

# Super-parameterization in Ocean Modelling: Application to Deep Convection.

Jean-Michel Campin, Chris Hill, Helen Jones, John Marshall

*Department of Earth, Atmospheric and Planetary Sciences  
Massachusetts Institute of Technology, Cambridge, MA 02139, USA*

---

## Abstract

We explore the efficacy of “super parameterization” (**SP**) in ocean modeling in which local 2-d non-hydrostatic plume-resolving fine-grained (**FG**) models are embedded at each vertical column of a coarse-grained (**CG**) hydrostatic model. A general multi-scale algorithm is described in which tendencies from the **FG** models are projected on to the **CG** model which in turn constrains the average state of the **FG** models, coupling the two models together. The approach is tested in the context of models of open ocean deep convection and compared with a pure hydrostatic, coarse resolution model using convective adjustment (**HYD**) and a full 3-d non-hydrostatic plume-resolving simulation (**NH**). The **SP** model is found to be greatly superior to **HYD** at much less computational cost than the fully non-hydrostatic calculation.

*Key words:*

---

## 1. Introduction

Turbulent mixing plays a central role in setting the stratification of the upper ocean both in open basins and in coastal areas. Models used to parameterize turbulence in the surface mixed-layer are based on “boundary layer” representations (e.g., Mellor and Yamada, 1974; Price et al., 1986; Large et al., 1997) and perform best if the stratification is not too strong and the flow remains highly turbulent. Models used to parameterize turbulence below the surface mixed layer are based on “wave-wave interaction models” (e.g., Müller et al., 1986) and work well in weakly turbulent environments, such as the ocean thermocline. However it is quite clear that such models are not appropriate in the near field of energetic forcing such as just below the mixed-layer and the benthic boundary layer. Contrary to the traditional notion of the mixed-layer base as a boundary between quiescent and turbulent regions, turbulent mixing does not immediately drop to the small interior values at the base of the mixed layer. Instead, there is a transition layer across which mixing rates decay with depth from high values at the surface to extremely low values in the interior. The

---

*Email addresses:* [jmc@mit.edu](mailto:jmc@mit.edu) (Jean-Michel Campin), [cnh@mit.edu](mailto:cnh@mit.edu) (Chris Hill), [helen@ocean.mit.edu](mailto:helen@ocean.mit.edu) (Helen Jones), [jmarsh@mit.edu](mailto:jmarsh@mit.edu) (John Marshall)

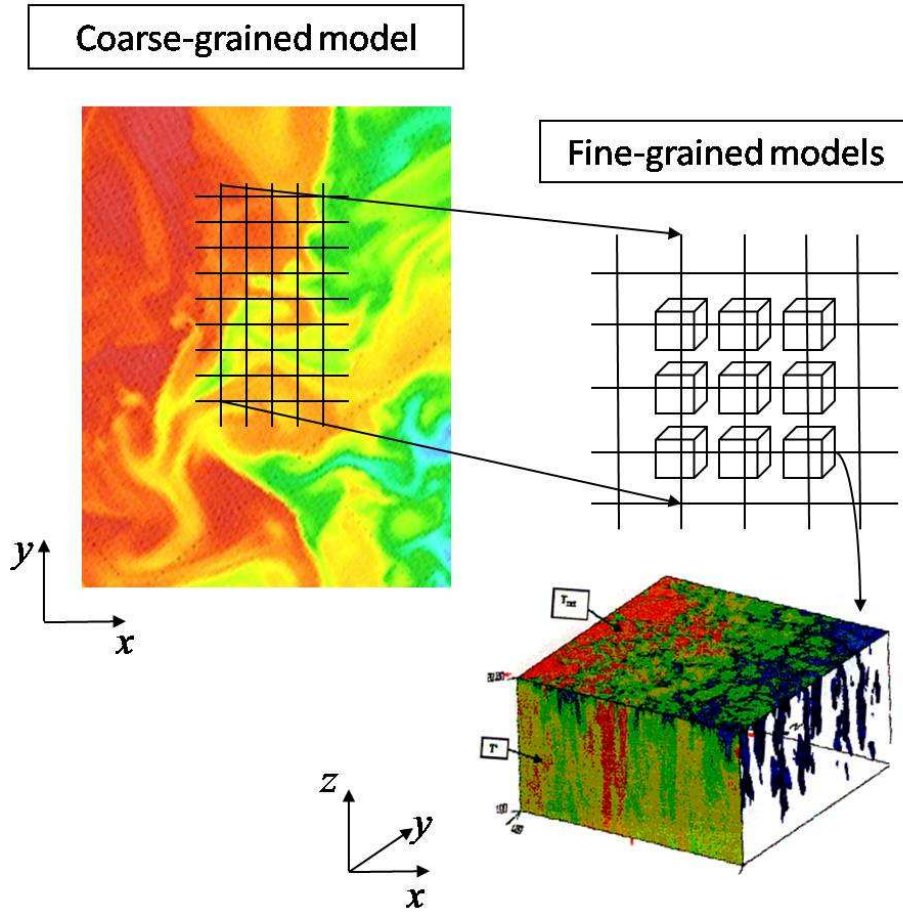


Figure 1: A schematic diagram showing a coarse-grained (**CG**) large-scale model on which the balanced motion is integrated forward with embedded fine-grained (**FG**) running at each column of the large-scale grid. The **FG** is non-hydrostatic and resolves the small-scale processes, as shown in the inset. The **FG**'s and the **CG** are integrated forward together and exchange information following the algorithmic set out in Section (3).

penetration depth is typically a few tens of meters, but can occasionally extend down a hundred meters or more. Such subtleties are extraordinarily difficult — perhaps impossible — to capture using conventional turbulence and mixing models.

Here we explore a new approach to the parameterization of subgridscale processes in ocean models which offers a route forward on from 1-d representations and attempts to resolve, rather than parameterize, small-scale processes. Fine-grid non-hydrostatic models (**FG**) are embedded in to a coarse-grid hydrostatic model (**CG**). Rather than employ a one-dimensional (1-d) parameterization of small-scale processes, the **CG** model includes tendencies from an array of **FG** models running at each horizontal grid-point of the **CG**, as sketched in Fig.1. The **FG** models resolve, rather than parameterize, the turbulent mixing processes. The coupling between the **FG** and the **CG** is two-way — the **FG**s receive information about the large-scale

shear and temperature/salinity ( $\theta/S$ ) environment from the **CG**, compute momentum and  $\theta/S$  tendencies by integrating forward **FG** submodels, and then return the tendencies to the **CG**. In this way we obviate the need for a conventional 1-d closure.

Our approach is motivated by the belief that in addition to traditional 1-d boundary layer approaches to the parameterization of turbulence in the surface mixed-layer (e.g., Kraus and Turner, 1967; Mellor and Yamada, 1982; Price et al., 1986; Large et al., 1994; Nurser, 1996) it is important to explore alternative routes that take advantage of modern massively parallel computers, permitting aspects of small-scale motions to be resolved rather than their transfer properties represented parametrically. However, a brute-force approach in which plume-resolving resolution is employed everywhere cannot yet be fully realized because of limitations in computational resources. Instead, here we experiment with high-resolution local sub-models that, initially at least, are run as vertical 2-d slices at each horizontal grid column of the **CG**.

In order to develop an appropriate algorithmic approach to the embedding of non-hydrostatic submodels in a hydrostatic large-scale model we focus on the interplay of plumes and baroclinic instability in the context of open ocean deep convection. We focus on the idealized configuration introduced by Jones and Marshall (1993). It is shown that the use of non-hydrostatic **FG** models embedded in a hydrostatic **CG** model is able to capture key aspects of the evolving flow at a computational cost which is orders of magnitude smaller than the 3-d plume resolving model over the entire domain. In particular the fidelity of the solutions greatly exceeds that obtained using a convective adjustment scheme.

The approach explored here is new to oceanography but has been and is being vigorously pursued in atmospheric modeling, where it goes under the name of “super-parameterization” (**SP=CG+FG**). Parameterization of convective cloud processes by overset grid explicit models is an active area of research following Grabowski (2001). Khairoutdinov and Randall (2001) explored the impact of introducing such a scheme into the NCAR Community Climate System Model. Subsequently several authors have reported on research that expands on these ideas in global, meso-scale and idealized atmospheric models (see for example Khairoutdinov et al., 2005; Wyant et al., 2006; Tao et al., 2009; Grabowski, 2006; Majda, 2007; Khairoutdinov et al., 2008). Recently Grabowski and collaborators have also examined multiscale approaches to cloud droplet growth in the presence of turbulence (Wang et al., 2005). Application of horizontal overset mesh ideas to coupling of land-surface models to atmospheric models has been explored by Molod et al. (2004). Subsequent work has looked at overset vertical meshes for planetary-boundary-layer physics (Molod, 2009). Other research (Freitas, 2006) has used the super-parameterization concept more directly to look at the interaction of land-surface fires and the atmosphere, again using overset methods.

