Inverse Modeling

Alonso Hernández Guerra

June 30, 2026

Non-dimensional Numbers

For typical interior ocean scales:

$$L \sim 10^6 \ \mathrm{m}, \qquad H \sim 10^3 \ \mathrm{m}$$

$$U \sim 10^{-1} \ \mathrm{m\,s^{-1}}, \qquad T \sim 10^7 \ \mathrm{s}$$

$$\Omega \sim 10^{-4} \ \mathrm{s^{-1}}, \qquad \nu \sim 10^{-2} \ \mathrm{m^2\,s^{-1}}$$

$$R_{o_T} = \frac{1}{\Omega T} \sim 10^{-3}$$

$$R_o = \frac{U}{\Omega L} \sim 10^{-3}$$

$$E_k = \frac{\nu}{\Omega H^2} \sim 10^{-7}$$

$$R_{o_T}, R_o, E_k \ll 1$$

Geostrophic Approximation

Since

$$R_{o_T} \ll 1, \qquad R_o \ll 1, \qquad E_k \ll 1,$$

the momentum equations become

$$\cancel{\frac{\partial u}{\partial t}} + \cancel{u\frac{\partial u}{\partial x}} + \cancel{v\frac{\partial u}{\partial y}} + \cancel{w\frac{\partial u}{\partial z}} - fv = -\frac{1}{\rho_0}\frac{\partial p}{\partial x} + \cancel{\nu\frac{\partial^2 u}{\partial z^2}}$$

$$\cancel{\frac{\partial v}{\partial t}} + \cancel{u\frac{\partial v}{\partial x}} + \cancel{v\frac{\partial v}{\partial y}} + \cancel{w\frac{\partial v}{\partial z}} + fu = -\frac{1}{\rho_0}\frac{\partial p}{\partial y} + \cancel{\nu\frac{\partial^2 v}{\partial z^2}}$$

$$0 = -\frac{\partial p}{\partial z} - \rho g$$

Geostrophic and Hydrostatic Balance

Horizontal momentum balance:

$$fv = \frac{1}{\rho_0}\frac{\partial p}{\partial x}$$

$$fu = -\frac{1}{\rho_0}\frac{\partial p}{\partial y}$$

Hydrostatic balance:

$$\frac{\partial p}{\partial z} = -\rho g$$

Key Idea

In the ocean interior, the dominant balance is between the Coriolis force and the horizontal pressure gradient.

Vertical Derivative of the Geostrophic Velocity

Starting from the geostrophic balance

$$fv = \frac{1}{\rho_0}\frac{\partial p}{\partial x}, \qquad fu = -\frac{1}{\rho_0}\frac{\partial p}{\partial y},$$

taking the derivative with respect to z gives

$$\frac{\partial v}{\partial z} = \frac{1}{\rho_0 f}\frac{\partial}{\partial z}\left(\frac{\partial p}{\partial x}\right),$$

$$\frac{\partial u}{\partial z} = -\frac{1}{\rho_0 f}\frac{\partial}{\partial z}\left(\frac{\partial p}{\partial y}\right).$$

Since x and y are independent of z, the order of differentiation can be exchanged:

$$\frac{\partial v}{\partial z} = \frac{1}{\rho_0 f}\frac{\partial}{\partial x}\left(\frac{\partial p}{\partial z}\right),$$

$$\frac{\partial u}{\partial z} = -\frac{1}{\rho_0 f}\frac{\partial}{\partial y}\left(\frac{\partial p}{\partial z}\right).$$

Thermal Wind Equations

Using the hydrostatic balance

$$\frac{\partial p}{\partial z} = -\rho g,$$

we obtain

$$\frac{\partial v}{\partial z} = \frac{1}{\rho_0 f}\frac{\partial(-\rho g)}{\partial x},$$

$$\frac{\partial u}{\partial z} = -\frac{1}{\rho_0 f}\frac{\partial(-\rho g)}{\partial y}.$$

Therefore,

$$\boxed{\frac{\partial v}{\partial z} = -\frac{g}{\rho_0 f}\frac{\partial \rho}{\partial x}}$$

$$\boxed{\frac{\partial u}{\partial z} = \frac{g}{\rho_0 f}\frac{\partial \rho}{\partial y}}$$

Integrating the thermal wind equations from a reference level z_0 to z:

$$\int_c^u du = \frac{g}{\rho_0 f} \int_{z_0}^{z} \frac{\partial \rho}{\partial y}\, dz$$

$$\int_b^v dv = -\frac{g}{\rho_0 f} \int_{z_0}^{z} \frac{\partial \rho}{\partial x}\, dz$$

which yields

$$u = c + \frac{g}{\rho_0 f} \int_{z_0}^{z} \frac{\partial \rho}{\partial y}\, dz$$

$$v = b - \frac{g}{\rho_0 f} \int_{z_0}^{z} \frac{\partial \rho}{\partial x}\, dz$$

Relative and Absolute Geostrophic Velocity

Defining the **relative geostrophic velocity**

$$u_r = \frac{g}{\rho_0 f} \int_{z_0}^{z} \frac{\partial \rho}{\partial y} \, dz$$

$$v_r = -\frac{g}{\rho_0 f} \int_{z_0}^{z} \frac{\partial \rho}{\partial x} \, dz$$

we obtain

$$u = c + u_r$$

$$v = b + v_r$$

The Need for an Inverse Model

Hydrographic observations provide:

$$u_r, \qquad v_r$$

through the thermal wind equations.

However, the integration constants

$$b, \qquad c$$

remain unknown.

Therefore, hydrography alone determines only the **relative geostrophic velocity**.

Classical Solutions to the Level-of-No-Motion Problem

To resolve the unknown reference velocity, classical dynamical oceanography proposed several empirical criteria for the **level of no motion**:

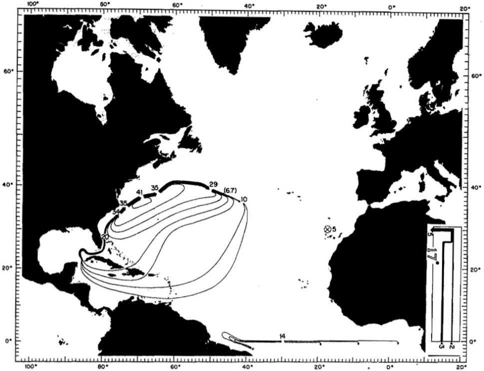
1. Assume a **negligible velocity at great depth**.
2. **Wüst (1935)**: if two water masses flow in opposite directions, a layer of no motion exists between them.
3. **Defant (1941)**: the level of no motion is chosen at the depth where the vertical shear of the geostrophic velocity is minimum.

Limitations of the Classical Approach

- None of these procedures is derived from physical principles.
- The circulation obtained does not satisfy the mass conservation equation.

Worthington's Drastic Remedy

Worthington (1962, 1976) proposed the most drastic remedy in his circulation scheme: it satisfies the mass conservation equation, but not the geostrophic equation.



Worthington's North Atlantic circulation scheme (volume transport, Sv).

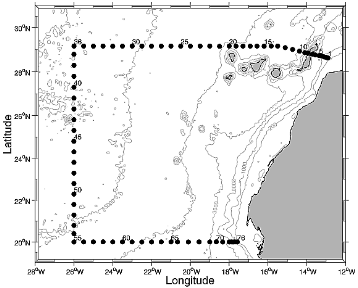
A Crisis – and a New Paradigm

Crisis in the use of the dynamic method.

New paradigm: the inverse models.

As Thomas Kuhn argued in *The Structure of Scientific Revolutions* (1962), persistent anomalies that a paradigm cannot resolve precipitate a **crisis**, which is ultimately resolved not by patching the old framework, but by a **paradigm shift**.

Why Inverse Models?



Hernández-Guerra et al. (2005)

Hydrographic section around the Canary Islands

- CTD stations define a closed box.
- Temperature and salinity measurements provide density fields.
- Geostrophic velocities can be estimated from thermal wind balance.
- However, the reference velocity at each station pair remains unknown.

Question:

Can we determine the absolute circulation that is consistent with the observations and physical conservation laws?

The Inverse Problem

Known:

- Density field $\rho(x, y, z)$
- Relative geostrophic velocities

Unknown:

- Reference velocities

Inverse models determine the unknown velocity field by imposing:

- Conservation of mass
- Conservation of salt
- Conservation of heat in deep layers
- Additional dynamical constraints

Mass Conservation in a Closed Box

Consider a hydrographic box bounded by a set of stations (e.g., Hernández-Guerra et al., 2005).

For a steady-state ocean, the amount of water contained within the box must remain constant.

Therefore, the net mass transport across the closed surface must vanish:

$$\oint_S \rho\, \vec{v} \cdot d\vec{S} = 0$$

where

- ρ is the seawater density,
- $\vec{v}$ is the velocity vector,
- $d\vec{S}$ is the outward surface element.

This equation states that the total inflow must balance the total outflow through the boundaries of the box.

From Mass Conservation to the Inverse Problem

Assuming that:

- vertical velocities are negligible,
- only the velocity component normal to each section is considered,

the conservation equation becomes

$$\oint_S \rho\, v\, dS = 0.$$

From the thermal wind equations, the velocity can be written as

$$v = v_r + b,$$

where

- v_r is the relative geostrophic velocity,
- b is the unknown reference velocity.

Substituting into the conservation equation,

$$\oint_S \rho\,(v_r + b)\, dS = 0.$$

From Mass Conservation to the Inverse Problem

$$\oint_S \rho \left(v_r + b \right) dS = 0.$$

The relative velocity v_r is known from hydrographic data, whereas the reference velocity b remains unknown.

Determining b is the central objective of the inverse model.

Discretization into Water-Mass Layers

Hydrographic stations are discrete and the water column is divided into Q neutral density layers, corresponding to different water masses.

The mass conservation equation can be approximated by

$$\sum_{j=1}^{N} \sum_{q=1}^{Q} \rho_{jq}(v_{r,jq} + b_j)\Delta a_{jq} \simeq 0$$

where

- j denotes the station pair,
- q denotes the neutral density layer,
- ρ_{jq} is the density,
- Δa_{jq} is the cross-sectional area.

Discretization into Water-Mass Layers

Rearranging,

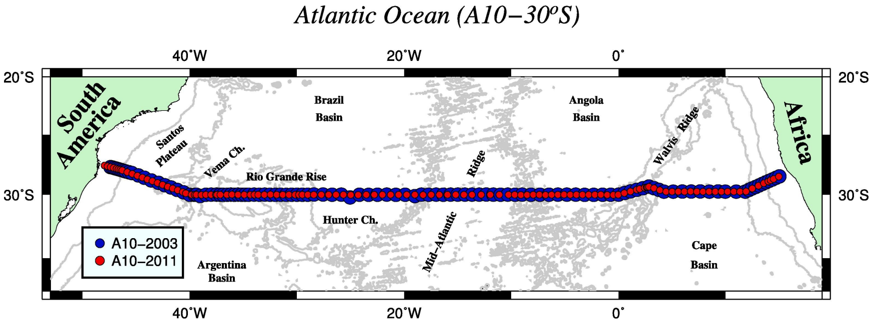
$$\sum_{j=1}^{N}\sum_{q=1}^{Q}\rho_{jq} b_j \Delta a_{jq} \simeq -\sum_{j=1}^{N}\sum_{q=1}^{Q}\rho_{jq} v_{r,jq} \Delta a_{jq}.$$

From Equations to Code: Mass, Heat and Salt Alike

The discretized sum derived above is exactly what the model script computes, layer by layer. The **same lines of code** work for mass, heat, salt – or any other property – only the property being summed changes.

