Model Green's Functions

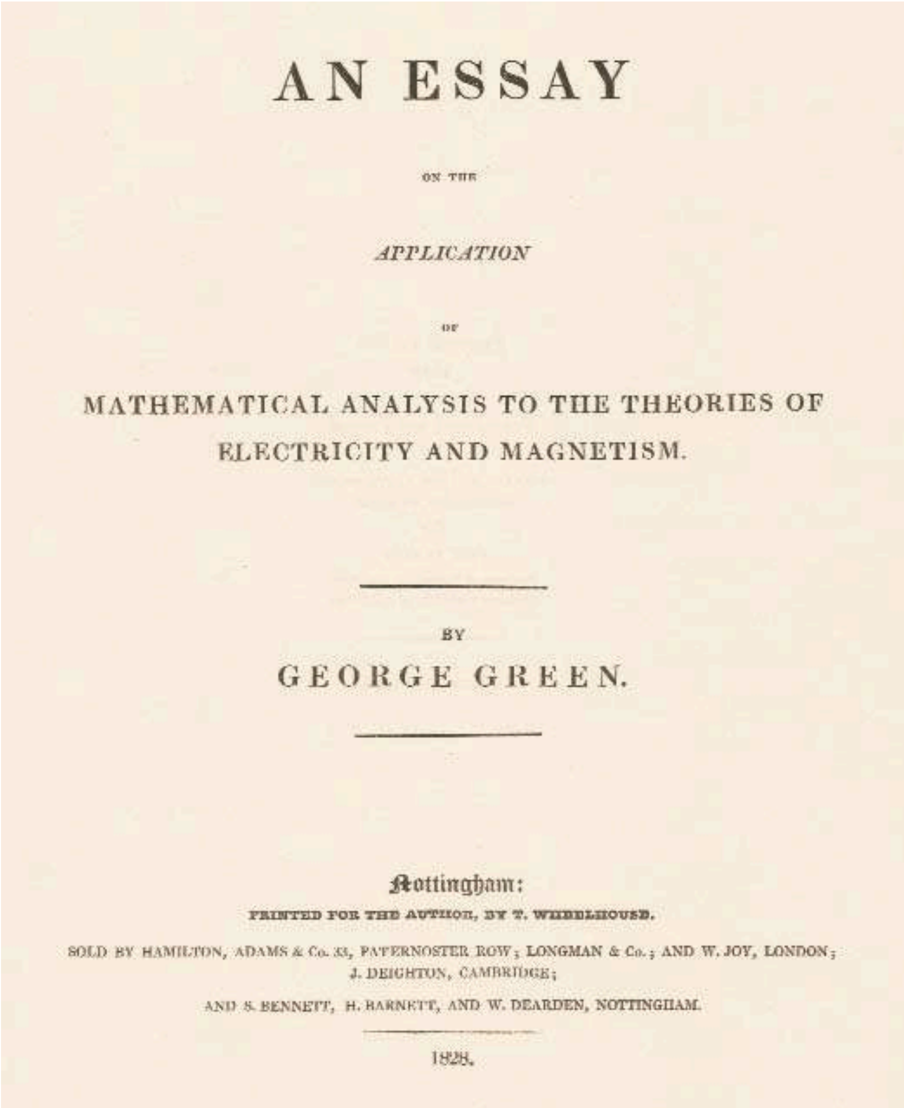
(a simple but effective way to adjust model parameters)

Dimitris Menemenlis
Jet Propulsion Laboratory
California Institute of Technology

ECCO Summer School 2019
May 19–31, 2019
Friday Harbor Laboratories
University of Washington
Friday Harbor, WA

- **The Green of Green’s Functions**
- **Green’s Functions for linear differential equations**
- **Model Green’s Functions estimation approach**
- **Some example applications**
- **Comparison with representer method**
- **Comparison with adjoint method**
- **Summary and concluding remarks**

The Green of Green's Functions (Challis & Shear, Physics Today, 2003)


AN ESSAY

ON THE

APPLICATION

OF

MATHEMATICAL ANALYSIS TO THE THEORIES OF ELECTRICITY AND MAGNETISM.

BY

GEORGE GREEN.

Nottingham:

PRINTED FOR THE AUTHOR, BY T. WHEELHOUSE.

SOLD BY HAMILTON, ADAMS & Co. 33, PATERNOSTER ROW; LONGMAN & Co.; AND W. JOY, LONDON;
J. DEIGHTON, CAMBRIDGE;
AND S. BENNETT, H. BARNETT, AND W. DEARDEN, NOTTINGHAM.

1828.

In 1828, an English miller from Nottingham published a mathematical essay that generated little response. George Green's analysis, however, has since found applications in areas ranging from classical electrostatics to modern quantum field theory.

Green's mill, near Nottingham, England

Bromley House, circa 1880, the location of the Nottingham Subscription Library.

Green’s Functions for linear differential equations

Let L be an arbitrary linear differential operator.

A Green's function, $G(x,y)$, is defined as the impulse response of this linear operator, that is:

$$LG(x,y) = \delta(x - y),$$

where $\delta(x-y)$ is the Dirac delta function applied at location y.

By linear superposition, Green’s functions can be used to solve a differential equation with arbitrary forcing term, $Lu(x) = f(x)$.

The solution is the convolution: $u(x) = \int G(x,y)\, f(y)\, dy$.

Model Green's Functions estimation approach

GCM: A General Circulation Model can be represented by a set of rules for time stepping a state vector $\mathbf{x}(t_i)$ one time step in the future:

$$\mathbf{x}(t_{i+1}) = M(\mathbf{x}(t_i), \boldsymbol{\eta})$$

where M represents the known time stepping rules and vector $\boldsymbol{\eta}$ represents perturbations to a set of model parameters. Vector $\boldsymbol{\eta}$ is assumed to be a noise process with zero mean and covariance matrix $\mathbf{Q}$.

Model Green's Functions estimation approach

GCM: A General Circulation Model can be represented by a set of rules for time stepping a state vector $\mathbf{x}(t_i)$ one time step in the future:

$$\mathbf{x}(t_{i+1}) = M(\mathbf{x}(t_i), \boldsymbol{\eta})$$

where M represents the known time stepping rules and vector $\boldsymbol{\eta}$ represents perturbations to a set of model parameters. Vector $\boldsymbol{\eta}$ is assumed to be a noise process with zero mean and covariance matrix **Q**.

Data: The state estimation problem aims to estimate parameters $\boldsymbol{\eta}$ given a set of observations:

$$\mathbf{y} = H(\mathbf{x}) + \boldsymbol{\varepsilon}$$

where H is the measurement function, and residual $\boldsymbol{\varepsilon}$ is a noise process assumed to have zero mean and covariance matrix **R**. For the Green's function approach, the data equation is rewritten:

$$\mathbf{y} = G(\boldsymbol{\eta}) + \boldsymbol{\varepsilon}$$

where G is the convolution of measurement function H with GCM dynamics M.

Model Green's Functions estimation approach

Cost function: Control parameters $\boldsymbol{\eta}$ can be estimated by minimizing a quadratic cost function:

$$J = \boldsymbol{\eta}^{T}\mathbf{Q}^{-1}\boldsymbol{\eta} + \boldsymbol{\varepsilon}^{T}\mathbf{R}^{-1}\boldsymbol{\varepsilon}$$

where superscript T is the transpose operator and superscript –1 denotes a matrix inversion.

Model Green's Functions estimation approach

Cost function: Control parameters $\boldsymbol{\eta}$ can be estimated by minimizing a quadratic cost function:

$$J = \boldsymbol{\eta}^{T}\mathbf{Q}^{-1}\boldsymbol{\eta} + \boldsymbol{\varepsilon}^{T}\mathbf{R}^{-1}\boldsymbol{\varepsilon}$$

where superscript T is the transpose operator and superscript –1 denotes a matrix inversion.

Linearization: To minimize this cost function, the GCM and data equations are linearized about a baseline simulation $\mathbf{x}_b$ ($\boldsymbol{\eta} = \mathbf{0}$). For "small" perturbations:

$$G(\boldsymbol{\eta}) \approx G(\mathbf{0}) + \mathbf{G}\boldsymbol{\eta}$$

where matrix $\mathbf{G}$ is an $n \times p$ matrix, n is the dimension of observation vector $\mathbf{y}$, and p is the dimension of parameter vector $\boldsymbol{\eta}$. Matrix $\mathbf{G}$ can be determined by performing a series of GCM sensitivity experiments. Specifically, each column of matrix $\mathbf{G}$ is obtained by perturbing the corresponding element in parameter vector $\boldsymbol{\eta}$ and then carrying out a GCM integration over the estimation period.

Model Green's Functions estimation approach

Minimization: The minimization of cost function J subject to the linearized model-data constraints has solution:

$$\eta_a = \mathbf{P}\mathbf{G}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{y}_d$$

where $\mathbf{y}_d$ is the model-data residual, that is, $\mathbf{y}_d \equiv \mathbf{y} - G(\mathbf{0})$, and $\mathbf{P}$ is the uncertainty covariance matrix:

$$\mathbf{P} = (\mathbf{Q}^{-1} + \mathbf{G}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{G})^{-1}$$

Model Green's Functions estimation approach

Minimization: The minimization of cost function J subject to the linearized model-data constraints has solution:

$$\eta_a = \mathbf{P}\mathbf{G}^T\mathbf{R}^{-1}\mathbf{y}_d$$

where $\mathbf{y}^d$ is the model-data residual, that is, $\mathbf{y}_d \equiv \mathbf{y}_o - G(\mathbf{0})$, and $\mathbf{P}$ is the uncertainty covariance matrix:

$$\mathbf{P} = (\mathbf{Q}^{-1} + \mathbf{G}^T\mathbf{R}^{-1}\mathbf{G})^{-1}$$

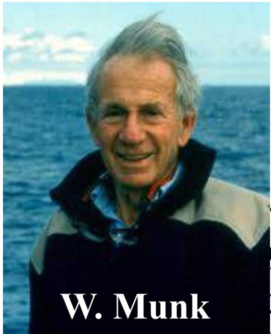
Solution: The optimized solution $\mathbf{x}_a$ is:

$$\mathbf{x}_a = \mathbf{x}_b + (\mathbf{G}^T\mathbf{R}^{-1}\mathbf{G})^{-1}\,\mathbf{R}^{-1}\mathbf{y}_d$$

where $\mathbf{x}_b = M(\mathbf{x}, \boldsymbol{\eta} = \mathbf{0})$ is from the baseline simulation and it is assumend that there is no prior information about control parameters, i.e., $\mathbf{Q}^{-1} \approx \mathbf{0}$.