Our paper is set out as follows. In section 2 we describe the numerical simulation of open ocean deep convection that provides our test bed to explore **SP**. In section 3 we introduce the numerical scheme used to couple **CG** and **FG** models. Section 4 evaluates the fidelity of our **SP** approach. In section 5 we discuss and conclude.

## 2. Target application: simulation of open-ocean deep convection

We have developed the super-parameterization algorithm in the context of the plume-resolving chimney problem of Jones and Marshall (1993). This has been the focus of many parameterization attempts reviewed in Marshall and Schott (1999). Figure 2 shows the development of open-ocean deep convection in a very high resolution (100 m in the horizontal and vertical) non-hydrostatic (**NH**) model which simultaneously resolves convective plumes and large-scale geostrophically balanced motion. The domain is doubly periodic and has a horizontal scale of  $40 \text{ km} \times 40 \text{ km}$  with uniform depth of 2 km. An  $800 \text{ Wm}^{-2}$  heat loss is applied at the surface over a centered circular region with radius of 10 km. In all cases, the fluid starts at rest with a uniform stratification of Brunt-Väisälä frequency  $N = 3 \times 10^{-4} \text{ s}^{-1}$ . In all models a linear equation of state is used with a thermal expansion coefficient of  $2 \times 10^{-4} \text{ K}^{-1}$ . There is no salt. A background rotation rate of  $f = 10^{-4} \text{ s}^{-1}$  is used. Simulations cover a four day elapsed period, early in which plumes of cooled water develop and create a roughly cylindrical region of higher density fluid (see Fig.2). The resulting radial density gradient induces a circulation which becomes baroclinically unstable, generating geostrophically balanced eddies that are shed from the convection region and carry the dense water away from the cooling patch. The plume dynamics processes are non-hydrostatic and require higher resolution than the geostrophically balanced eddies that emerge later in the simulation.

With **NH** as our reference, we now wish to evaluate the difference between (i) a conventional hydrostatic model (**HYD**) model that resolves the geostrophic eddy scale, but relies on a convective adjustment scheme (which mixes unstable water through an enhanced vertical diffusivity — see Klinger et al., 1996) to represent the effects of plume dynamics and (ii) a model employing a super-parameterization approach. The “super-paramterized” model introduces a two-dimensional (vertical slice) plume-resolving model, **FG**, which is coupled to the coarse-grid part of the hydrostatic model, **CG**. One instance of the **FG** model is located in each grid column of the **CG**. The **FG** scheme takes the place of the convective adjustment scheme of **HYD**. A time-step of the **SP** system involves stepping forward the hydrostatic **CG** equations and the non-hydrostatic **FG** equations. However, because the **FG** model is two-dimensional and local to a grid column it is computationally cheaper than the full **NH** model (see below).

Key mesh parameters of the three model setups are given in table 1. All models have the same vertical resolution. The **NH** model requires 400 times the number of grid cells as the **HYD** model due to its plume resolving horizontal resolution. The **SP=CG+FG** scheme fall between the two, with 21 times the number of grid cells as **HYD**.

Results from the three models are shown in figure 2. The top two panels show the high-resolution non-hydrostatic reference simulation (**NH**) of chimney convection by Jones and Marshall (1993). Cooling of an ocean at rest over a disc (shown in the plan view of the surface temperature field on the top left) generates convection (as seen in the vertical section of temperature shown on the right). The middle two

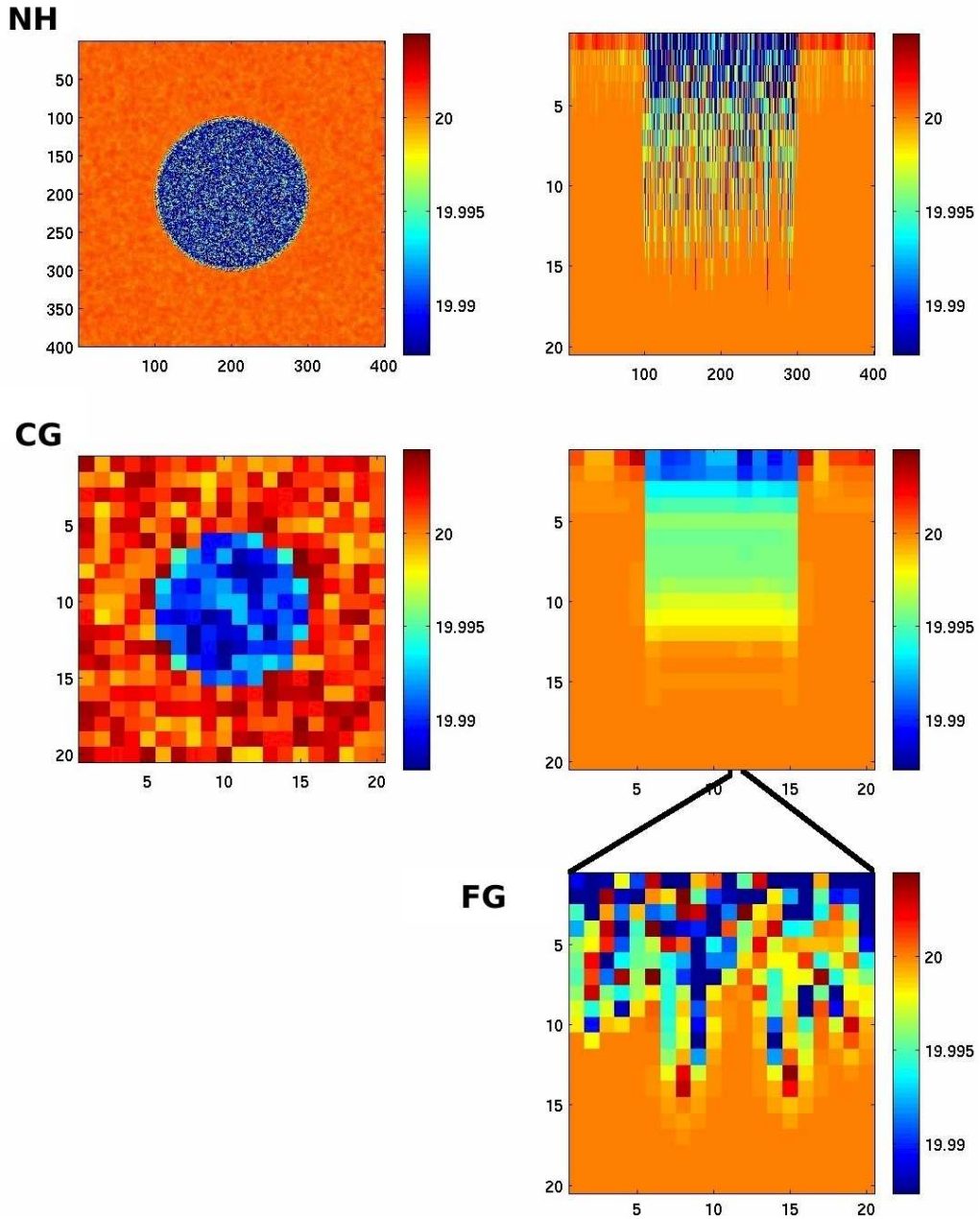


Figure 2: The top two panels show the temperature field in a high-resolution non-hydrostatic simulation of chimney convection by Jones and Marshall (1993). Cooling of a weakly stratified ocean at rest over a disc (shown in the plan view of the surface temperature field on the left) generates convection (as seen in the vertical section of temperature shown on the right). The axes are labelled in terms of grid-points. Since the horizontal resolution is 100m, the  $400 \times 400$  horizontal domain has a dimension of  $(40 \times 40)$  km. The middle two panels show the same fields from the **CG** model run at a horizontal resolution of  $(2 \times 2)$  km with embedded **FG** models at each horizontal grid point. The panel at the bottom shows plumes resolved in the embedded non-hydrostatic **FG** model which has a resolution of  $(100 \times 100)$  m.

Table 1: The three model configurations used in this study: coarse resolution hydrostatic model (**HYD**), high resolution non-hydrostatic model (**NH**) and super-parameterization model (**SP=CG+FG**).