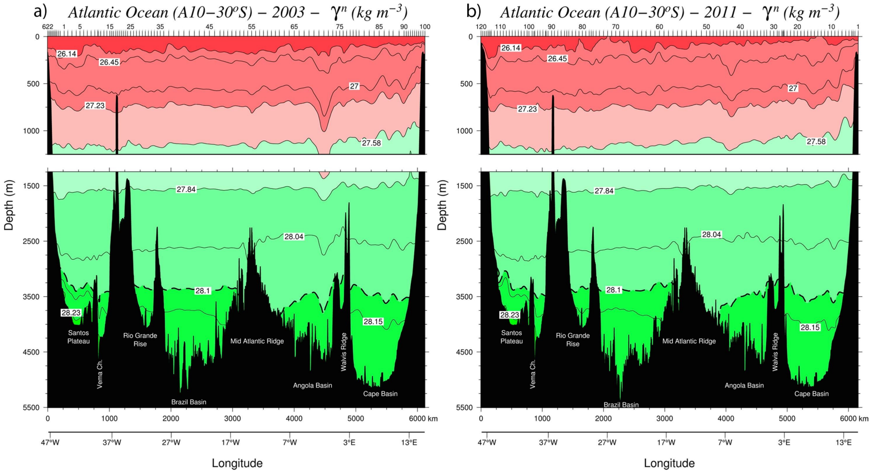
```
for c = 1:length(cuts)-1
    fs = ll >= cuts(c) & ll < cuts(c+1);
    % now do sums of quatities for each layer
    area(c,s) = sum(starea.*fs);
    mass(c,s) = sum(starea.*fs.*crho);
    tri_area(c,s) = sum(starea.*tri_idx.*fs);
    vol_trans(c,s)=sum(starea.*stvel.*fs);
    tri_vol_trans(c,s)=sum(starea.*stvel.*tri_idx.*fs);
    mass_trans(c,s)=sum(starea.*stvel.*crho.*fs);
    tri_mass_trans(c,s)=sum(starea.*stvel.*crho.*tri_idx.*fs);

    %now calculate transport properties
    trans1(c,s) = sum(heat.*stvel.*fs);
    trans2(c,s) = sum(csalt.*stvel.*fs);
    trans3(c,s) = sum(cox.*stvel.*fs);
    prop1(c,s) = sum(heat.*fs);
    prop2(c,s) = sum(csalt.*fs);
    prop3(c,s) = sum(cox.*fs);
```

Excerpt from the inverse model script: for each layer `c`, `fs` selects the data within that density range, and `mass_trans` computes $\sum \rho_{jq}\, v_{jq}\, \Delta a_{jq}$ – the same sum just derived. `trans1--3` compute the identical sum for heat, salt and a third tracer (e.g. oxygen), simply by replacing the density `crho` with `heat`, `csalt` or `cox`.

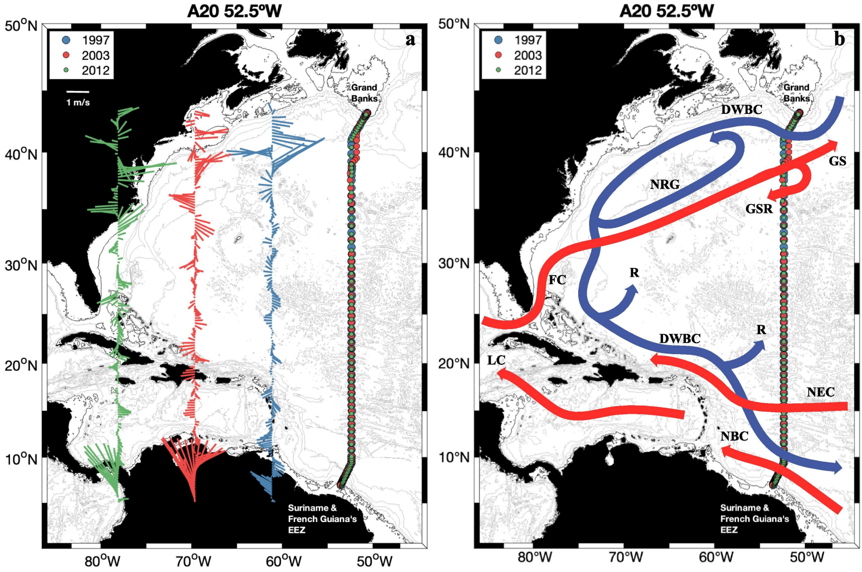


Hernández-Guerra et al. (2019)

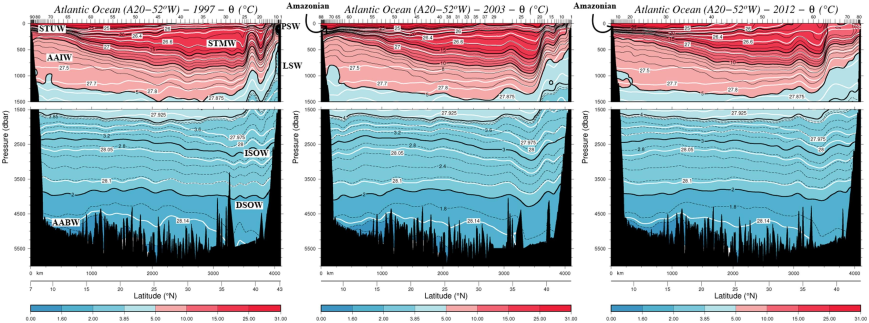
The ocean may divided into Q water-mass layers, in this case, in neutral density layers.



Hernández-Guerra et al. (2019)



Santana-Toscano et al. (2023)



Santana-Toscano et al. (2023)

If mass conservation is imposed separately in each water-mass layer,

$$\sum_{j=1}^{N} \rho_{jq} b_j \Delta a_{jq} \simeq -\sum_{j=1}^{N} \rho_{jq} v_{r,jq} \Delta a_{jq}, \qquad q = 1, \ldots, Q.$$

Defining

$$e_{jq} = \rho_{jq} \Delta a_{jq},$$

and

$$y_q = -\sum_{j=1}^{N} \rho_{jq} v_{r,jq} \Delta a_{jq},$$

Inverse Model Formulation

the system becomes

$$\sum_{j=1}^{N} e_{jq} b_j + n_q = y_q, \qquad q = 1, \ldots, Q,$$

where

- y_q is the baroclinic transport contribution for layer q,
- n_q represents observational and model errors.

Matrix Formulation of the Inverse Problem

The conservation equations for all water-mass layers can be combined into a single linear system:

$$E\,x + n = y$$

where

- x is an $N \times 1$ vector containing the unknown reference velocities,
- E is a $Q \times N$ coefficient matrix,
- n is a $Q \times 1$ error vector,
- y is a $Q \times 1$ vector containing the baroclinic transport contributions.

An additional equation is introduced to impose conservation of the total mass within the box.
For each station pair,

$$e_{tj} = \sum_{q=1}^{Q} e_{jq}, \qquad j = 1, \ldots, N,$$

and the corresponding transport is

$$y_t = \sum_{q=1}^{Q} y_q.$$

Structure of the Linear System

Including the total mass conservation equation, the system becomes

$$\begin{pmatrix} e_{11} & e_{12} & \cdots & e_{1N} \\ e_{21} & e_{22} & \cdots & e_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ e_{Q1} & e_{Q2} & \cdots & e_{QN} \\ e_{t1} & e_{t2} & \cdots & e_{tN} \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_N \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_Q \\ y_t \end{pmatrix}.$$

The last row corresponds to the total mass conservation constraint.

In general,

$$Q + 1 < N,$$

so the number of equations is smaller than the number of unknowns.

Therefore, the system is underdetermined and must be solved using inverse methods.

Ekman Transport Across Hydrographic Sections

The vertically integrated Ekman transport is

$$\mathbf{T}_{Ek} = \frac{1}{\rho_0 f} \mathbf{k} \times \boldsymbol{\tau},$$

where $\boldsymbol{\tau} = (\tau_x, \tau_y)$ is the wind stress.

Therefore,

$$T_{Ek,x} = \frac{\tau_y}{\rho_0 f},$$

$$T_{Ek,y} = -\frac{\tau_x}{\rho_0 f}.$$

For a hydrographic section:

- **Zonal section** (east-west): use the meridional component

$$T_{Ek,y} = -\frac{\tau_x}{\rho_0 f}.$$

- **Meridional section** (north-south): use the zonal component

$$T_{Ek,x} = \frac{\tau_y}{\rho_0 f}.$$

Including Ekman Transport

However, uncertainties in the wind stress may lead to errors in the estimated transport.

Therefore, the inverse model can treat the Ekman transport as an additional unknown and adjust it together with the reference velocities.

$$x = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_N \\ \Delta T_{Ek} \end{pmatrix}$$

where ΔT_{Ek} is an adjustment to the prescribed Ekman transport.

Inverse Model with Ekman Adjustment

The unknown vector is expanded to include the Ekman correction:

$$x = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_N \\ \Delta T_{Ek} \end{pmatrix}.$$

The conservation equations become

$$\begin{pmatrix} e_{11} & \cdots & e_{1N} & 1 \\ e_{21} & \cdots & e_{2N} & 0 \\ \vdots & & \vdots & \vdots \\ e_{Q1} & \cdots & e_{QN} & 0 \\ e_{t1} & \cdots & e_{tN} & 1 \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_N \\ \Delta T_{Ek} \end{pmatrix} = \begin{pmatrix} y_1 + T_{Ek} \\ y_2 \\ \vdots \\ y_Q \\ y_t + T_{Ek} \end{pmatrix}.$$

The Ekman transport contributes only to the surface layer and to the total mass conservation equation.

Ekman Transport When an Isopycnal Outcrops

The coefficient of 1 assigned to layer 1 assumes that layer 1 is the only layer that reaches the surface across the whole section.

If the isopycnal separating layers 1 and 2 **outcrops** within the array, layer 2 also reaches the surface over part of the section, and the Ekman transport must be shared between the two layers:

$$\begin{pmatrix} e_{11} & \cdots & e_{1N} & 0.5 \\ e_{21} & \cdots & e_{2N} & 0.5 \\ e_{31} & \cdots & e_{3N} & 0 \\ \vdots & & \vdots & \vdots \\ e_{Q1} & \cdots & e_{QN} & 0 \\ e_{t1} & \cdots & e_{tN} & 1 \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_N \\ \Delta T_{Ek} \end{pmatrix} = \begin{pmatrix} y_1 + 0.5\, T_{Ek} \\ y_2 + 0.5\, T_{Ek} \\ y_3 \\ \vdots \\ y_Q \\ y_t + T_{Ek} \end{pmatrix}.$$

The 0.5/0.5 split assumes the outcropping divides the section evenly between the two layers. In general this need not be the case: the fraction assigned to each layer should reflect the proportion of the section width (or area) over which each layer actually outcrops, and may differ from 0.5.