If linearization assumption holds, we will have: $\mathbf{x}_a \approx M(\mathbf{x}, \boldsymbol{\eta}_a)$.

State estimation: Formally combining the two knowledge reservoirs (an early vision, ca. 1982)



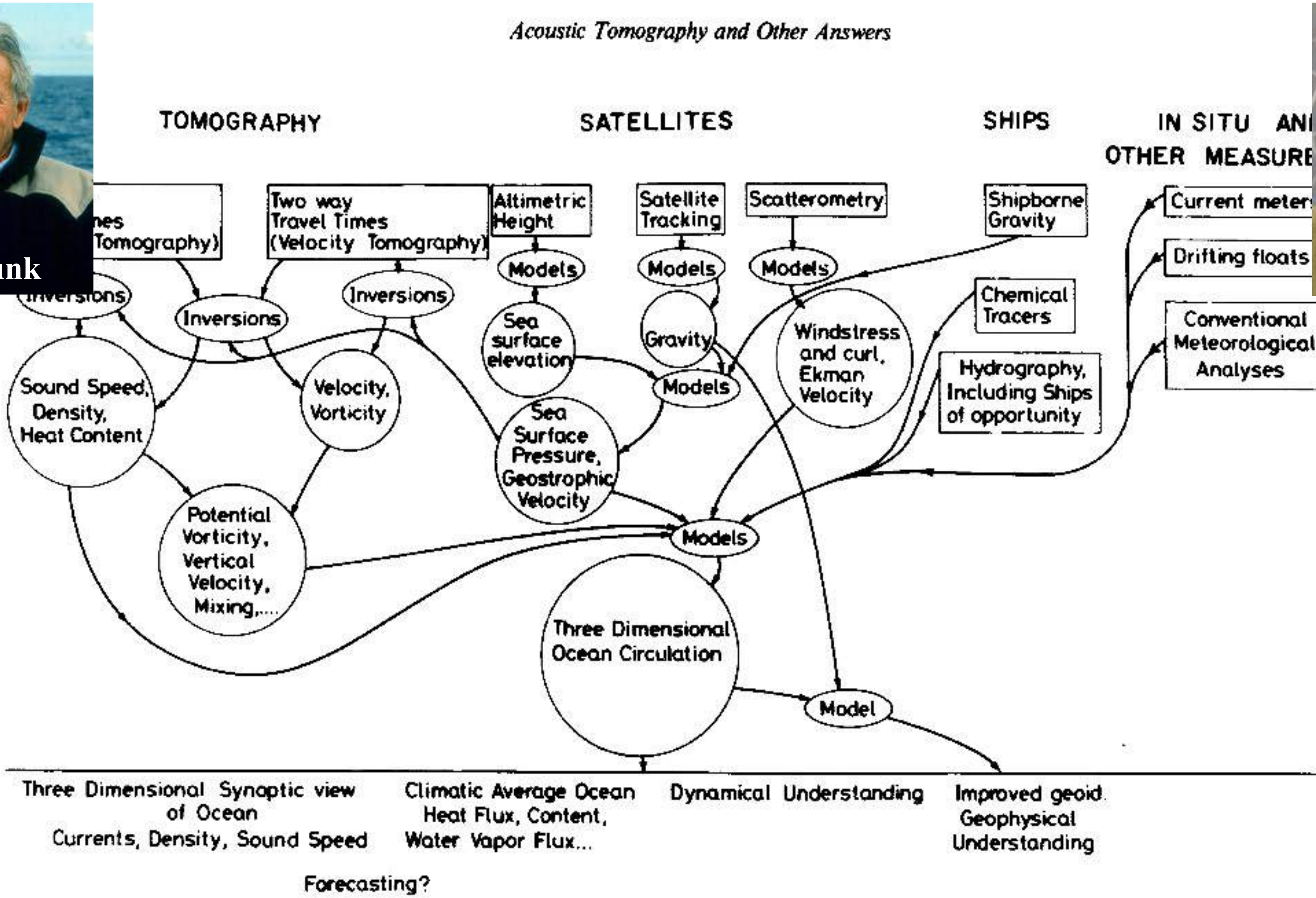


Figure 26. All measurements and models of the ocean can be interconnected to provide global estimates of the state of the three-dimensional ocean. Some side benefits accrue — e.g. improved estimates of the earth's gravity field.

Taken from: ***C. Wunsch***, in *"A Celebration in Geophysics and Oceanography 1982. In Honor of Walter Munk on his 65th birthday."*

Example application:

Large-Scale Circulation of the Pacific Ocean from Satellite Altimetry (Stammer and Wunsch, 1996)

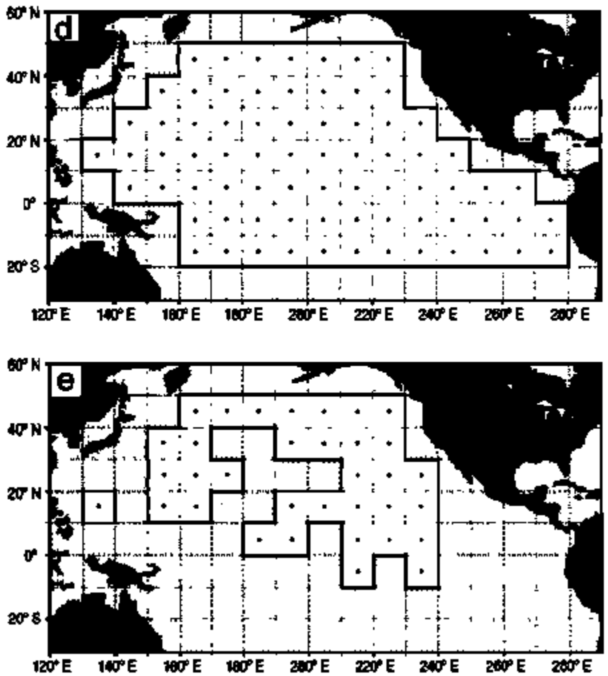


Figure 1d,e. (d) Coarse 10° by 10° grid, on which the Green's functions where calculated in layers 1 through 3. (e) Owing to geographical shoaling, the bottom layer has a smaller geographical extent on the coarse grid.

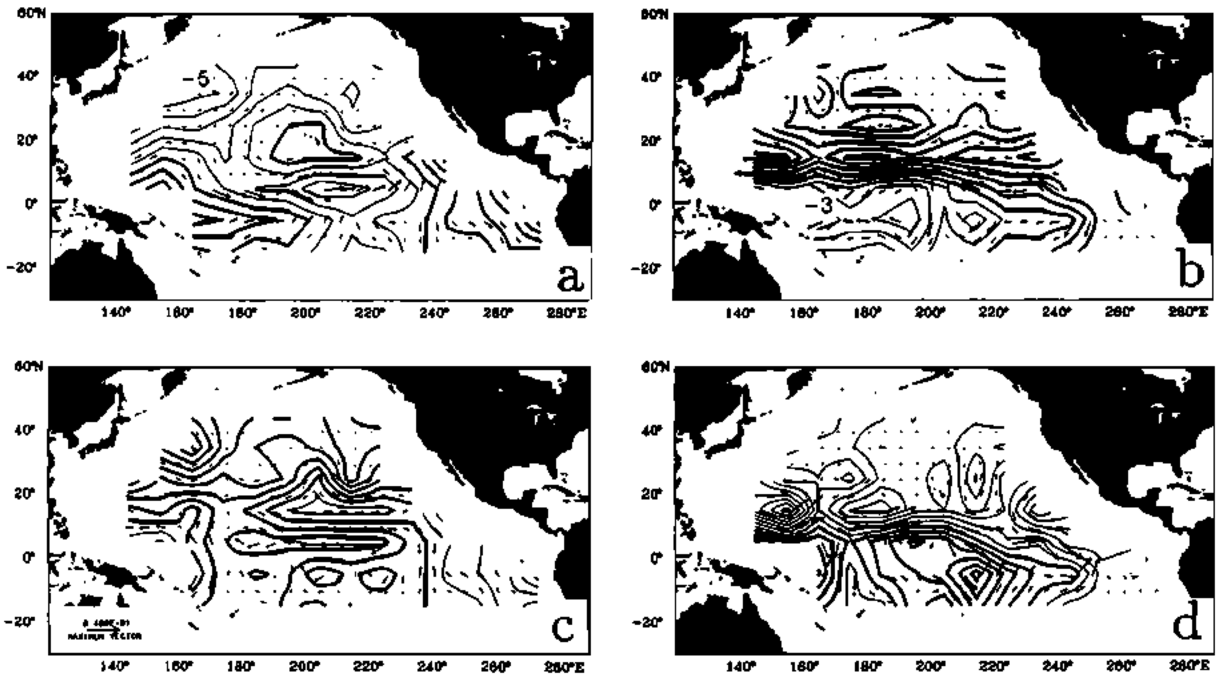


Figure 23. Estimates of seasonal surface elevation anomalies relative to the 1-year mean and related geostrophic currents. Fields represent (a) spring, (b), summer, (c) fall, and (d) winter, with spring starting at the beginning of March. Positive and negative values are drawn by bold, and thin lines, respectively. Contour increment is 1 cm. The reference vector represent 4 cm/s.