	<b>HYD</b>	<b>SP=CG+FG</b>	<b>NH</b>
Horizontal resolution	2 km	<b>CG</b> : 2 km <b>FG</b> : 100 m	100 m
Vertical resolution	100 m	<b>CG</b> : 100 m <b>FG</b> : 100 m	100 m
Total number of grid cells	8000	168,000	3,200,000

panels show the same fields from the **CG** — the coarse resolution model run at a resolution of  $(2 \times 2)$  km — with embedded **FG** models at each horizontal grid point. Note that the temperature field in the **CG** has evolved because **CG** and **FG** exchange information. The panel at the bottom shows plumes resolved in the embedded plume model.

A test of the super-parameterization approach is the extent to which the large-scale evolution of the chimney, as revealed in figure 2, is captured in the **CG** with embedded plume-resolving **FG** models. This will be evaluated in Section 4.

### 3. **FG** ↔ **CG** algorithm

#### 3.1. Overview

The idea behind super-parameterization is that the coarse-grained model (the **CG**) carries information about the balanced dynamics, represented by variables with the subscript ‘**c**’. Thus, for example, **CG** might be a primitive equation model running on a grid with a horizontal spacing of a few kilometers. The dynamics of small-scale unbalanced motions — the scales we wish to parameterize — is represented by local fine scale models (the **FGs**), represented by variables with subscript ‘**f**’, and run at each vertical grid column of the **CG**.

The algorithm employed can be decomposed into four steps:

1. integrate each **FG** forward to compute tendencies on the fine grid
2. average **FG** tendencies to the coarse grid ‘**c**’
3. integrate **CG** forward (including coarse grid averaged tendencies from **FG**)
4. adjust state variables ( $\mathbf{v}, \theta, \dots$ ) of each **FG** model to make it consistent with the corresponding coarse grained vertical profile.

An adjustment is used in step 4 to ensure that fine-scale variables interpolated to the coarse model are the same as coarse-grained variables, thus keeping the ‘**c**’ and ‘**f**’ variables consistent with one-another. This is similar to the “gridalt” technology described in Molod (2009). Note that both momentum and tracer ( $\theta$ ) variables are treated in the super-parameterization, i.e. steps 1 to 4 above are also applied to momentum.

Because the embedded 2-d non-hydrostatic models can be run very efficiently, one can integrate a large array of such models and so afford a **FG** at each horizontal grid-point of the **CG**. One could also contemplate running small 3-d non-hydrostatic submodels, computational resources permitting. We will see that super-parameterization increases the cost of simulations by a factor of several hundred, but can make efficient use of massively parallel computers. In addition, super-parameterizations makes it possible for an ocean model to converge to a fully 3-d non-hydrostatic model as the horizontal grid spacing of the model decreases to a few 100 m.

### 3.2. Implementation details

The **HYD** and **NH** model are standard configurations of the MITgcm (Marshall et al., 1997a,b, 1998). Here we focus on outlining the approach used in the **SP** (= **CG**+**FG**) model. All models are based on configurations of the MITgcm software. Each model steps forward in time prognostic equations for potential temperature ( $\theta$ ), two horizontal components of velocity ( $u$  and  $v$ ) and solves a vertically integrated implicit equation for the surface elevation field ( $\eta$ ). In addition the non-hydrostatic model, **NH**, and non-hydrostatic sub-model, **FG**, step forward a prognostic equation for vertical velocity ( $w$ ) and solve for a non-hydrostatic pressure field ( $P_{nh}$ ). Hereafter the vector notation  $\mathbf{v} = \mathbf{u} + w\hat{\mathbf{z}} = u\hat{\mathbf{x}} + v\hat{\mathbf{y}} + w\hat{\mathbf{z}}$  is used to represent the three component velocity ( $u, v, w$ ) along three orthogonal axes x,y and z.

#### 3.2.1. Tracer equations

We formulate **SP** as a coupled system in which two sets of equations are stepped forward, one set for variables ( $\theta_c, u_c, v_c, \eta_c$ ) in **CG** and one set for variables ( $\theta_f, u_f, v_f, \eta_f, (P_{nh})_f$ ) in the **FG** sub-models. Like **NH**, **HYD** and **CG**, the fine-grid model **FG** is configured in a doubly periodic domain on an f-plane, although with an x-z domain only, so that any gradient in the Y direction vanishes. In our scheme the coupling between **CG** and **FG** occurs through the prognostic equations for  $\theta, u, v$ . For example, the temperature equation has the form

$$\text{Fine : } \quad \frac{\partial \theta_f}{\partial t} = -\mathbf{v}_f \cdot \nabla \theta_f \quad (1)$$

$$\text{Coarse : } \quad \frac{\partial \theta_c}{\partial t} = -\mathbf{v}_c \cdot \nabla \theta_c + F_\theta^{SGS} \quad (2)$$

The term  $F_\theta^{SGS}$  represents sub-grid scale (SGS) forcing effects that are calculated from **FG** components (1) and then averaged to the corresponding **CG** column:

$$F_\theta^{SGS} = \left[ \frac{\partial \theta_f}{\partial t} \right]_c \quad (3)$$

In addition, each **FG** sub-model is subject to the constraint:

$$[\theta_{\mathbf{f}}]_{\mathbf{c}} = \theta_{\mathbf{c}}(z) \quad (4)$$

For each **FG** sub-model, the  $[\ ]_{\mathbf{c}}$  operator is defined as the horizontal average over the small **FG** domain and maps to the corresponding **CG** water column in which it is embedded. In our comparisons (see section 4) the surface forcing is directly applied to each **FG** sub-model and is either prescribed or computed from local surface condition, depending on which tracer is considered (temperature or passive tracer). The average surface forcing is transmitted to the **CG** component through the **SP** model mapping operator  $[\ ]_{\mathbf{c}}$ , by  $F_{\theta}^{SGS}$  as defined in Eq.3.

In addition to prognostic equation (1), the constraint (4) is applied at the beginning of a new time-step by adjusting the **FG** model  $\theta_{\mathbf{f}}$  field to a new value  $\theta_{\mathbf{f}}^*$ , to ensure that the mean vertical profiles in the **FG** sub-models match the vertical profile of the corresponding **CG** water column:

$$\theta_{\mathbf{f}}^*(x) = \text{Fct}(\theta_{\mathbf{f}}(x), [\theta_{\mathbf{f}}]_{\mathbf{c}}, \theta_{\mathbf{c}}) \quad (5)$$

where ‘‘Fct’’ represents an appropriate mapping function.

Use of the simplest mapping function

$$\theta_{\mathbf{f}}^*(x) = \theta_{\mathbf{f}}(x) - [\theta_{\mathbf{f}}]_{\mathbf{c}} + \theta_{\mathbf{c}}$$

can produce unphysical extrema in the **FG** temperature field  $\theta_{\mathbf{f}}^*$  which lie outside the range of the original temperature bounds of both **FG** and **CG** models. Instead, a linear mapping that prevents false extrema is used:

$$\theta_{\mathbf{f}}^*(x) = A + (\theta_{\mathbf{f}}(x) - A) \times \frac{\theta_{\mathbf{c}} - A}{[\theta_{\mathbf{f}}]_{\mathbf{c}} - A} \quad \text{with } A = \min(\theta_{\mathbf{c}}, \min_x \theta_{\mathbf{f}}) \text{ if } \theta_{\mathbf{c}} < [\theta_{\mathbf{f}}]_{\mathbf{c}}$$

$$\text{and } A = \max(\theta_{\mathbf{c}}, \max_x \theta_{\mathbf{f}}) \text{ if } \theta_{\mathbf{c}} > [\theta_{\mathbf{f}}]_{\mathbf{c}}$$

In the algorithm described by Grabowski (2001), the coupling between **CG** and **FG** appears as a relaxation term in equations (1) and (2) (his equations 3a and 3b), which becomes similar to the present formulation (our equations 3 and 4) once the relaxation time-scale is set equal to the model time step, as discussed in Khairoutdinov et al. (2005).