Generalization of the Inverse Model

The system

$$E\,x + n = y$$

is underdetermined. Therefore, additional conservation constraints can be introduced to increase the number of equations.

Besides mass conservation, inverse models may impose the conservation of:

- Salt
- Heat in deep layers
- Silicate (important in the Indian Ocean)
- Other tracers or chemical properties

Generalization of the Inverse Model

These equations have the same form as the mass conservation equations, but weighted by the concentration of the considered property.

For a generic property C:

$$\sum_{j=1}^{N} C_{jq}\,\rho_{jq}\,b_j\,\Delta a_{jq} \approx -\sum_{j=1}^{N} C_{jq}\,\rho_{jq}\,v_{r,jq}\,\Delta a_{jq}, \qquad q = 1, \ldots, Q.$$

Each additional conservation equation provides new constraints on the unknown reference velocities.

Adding Salt and Heat Conservation

Additional conservation equations are incorporated.

The inverse system becomes

$$\begin{pmatrix} E_M \\ E_S \\ E_H \\ E_T \end{pmatrix} x = \begin{pmatrix} y_M \\ y_S \\ y_H \\ y_T \end{pmatrix},$$

where

- E_M contains the mass conservation equations,
- E_S contains the salt conservation equations,
- E_H contains the heat conservation equations,
- E_T contains the conservation equations for tracers such as silicate.

The number of constraints increases while the number of unknowns remains unchanged.

Additional Constraints

Although additional constraints are included, the unknown reference velocities

$$x = (b_1, b_2, \ldots, b_N)^T$$

remain numerous, since one reference velocity must be determined for each station pair. Consequently, the system does not generally admit a unique solution.

Regional Transport Constraints

Direct observations can be used to constrain transport through a subset of station pairs and water-mass layers.

For a region defined by station pairs J and layers Q,

$$T = \sum_{j \in J} \sum_{q \in Q} e_{jq} b_j + T_r.$$

If an independent estimate

$$T_{obs} \pm \sigma_T$$

is available, the corresponding inverse constraint is

$$\sum_{j \in J} \sum_{q \in Q} e_{jq} b_j + n_T = T_{obs} - T_r.$$

Examples include mooring arrays, current-meter measurements, or transport estimates in narrow passages.

Including Regional Constraints

Additional observations are incorporated as extra rows in the inverse system:

$$\begin{bmatrix} E_M \\ E_S \\ E_H \\ E_T \\ E_R \end{bmatrix} x = \begin{bmatrix} y_M \\ y_S \\ y_H \\ y_T \\ y_R \end{bmatrix}$$

where

- E_M: mass conservation equations,
- E_S: salt conservation equations,
- E_H: heat conservation equations,
- E_T: tracer conservation equations,
- E_R: regional observational constraints.

Examples of regional constraints:

- Florida Current transport,
- Hunter Channel transport,
- Mooring-derived transports.

Structure of the Inverse Model

$$\underbrace{\begin{bmatrix} E_M \\ E_S \\ E_H \\ E_T \end{bmatrix}}_{\text{Conservation Laws}}$$

$$+$$

$$\underbrace{\begin{bmatrix} E_R \end{bmatrix}}_{\text{Regional Constraints}}$$

$$\Longrightarrow \qquad E\,x = y$$

The inverse model combines large-scale conservation laws with local observations to estimate the reference velocities.

Example of a Regional Constraint

A transport measurement of a regional constraint only affects a subset of station pairs:

$$\begin{pmatrix} 0 & 0 & 0 & e_4 & e_5 & e_6 & 0 & 0 \end{pmatrix} x = T_{obs} - T_r.$$

For example, a mooring array in a narrow passage constrains only the station pairs crossing that region.

Initial Estimate of the Reference Velocity

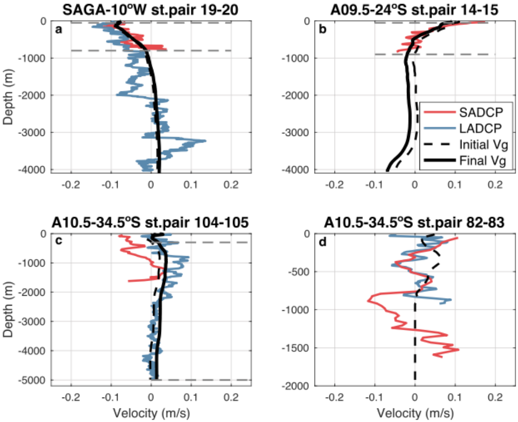
Whenever direct velocity observations are available, they can be used to obtain an initial estimate of the reference velocity b_j.

Examples include:

- LADCP (Lowered Acoustic Doppler Current Profiler),
- SADCP (Shipboard Acoustic Doppler Current Profiler),

Initial Estimate of the Reference Velocity

Whenever direct velocity observations are available, they can be used to obtain an initial estimate of the reference velocity.



Example of initial and adjusted geostrophic velocities. From Arumí-Planas et al. (2023)

Adjustment of the Velocity Field

Using these estimates, absolute velocities are obtained:

$$v = v_r + b^*_{ADCP}.$$

The corresponding transport of any conserved property ϕ can be written as

$$T_\phi = \sum_{j=1}^{N} \sum_{q=1}^{Q} \phi_{jq}\, \rho_{jq}\, v_{jq}\, \Delta a_{jq},$$

where

$$\phi = 1,\ S,\ c_p\theta,\ C$$

for mass, salt, heat and tracers, respectively, where c_p is the specific heat capacity of seawater at constant pressure.

These transports provide the first guess for the inverse model.

Beyond Conservation in Individual Layers

So far, the inverse model assumes that each water-mass layer is conserved independently.

However, real oceanic layers exchange properties through diapycnal processes.

Examples include:

- Vertical advection,
- Turbulent mixing.

Therefore, conservation equations must account for fluxes across layer interfaces.

Conservation of a Generic Property

Consider a conserved property C (salinity, silicate, etc.).

Its conservation equation can be written as

$$\nabla \cdot (\rho C \mathbf{v}) = \nabla \cdot (K \nabla (\rho C)),$$

where

- $\mathbf{v}$ is the velocity field,
- K is the turbulent diffusivity tensor.

The left-hand side represents advection, while the right-hand side represents turbulent diffusion.

Diapycnal Fluxes Between Layers

After integrating over the hydrographic box, the conservation equation becomes

$$\int_S \rho C(v_r + b)\, dS \approx \int_{S_z} \left(K_z \frac{\partial(\rho C)}{\partial z} - \rho C w \right) dS_z.$$

The right-hand side represents fluxes through the interfaces between adjacent layers:

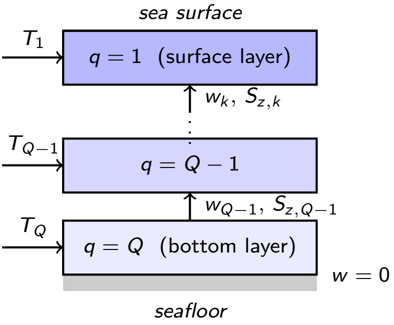
$$\underbrace{\rho C w}_{\text{Vertical Advection}}$$

$$\underbrace{K_z \frac{\partial(\rho C)}{\partial z}}_{\text{Vertical Diffusion}}$$

These terms describe exchanges between water masses that are not captured by the reference velocities alone.

Diagnosing w: A Bottom-Up Recursion

Once b_j is known, w can be obtained **diagnostically**, layer by layer, starting at the **bottom** of the box.



$T_q \equiv \int_S \rho\,(v_r + b)\,dS\Big|_q$: net horizontal mass transport into layer q.

Boundary condition: seafloor impermeable $\Rightarrow w = 0$ below layer Q. **Recursion**, moving one layer up at a time:

$$\rho_k\, w_k\, S_{z,k} \approx \rho_{k+1}\, w_{k+1}\, S_{z,k+1} + T_{k+1}, \qquad k = Q-1, \ldots, 1.$$

Diagnosing K_z Directly

With w now known at every interface, the heat balance can be solved for K_z, again working layer by layer from the bottom upward. In discretized form, for layer q and its upper interface k:

$$\sum_{j=1}^{N} \rho_{jq}\, c_p\theta_{jq}\,(v_{r,jq} + b_j)\,\Delta a_{jq} \approx K_{z,k}\, \left.\frac{\partial(\rho c_p\theta)}{\partial z}\right|_k S_{z,k} - \rho_k\, c_p\theta_k\, w_k\, S_{z,k}.$$

Since w_k, $S_{z,k}$ and the property gradient are already known, this is a single equation in the single unknown $K_{z,k}$:

$$K_{z,k} \approx \frac{\displaystyle\sum_{j=1}^{N} \rho_{jq} c_p\theta_{jq}(v_{r,jq} + b_j)\Delta a_{jq} + \rho_k c_p\theta_k w_k S_{z,k}}{\left.\dfrac{\partial(\rho c_p\theta)}{\partial z}\right|_k S_{z,k}}.$$

The Layer Areas S_z: A WOA-Based Estimate

The recursion above also requires S_z, the horizontal area of each neutral density interface **inside** the box – a quantity the hydrographic section itself does not provide, since it only samples a 1-D line of stations. A practical estimate can be built from a gridded climatology such as the **World Ocean Atlas (WOA)**:

1. Take WOA temperature and salinity profiles at every grid point inside the box,
2. compute the neutral density profile $\gamma_n(z)$ at each grid point,
3. check whether that profile contains a neutral density *greater* than the neutral density of that layer,
4. sum the area of the grid cells where it is present.

Advise: the WOA horizontal resolution (e.g. 1° or 0.25°) sets the precision of $S_{z,k}$ – coarser grids smooth out the true area where the interface is steep, such as near fronts or topography.

Two Ways to Obtain w and K_z

This sequential, diagnostic calculation is simpler and is the approach used in practice: w from mass conservation, then K_z as the residual of the heat balance.

Alternatively, the box inverse model itself can be generalized by introducing w and K_z as additional unknowns, solved **jointly** with the reference velocities b_j.

This second option is more complex, but yields a fully consistent solution – and its uncertainty – for circulation and diapycnal exchange together.

Generalized Inverse Model

The inverse model can estimate not only the reference velocities, but also diapycnal exchanges.