Example application:

Linearization of an Oceanic General Circulation Model for Data Assimilation and Climate Studies (Menemenlis and Wunsch, 1997)

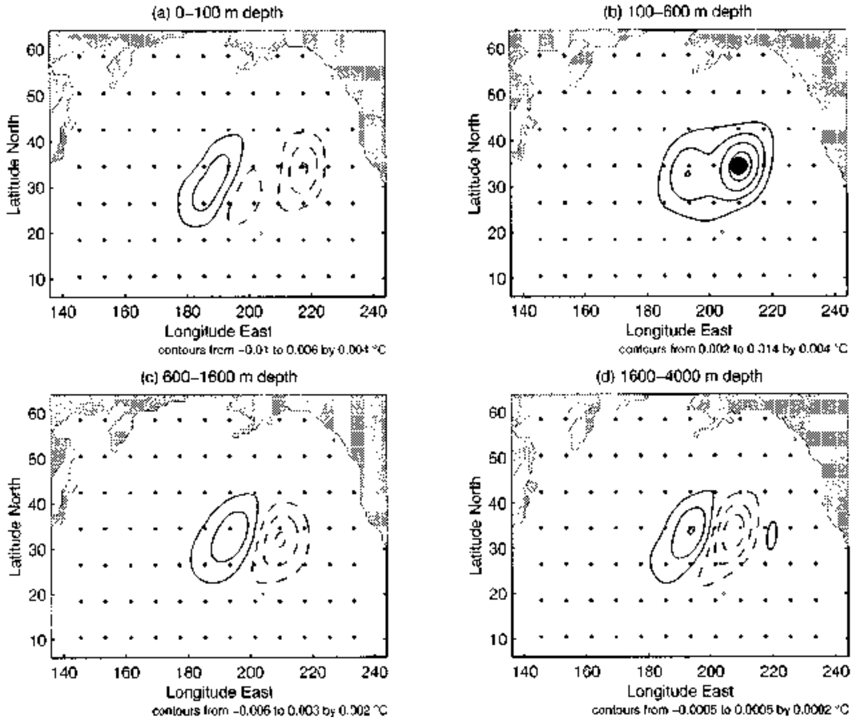


FIG. 9. Response of the four-level GFDL model to a 0.05°C perturbation, between 100- and 600-m depth, at the end of month 16. A two-dimensional low-pass spatial filter with cutoff wavelength of 16° has been applied to smooth scales not resolved by the reduced-order linear model. The heavy dot indicates the initial location of the disturbance.

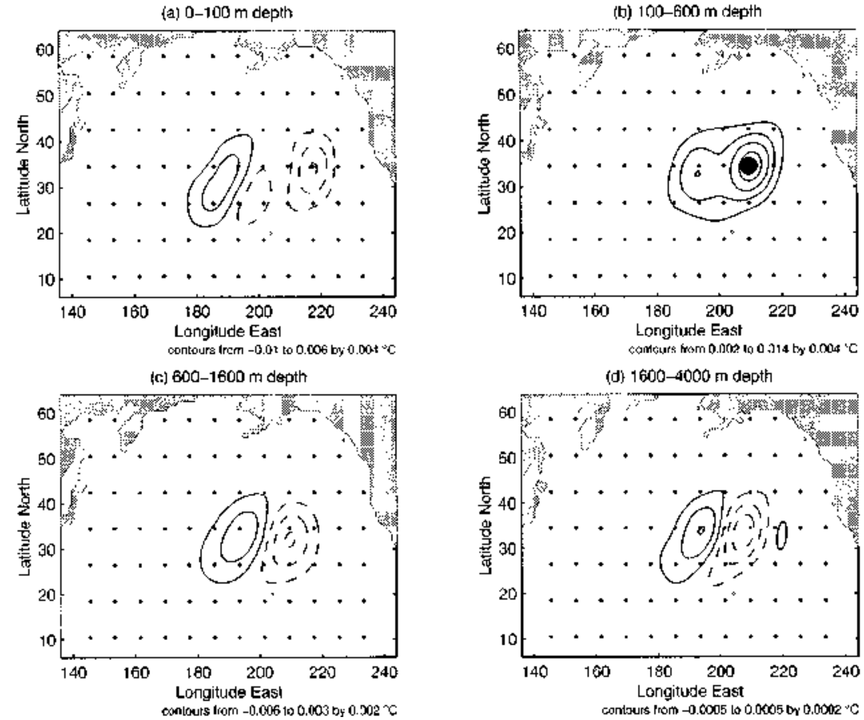


FIG. 9. Response of the four-level GFDL model to a 0.05°C perturbation, between 100- and 600-m depth, at the end of month 16. A two-dimensional low-pass spatial filter with cutoff wavelength of 16° has been applied to smooth scales not resolved by the reduced-order linear model. The heavy dot indicates the initial location of the disturbance.

Example application:

Linearization of an Oceanic General Circulation Model for Data Assimilation and Climate Studies (Menemenlis and Wunsch, 1997)

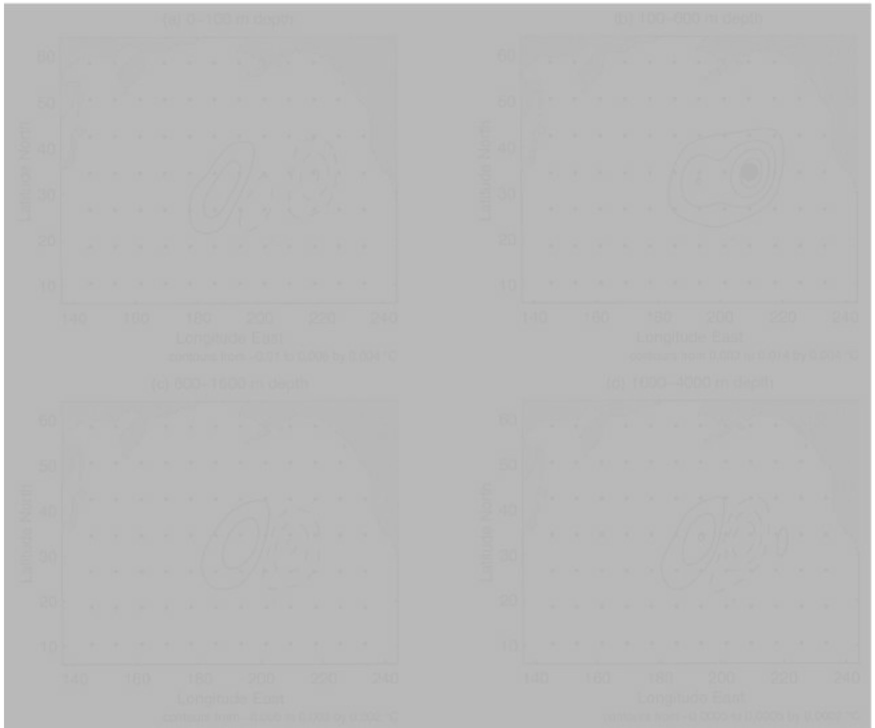


FIG. 9. Response of the four-level GFDL model to a 0.05°C perturbation, between 100- and 600-m depth, at the end of month 16. A two-dimensional low pass spatial filter with cutoff wavelength of 16° has been applied to smooth scales not resolved by the reduced-order linear model. The heavy dot indicates the initial location of the disturbance.

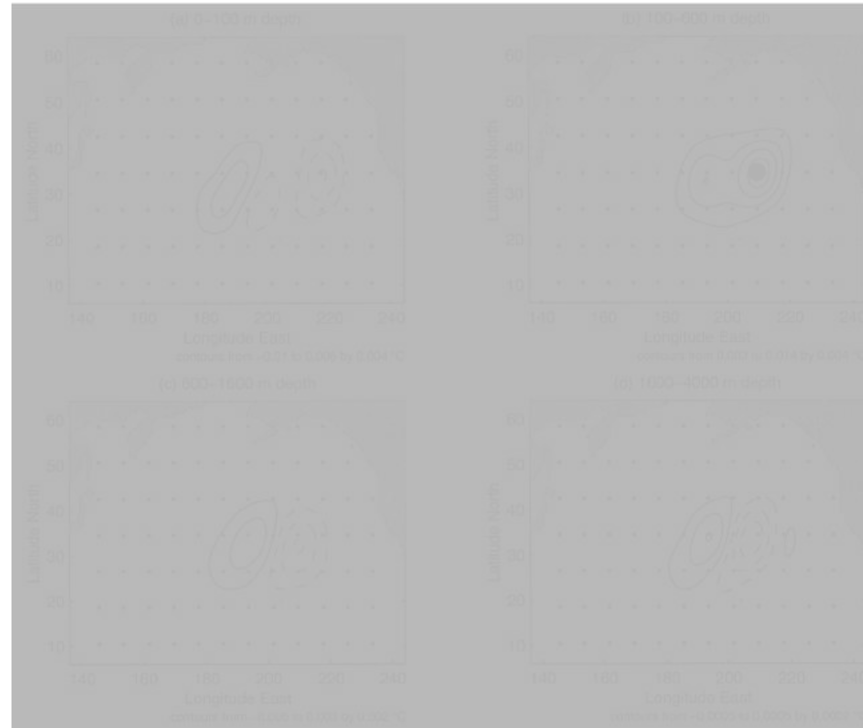


FIG. 9. Response of the four-level GFDL model to a 0.05°C perturbation, between 100- and 600-m depth, at the end of month 16. A two-dimensional low pass spatial filter with cutoff wavelength of 16° has been applied to smooth scales not resolved by the reduced-order linear model. The heavy dot indicates the initial location of the disturbance.

3. Model description

The current study was initiated using the GFDL numerical code and model output from a global eddy-resolving integration by Semtner and Chervin (1992). These results are reported in sections 6 and 7. We have now switched over to the newly developed MIT GCM. This model is used to carry out the perturbation analysis reported in section 4 and will be the focus of our future assimilation efforts. The above models and their configurations are briefly described below.

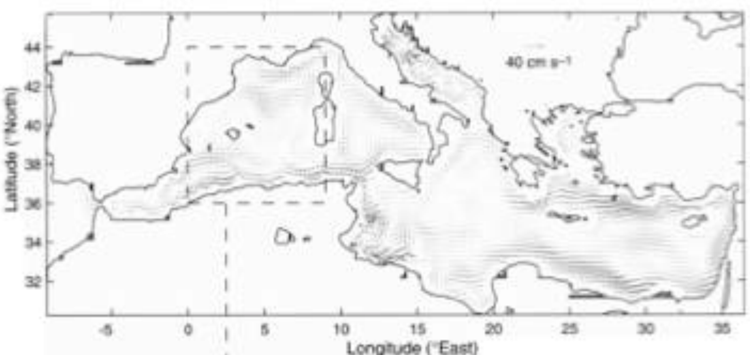
a. MIT model

In its current configuration, the MIT GCM (Marshall et al. 1997a,b) solves the incompressible Navier–Stokes

!!!