### 3.2.2. Momentum equations and orientation of 2-d FG models

The algorithm that applies to the momentum equations is similar to that described above for temperature, but there are several differences. Since only one horizontal dimension ( $\underline{x}_{\mathbf{f}}$ ) is represented in the **FG** model, the horizontal momentum equation is much simpler in the cross plane direction ( $\underline{y}_{\mathbf{f}}$ ) (zero gradient) than along  $\underline{x}_{\mathbf{f}}$  direction. However, the orientation of the **FG** x-axis does not need to coincide with the **CG** grid axes, and can be selected in a favorable way. In order to capture the large scale vertical shear present in

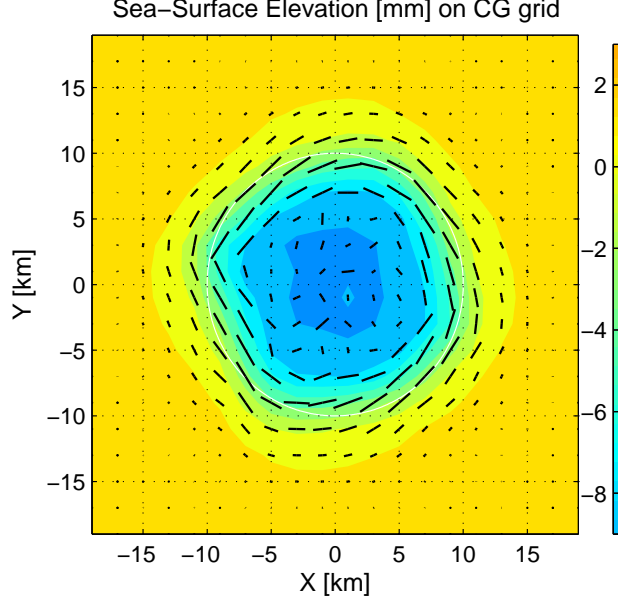


Figure 3: Sea surface height (in mm) and **FG** orientation vector defined in Eq.6 after 4 days of simulation. The ocean is cooled within the white circle.

the **CG** solution, the orientation of the **FG** model is allowed to evolve and to align along the direction of maximum vertical shear. Practically, the orientation is relaxed towards a target direction  $\alpha_{tg}$ , which is defined to maximize the vertical standard deviation of velocity ( $\text{RMS}(\alpha)$ ) in the corresponding **CG** column: If  $\alpha$  is the angle relative to the **CG** x-axis, the function:

$$\text{RMS}(\alpha)^2 = \frac{1}{H} \int_H [(u_c - \overline{u_c^z}) \cos(\alpha) + (v_c - \overline{v_c^z}) \sin(\alpha)]^2 dz$$

reaches a maximum for  $\alpha = \alpha_{tg}$  corresponding to the target orientation  $\alpha_{tg}$ . Here  $\mathbf{u}_c = (u_c, v_c)$  is the horizontal **CG** velocity and  $(\cdot^z)$  is the vertical averaging operator. The **FG** orientation vector  $\mathcal{V}_f$  along  $\mathbf{x}_f$  is then relaxed towards the target orientation vector  $\mathcal{V}_{tg}$  with a time scale  $\tau_{rot}$ :

$$\frac{\partial}{\partial t} \mathcal{V}_f = (\mathcal{V}_{tg} - \mathcal{V}_f) / \tau_{rot} \quad \text{where} \quad \mathcal{V}_{tg} = \text{RMS}(\alpha_{tg}) \left[ \cos(\alpha_{tg}) \hat{\mathbf{x}}_c + \sin(\alpha_{tg}) \hat{\mathbf{y}}_c \right] \quad (6)$$

Using a relaxation time scale  $\tau_{rot}$  shorter than the typical time-scale of coarse-grid flow adjustment (e.g., an inertial period, in our case  $\sim 17$  h) allows **FG** models to track the direction of **CG** large vertical shear. A time scale much longer than the model time step (in our case  $\Delta t = 60$  s) prevents the occurrence of sudden changes of orientation due to numerical noise in the **CG** model. A time scale  $\tau_{rot} = 1$  h was found to be satisfactory in **SP** simulations; figure 3 illustrates the **FG** orientation vector,  $\mathcal{V}_f$ , field after four days of simulation.

The remaining aspect of the momentum coupling closely follows the tracer algorithm: the averaging

operator  $[ \ ]_{\mathbf{c}}$  accounts for the orientation ( $\alpha$ ) of the **FG** model:

$$[\mathbf{u}_{\mathbf{f}}]_{\mathbf{c}} = \mathbf{u}_{\mathbf{c}}(z) \quad (7)$$

which can be written:

$$[\cos \alpha u_{\mathbf{f}} - \sin \alpha v_{\mathbf{f}}]_{\mathbf{c}} = u_{\mathbf{c}}(z)$$

$$[\sin \alpha u_{\mathbf{f}} + \cos \alpha v_{\mathbf{f}}]_{\mathbf{c}} = v_{\mathbf{c}}(z)$$

The SGS forcing term is computed from the advective tendency only:

$$F_{\mathbf{u}}^{SGS} = [-\mathbf{v}_{\mathbf{f}} \cdot \nabla \mathbf{v}_{\mathbf{f}}]_{\mathbf{c}} \quad (8)$$

since the other (linear) terms are explicitly represented in the **CG** model.

$$\text{Fine : } \frac{\partial \mathbf{v}_{\mathbf{f}}}{\partial t} = -\mathbf{v}_{\mathbf{f}} \cdot \nabla \mathbf{v}_{\mathbf{f}} - 2\mathbf{\underline{\Omega}} \times \mathbf{v}_{\mathbf{f}} - \frac{1}{\rho_0} \nabla (P_h + P_{nh})_{\mathbf{f}} + D_{\mathbf{f}} \quad (9)$$

$$\text{Coarse : } \frac{\partial \mathbf{u}_{\mathbf{c}}}{\partial t} = -\mathbf{v}_{\mathbf{c}} \cdot \nabla \mathbf{u}_{\mathbf{c}} - 2\mathbf{\underline{\Omega}} \times \mathbf{u}_{\mathbf{c}} - \frac{1}{\rho_0} \nabla_h (P_h)_{\mathbf{c}} + D_{\mathbf{c}} + F_{\mathbf{u}}^{SGS} \quad (10)$$

where  $D$  is the dissipation term and  $P_h$  is the hydrostatic pressure including surface pressure contribution.

### 3.2.3. Time stepping

The coarser horizontal resolution of the **CG** model allows one to use a longer time-step than in the **FG** component. However, to simplify the comparison of the different simulations, the same time-step is used in all models (**NH, HYD** and **SP**) and components (**CG, FG**).

In the present implementation, the two components are stepped forward sequentially: first the orientation of the **FG** slice is updated (equation 6) using the current **CG** velocity field and then each of the **FG** instances is advanced in time (equations 1 and 9). This allows one to compute the resulting sub-grid scale contribution (equations 3 and 8) required to step forward the **CG** prognostic variables (equations 2 and 10). A global budget of volume and tracer in the **CG** model shows perfect conservation.

### 3.3. Super-parameterization coupled computation — design and cost

We use an acyclic graph coupling approach to link the **CG** and **FG** sub-models in **SP**. In this approach both the **CG** and **FG** sub-models are viewed as “components” at the leafs of a two branch tree, or acyclic graph, that are distributed over multiple processors using the Earth System Modeling Framework (Hill et al., 2004; Suarez et al., 2007) software library. The arrangement is illustrated in figure 4.

The overall computation is supervised by a root component,  $r$ , that spans a set of one or more processes and/or threads  $\{p_r\}$  and executes concurrently over  $p_r$ . Child “components”,  $c$  and  $f$ , that correspond to