The unknown vector becomes

$$x = \begin{pmatrix} b_1 \\ \vdots \\ b_N \\ w_1 \\ \vdots \\ w_{M-1} \\ K_{z,1} \\ \vdots \\ K_{z,M-1} \\ \Delta T_{Ek} \end{pmatrix}.$$

Generalized Inverse Model

The coefficient matrix can be written schematically as

$$E = \left[\begin{array}{cccc} E_b & E_w & E_K & E_{Ek} \end{array}\right].$$

where, for property C in layer q and interface k at depth z_k, with $\rho_k \equiv \rho(z_k)$ and $C_k \equiv C(z_k)$ the density and property value evaluated at that interface,

- E_b: $\quad e^b_{qj} = \rho_{jq}\, C_{jq}\, \Delta a_{jq}$ (horizontal transport coefficients),
- E_w: $\quad e^w_{qk} = -\,\rho_k C_k\, \Delta S_{z,k}$ (vertical advection coefficients),
- E_K: $\quad e^K_{qk} = \left.\dfrac{\partial(\rho C)}{\partial z}\right|_k \Delta S_{z,k}$ (vertical diffusion coefficients),
- E_{Ek}: $\quad e^{Ek}_q = \begin{cases} 1, & q \text{ is the surface layer or total mass row} \\ 0, & \text{otherwise} \end{cases}$
  (Ekman transport corrections),

Block Structure of the Matrix

$$
\left[\begin{array}{ccccc}
 & \color{red}{E_b} & \color{red}{E_w} & \color{red}{E_K} & \color{red}{E_{Ek}} \\
\text{Mass} & \times & \times & 0 & \times \\
\text{Salt} & \times & \times & \times & \times \\
\text{Heat} & \times & \times & \times & 0 \\
\text{Silicate} & \times & \times & \times & \times \\
\text{Regional} & \times & 0 & 0 & 0
\end{array}\right]
\begin{pmatrix} b_j \\ w_k \\ K_{z,k} \\ \Delta T_{Ek} \end{pmatrix} = y
$$

- Horizontal transports constrain b_j.
- Property conservation constrains w_k and $K_{z,k}$.
- Surface constraints affect Ekman term.
- In the Salt and Silicate rows, the E_{Ek} entry is the property's concentration in the surface layer multiplying ΔT_{Ek}, not a unit coefficient as in the Mass row.
- Heat conservation is not imposed in the surface layer, due to air–sea exchange, so its E_{Ek} entry is zero.

Why a Gauss-Markov Estimator?

Despite the increased complexity, the inverse problem retains the same form:

$$E\,x + n = y,$$

where

- x contains the unknown reference velocities and diapycnal exchange terms,
- y contains the observations and constraints,
- n represents observational and model errors.

The system generally does not admit a unique solution, so we seek the solution that is most consistent with both the observations and our prior information.

The Gauss–Markov estimator can therefore be applied without modification to obtain the optimal solution and its uncertainty.

Prior Statistical Information

The Gauss–Markov estimator requires prior information about the first- and second-order moments of these variables.

For the unknowns, we assume that the mean and covariance matrix are:

$$\langle x \rangle = x_0,$$

$$P = \left\langle (x - x_0)(x - x_0)^T \right\rangle,$$

where x_0 is the prior estimate of the solution.

For the errors, we assume the mean and the covariance matrix as:

$$\langle n \rangle = 0,$$

$$R_{nn} = \langle nn^T \rangle.$$

The matrices R_{xx} and R_{nn} quantify the uncertainty associated with the prior estimate and the observations, respectively.

Gauss-Markov Solution

Assuming a linear estimator,

$$\tilde{x} = By,$$

the estimation error covariance is minimized.

The resulting Gauss-Markov solution is

$$\boxed{\tilde{x} = R_{xx}E^T\left(ER_{xx}E^T + R_{nn}\right)^{-1} y}$$

$R_{xx} = \langle xx^T \rangle$

This solution provides the set of reference velocities that best satisfies all conservation equations while accounting for observational uncertainties. The solution depends on two covariance matrices: one describing our confidence in the prior estimate and another describing the uncertainty of the observations.

Uncertainty of the Solution

The covariance of the estimated solution is

$$\boxed{P = R_{xx} - R_{xx}E^T\left(ER_{xx}E^T + R_{nn}\right)^{-1}ER_{xx}}$$

The residuals are

$$\tilde{n} = y - E\tilde{x},$$

with covariance

$$P_{nn} = EPE^T.$$

Therefore, the Gauss-Markov estimator provides:

- the optimal solution $\tilde{x}$,
- the uncertainty of the solution (P),
- the unexplained residuals ($\tilde{n}$).

Key Idea

The Gauss-Markov estimator combines

$$\text{Observations} + \text{Conservation Laws} + \text{Prior Information}$$

to obtain the most probable ocean circulation and its associated uncertainty.

The Gauss-Markov Estimator - Extended Formulation

The system of equations (2.22) is **underdetermined**, so we use the Gauss-Markov estimator to determine the solution $\tilde{x}$.

The estimator $\tilde{x}$ must **minimise the variance** of its true value, i.e. minimise:

Error covariance matrix (eq. 1)

$$P = \langle (\tilde{x} - x)(\tilde{x} - x)^T \rangle$$

We assume that the relationship between the data and the estimates is **linear**:

Linear estimator (eq. 2)

$$\tilde{x} = By$$

Derivation of the Covariance Matrix

Substituting $\tilde{x} = By$ into the expression for P:

$$\begin{aligned} P &= \langle (By - x)(By - x)^T \rangle \\ &= B\langle yy^T \rangle B^T - b\langle yx^T \rangle - \langle xy^T \rangle B^T + \langle xx^T \rangle \\ &= BR_{yy}B^T - BR_{xy}^T - R_{xy}B^T + R_{xx} \end{aligned}$$

where we have used $R_{xy} = R_{yx}^T$.

Applying the algebraic identity:

$$ACA^T - AB^T - BA^T = (A - BC^{-1})C(A - BC^{-1})^T - BC^{-1}B^T$$

we obtain:

$$P = (B - R_{xy}R_{yy}^{-1})R_{yy}(B - R_{xy}R_{yy}^{-1})^T - R_{xy}R_{yy}^{-1}R_{xy}^T + R_{xx}$$

Optimal Solution: The Gauss-Markov Estimator

Since R_{xx} and R_{yy} are **positive definite**, the diagonal elements of P are minimised by choosing B such that the first term vanishes:

Optimal matrix B (eq. 3)

$$B = R_{xy} R_{yy}^{-1}$$

The **estimated solution** is then (eq. 4):

Gauss-Markov estimator (eq. 4)

$$\tilde{x} = R_{xy} R_{yy}^{-1} y$$

and the **minimum variance** achieved is (eq. 5):

Minimum error covariance (eq. 5)

$$P = R_{xx} - R_{xy} R_{yy}^{-1} R_{xy}^{T}$$

Why is $B = R_{xy} R_{yy}^{-1}$ the Optimal Choice?

Recall the expanded expression for P (eq. 2.42):

$$P = \underbrace{(B - R_{xy} R_{yy}^{-1})\, R_{yy}\, (B - R_{xy} R_{yy}^{-1})^T}_{\text{Term 1}} - \underbrace{R_{xy} R_{yy}^{-1} R_{xy}^T}_{\text{Term 2}} + \underbrace{R_{xx}}_{\text{Term 3}}$$

- **Terms 2 and 3** do not depend on B — they are fixed given the data.
- **Term 1** is the only part we can control through our choice of B.

Term 1 has the form $A\, R_{yy}\, A^T$, where $A = (B - R_{xy} R_{yy}^{-1})$.

Key argument

Since R_{yy} is **positive definite**, $A\, R_{yy}\, A^T \geq 0$ always (diagonal elements are non-negative).

The **minimum is zero**, achieved only when $A = 0$, i.e.:

$$B - R_{xy} R_{yy}^{-1} = 0 \implies B = R_{xy} R_{yy}^{-1}$$

We minimise P by **eliminating the only term we can control**.

Application to the Noisy Observation Model

Our observations are related to the unknown vector x by:

Observation equation (eq. 6)

$$Ex + n = y$$

We compute the required covariance matrices, assuming noise n and unknowns x are **independent** ($R_{xn} = R_{nx}^T = 0$):

Data covariance (eq. 7):

$$\begin{aligned} R_{yy} &= \langle (Ex + n)(Ex + n)^T \rangle \\ &= E R_{xx} E^T + R_{nn} \end{aligned}$$

Cross-covariance (eq. 8):

$$\begin{aligned} R_{xy} &= \langle x(Ex + n)^T \rangle \\ &= R_{xx} E^T \end{aligned}$$

Estimator and Variance for the Noisy Model

Substituting into equations 2.44 and 2.45:

Estimate of x (eq. 9)

$$\tilde{x} = R_{xx} E^T \left(E R_{xx} E^T + R_{nn} \right)^{-1} y$$

Estimator variance (eq. 10)

$$P = R_{xx} - R_{xx} E^T \left(E R_{xx} E^T + R_{nn} \right)^{-1} E R_{xx}$$

This allows us to estimate the **velocity in the reference layer** and its variance.

Noise Estimate and Its Variance

From the observation equation (2.46), we obtain an estimate of the **noise** (eq. 11):

$$\tilde{n} = y - E\tilde{x}$$

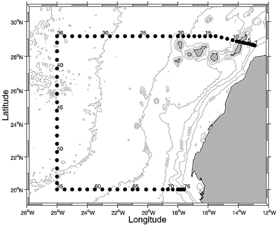
And its **variance** (eq. 12):

$$\begin{aligned} P_{nn} &= \langle (n - \tilde{n})(n - \tilde{n})^T \rangle \\ &= \langle (y - Ex - y + E\tilde{x})(y - Ex - y + E\tilde{x})^T \rangle \\ &= \langle E(\tilde{x} - x)(\tilde{x} - x)^T E^T \rangle \\ &= EPE^T \end{aligned}$$

Application Example: An Inverse Box Model in the Canary Current

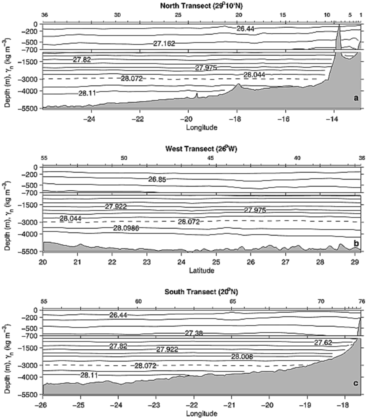
As a worked application of the Gauss–Markov estimator, we revisit the box inverse model of

Hernández-Guerra, A., E. Fraile-Nuez, F. López-Laatzen, A. Martínez, G. Parrilla & P. Vélez-Belchí (2005), *Canary Current and North Equatorial Current from an inverse box model*, J. Geophys. Res., 110, C12019, doi:10.1029/2005JC003032.



Hydrographic box bounded by the North, West and South transects around the Canary Islands. Numbers indicate station pairs j.

The water column along each transect is divided into neutral density (γ_n) layers, following the discretization introduced earlier ($q = 1, \ldots, Q$).



Neutral density sections along the North, West and South transects.

Water-Mass Layer Definitions

$Q = 14$ neutral density layers

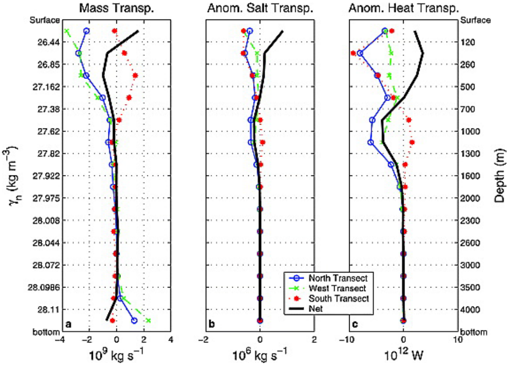
Table 1. Layer Definitions and Approximate Equivalences With Water Masses

Layer	Lower Interface γ_n, kg m^{-3}	Water Wass
1	26.44	surface water
2	26.85	NACW/SACW
3	27.162	NACW/SACW
4	27.38	NACW/SACW
5	27.62	MW/AAIW
6	27.82	MW/AAIW
7	27.922	MW/AAIW
8	27.975	UNADW
9	28.008	MNADW
10	28.044	MNADW
11	28.072	MNADW
12	28.0986	LNADW
13	28.11	LNADW
14	bottom	diluted AABW

Layer definitions and approximate equivalences with water masses

Transport Imbalance Before the Inverse Model

Using only the relative geostrophic velocity ($x = x_0$, i.e. $b_j = 0$), the layer-by-layer conservation equation $\sum_{j=1}^{N} e_{jq}\, b_j + n_q = y_q$ is **not** satisfied: the net transport in each layer departs from zero.