Example application:

Basin-Scale Ocean Circulation from Combined Altimetric, Tomographic and Model Data (Menemenlis et al., 1997)



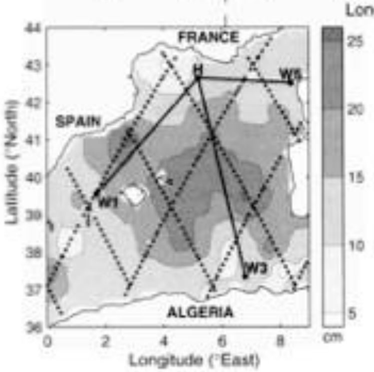


Figure 1 Top panel, the integration domain and the 40-m horizontal velocity on 1 March, 1994 from the Marshall et al.[8,9] general circulation model (GCM). Bottom panel, the estimation domain and (contours) the sea level in September 1994 minus the sea level in March 1994, as measured by the TOPEX/POSEIDON altimeter. Dotted lines are altimeter tracks and solid lines are tomographic sections joining an acoustic source H, and acoustic receivers W1, W3, W5 deployed during the THETIS-2 experiment. Mapping of the altimeter data is from a multiresolution optimal interpolation algorithm[7]. Estimates have uncertainties varying from ±2.3 cm near the altimeter ground tracks to a maximum of ±4.5 cm between tracks. The mean sea surface height increase is indicative of seasonal warming, but cannot be used to measure absolute heating because of mass fluxes in and out of the basin (flows through the Straits of Sicily and Gibraltar[1], evaporation minus precipitation, and river runoff). Relative changes, for example, the band of higher elevation centred at 40° N, contain contributions from changes in wind forcing and from differential heating and cooling of the water column. The GCM and the acoustic-tomography data are used to separate the respective contributions to sea-level anomaly of wind and heat content.

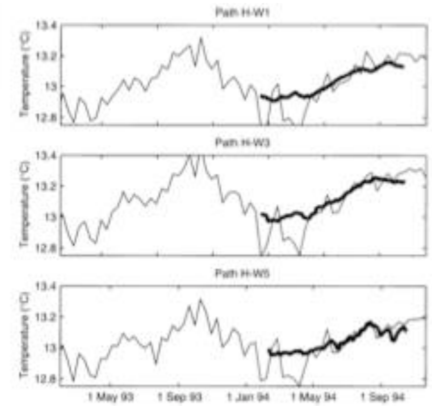


Figure 2 Depth-averaged (0–2,000 m) and path-averaged potential temperature measured using ocean acoustic tomography (thick lines) and estimated from the altimeter data (thin lines) along the three sections marked on Fig. 1. Tomographic heat content estimates are from inversions using a single deep-diving acoustic ray. The estimated uncertainty is ±0.013 °C. Section H–W3 was sampled on a regular basis with expendable bathythermographs during the duration of the experiment and the agreement with the inversions is to within measurement uncertainty. The data show a strong seasonal heating which is mainly confined to the top 50 m and is consistent with ECMWF heat flux analyses over the western Mediterranean[1]. Altimeter heat content estimates are made under the assumption that the sea-level anomaly signal results from uniform heating or cooling of the top 50 m. Differences between the tomographic and the altimetric heat content estimates result from physical processes other than heating or cooling of the surface layer, for example, the response of the sea-level anomaly to changes in atmospheric pressure, wind stress, and freshwater fluxes. A key contribution of the tomographic data is the ability to estimate the heat content part of the altimetric sea-level anomaly signal.

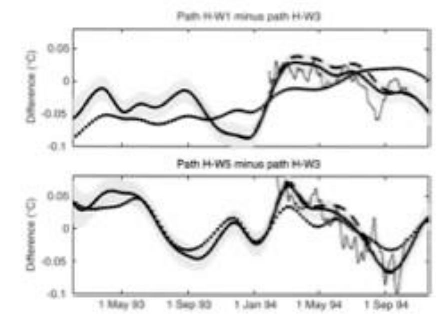


Figure 3 Difference between depth-averaged (0–2,000 m) and path-averaged potential temperature along the tomographic sections. The dotted line is the GCM estimate; the dashed line indicates the GCM/altimeter combination; the thick solid line represents the GCM/altimeter/tomography combination with the shaded area indicating the standard error of this last estimate; the thin solid line is the tomographic data; and circles represent altimeter data converted to depth-integrated temperature as in Fig. 2. GCM and data biases have been removed as discussed in the text. The GCM predictions differ significantly from the tomographic data because of the lack of realistic buoyancy forcing at the surface. However, the GCM/altimeter combination successfully recovers the time-evolving spatial anomaly of heat content, as seen through the calibration against tomographic data. The addition of the tomographic data does not change the estimates significantly, but the uncertainty is reduced by 17%, from ±0.018 °C when there are no tomographic data to ±0.015 °C during the period when tomographic data are available. Remaining differences between the estimates and the tomographic data are believed real—representing the still-inadequate resolution of the GCM and of the estimation method.

Example application:

Ocean Climate Change: Comparison of Acoustic Tomography, Satellite Altimetry and Modeling (ATOC Consortium, 1998)

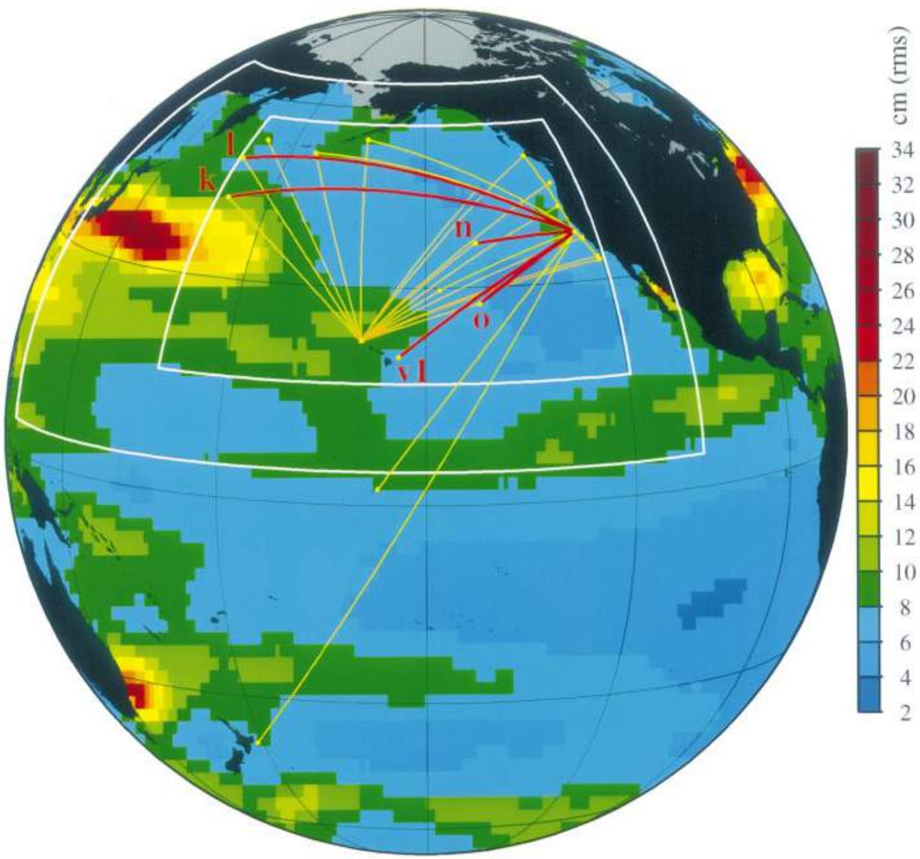


Fig. 1. The ATOC acoustic array is superimposed on a map of the root-mean-square (rms) sea level anomaly from 4 years (January 1993 to December 1996) of TOPEX/POSEIDON altimetric measurements. Red lines indicate the sections used in the present study and are referenced by letter labels. Yellow lines show additional sections along which the acoustic propagation has been observed, but for which the data were not used here. Data assimilation was carried out in the region bounded by the outer white rectangle, and heat content estimates were obtained inside the inner white rectangle. Much, but not all, of the elevation anomalies represent seasonal thermal changes within the ocean, with the acoustic data providing a stable spatial average that is otherwise difficult to obtain. The ATOC region, being on the eastern side of the ocean, shows comparatively weak variability. Nevertheless, it is evident that the different acoustic sections will, during any 10-day period, have potentially very different anomalies.