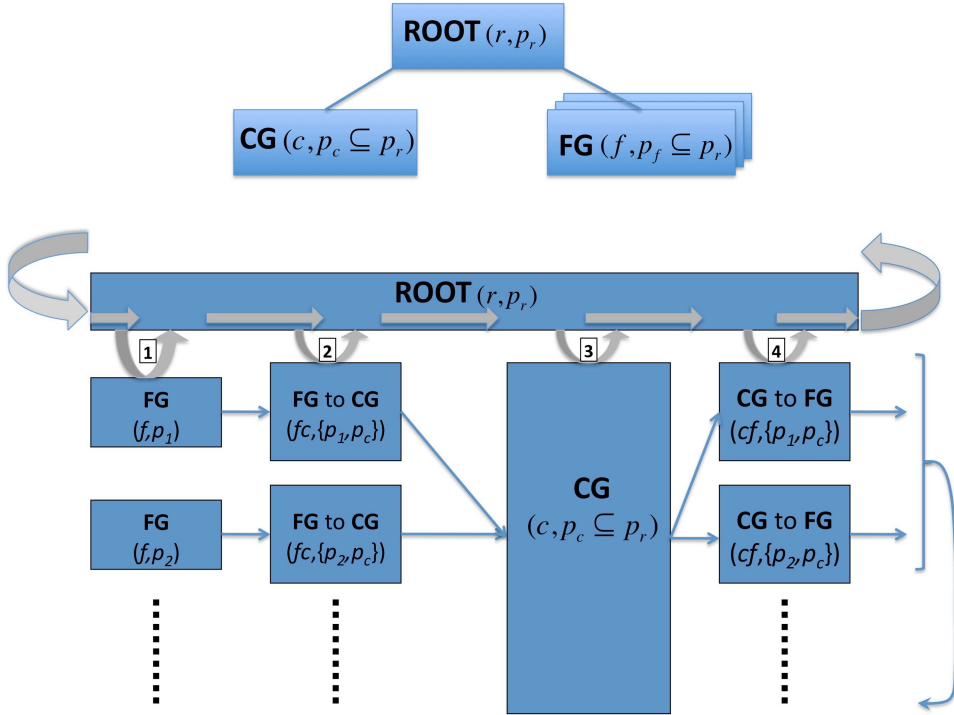


Figure 4: TOP PANEL: The computational elements of the SP scheme involving **CG** and **FG** models organized as components under a parent component **ROOT** using the Earth System Modeling Framework. The **ROOT** acts to schedule **CG** and **FG**s pairing in a set of computational threads/processes  $p_r$  that includes all the **CG** threads/processes  $p_c$  and all the **FG** threads/processes  $p_f$ . BOTTOM PANEL: The sequence of events in a model timestep of **SP** coordinated by **ROOT**. The numbers 1-4 correspond to the algorithm steps 1-4 described in section 3.1. In step (1) the **FG** models are integrated forward. There are  $n_x \times n_y$  independent **FG** models, one for each grid cell. These models are spread over threads/processes  $p_n, n = 1, m$  within the set  $p_f$ . In step (2) updated tendencies from the **FG** models are mapped to the **CG** model grid and transferred to the **CG** model threads/processes  $p_c$ . In step (3) the **CG** model executes on a set of threads/processes  $p_c$ . Finally, in step (4), **CG** model state information is distributed to the **FG** models. In between each step control returns to the **ROOT** component, allowing it to coordinate **SP** model integration.

submodels **CG** and **FG** are controlled by the root component  $r$ . The  $c$  and  $f$  components are designed to execute concurrently, under the overall control of  $r$ , on process/thread sets  $\{p_c\}$  and  $\{p_f\}$ . The sets  $\{p_c\}$  and  $\{p_f\}$  are subsets of  $\{p_r\}$ . In the experiments described here we define the  $c$  component on a single process that integrates forward the **CG** terms and  $(n_x \times n_y)$   $f$  components, each of which integrates forward the **FG** terms for a single grid column.

The assignment of process/thread subsets  $\{p_c\}$  and passing of information between the  $c$  and  $f$  components is orchestrated by the  $r$  component. Using this approach the  $r$  component can flexibly map the computations to available compute processes/threads. In the experiments described here we execute the coarse component on a single process. The work of the  $(n_x \times n_y)$   $f$  components is spread over 20 processes, so that each process handles  $(n_x \times n_y)/20$   $f$  components. Grouping  $f$  components provides a way to amortize

data copy costs between  $c$  and  $f$  components. Map and inverse map functions  $R : c \rightarrow f$ ,  $R^{-1} : f \rightarrow c$  are defined using the ESMF regrid library. For this computation the  $R$  map is a general scatter that decimates the  $c$  index space and communicates different parts of it to different members of  $p_f$  that control different  $f$  sub-models. The  $R^{-1}$  map is a general gather that collects information from the members of  $p_f$  and passes an assembled  $c$  index space to  $p_c$ . The necessary routing and data transfer for this is handled by the ESMF library.

### 3.3.1. Computational Cost

Computational costs for the different simulations **HYD**, **SP** and **NH** can be understood in terms of table 1. The **HYD** simulation requires solving the hydrostatic equations of motion on a grid of size  $n_x \times n_y \times n_z$  at an associated computational cost per grid cell of  $\alpha_{\text{dyn}} + \alpha_{\text{conv}}$ , where  $\alpha_{\text{dyn}}$  is the cost of hydrostatic dynamics and  $\alpha_{\text{conv}}$  is the cost of the convective adjustment scheme. The total cost of a time-step,  $C_{\text{HYD}}$  is then

$$\frac{C_{\text{HYD}}}{n_z} = \alpha_{\text{dyn}} N_h + \alpha_{\text{conv}} N_h \quad (11)$$

where  $N_h = n_x \times n_y$ . The **SP** simulation cost per time-step can be similarly expressed as

$$\frac{C_{\text{SP}}}{n_z} = \alpha_{\text{dyn}} N_h + \alpha_{\text{nh}} N_h s \quad (12)$$

where  $\alpha_{\text{nh}}$  is the cost per grid cell of the non-hydrostatic algorithm and  $s$  is the scale ratio of the embedded model resolution to the coarse model resolution. The cost of the **NH** model is then

$$\frac{C_{\text{NH}}}{n_z} = \alpha_{\text{nh}} N_h s^2 \quad (13)$$

The non-hydrostatic algorithm involves all the hydrostatic terms in  $\alpha_{\text{dyn}}$  and some additional terms (notably an elliptic problem for the three-dimensional pressure field), so that  $\alpha_{\text{dyn}} < \alpha_{\text{nh}}$ . In the experiments reported here it is also true that  $\alpha_{\text{conv}} \ll \alpha_{\text{dyn}}$  - the experiments use a simple elevated vertical diffusivity to convectively mix water parcels that are found to be statically unstable with respect to a common reference level, as described in Klinger et al. (1996). We therefore see that the ratio of compute cost of **HYD:SP:NH** is, to first order, given by  $1 : s : s^2$ . From table 1,  $s = 20$ , so that given the same timestep is used in all experiments, the computational cost **HYD:SP:NH** is  $1 : 20 : 400$  per timestep.

In practice we can offset some of the increased computational cost by exploiting parallelism in the **SP** and **NH** simulations that is greater than in the **HYD** simulations. Execution time is a function of both computational cost and the degree to which concurrent computations can be executed efficiently in parallel. In all of **HYD**, **SP** and **NH** there is abundant data parallelism (of the order of the number of grid cells). However, in practice process to process and/or thread to thread synchronization, communication and computer resource contention limits mean that the **HYD** calculation is performed as a single process computation.