Mass, anomalous salt and anomalous heat transport per layer, before the inverse model is applied.

Prior Statistical Information for This Box

Repeating the definitions introduced earlier,

$$\langle x \rangle = x_0, \qquad P = R_{xx} = \left\langle (x - x_0)(x - x_0)^T \right\rangle .$$

Errors in different unknowns are usually assumed uncorrelated, so R_{xx} is diagonal:

$$R_{xx} = \begin{pmatrix} \underline{\qquad\qquad} & 0 & \cdots & 0 \\ 0 & \underline{\qquad\qquad} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \underline{\qquad\qquad} \end{pmatrix}$$

- b_j ($N = 74$ station pairs): $(0.05\ \mathrm{m\,s^{-1}})^2$ near the edges of the array, $(0.02\ \mathrm{m\,s^{-1}})^2$ for the interior station pairs,
- w_k (13 interfaces): $10^{-12}\ (\mathrm{m\,s^{-1}})^2$,
- $K_{z,k}$ (13 interfaces): $10^{-8}\ (\mathrm{m^2\,s^{-1}})^2$,
- ΔT_{Ek} (one per open boundary): $(0.5\ T_{Ek})^2$ – a 50% relative uncertainty on the Ekman transport itself.

Observational Error Covariance for This Box

$$\langle n \rangle = 0, \qquad R_{nn} = \left\langle nn^T \right\rangle .$$

Assuming uncorrelated errors between conservation equations, R_{nn} is diagonal:

$$R_{nn} = \begin{pmatrix} \_\_\_\_\_\_ & 0 & \cdots & 0 \\ 0 & \_\_\_\_\_\_ & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \_\_\_\_\_\_ \end{pmatrix}$$

- Mass (14 layers + 1 total-mass row): $(1.5\,\mathrm{Sv})^2$, $(1\,\mathrm{Sv})^2$ or $(0.5\,\mathrm{Sv})^2$ depending on the layer, and $(4\,\mathrm{Sv})^2$ for the total-mass row, with $1\,\mathrm{Sv} \equiv 10^6\,\mathrm{m}^3\,\mathrm{s}^{-1}$ converted to $(\mathrm{kg\,s}^{-1})^2$,
- Salt, and heat (deep layers only): scaled from the mass variance of the same layer following Ganachaud (2003), $4\,\sigma_C^2\,R_{nn}^{\mathrm{mass}}$, with σ_C the layer's climatological standard deviation of salinity or temperature.
- No separate regional (E_R) constraint was used in this particular box.

With E, y, R_{xx} and R_{nn} specified, the Gauss–Markov estimator gives

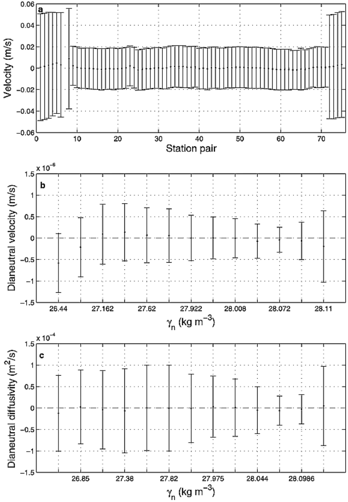
$$\boxed{\tilde{x} = R_{xx}E^T\left(ER_{xx}E^T + R_{nn}\right)^{-1}y}$$

and its uncertainty

$$\boxed{P = R_{xx} - R_{xx}E^T\left(ER_{xx}E^T + R_{nn}\right)^{-1}ER_{xx}}$$

The solution $\tilde{x}$ contains the reference velocities b_j and, in the generalized model, the diapycnal terms w_k, $K_{z,k}$ and the Ekman correction ΔT_{Ek}.

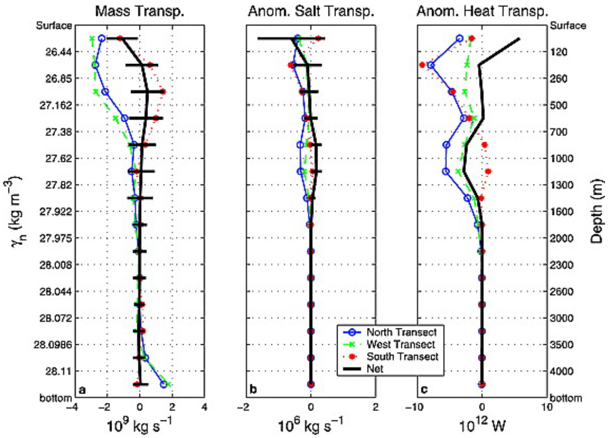
Estimated b_j, w_k and $K_{z,k}$ for This Box



(a) Reference velocity b_j per station pair; (b) dianeutral velocity w_k and (c) dianeutral diffusivity $K_{z,k}$, both as a function of neutral density γ_n. Error bars show the solution uncertainty, $\sqrt{P_{ii}}$.

Transport Balance After the Inverse Model

Once $\tilde{x}$ is obtained, the layer-by-layer conservation equation $\sum_{j=1}^{N} e_{jq}\, b_j + n_q = y_q$ is satisfied, up to the assigned error variance: the net transport in each layer is now close to zero.

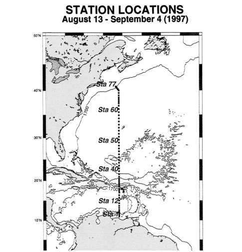


Mass, anomalous salt and anomalous heat transport per layer, after the inverse model is applied.

A Second Application: The NW Atlantic and Caribbean at 66°W

As a second worked example, illustrating mass and **silica** conservation together, we use the box inverse model of

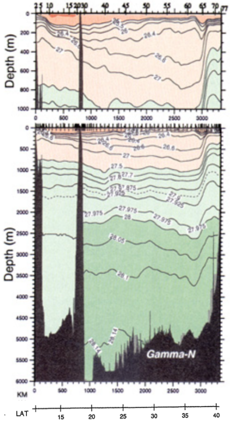
Joyce, T. M., A. Hernández-Guerra & W. M. Smethie (2001), *Zonal circulation in the NW Atlantic and Caribbean from a meridional World Ocean Circulation Experiment hydrographic section at 66° W*, J. Geophys. Res., 106, 22,095–22,113.



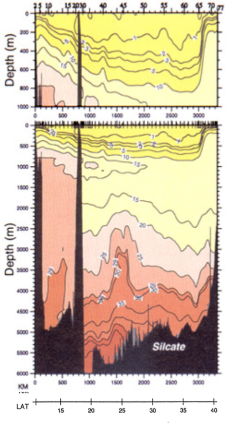
Station locations, 13 August – 4 September 1997 (stations 1–77). The section runs from the Caribbean, through the Jungfern Passage, into the open North Atlantic.

Hydrographic Section and Neutral Density Layers

The water column is divided into $Q = 17$ neutral density (γ_n) layers along the section, from the surface to the bottom.



Neutral density section along the 66°W line.



Silica section along the same line.

Water-Mass Layer Definitions

$Q = 17$ neutral density layers

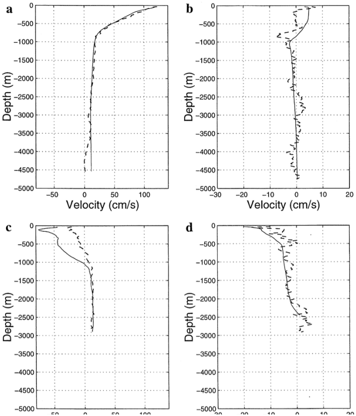
Table 1. Neutral Density Levels Used in Analysis

γ_n	Layer	Water Mass
20		
	1	surface layer
25		
	2	upper thermocline
25.5		
	3	SUW
26		
	4	upper thermocline
26.4		
	5	STMW
26.6		
	6	lower thermocline
27		
	7	AAIW
27.5		
	8	AAIW
27.7		
	9	AAIW
27.7		
	9	AAIW
27.8		
	10	ULSW
27.875		
	11	LSW
27.925		
	12	CLSW
27.975		
	13	ISOW
28		
	14	ISOW
28.05		
	15	DSOW
28.1		
	16	LDW
28.14		
	17	AABW

Neutral density levels and corresponding water masses, from the surface layer down to Antarctic Bottom Water (AABW).

Comparison with LADCP Velocities

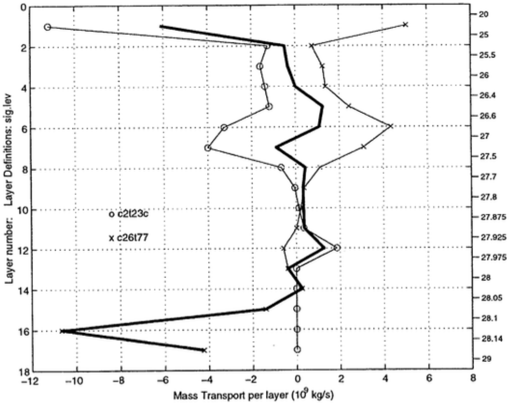
At a subset of station pairs, LADCP profiles provide an independent estimate of the absolute velocity, used to inform the reference-velocity prior.



Geostrophic (solid) and LADCP (dashed) velocity profiles: (a) Gulf Stream; (b) Sargasso Sea; (c,d) Deep Western Boundary Current.

Transport Imbalance Before the Inverse Model

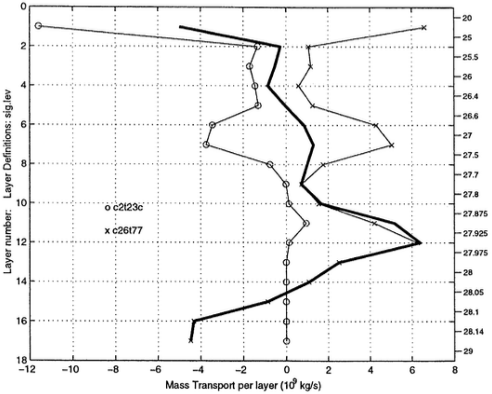
Using only the relative geostrophic velocity, mass transport is integrated separately for the Caribbean and Atlantic portions of the section.



Mass transport per layer for the Caribbean (circles) and Atlantic (crosses) portions of the section, and their sum (heavy line), before the inverse model is applied.

Incorporating LADCP into the First Guess

Adjusting the geostrophic velocity to the LADCP data, where available, provides an improved first guess x_0 before solving the inverse problem.



Mass transport per layer, as before, but using the geostrophic flow adjusted to the LADCP data.

Additional Constraints for This Section

In addition to mass conservation, this application imposes conservation of **silica**, following the generic property equation introduced earlier ($C =$ silica concentration).