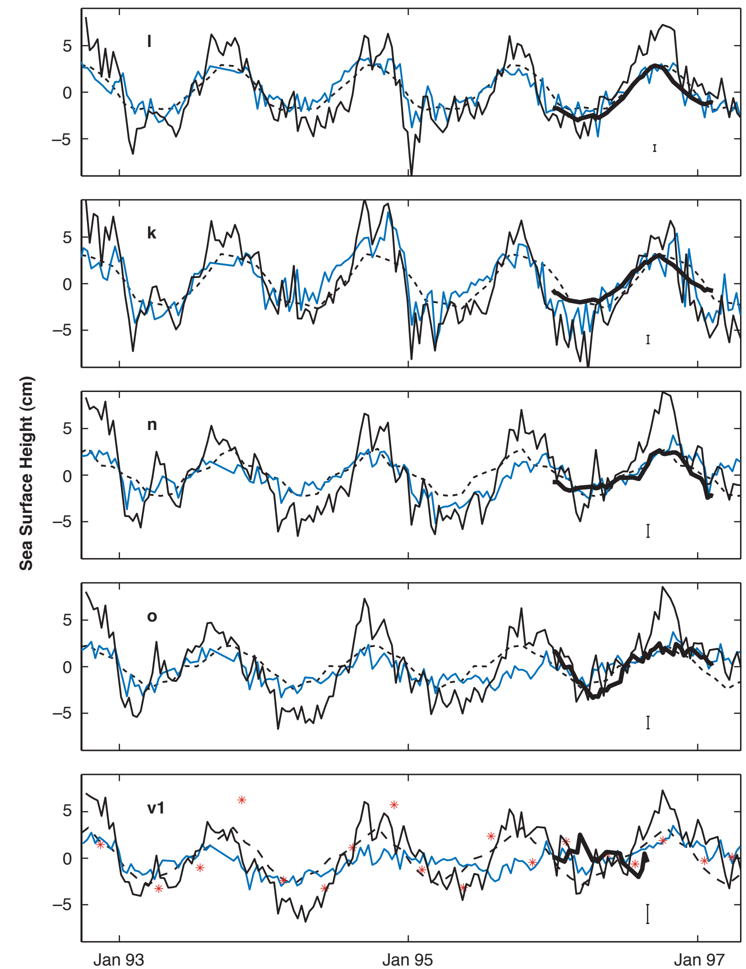


Fig. 3. The range-averaged sea level anomaly along the acoustic sections inferred by several independent methods: (i) thick black lines indicate the ATOC acoustic measurements converted to equivalent sea surface height for comparison with the altimeter data, (ii) thin black lines are from the TOPEX/POSEIDON altimeter data, (iii) dashed lines represent the climatological thermal anomaly converted to sea surface height, (iv) blue lines are the GCM estimates, and (v) the asterisks along section v1 are the XBT data. Uncertainties are indicated for the acoustic estimates: the possible errors are largest along section v1 because the upper ocean variability is unresolved due to a lack of surface-reflecting rays near the receiver.

Example application:

Using Green's Functions to Calibrate an Ocean General Circulation Model (Menemenlis et al., 2005)

TABLE 4. List of sensitivity experiments and optimized parameters for the second Green's function optimization. For experiment 6, the optimized parameter is indicated as a factor multiplying the $\partial Q/\partial T$ fields of Barnier et al. (1995).

Expt	Parameter	Baseline	Optimized
1	Vertical diffusivity (10^{-6} m^2 s^{-2})	5	15.1 ± 12
2	Vertical viscosity (10^{-6} m^2 s^{-2})	100	17.7 ± 3.0
3	Ri_c, boundary layer depth	0.300	0.354 ± 0.004
4	Ri_0, shear instability	0.700	0.699 ± 0.008
5	Salinity relaxation (days)	60	44.5 ± 1.2
6	Temperature relaxation ($\partial Q/\partial T$)	1.000	1.630 ± .008
7–10	Isopycnal diffusivity (m^2 s^{-2})	500	Linear combination
11–14	Surface wind stress	NCEP/COADS	Linear combination
15–20	Initial conditions	SPINUP	Linear combination

- 43% decrease in cost function
- significant reduction in model bias and drift
- 10–30% increase in explained variance

Example application:

Using Green's Functions to Calibrate an Ocean General Circulation Model (Menemenlis et al., 2005)

TABLE 2. Optimized parameters for case 3 (Table 1) are compared to parameters estimated one at a time. The last row displays the cost function reduction in percent assuming that the problem is linear. Because the parameter estimates are linearly dependent, the one-at-a-time estimates differ substantially from those of case 3.

Parameter	Case 3	Case 4	Case 5	Case 6	Case 7	Case 8	Case 9
Vertical diffusivity (10^{-6} m^2 s^{-2})	15.4	17.4	—	—	—	—	—
Vertical viscosity (10^{-6} m^2 s^{-2})	46	—	348	—	—	—	—
Isopycnal diffusivity (m^2 s^{-2})	572	—	—	399	—	—	—
Time-mean wind stress	0.43	—	—	—	0.72	—	—
Initial temperature	0.11	—	—	—	—	0.60	—
Initial temperature and salt	0.71	—	—	—	—	—	2.5
Cost function reduction (%)	29.8	19.4	0.58	0.14	5.42	6.46	14.2

Example application:

Using Green's Functions to Calibrate an Ocean General Circulation Model (Menemenlis et al., 2005)

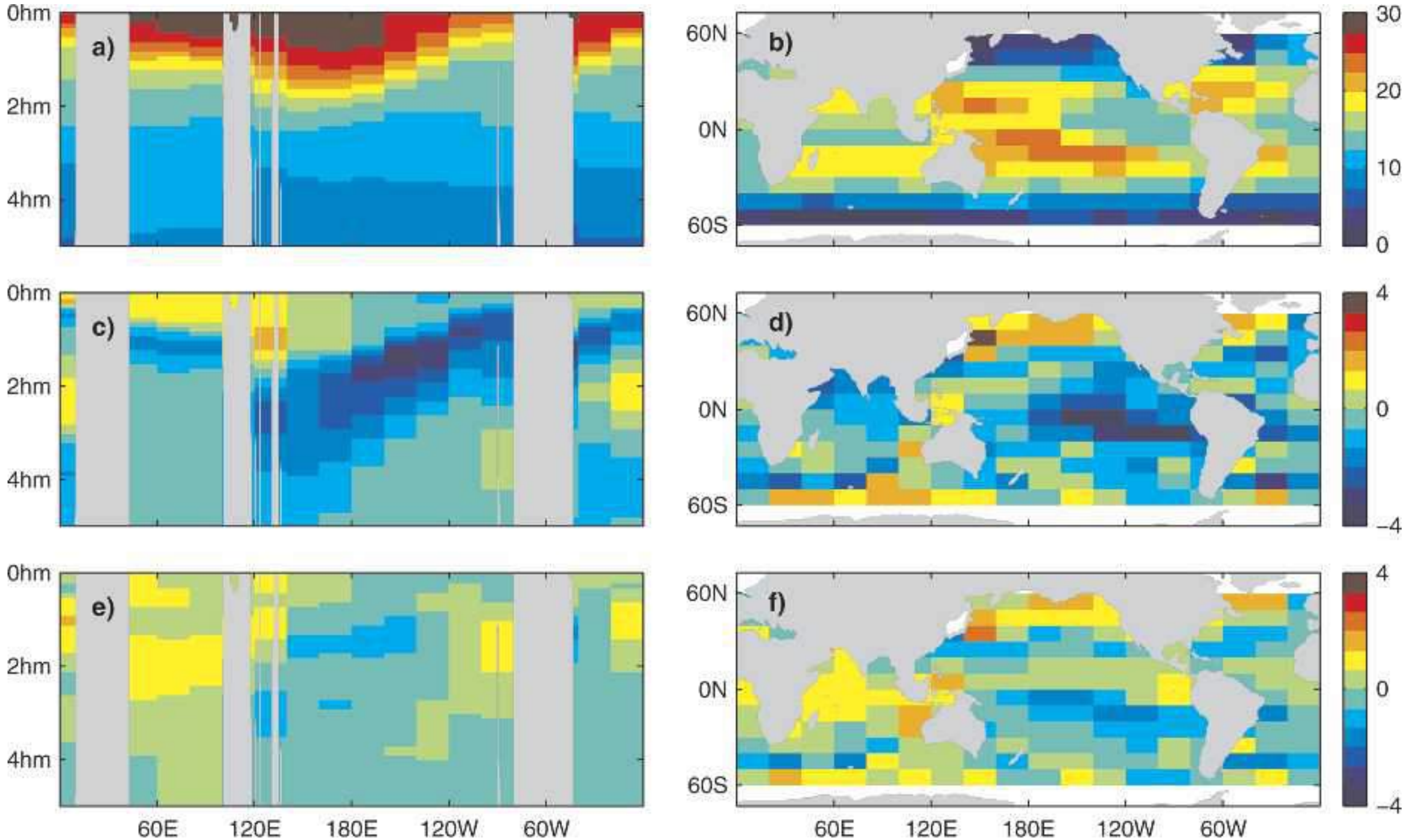


FIG. 13. Time-mean potential temperature, 1993–2000: (a) Green's function estimate at the equator down to 500-m depth; (b) Green's function estimate at the 156-m depth; (c) smoother bias relative to data at the equator; (d) smoother bias relative to data at the 156-m depth; (e) Green's function bias relative to data at the equator; and (f) Green's function bias relative to data at the 156-m depth. Units are °C.

Example application:

Using Green's Functions to Calibrate an Ocean General Circulation Model (Menemenlis et al., 2005)