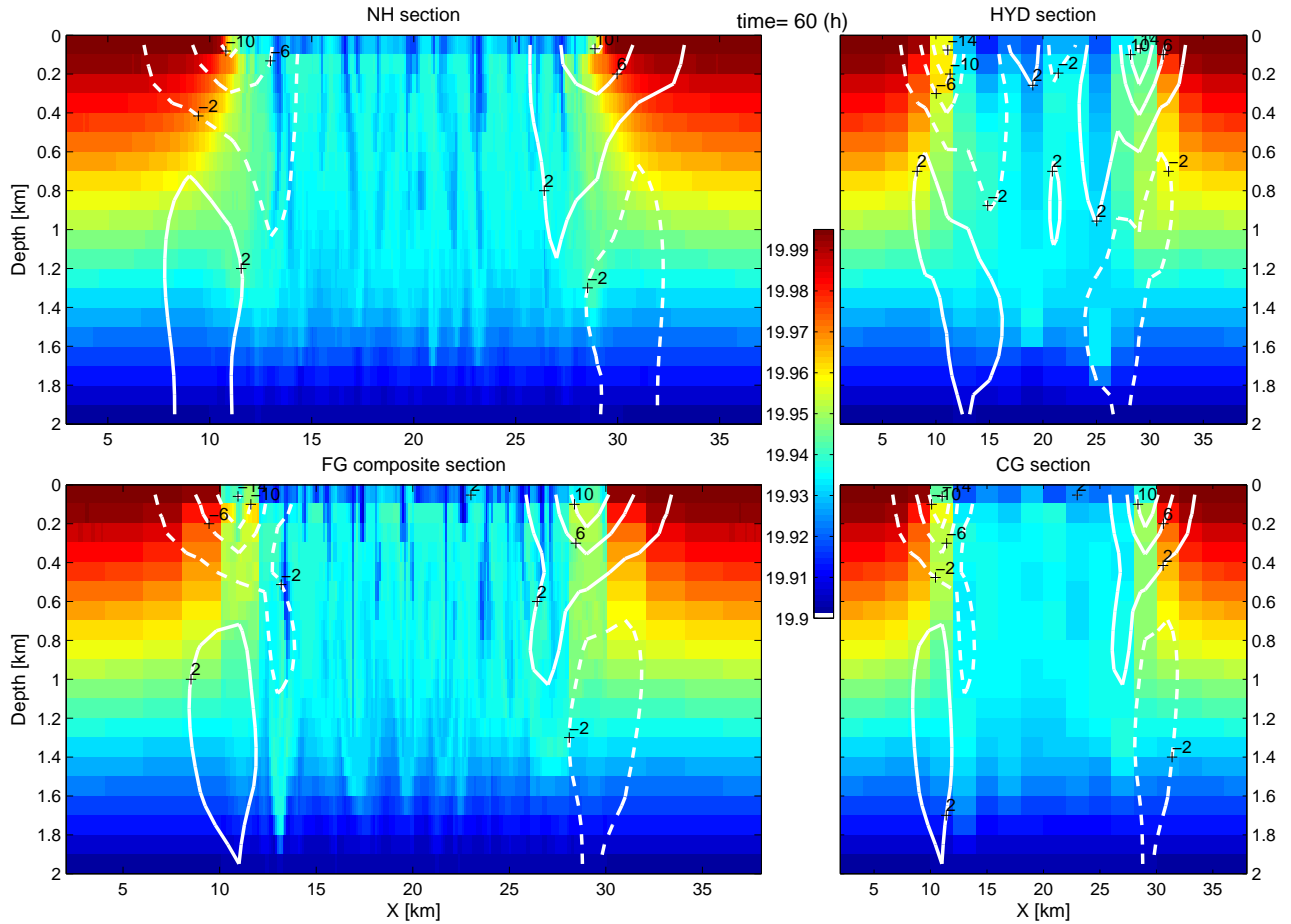


Figure 5: Temperature distribution in one vertical radial section after 60.h of simulation, for the 3 cases: Top left: **NH** ; Top right: **HYD** ; Bottom left: **SP** showing a composite of all **FG** instances along the radial section. Bottom right: same as bottom left but averaged on the **CG** grid. The velocity component normal to the corresponding section is represented by white isolines.

In contrast, the **SP** computation is executed as a single process/thread **CG** dynamics computation that executes concurrently with multiple two-dimensional **FG<sub>2d</sub>** computations. The full, three-dimensional, **NH** computation is also executed in parallel. This brings the execution time ratios to less than the computational cost ratios (by utilizing more compute resources).

#### 4. Results and analysis

For the purposes of comparing the different algorithms we treat the **NH** as our reference solution. To aid in the comparison, a passive tracer has been added in the 3 models, and follows the same algorithm as temperature. The initial tracer distribution is zero everywhere except a concentration of unity in the surface level. Moreover, in contrast to temperature, a strong restoring of the tracer to a concentration of unity is applied at the surface. The **HYD** model uses a simple convective adjustment in which vertical diffusivity

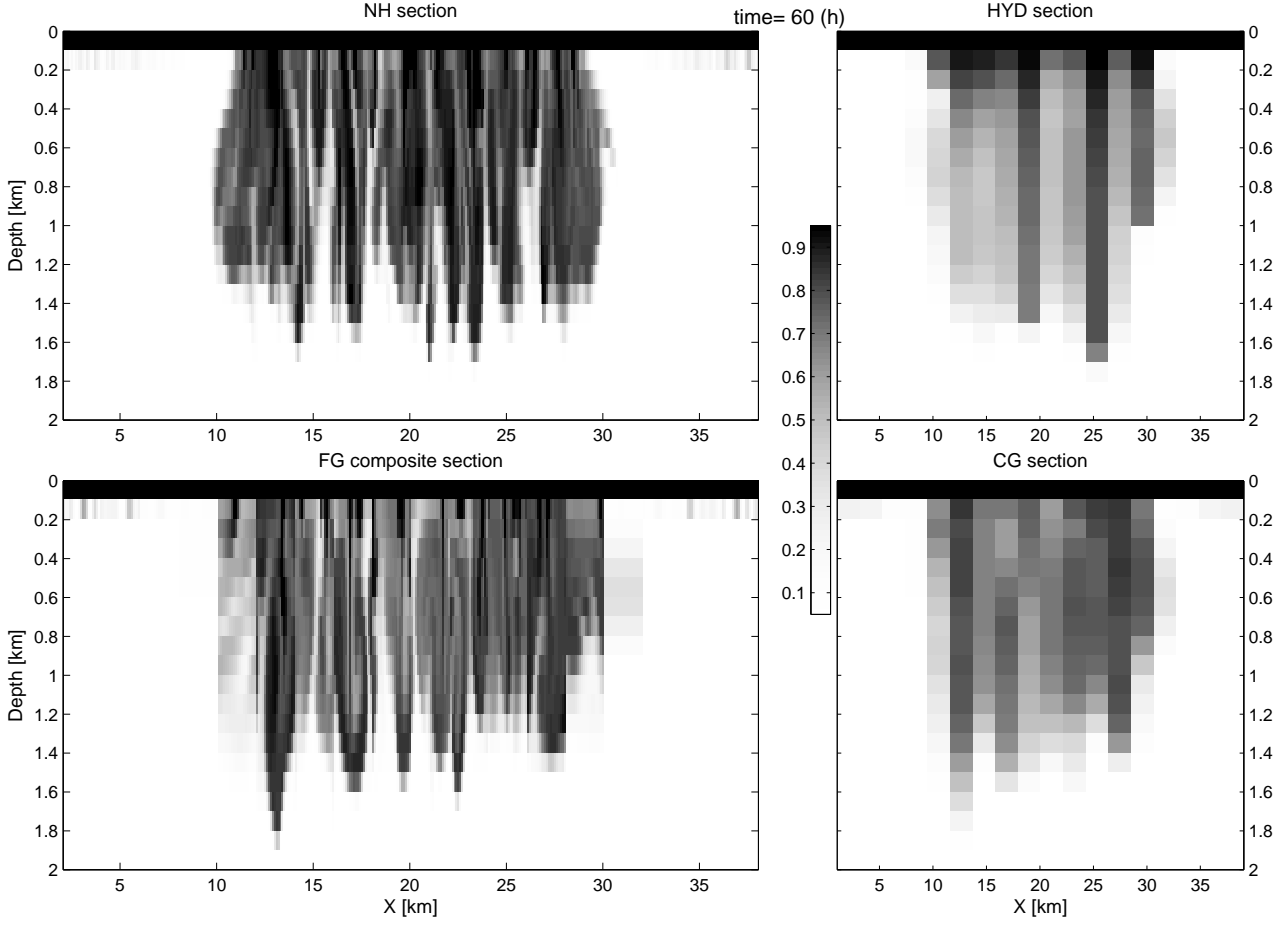


Figure 6: Tracer distribution along the section shown in Fig.5 after 60.h of simulation.

is set to a large value to mix water parcels that are found to be statically unstable. The magnitude of the enhanced convective mixing in **HYD** has been adjusted to  $K_{conv} = 2.5 \text{ m}^2/\text{s}$  to ensure that the mean tracer content matches the **NH** simulation after 3 days of simulation.

Qualitatively, the **SP** model captures several features of the **NH** simulation which **HYD** poorly represents:

- The evolution of the convective plumes in the **NH** and **FG** simulations is quite similar, both in term of timing and form, as seen in temperature and tracer sections (figures 5 and 6). In particular, since the 2-d **FG** models feel the large scale vertical shear present in the **CG** simulation, they capture the tilt of the plumes observed in the **NH** simulation.
- Convective plumes can penetrate deeper than their neutrally buoyant level, burrowing in to the stratified (colder) water below (seen on bottom left figures 5 and 6, around  $x = 13 \text{ km}$ ). This is a non-hydrostatic process which is difficult to capture in parameterization schemes.

- The convective plumes overturn the water column and bring cold and tracer depleted water upward, as shown on figure 6 in both **NH** and **FG** sections. As a consequence, the resulting coarse-grid average vertical profile is not always monotonic, tracer content being sometimes lower near the surface with patches of higher concentration at depth (Fig.6, bottom right). Use of a diffusive convective adjustment scheme, in contrast, maintains decreasing tracer concentration with depth in each column at all times (see Fig.6, top right). This feature has motivated development of non-local convective parameterization schemes (e.g., KPP, Large et al., 1994).
- The horizontal variance inside the convective patch is higher in **HYD** than in both **SP** (Figs 5 and 6) and in **NH** (as seen from Fig.10). The nature of the convective adjustment (on-off, depending on a single threshold criteria) is likely to be responsible for too high a sensitivity and could explain the large horizontal variability observed within the patch. Stochastic convective parameterizations (e.g., Lin and Neelin, 2003) attempt to address this issue.
- In the convective patch, the level just below the surface remains warmer than above and immediately below, in both **NH** and **SP** simulations (Fig.5), indicating a reduced vertical exchange at the first vertical interface. This is consistent with turbulence scaling where vertical velocity decreases toward zero at the surface. By contrast, no inversion develops in our **HYD** simulation and the sub-surface level is significantly colder (Fig.10). However, these differences are likely an artifact of the low vertical resolution near the surface and are expected to become smaller with increasing vertical resolution.