- Within the **Caribbean Basin**, the transport in deep layers is conserved independently, consistent with the shallow sill of the Jungfern Passage limiting the inflow of deep water (LDW, AABW) from the open Atlantic.

Building the Regional Row: Splitting E_M by Sub-Basin

For most layers, mass conservation is a single row summed over all N station pairs. For the Caribbean-restricted deep layers (LDW, AABW), that single row is replaced by two regionally-restricted equations instead:

$$\underbrace{\sum_{j \in J_{Carib}} e_{jq}\, b_j + n_q^{Carib}}_{\text{Caribbean station pairs only}} = y_q^{Carib}, \qquad \underbrace{\sum_{j \in J_{Atl}} e_{jq}\, b_j + n_q^{Atl}}_{\text{Atlantic station pairs only}} = y_q^{Atl},$$

with $J_{Carib} \cup J_{Atl} = \{1, \dots, N\}$ and $J_{Carib} \cap J_{Atl} = \varnothing$.

In matrix form, the single combined row used elsewhere is replaced, for these layers, by two sparse rows:

$$\begin{pmatrix} e_1 & \cdots & e_k & 0 & \cdots & 0 \end{pmatrix} x = y_q^{Carib},$$
$$\begin{pmatrix} 0 & \cdots & 0 & e_{k+1} & \cdots & e_N \end{pmatrix} x = y_q^{Atl},$$

where station pairs $1, \dots, k$ lie in the Caribbean and $k+1, \dots, N$ in the open Atlantic. This is exactly the Caribbean/Atlantic split already shown in the transport-imbalance figures of this example.

Prior Statistical Information for This Section

$$\langle x \rangle = x_0, \qquad P = R_{xx} = \left\langle (x - x_0)(x - x_0)^T \right\rangle .$$

$$R_{xx} = \begin{pmatrix} \underline{\hspace{2cm}} & 0 & \cdots & 0 \\ 0 & \underline{\hspace{2cm}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \underline{\hspace{2cm}} \end{pmatrix}$$

- b_j: $(0.05\ \mathrm{m\,s^{-1}})^2$ for station pairs with LADCP measurements, $(0.10\ \mathrm{m\,s^{-1}})^2$ for the rest.

Observational Error Covariance for This Section

$$\langle n \rangle = 0, \qquad R_{nn} = \left\langle nn^T \right\rangle .$$

$$R_{nn} = \begin{pmatrix} \rule{2em}{0.4pt} & 0 & \cdots & 0 \\ 0 & \rule{2em}{0.4pt} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \rule{2em}{0.4pt} \end{pmatrix}$$

- Mass: $(2\,\mathrm{Sv})^2$ for the first four layers, $(0.5\,\mathrm{Sv})^2$ for the last three constraints, and $(1\,\mathrm{Sv})^2$ for the rest.
- Silica: $(20\ \mathrm{kmol\,s^{-1}})^2$ for the first four layers, $(5\ \mathrm{kmol\,s^{-1}})^2$ for the last three constraints, and $(10\ \mathrm{kmol\,s^{-1}})^2$ for the rest of the equations.

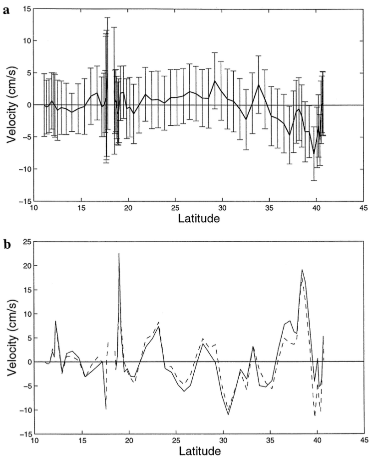
Gauss-Markov Solution for This Section

$$\boxed{\tilde{x} = R_{xx}E^T\left(ER_{xx}E^T + R_{nn}\right)^{-1}y}$$

$$\boxed{P = R_{xx} - R_{xx}E^T\left(ER_{xx}E^T + R_{nn}\right)^{-1}ER_{xx}}$$

The solution $\tilde{x}$ contains the reference velocities b_j consistent with the geostrophic, LADCP, mass and silica information described above.

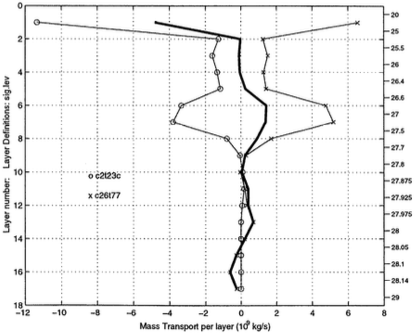
Estimated Velocities Along the Section



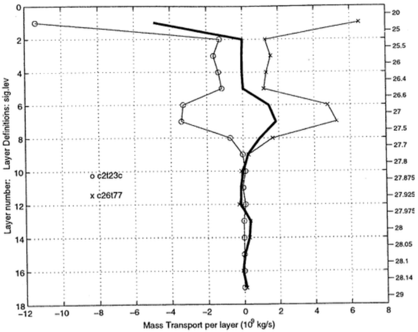
(a) Velocity from the inverse solution using mass and silica constraints, with uncertainty; (b) LADCP-adjusted velocity (solid) and LADCP plus inverse-derived velocity (dashed), as a function of latitude.

Transport Balance After the Inverse Model

Once $\tilde{x}$ is obtained, the net transport in each layer is much closer to zero in both portions of the section – shown here without (left) and with (right) the silica constraint.

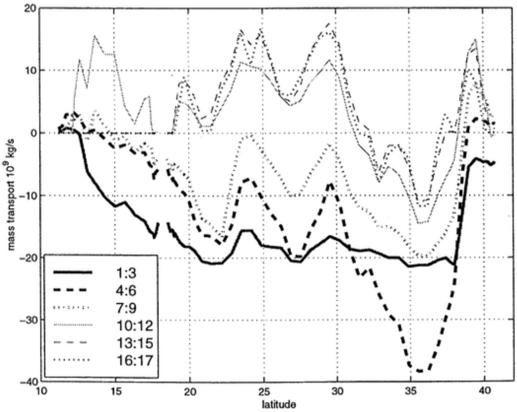


Mass conservation only (no silica).



Mass and silica constraints.

Meridional Transport by Water-Mass Group



Meridional mass transport stream function for groups of layers: near surface (1–3), thermocline (4–6), AAIW (7–9), LSW (10–12), and Nordic overflow waters, LDW and AABW (13–17).

A Third Application: Meridional Overturning at 7.5N and 24.5N

As a third worked example, illustrating the use of **regional transport constraints**, we use the box inverse model of

Hernández-Guerra, A., J.L. Pelegrí, E. Fraile-Nuez, M. Benítez-Barrios, M. Emilianov, M.D. Pérez-Hernández & P. Vélez-Belchí (2014), *Meridional overturning transports at 7.5N and 24.5N in the Atlantic Ocean during 1992–93 and 2010–11*, Progress in Oceanography, 128, 98–114, doi:10.1016/j.pocean.2014.08.016.

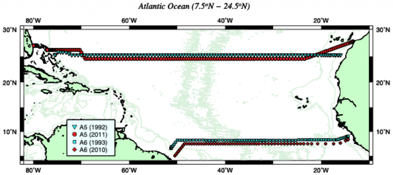


Fig. 1. Map showing the hydrographic sections carried out in 1993 and 2010 at 7.5N (section A06) and in 1992 and 2011 at 24.5N (section A05). Note that cruise tracks in 1993 at 7.5N and in 1992 at 24.5N are offset 0.5 degree in latitude for clarity.

The water column along each section is divided into $Q = 17$ neutral density (γ_n) layers, following the same discretization introduced earlier.

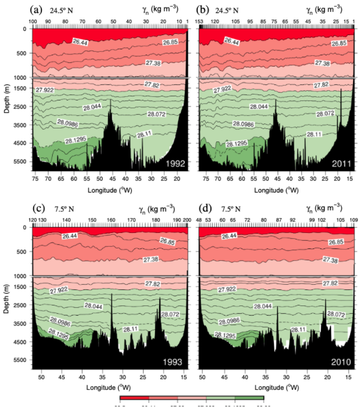


Fig. 5. Vertical sections of neutral density (kg m^{-3}) along 24.5N in (a) 1992 and (b) 2011, and along 7.5N in (c) 1993 and (d) 2010.

Transport Imbalance Before the Inverse Model

Using only the relative geostrophic velocity ($x = x_0$, i.e. $b_j = 0$), the zonally integrated mass transport per layer departs from zero across both sections, in both periods.

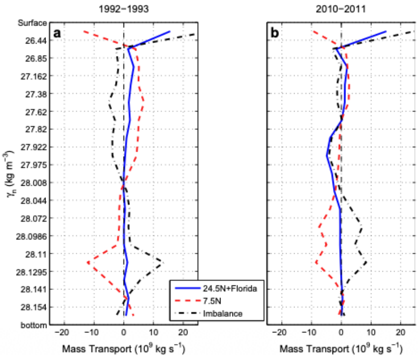


Fig. 6. Initial solution for the zonally integrated mass transports (Sv) per layer across 24.5N and 7.5N, during (a) 1992–1993 and (b) 2010–2011.

Regional Transport Constraints for This Application

Independent transport estimates constrain specific station pairs, exactly as in the regional constraint E_R introduced earlier:

- the **Florida Current** transport through the Florida Strait,
- the **Bering Strait** inflow,
- the **Deep Western Boundary Current (DWBC)** and its **recirculation (RECIR)** at 7.5N,
- **Antarctic Bottom Water (AABW)** transport,
- the **Antilles Current (AC)**.

	Long. (W)	Layers	Constraint	1992–3 Initial	1992–3 Final	2010–1 Initial	2010–1 Final
7.5N (A06)							
Bering St.	All	All	−0.8 ± 0.6	11.2	−0.8 ± 3.5	−26.3	−0.8 ± 3.9
DWBC	−52: −45	6: 11	−15 ± 7.5	−17.8	−15.2 ± 2.6	3.6	−11.7 ± 2.7
RECIR	−45: −35	6: 11	6 ± 3	−9.6	3.6 ± 1.3	−9.1	5.0 ± 1.3
AABW	−49: −36	16: 17	1.7 ± 0.6	0.6	1.6 ± 0.9	0.3	1.5 ± 0.8
24.5N (A05)							
Florida Current	Florida Strait	All	31.0 ± 1.0 32.0 ± 2.8	23.0	31.0 ± 0.2	23.0	31.8 ± 0.3
Bering St.	All	All	−0.8 ± 0.6	21.4	−0.8 ± 5.7	−4.1	−0.8 ± 5.4
AC	Diff.	1: 6	6.0 ± 8.1	19.6	18.8 ± 1.6	15.6	12.8 ± 1.0
DWBC	−77: −72	7: 14	−26.5 ± 13.6	16.0	−22.1 ± 4.2	−15.8	−20.1 ± 3.7

Table 4. Specific constraints (Sv) for sections A06 (along 7.5N) and A05 (along 24.5N), for a range of longitudes and layers. Positive transports are to the north. The initial and final transports (respectively calculated as described in the text and as deduced with the inverse model) are shown. DWBC stands for Deep Western Boundary Current, RECIR for the recirculation of DWBC at 7.5N, AABW for Antarctic Bottom Water, and AC for Antilles Current. The eastern end of the Antilles Current was located at about 71.5W during 1992 and near 75.5W in 2011. The Florida Current uses two different constraints, depending on the year: the top value corresponds to 1992 and the bottom one to 2011.

