)$  }
  - $x^a = (x^f c^f) + K \{(y b^f) (x^f c^f)\}$   $K = (B+P)[B+P+O+R]^{-1}$
  - $b^{a} = b^{f} + F \{(y-b^{f}) (x^{f}-c^{f})\}$  F = O[B+P+O+R]<sup>-1</sup>
  - $c^a = c^f$  + G {(y-b^f) (x^f-c^f)} G = P[B+P+O+R]^{-1}
  - If O,P i.e. F,G are small => may hope to converge to ~ constant b,c
  - Simplifications also arise if  $P=\alpha B$ ;  $O=\beta R \Rightarrow$  all Innovations proportional
- Attribution of Bias: When are O,P sufficiently different to allow identification of misfits  $\{(y-b^f) (x^f-c^f)\}$ ?
- Should always check total misfits are consistent with B+P+O+R





#### Drift in ocean temperatures in tropical Pacific during 6 month free runs of ECMWF coupled model



Drift must be removed to interpret ENSO forecasts





#### Accounting for bias during data assimilation

 $x_{k+1} = M(x_k, u_k)$  deterministic model  $x_k$  variables,  $u_k$  parameters, at time k

For a general biased model  $x_{k+1} = M^t (x_k, u_k) + T(b_k)$  $M^t$  true model,  $b_k$  bias variables

Now define new State vector  $\{x_{k,}b_{k}\}$ with model for bias evolution  $b_{k+1} = W(b_{k}, x_{k}) + \zeta_{k},$  $\zeta_{k}$  white noise Mean assimilation T increments in Met Office assimilation in Equatorial Pacific, Bell et al 2002



In sequential assimilation  $\mathbf{x}_{\mathbf{k}}^{-5.5}$  will converge to  $\mathbf{x}_{\mathbf{k}}^{t^{2.5}}$ provided bias model W is correct



ESA Summer School Frascati August 2006



4.5

#### Bell et al 2002 assumed Equatorial T bias due to wrong wind stress $\tau(x,y,t)$



But they modelled the bias with pressure field  $p^{-5.5}$   $p^{-3.5}$   $p^{-1.5}$   $p^{0.5}$   $p^{0$ 

 $\partial \mathbf{y}$ 



∂t



## **Example of Bias Modelling in Seasonal Forecasting**



- Method reduces undesirable transients while allowing T to approach T<sup>true</sup>
- Could one recover the cause of the bias (probably wind stress error)?
- Similar method reduces misfits in Ocean Reanalysis; Chepurin et al (2005)





- No alteration of data assimilation (DA) procedure
- Aim is to diagnose mean misfits/innovations directly as biases in physical processes
- **1** Assimilation impacts on *Local* Heat budget (or other tracers)
- 2 Assimilation impacts on water volumes in each temperature class within an *Extended Region*, eg. N. Atlantic (c.f. ocean inverse theory)
- Consider a model 'held' close to observational trajectory by DA against a drift tendency
  - (a) How do we quantify role of DA in preventing drift?
  - (b) How do we identify drift with inadequacies of physics?





## **Conclusions**

- Lagrangian conservation of water properties provides useful constraints for both steady state ocean inverse problems and time evolving ocean data assimilation
- Generates state-dependent multivariate covariances in a natural way
- Provides a framework for obtaining climate quality ocean reanalyses using sequential data assimilation
- Useful in operational oceanography and seasonal forecasting when error covariances poorly defined empirically
- Allows improved assimilation of salinity (and potentially other tracer) data by reducing "Representivity" errors and increasing Kalman gain
- Model bias should also be accounted for correctly in order to correctly weight assimilated data.





### End of second Lecture