Despite these differences, the depth of the convective patch in the three simulations are rather similar after a few days of simulation (Figs 5, 6). However, early on the **HYD** model shows larger departures from the two others. For example after 1 day, convection is shallower and more uniform in **HYD** than in the **NH** and **SP**, as is evident from the tracer section shown in figure 7.

The time evolution of the mean vertical profiles of temperature and passive tracer (Fig.8) allows one to quantify the differences between the three simulations. In both cases, **HYD** is much further away from our reference true solution **NH** than is **SP**, especially during the first 2 days of simulation. At this time the convective patch in **HYD** is too shallow and cold with rather too high a concentration of tracer in the upper part of the patch, and too warm with too low a concentration beneath. Also noticeable are differences in the initial response to buoyancy loss: the convective adjustment in **HYD** begins to mix immediately whereas the plume dynamics spin up gradually resulting in significant tracer differences during the first 6 hours (Fig.10, upper panel). More elaborate mixing parameterizations attempt to include additional prognostic variables (e.g., Mellor and Yamada, 1982) and are likely to better capture such subtleties in the time evolution of deepening convection.

The **SP** simulation experience some difficulties in reproducing the spatial variability observed inside the convective patch of the **NH** model (Fig.9), but still outperforms **HYD** (see also Fig.10, bottom panel). The

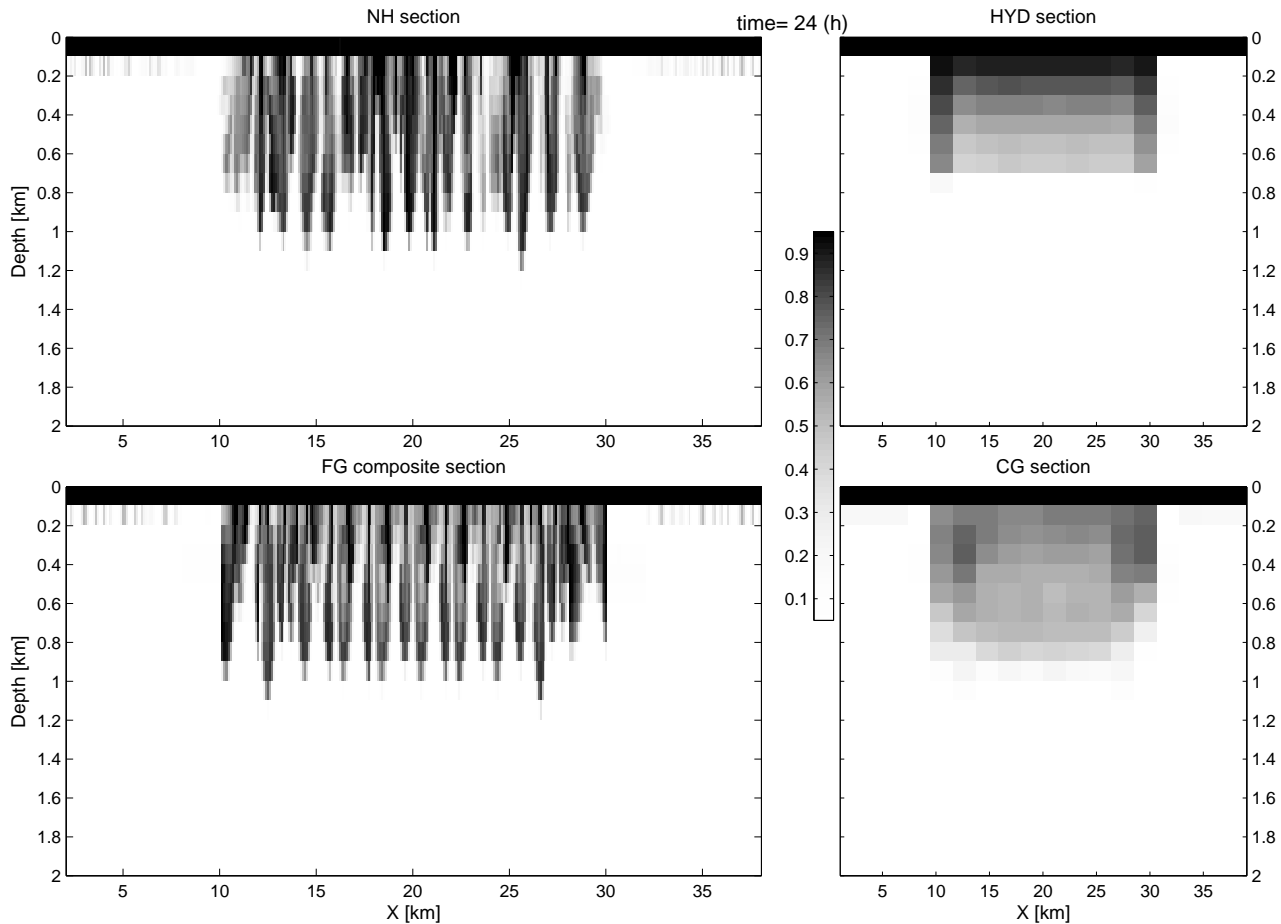


Figure 7: Tracer distribution along the section shown in Fig.5 and 5 after 24.h of simulation.

use of 2-d **FG** models and an imperfect scale separation are likely to contribute to this.

## 5. Conclusions

We have described a multi-scale, super-parameterization model applicable to open-ocean deep convection simulation. By coupling a two-dimensional, plume resolving, non-hydrostatic sub-model at each grid cell of a coarser resolution hydrostatic model, the resulting **SP** model captures many of the features of a full non-hydrostatic simulation. These features (for example tilted plumes due to vertical shear and transient plume overshoots past neutral buoyancy) are not found in the pure hydrostatic, **HYD**, simulation. More sophisticated one-dimensional vertical mixing schemes, such as KPP, can represent some aspects of the processes resolved explicitly by the non-hydrostatic 2-d models. However, phenomena such as temporary individual plume overshoots past the level of neutral buoyancy are not captured by these schemes.

The temporal and spatial inhomogeneities in the 2-d **FG** sub-model impact the transient evolution of an