Building the Regional Constraint Row: Two Examples

Each regional constraint adds one sparse row to E, nonzero only for the station pairs and layers crossed by that current:

$$\sum_{j \in J} \sum_{q \in Q} e_{jq} \, b_j + n_T = T_{obs} - T_r$$

Florida Current (Florida Strait, all layers, 1992):

$$\begin{pmatrix} e_j & \cdots & e_j & 0 & \cdots & 0 \end{pmatrix} x = T_{obs} - T_r$$

$$31.0 - 23.0 = 8.0 \text{ Sv}$$

(J = station pairs across the Florida Strait, Q = all layers)

DWBC at 7.5N (long. -52 to -45, layers 6–11, 1992):

$$\begin{pmatrix} 0 & \cdots & 0 & e_j & \cdots & e_j & 0 & \cdots & 0 \end{pmatrix} x = T_{obs} - T_r$$

$$-15 - (-17.8) = 2.8 \text{ Sv}$$

(J = station pairs in that longitude range, Q = layers 6–11 only)

Reference Velocities from the Inverse Solution

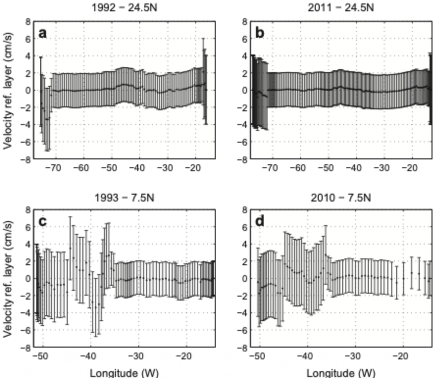


Fig. 7. Reference velocities (cm s^{-1}), with error bars, along 24.5N in (a) 1992 and (b) 2011, and along 7.5N in (c) 1993 and (d) 2010.

Diapycnal Velocities Between Layers

Combining mass conservation with the regional constraints above also yields the dianeutral velocities w_k across each interface, for the basin enclosed by the two sections.

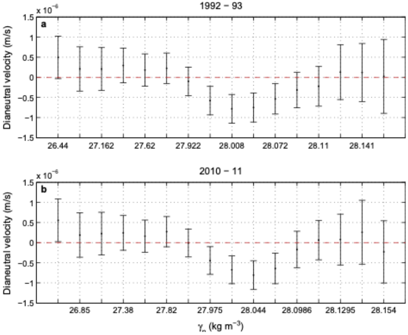


Fig. 8. Dianeutral velocities between layers, corresponding to the mean values over the basin enclosed by the 7.5N and 24.5N sections, with the layers specified as indicated in Table 2.

Transport Balance After the Inverse Model

Once $\tilde{x}$ is obtained, the net transport per layer is much closer to zero in both periods, allowing the meridional overturning to be compared between 1992–93 and 2010–11.

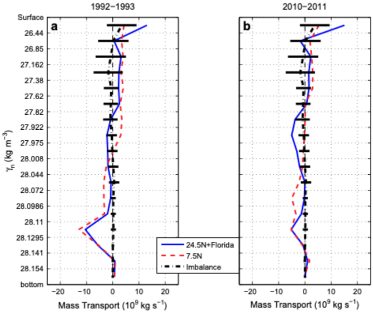


Fig. 9. Zonally integrated mass transport (Sv) per layer across 24.5N and 7.5N, during (a) 1992–1993 and (b) 2010–2011; it is analogous to Fig. 6 but with the reference and dianeutral velocities as estimated from the inverse model. The uncertainties per layer are shown as error bars in the imbalance.

A Fourth Application: Indian and Pacific Oceans

As a fourth worked example, again focusing on **regional transport constraints**, we use the box inverse model of

Hernández-Guerra, A. & L.D. Talley (2016), *Meridional overturning transports at 30°S in the Indian and Pacific Oceans in 2002–2003 and 2009*, Progress in Oceanography, 146, 89–120, doi:10.1016/j.pocean.2016.06.005.

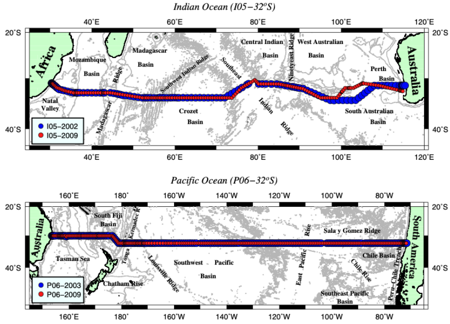


Fig. 1. Station locations for the four cruises (Table 2).

Hydrographic Sections: Salinity Along 32°S

As in the previous examples, the inverse model is built from hydrographic sections occupied across each basin.

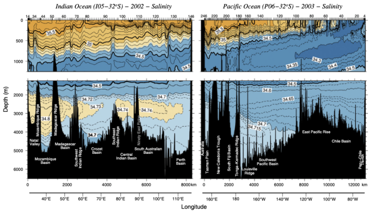


Fig. 4. Vertical sections of salinity in 2002–2003 for the (a) Indian Ocean and (b) Pacific Ocean.

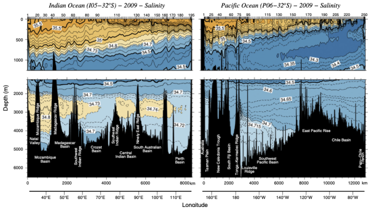


Fig. 5. Vertical sections of salinity in 2009 for the (a) Indian Ocean and (b) Pacific Ocean.

Regional Transport Constraints for This Application

Beyond mass and silicate conservation, independent transport estimates constrain specific boundary currents and deep passages in each basin, exactly as in the regional constraint E_R introduced earlier:

- **Indian Ocean (I05)**: the **Agulhas Current**, the **Leeuwin Current**, and deep transport through the Natal Valley, Mozambique Basin, Madagascar Basin and Central Basin,
- **Pacific Ocean (P06)**: the **Tonga–Kermadec Ridge**, the **East Australian Current**, the **Peru–Chile Current** and **Undercurrent**, and deep transport through the Tasman Sea and the Eastern Basin,
- both basins: the **Indonesian Throughflow (ITF)** and Bering Strait transport.

	Property	Longitude	Layers	Constraint	2002 Initial	2002 Final	2009 Initial	2009 Final	Model
I05									
Total mass	ITF transport (Sv)	All	1:11	−14 ± 5	5.2	−8.0 ± 4.5 −10.8 ± 4.7 −14.1 ± 8.6	−20.4	−10.0 ± 3.9 −12.0 ± 3.9 −13.7 ± 6.3	A B C
Silicate conservation	Total Silicate (kmol/s)	All	1:11	0 ± 100	555	8.2 ± 449 −28.1 ± 498 −11.2 ± 954	445	40 ± 373 20 ± 405 17 ± 679	A B C
Deep transport[a]	Natal Valley (Sv)	30–34E	10:11	0 ± 0.5	0.0	−0.3 ± 0.3 −0.1 ± 0.3 0.0 ± 0.4	0.1	0.6 ± 0.3 0.2 ± 0.3 0.1 ± 0.3	A B C
Deep transport[a]	Mozamb. Basin (Sv)	35–46E	9:11	1 ± 2	−1.1	−1.5 ± 1.2 −0.5 ± 1.2 0.7 ± 1.3	−0.9	1.5 ± 1.2 1.4 ± 1.2 1.0 ± 1.3	A B C
Deep transport[a]	Madag. Basin (Sv)	46–56E	10:11	0 ± 0.5	−0.7	−0.2 ± 0.5 −0.2 ± 0.5 −0.1 ± 0.5	−1.3	0.2 ± 0.5 0.2 ± 0.5 0.1 ± 0.5	A B C
Deep transport[a]	Central Basin (Sv)	78–89E	9:11	0 ± 1	0.5	0.7 ± 0.9 −0.2 ± 0.9 −0.1 ± 1.0	0.4	0.4 ± 0.9 −0.1 ± 0.9 0.0 ± 1.0	A B C
Boundary current constraint[c]	Agulhas Current (Sv)	~30–32E	1:11	−55.1 ± 10 (for 2002 LADCP stas. 1–18) −91.2 ± 10 (for 2009 LADCP stas. 1–9)	−65.4	−64.6 ± 2.5 −61.4 ± 2.5 −61.9 ± 2.6	−87.7	−84.4 ± 2.2 −86.9 ± 2.1 −87.2 ± 2.3	A B C
Boundary current	Leeuwin Current	114.5E to coast	NA	NA	NA	−1.0 ± 0.1	NA	−5.4 ± 0.1	A, B, C

Table 5 (I05). Regional transport constraints and inverse model results for the Indian Ocean (I05) at 32°S. Positive transports are northward. Initial transports are relative to the ZVS at $\gamma^n = 28.10$ kg m^{-3}; final transports are given for each model (A, B, C).

Regional Constraints: Pacific Ocean (P06)

	Property	Long.	Layers	Constraint	2003 Initial	2003 Final	2009 Initial	2009 Final	Model
P06									
Total mass	ITF and Bering Strait transport (Sv)	All	1:11	15 ± 5	11.8	6.8 ± 4.1	−4.3	8.8 ± 4.0	A
						10.0 ± 4.7		10.7 ± 4.6	B
						13.2 ± 10.7		12.6 ± 10.0	C
Silicate conservation	Total Silicate (kmol/s)	All	1:10	0 ± 100	−202	68 ± 420		61 ± 419	A
						−3 ± 509	−422	72 ± 509	B
						0 ± 1213		28 ± 1139	C
Deep transport[a]	Tasman Sea (Sv)	150–161E	9:10	0.5 ± 0.5	0.6	0.5 ± 0.5		0.7 ± 0.5	A
						0.6 ± 0.5	1.7	0.7 ± 0.5	B
						0.5 ± 0.5		0.6 ± 0.5	C
Deep transport[a]	Tasman Sea (Sv)	161–173E	7:10	0 ± 0.5	0.5	0.0 ± 0.5		0.0 ± 0.5	A
						0.0 ± 0.5	0.4	0.0 ± 0.5	B
						0.0 ± 0.5		0.0 ± 0.5	C
Deep transport[a]	Tasman Sea (Sv)	173–181E	7:10	0 ± 0.5	0.3	0.0 ± 0.5		0.0 ± 0.6	A
						0.0 ± 0.5	1.1	0.0 ± 0.6	B
						0.0 ± 0.5		0.0 ± 0.6	C
Deep transport[a]	Eastern Basin (Sv)	110–70 W	8:10	0.75 ± 0.75	−0.4	0.9 ± 0.7		0.6 ± 0.7	A
						1.3 ± 0.7	−0.9	0.7 ± 0.7	B
						0.7 ± 0.7		0.7 ± 0.7	C
Boundary current constraint[c]	Tonga-Kermadec Ridge (Sv)	180–168 W	8:10	15 ± 4 (Whitworth et al., 1999)	7.1	14.5 ± 3.0		10.8 ± 2.9	A
				12.4 ± 7 (2009 LADCP data)		14.6 ± 3.0	14.0	11.4 ± 2.9	B
						14.6 ± 3.1		11.8 ± 3.0	C
Boundary current result	East Australian Current (Sv)	Coast to 156.7E (2003) 158.9E (2009)	NA	NA	NA	−51.1 ± 1.4	NA	−46.1 ± 2.1	A
						−53.0 ± 2.1		−47.9 ± 2.0	B
						−51.1 ± 2.0		−49.9 ± 2.1	C
Boundary current result	Peru–Chile Current	85 W to coast	NA	NA	NA	3.3 ± 0.9	NA	2.3 ± 0.8	A, B, C
Boundary current result	Peru–Chile Undercurrent	75 W to coast	NA	NA	NA	−1.6 ± 1.2	NA	−5.0 ± 1.2	A
						−1.3 ± 1.2		−3.8 ± 1.2	B

Table 5 (P06, continued). Regional transport constraints and inverse model results for the Pacific Ocean (P06) at 32°S. Positive transports are northward. Initial transports are relative to the ZVS at $\gamma^n = 28.10$ kg m^{-3}; final transports are given for each model (A, B, C).