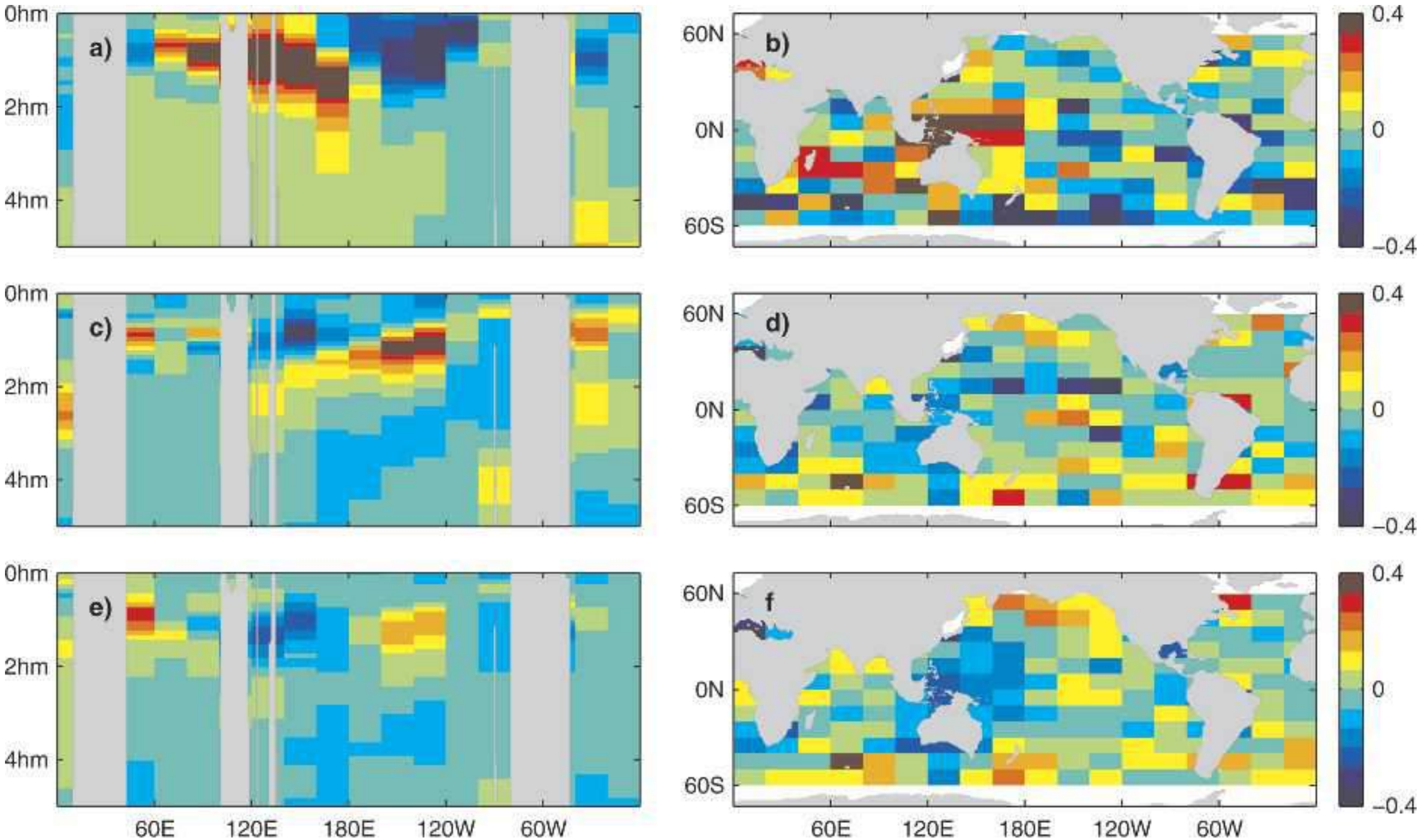


FIG. 14. Potential temperature trend, 1993–2000: (a) Green's function estimate at the equator down to 500-m depth; (b) Green's function estimate at the 156-m depth; (c) smoother drift relative to data at the equator; (d) smoother drift relative to data at the 156-m depth; (e) Green's function drift relative to data at the equator; and (f) Green's function drift relative to data at the 156-m depth. Units are °C yr^{-1}.

Example application:

Using Green's Functions to Calibrate an Ocean General Circulation Model (Menemenlis et al., 2005)

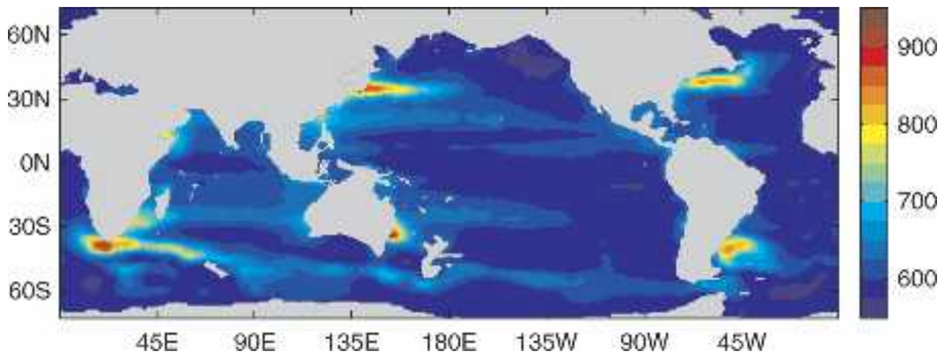


FIG. 8. Estimated isopycnal diffusivity in $m^2\ s^{-1}$ at the 1000-m depth.

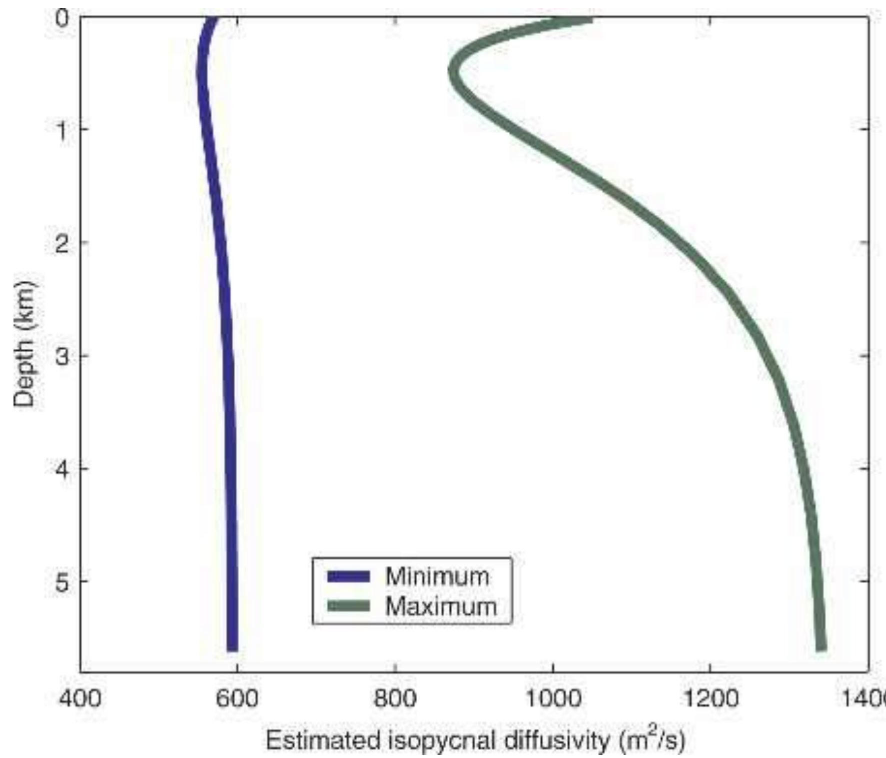


FIG. 9. Vertical profile of estimated isopycnal diffusivity.

Example application:

Ocean Carbon-cycle Model Intercomparison Project 3 (OCMIP-3)
(Mikaloff Fletcher et al. 2006, 2007; Gruber et al. 2009)

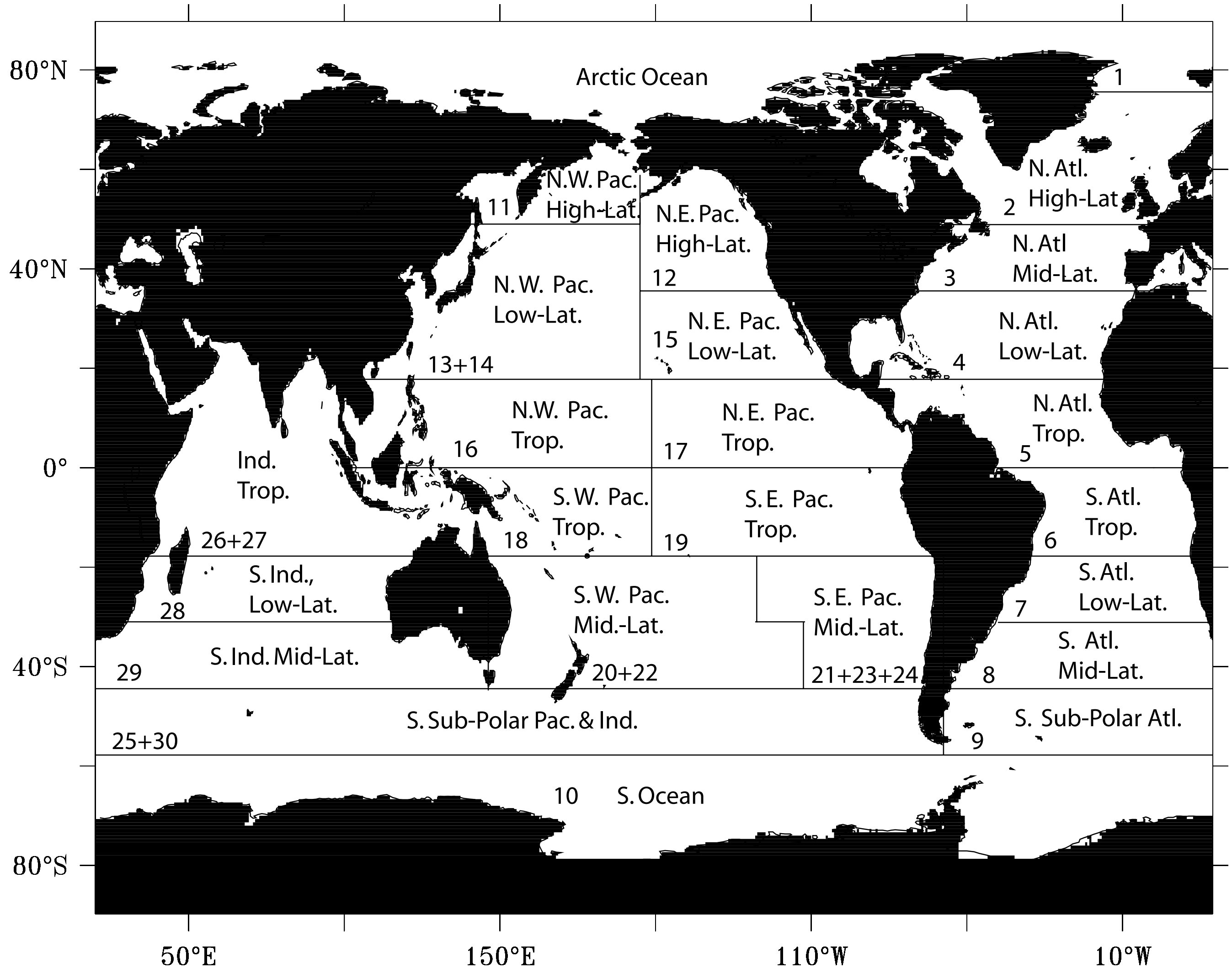


Figure 2. The 24 regions used for the ocean inversion. The region numbers show the aggregation from the original 30 regions [*Mikaloff Fletcher et al.*, 2003] to the 24 regions used in this study.

Example application:

Ocean Carbon-cycle Model Intercomparison Project 3 (OCMIP-3) (Mikaloff Fletcher et al. 2006, 2007; Gruber et al. 2009)

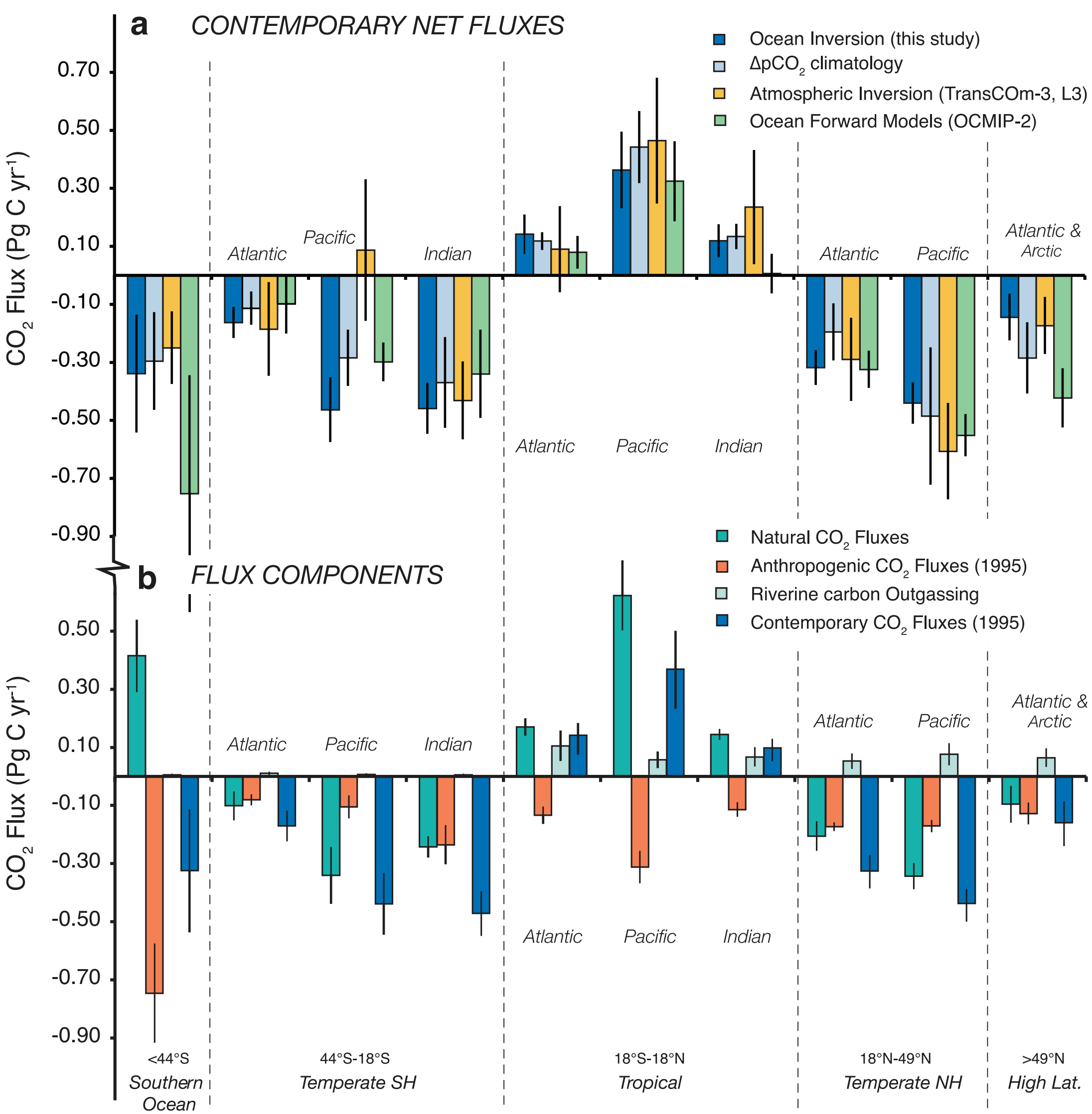


Figure 1. Air-sea CO2 fluxes for 10 regions, ordered by latitude and Ocean basin (positive: outgassing; negative: uptake). (a) Comparison of contemporary air-sea fluxes of CO2. Shown are the ocean inversion estimates (this study), the new pCO2-based estimates of Takahashi et al. [2008], the mean estimates based on results from the 13 ocean biogeochemistry models that participated in the second phase of the Ocean Carbon-cycle Model Intercomparison Project (OCMIP-2) [Watson and Orr, 2003], and the mean estimates from the TransCom-3 project based on the interannual (level 3) inversions of atmospheric CO2 [Baker et al., 2006]. The uncertainties for the OCMIP-2 estimates reflect the (unweighted) standard deviation across the 13 models, while the uncertainties for the TransCom estimates were obtained by quadrature of the within and between model errors reported by Baker et al. [2006]. (b) Weighted mean estimates of the natural, anthropogenic, river-induced, and contemporary air-sea fluxes of CO2 based on our ocean inversion [Mikaloff Fletcher et al., 2006, 2007]. The results are aggregated to 10 regions from the 23 regions solved for in the inversion for reasons of clarity. Error bars denote the cross-model weighted standard deviation of the mean. The anthropogenic and contemporary CO2 fluxes are for a nominal year of 1995.

Example application:

Ocean Carbon-cycle Model Intercomparison Project 3 (OCMIP-3)

(Mikaloff Fletcher et al. 2006, 2007; Gruber et al. 2009)

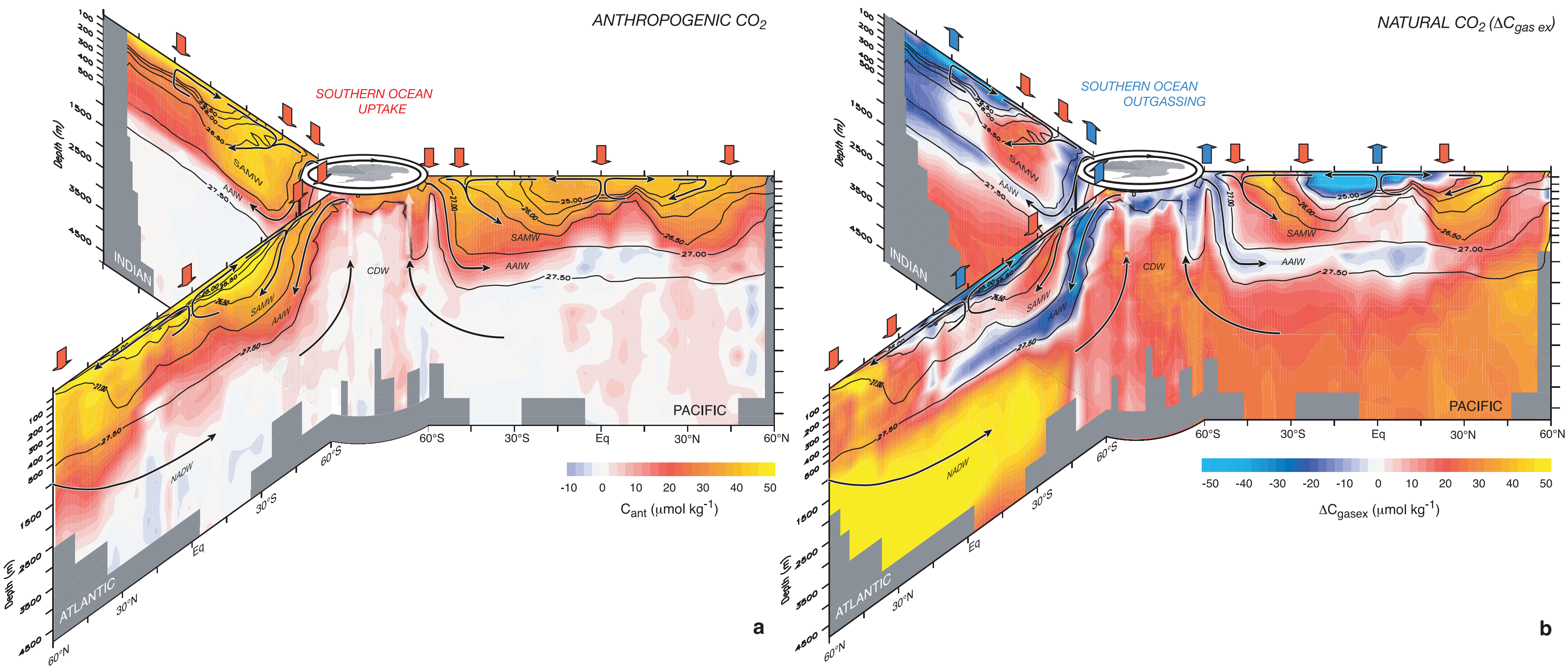


Figure 5. Ocean interior distributions of the tracers reflecting the exchange of CO_2 across the air-sea interface, displayed as global-scale section plots organized around the Southern Ocean in the center. (a) Distribution of anthropogenic CO_2, C_{ant}, estimated using the ΔC^* method of *Gruber et al.* [1996]. (b) Distribution of the gas exchange component of natural CO_2, $\Delta C_{\mathrm{gas\ ex}}$, following *Gruber and Sarmiento* [2002]. The inversion interprets these distributions by determining, given ocean circulation and mixing, a set of surface ocean fluxes that most closely matches these observations. Also shown are isolines of potential density anomalies, σ_θ (density referenced to the ocean surface minus 1000 kg m^{-3}), along which most of the oceanic flow occurs. Major ocean circulation features are indicated by schematic arrows. Figure 5 is based on data taken from GLODAP [*Key et al.*, 2004]. NADW: North Atlantic Deep Water, CDW: Circumpolar Deep Water; SAMW: Subantarctic Mode Water; AAIW: Antarctic Intermediate Water.