Building the Regional Constraint Row: Two Examples

As in the previous example, each regional constraint adds one sparse row to E, nonzero only for the station pairs and layers crossed by that current:

$$\sum_{j \in J} \sum_{q \in Q} e_{jq}\, b_j + n_T = T_{obs} - T_r$$

Agulhas Current (I05, long. -30 to -32, all layers, 2002 LADCP):

$$\begin{pmatrix} 0 & \cdots & 0 & e_j & \cdots & e_j & 0 & \cdots & 0 \end{pmatrix} x = T_{obs} - T_r$$

$$-55.1 - (-65.4) = 10.3 \text{ Sv}$$

Tonga–Kermadec Ridge (P06, long. 180–168W, layers 8–10, 2003):

$$\begin{pmatrix} e_j & \cdots & e_j & 0 & \cdots & 0 & 0 & \cdots & 0 \end{pmatrix} x = T_{obs} - T_r$$

$$15 - 7.1 = 7.9 \text{ Sv}$$

Overturning Mass Transport Streamfunction

Once the reference velocities are estimated, the overturning streamfunction is obtained by integrating the zonally-integrated layer transports from the bottom to the top.

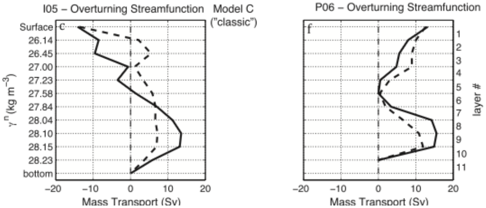
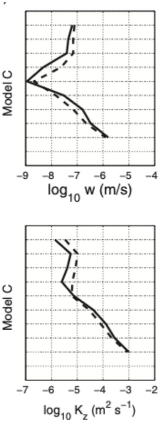


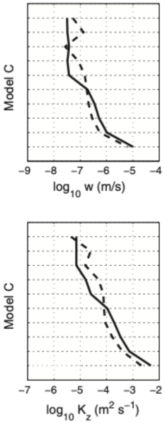
Fig. 16. Overturning mass transport streamfunction across 32°S for the Indian Ocean (left subplots) and Pacific Ocean (right subplots) from each model and year of the cruise. This function corresponds to the zonally-integrated mass transport in isoneutral layers, integrated from the bottom to the top.

Diapycnal Velocity and Diffusivity

The generalized inverse model also yields the dianeutral velocity w and diffusivity K_z profiles for each basin (Model C shown).



Pacific Ocean (P06): diapycnal velocity (top) and diffusivity (bottom), Model C.



Indian Ocean (I05): diapycnal velocity (top) and diffusivity (bottom), Model C.

Context and objective

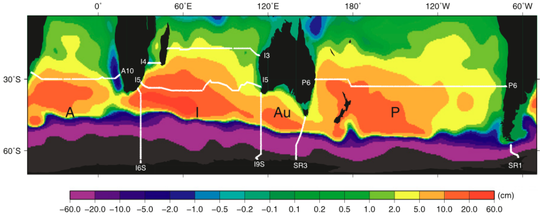
As a fifth worked example, we use the box inverse model of

Katsumata, K., B.M. Sloyan & S. Masuda (2013), *Diapycnal and Isopycnal Transports in the Southern Ocean Estimated by a Box Inverse Model*, Journal of Physical Oceanography, 43, 2270–2287, doi:10.1175/JPO-D-12-0210.1.

- The Southern Ocean is one of only two latitude bands where surface waters gain density to become bottom waters, while also acting as a return pathway of deep water back to the surface.
- **Objective:** estimate the *isopycnal* transports (along the sections) and the *diapycnal* transports (across the isopycnals), distinguishing those that occur in the **mixed layer** from those in the ocean **interior**.
- **Strategy:** combine two occupations of eight hydrographic sections and the mid-depth velocity derived from float drift, within a **box inverse model**.

Data: two occupations and eight sections

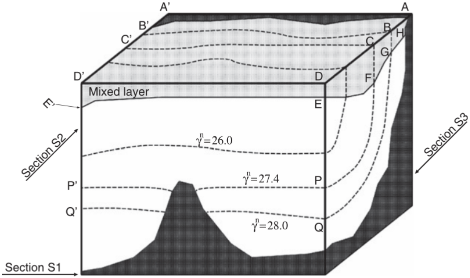
- **Two datasets:**
  - WOCE/WHP (Nov 1987 – Oct 1998)
  - GO-SHIP / Revisit (Aug 2003 – May 2009)
- Eight sections define the box boundaries.
- Neutral density γ^n as the vertical coordinate (sea floor to surface).
- Geostrophic velocity with initial zero-velocity surfaces; the inverse determines the barotropic correction.



Hydrographic sections defining the boxes (white lines). Black labels mark the sectors of the Southern Ocean: Atlantic (A), Indian (I), Australian (Au) and Pacific (P). The color scale shows the geostrophic pressure (cm of water column) at 1000 dbar, estimated from WOCE subsurface (Davis 1998, 2005) and Argo (Katsumata and Yoshinari 2010) drift data.

Box geometry: the mixed-layer box

- Each basin is split into a **surface box** (mixed layer) and **interior boxes**.
- The base of the mixed layer is defined by density: where σ is $0.03\ \mathrm{kg\,m^{-3}}$ heavier than at the surface.
- **Subduction/upwelling** through the mixed-layer base is estimated as part of the inverse solution.



Box inverse model with a surface mixed-layer box, bounded by the Antarctic continent, two meridional sections (S2, S3) and one zonal section (S1). The isopycnals (dashed) outcrop and cross the base of the mixed layer at various points; e.g. $\gamma^n = 27.4$ outcrops at C and C′ and crosses the base of the mixed layer at F and F′.

Mixed-layer formulation

Volume budget of the mixed layer (following Sallée *et al.* 2010):

$$S = \frac{\partial H}{\partial t} + \nabla \cdot \int_{-H}^{0} \mathbf{u}\, dz,$$

where S is the rate at which water crosses the mixed-layer base (positive is upward), H its thickness and $\mathbf{u}$ the horizontal velocity.

- Integrating over the area between two sections and two isopycnals yields the constraint that links subduction/upwelling to the horizontal transport (Ekman, geostrophic and residual "eddy") within the layer.
- Volume conservation of the interior boxes closes the system through the diapycnal fluxes W_{γ^n} across each isopycnal.

Discretizing the Mixed-Layer Budget

Starting from $S = \dfrac{\partial H}{\partial t} + \nabla \cdot \int_{-H}^{0} \mathbf{u}\, dz$, integrate over the area A of one box and apply the **divergence theorem**: the area integral of a horizontal divergence equals the line integral of the transport normal to the box boundary – exactly the step already used to turn every other layer budget into a sum over station pairs.

$$\underbrace{\int_A \nabla \cdot \int_{-H}^{0} \mathbf{u}\, dz\; dA}_{\text{divergence theorem}} = \sum_i (v_i + b_i)\, a_i^{ML}, \qquad a_i^{ML} = \rho\, \Delta a_i^{ML},$$

where a_i^{ML} is the cross-sectional area of station pair i **from the surface down to the mixed-layer base** H – the same kind of coefficient as $e_{jq} = \rho_{jq} \Delta a_{jq}$ used for every isopycnal layer, just integrated over a different (shallower, moving) depth range.

The two other area integrals are:

$$S_{box} \equiv \int_A S\, dA \qquad \text{(unknown)},$$

$$H'_{box} \equiv \int_A \frac{\partial H}{\partial t}\, dA \quad \text{(known, from MLD climatology)}.$$

Identifying E, x and y for the Mixed-Layer Row

Substituting and collecting unknowns (b_i, S_{box}) on the left and known quantities on the right gives one new row of $E\,x = y$:

$$\underbrace{\begin{pmatrix} \cdots & a_i^{ML} & \cdots & -1 \end{pmatrix}}_{E\ \text{(this row)}} \underbrace{\begin{pmatrix} \cdots & b_i & \cdots & S_{box} \end{pmatrix}^T}_{x} = \sum_i b_i\, a_i^{ML} - S_{box}$$

$$\underbrace{-\sum_i v_i\, a_i^{ML}}_{\text{relative transport, as in } y_q = -\sum \rho\, v_r\, \Delta a} \qquad - H'_{box} \equiv y_{ML}.$$

- **Row of E** (above): entry a_i^{ML} under each b_i for the station pairs bounding this box, entry -1 under S_{box}, zero elsewhere.
- x: the same vector as before, extended by one entry, S_{box} – the shallowest “interface” unknown, alongside every b_i and every interior W_j.
- y: $y_{ML} = -\sum_i v_i\, a_i^{ML} - H'_{box}$, i.e. the usual hydrography-only transport, **plus** the one genuinely new term, the climatological mixed-layer-depth tendency H'_{box} (computed independently, e.g. from Argo).

The inverse problem

Conservation of volume, heat and salt (silicate in the Indian) is imposed in every box and layer this way – one row of $E\,x = y$ per box/layer, exactly as just derived for the mixed-layer box.

Solution vector: $\boldsymbol{x} = (b_i,\ S_{box},\ W_j)$ – barotropic corrections, mixed-layer subduction/upwelling and interior diapycnal fluxes, all solved for jointly.

Two ways to combine the occupations

Direct method

- The two occupations are *averaged first* and the inverse is then applied to the mean section.
- Uses the time-mean Ekman transport.

Merging method

- The inverse is applied to each occupation separately and the *solutions are averaged*.
- Lower uncertainty than the direct method.

When they agree

Both agree where the cross-section eddy transport is small (Tziperman 1988). They differ in the **Atlantic** and **Indian** sectors, where eddies are strong (Agulhas retroflection at I6S, Benguela Current at A10).

Summary

- The **central novelty** of the model is the explicit separation between the **mixed-layer box** and the ocean **interior**.
- It enables an indirect estimate, from hydrographic data, of subduction/upwelling through the mixed-layer base and of interior diapycnal mixing, including the density range under winter sea ice ($\gamma^n > 27.5$).
- The merging method reduces the uncertainty relative to the direct method.
- Uncertainties remain large because of data scarcity, especially in austral winter and under sea ice: more high-latitude observations are needed.