Example application:

Ocean Carbon-cycle Model Intercomparison Project 3 (OCMIP-3) (Mikaloff Fletcher et al. 2006, 2007; Gruber et al. 2009)

Table 1. Evaluation of Model Skill Based on Comparisons Between CFC-11 Model Simulations and the GLODAP Gridded CFC Data Set[a]

	Correlation	Normalized Std. Dev.[b]	Model Skill[c]	Inverse Anthropogenic CO_2 Uptake, Pg C yr^{-1}	Forward Anthropogenic CO_2 Uptake, Pg C yr^{-1}
BERN	0.89	1.04	0.81	2.05	N.A.
ECCO	0.96	0.89	0.91	2.01	N.A.
MIT	0.91	1.00	0.85	2.22	N.A.
NCAR	0.95	0.98	0.91	2.18	2.36
PRINCE-LL	0.90	1.18	0.80	1.85	1.90
PRINCE-HH	0.93	1.05	0.87	2.33	2.43
PRINCE-LHS	0.93	1.04	0.86	1.99	2.04
PRINCE-2	0.93	1.03	0.87	2.17	2.24
PRINCE-2a	0.91	1.05	0.85	2.25	2.35
UL	0.87	1.0	0.77	2.81	2.95
Mean	0.92	1.02	0.85	2.18	2.32

[a]Also tabulated are forward and inverse estimates of the global total anthropogenic CO_2 uptake (Pg C yr^{-1}, scaled to 1995). Forward results are from OCMIP-2 [*Dutay et al.*, 2002; *Watson and Orr*, 2003].

[b]Normalized Std. Dev. is defined as the standard deviation of the modeled field divided by the corresponding standard deviation of the observed field.

[c]Following *Taylor* [2001].

Tracer Green's Functions from old 2-deg ECCO solution was among solutions with highest correlation, lowest standard error, and highest model skill relative to CFC-11 observations!

References

George Green

Challis, Lawrie, and Fred Sheard. 2003. “The Green of Green Functions.” *Physics Today* 56 (12): 41–46.

Atmospheric tracer inversions

Enting, I. G., and J. V. Mansbridge. 1989. “Seasonal Sources and Sinks of Atmospheric CO 2 Direct Inversion of Filtered Data.” *Tellus B* 41B (2): 111–26.

Tans, Pieter P., Inez Y. Fung, and Taro Takahashi. 1990. “Observational Constraints on the Global Atmospheric CO2 Budget.” *Science*.

Bousquet, P. 2000. “Regional Changes in Carbon Dioxide Fluxes of Land and Oceans Since 1980.” *Science* 290 (5495): 1342–46.

Ocean circulation estimates

Stammer, Detlef, and Carl I. Wunsch. 1996. “The Determination of the Large-Scale Circulation of the Pacific Ocean from Satellite Altimetry Using Model Green ’ s Functions.” *J. Geophys. Res.* 101 (1): 18,409-418,432.

Menemenlis, Dimitris, and Carl I. Wunsch. 1997. “Linearization of an Oceanic General Circulation Model for Data Assimilation and Climate Studies.” *Journal of Atmospheric and Oceanic Technology* 14 (6): 1420–43.

Menemenlis, Dimitris, Tony Webb, Carl I. Wunsch, Uwe Send, and Christopher N. Hill. 1997. “Basin-Scale Ocean Circulation from Combined Altimetric, Tomographic and Model Data.” *Nature*.

ATOC Consortium. 1998. “Ocean Climate Change: Comparison of Acoustic Tomography, Satellite Altimetry and Modeling.” *Science (New York, N.Y.)* 281 (5381): 1327–32.

Menemenlis, Dimitris, Ichiro Fukumori, and Tong Lee. 2005. “Using Green’s Functions to Calibrate an Ocean General Circulation Model.” *Monthly Weather Review* 133 (5): 1224–40.

Nguyen, An T., Dimitris Menemenlis, and Ronald Kwok. 2011. “Arctic Ice-Ocean Simulation with Optimized Model Parameters: Approach and Assessment.” *Journal of Geophysical Research: Oceans* 116 (4).

Ocean carbon inversions

Gloor, Manuel, Nicolas Gruber, J. Sarmiento, Christopher L. Sabine, Richard A. Feely, and Christian Rödenbeck. 2003. “A First Estimate of Present and Preindustrial Air-Sea CO2 Flux Patterns Based on Ocean Interior Carbon Measurements and Models.” *Geophysical Research Letters* 30 (1): 101–104.

Mikaloff Fletcher, Sara E., Nicolas Gruber, Andrew R. Jacobson, Scott C. Doney, Stephanie Dutkiewicz, Markus Gerber, M. Follows, et al. 2006. “Inverse Estimates of Anthropogenic CO2 Uptake, Transport, and Storage by the Ocean.” *Global Biogeochemical Cycles* 20 (2).

Mikaloff Fletcher, Sara E., Nicolas Gruber, Andrew R. Jacobson, Manuel Gloor, Scott C. Doney, Stephanie Dutkiewicz, Markus Gerber, et al. 2007. “Inverse Estimates of the Oceanic Sources and Sinks of Natural CO2 and the Implied Oceanic Carbon Transport.” *Global Biogeochemical Cycles* 21 (1): 1–19.

Gruber, Nicolas, Manuel Gloor, Sara E. Mikaloff Fletcher, Scott C. Doney, Stephanie Dutkiewicz, Michael J. Follows, Markus Gerber, et al. 2009. “Oceanic Sources, Sinks, and Transport of Atmospheric CO 2.” *Global Biogeochemical Cycles* 23 (1): GB1005.

Brix, Holger, Dimitris Menemenlis, Christopher N. Hill, Stephanie Dutkiewicz, Oliver Jahn, D. Wang, Kevin W. Bowman, and H. Zhang. 2015. “Using Green’s Functions to Initialize and Adjust a Global, Eddying Ocean Biogeochemistry General Circulation Model.” *Ocean Modelling* 95: 1–14.

Joint ocean-atmosphere carbon dioxide inversions

Jacobson, Andrew R., Sara E Mikaloff Fletcher, Nicolas Gruber, Jorge L. Sarmiento, and Manuel Gloor. 2007. “A Joint Atmosphere-Ocean Inversion for Surface Fluxes of Carbon Dioxide: 1. Methods and Global-Scale Fluxes.” *Global Biogeochemical Cycles* 21 (1).

Jacobson, Andrew R., Sara E Mikaloff Fletcher, Nicolas Gruber, Jorge L. Sarmiento, and Manuel Gloor. 2007. “A Joint Atmosphere-Ocean Inversion for Surface Fluxes of Carbon Dioxide: 2. Regional Results.” *Global Biogeochemical Cycles* 21 (1): 1–15.

Comparison with representer method

The representer method (see Andrew Bennett's books and publications) was developed for data-sparse inverse modeling problems.

Both the Green's Functions and representer approaches provide a reduced orthogonal basis sets for inversions. The two methods are mirror images of each other.

The representer method should be used when the number of available observations is small. The optimized solution is projected on the "observable" parameter space.

The Green's Functions approach should be used when the number of control parameters is small. The optimized solution is projected on the "controllable" parameter space.

Comparison with adjoint method

The Green's function approach has been called a poor-man's adjoint.

Advantages relative to the adjoint method are simplicity of implementation, the possibility of offline experimentation with different cost functions, improved robustness in the presence of nonlinearities, and complete a posteriori error statistics for the parameters being estimated.

The major drawback of the Green's function approach is that computational cost increases linearly with the number of control parameters. By comparison, the cost of the adjoint method, while substantial, is largely independent from the number of control parameters.

Summary and concluding remarks

Green's functions provide a simple yet effective method to test and to calibrate general circulation model parameterizations, to study and to quantify model and data errors, to correct model biases and trends, and to blend estimates from different solutions and data products.

They can be applied to pretty much any general circulation model since all that is required is forward-model sensitivity experiments.

They are a better way to adjust uncertain model parameterizations than ad-hoc or one-at-a time parameter adjustments.

In the absence of adjoint model, or for strongly nonlinear systems, they can be used for preliminary model adjustments.

Model Green's Functions cheat sheet

Least squares method based on computation of model Green's functions.

Used for, e.g., atmospheric tracer inversions (Enting and Mansbridge, 1989; Tans et al., 1990; Bousquet et al., 2000), ocean circulation estimates (Stammer and Wunsch, 1996; Menemenlis et al., 1997a, b; ATOC 1998, 2005; Nguyen et al., 2011), ocean carbon inversions (Gloor et al., 2003; Mikaloff Fletcher et al., 2006; 2007; Gruber et al., 2009; Brix et al., 2015), and joint ocean-atmosphere carbon dioxide inversions (Jacobson et al., 2007a; 2007b).

GCM:

$$\mathbf{x}(t_{i+1}) = M(\mathbf{x}(t_i), \boldsymbol{\eta})$$

$\mathbf{x}(t_i)$ is the ocean model state vector at time t_i
M represents the numerical model
$\boldsymbol{\eta}$ is a set of control parameters.

Data:

$$\mathbf{y} = H(\mathbf{x}) + \boldsymbol{\varepsilon} = G(\boldsymbol{\eta}) + \boldsymbol{\varepsilon}$$

$\mathbf{y}$ is the available observations
H is the measurement model
G is a function of M and H
$\boldsymbol{\varepsilon}$ is additive noise

Cost function:

$$J = \boldsymbol{\varepsilon}^{\mathrm{T}} \mathbf{R}^{-1} \boldsymbol{\varepsilon}$$

J is quadratic cost function
$\mathbf{R}$ is estimate of covariance matrix of $\boldsymbol{\varepsilon}$

Linearization:

$$G(\boldsymbol{\eta}) \approx G(\mathbf{0}) + \mathbf{G}\boldsymbol{\eta}$$

$\mathbf{G}$ is a kernel matrix whose columns are computed using a GCM sensitivity experiment for each parameter in vector $\boldsymbol{\eta}$.
$G(\mathbf{0})$ is from baseline GCM integration.

Solution:

$$\mathbf{x}_a = \mathbf{x}_b + (\mathbf{G}^{\mathrm{T}} \mathbf{R}^{-1} \mathbf{G})^{-1} \mathbf{R}^{-1} (\mathbf{y} - G(\mathbf{0}))$$

$\mathbf{X}_a$ is optimized solution that minimizes cost function J.
$\mathbf{X}_b$ is the solution of the baseline simulation