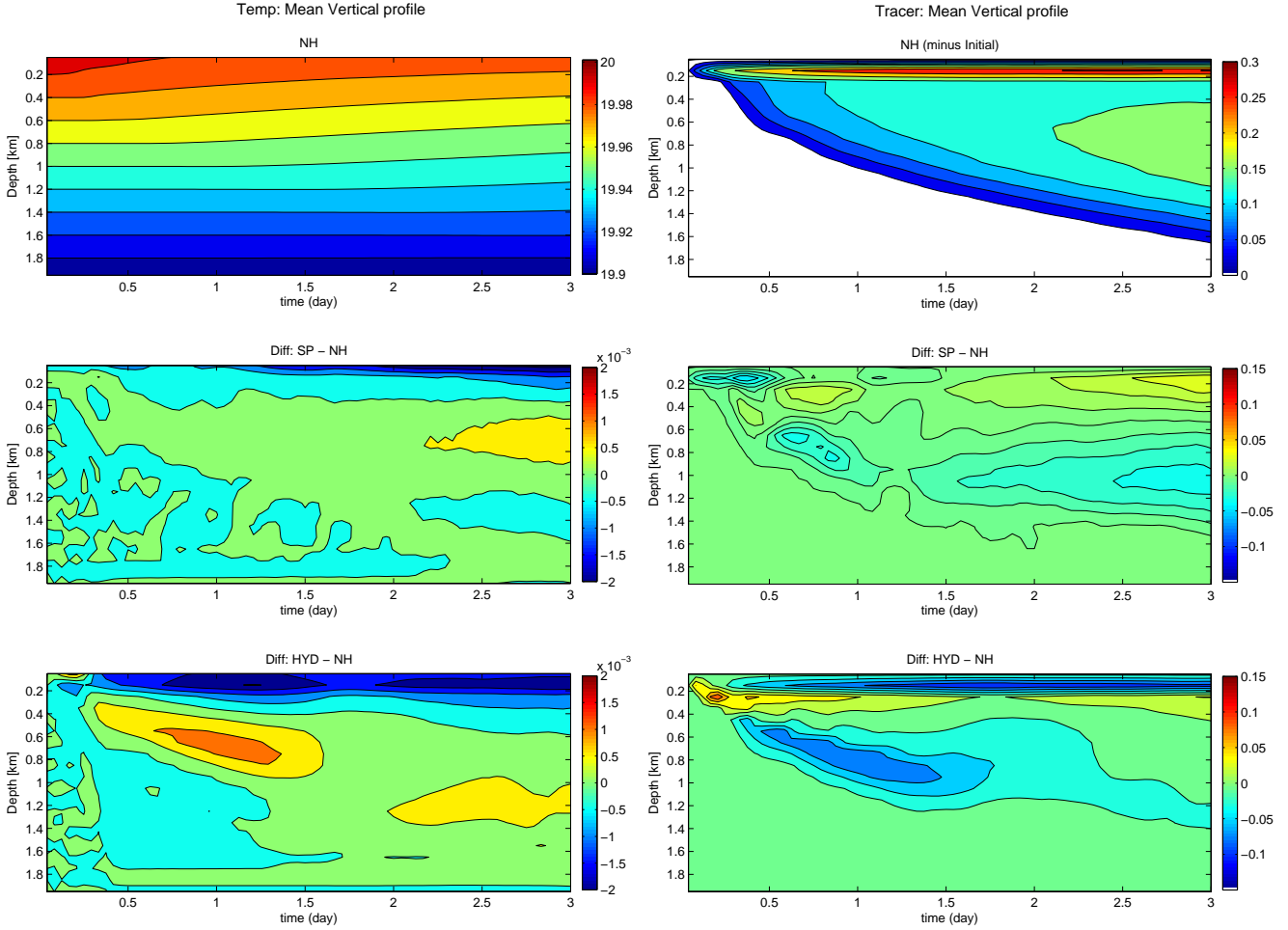


Figure 8: Evolution of the spatial mean vertical mean profile of Temperature (left) and surface passive tracer (right), with time (in days) as horizontal axis and depth (in 100.m) as vertical axis. Top: **NH** simulation; middle: Difference **SP** minus **NH**; bottom: Difference **HYD** minus **NH**.

inert tracer and could be important in rectifying biological processes involving chemical tracers and ocean microbial species, as well as in modulating buoyancy and momentum vertical transports.

The computational cost of the **SP** approach is significantly less than that of a full 3-d **NH** model. Moreover the independent 2-d plume models provide a rich source of parallelism. The scheme outlined here could prove beneficial to emerging petascale ocean applications which are targeting basin and global scale simulation at a few kilometer resolution. Embedding a 2-d non-hydrostatic special purpose model in such integrations would provide a computationally tractable way to incorporate non-hydrostatic effects in the relatively near future.

The recipe we have outlined should also apply to other processes where there is a relatively clean separation of scales and where approximate parameterizations are currently employed. For example, various

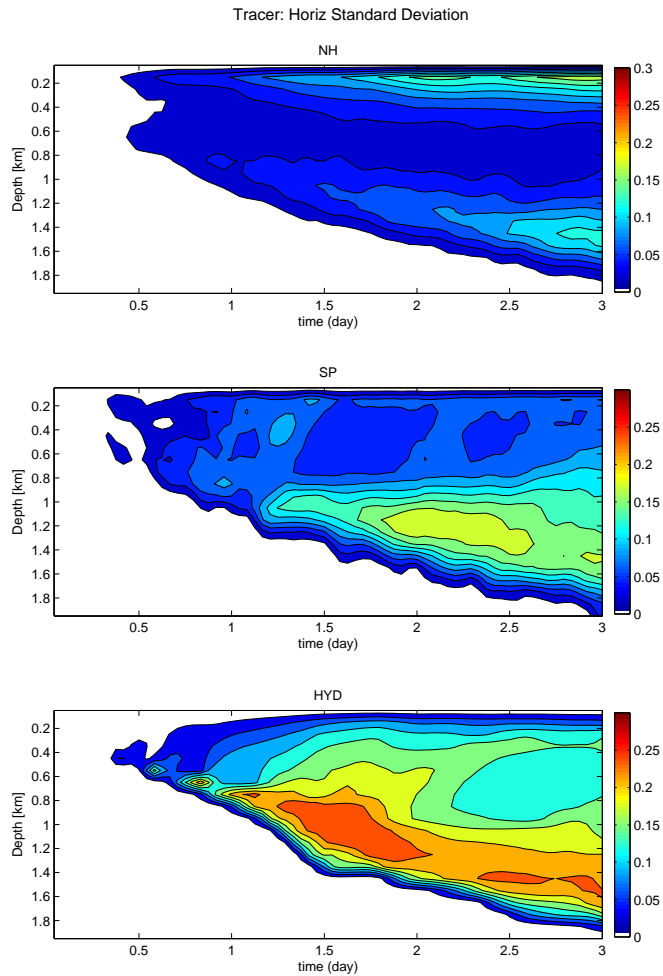


Figure 9: Evolution of standard deviation of tracer at each vertical level inside the convective region (where surface cooling is applied), with time (in days) as horizontal axis and depth (in 100.m) as vertical axis. All statistics are computed from the average coarse grid (ignoring sub-grid/fine grid variance). Top: **NH** simulation; middle: **SP** simulation; bottom: **HYD** simulation.

parameterizations have been devised to capture the sub-grid scale sinking of dense, salty water from ice formation. A two-dimensional, non-hydrostatic model of the sort used here could provide an alternative rooted in fundamental principles. This could obviate the need for tunable vertical depth scales in such parameterizations (e.g., Nguyen et al., 2009).

More generally the ESMF machinery we have used allows coupling of other overset models. For example, it would be possible to couple a non-eddy ocean circulation model to a limited-area quasi-geostrophic model providing a route to represent geostrophic eddy fluxes in extended global simulations. In the near term we are beginning to examine the inclusion of momentum forcing into the testbed problem we have outlined. This is a precursor to the more elaborate applications outline above.

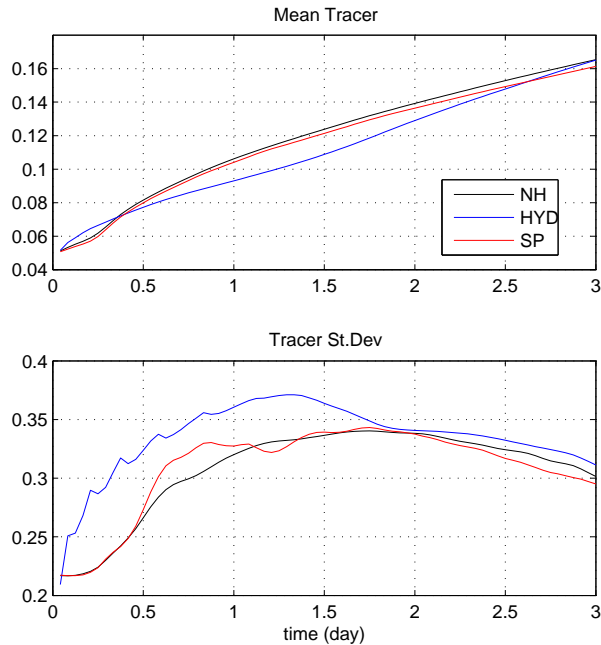


Figure 10: Evolution of global mean tracer concentration (top) and standard deviation within the convective region (bottom) for the three simulations: **NH** (black line), **SP** (red line) and **HYD** (blue line).

## Acknowledgements

This work was supported by a grant from ONR and NASA (ECCO-2, MAR Program).

## References

- Freitas, S., 2006. Using the super-parameterization concept to include the sub-grid plume-rise of vegetation fires in low resolution atmospheric chemistry-transport models. Proceedings of 8 ICSHMO, 109–113.
- Grabowski, W., 2006. Comment on “Preliminary Tests of Multiscale Modeling with a Two-Dimensional Framework: Sensitivity”. Monthly Weather Review 134, 2021–2026.
- Grabowski, W. W., 2001. Coupling cloud processes with the large-scale dynamics using the Cloud-Resolving Convection Parameterization (CRCP). Journal of the Atmospheric Sciences 58, 978–997.
- Hill, C., DeLuca, C., Balaji, V., Suarez, M., Da Silva, A., 2004. The architecture of the Earth System Modeling Framework. Computing in Science and Engg. 6 (1), 18–28.
- Jones, H., Marshall, J., 1993. Convection with rotation in a neutral ocean; a study of open-ocean deep convection. Journal of Physical Oceanography 23, 1009–1039.
- Khairoutdinov, M., DeMott, C., Randall, D. A., 2008. Evaluation of the simulated interannual and subseasonal variability in an AMIP-style simulation using the CSU Multiscale Modeling Framework. Journal of Climate 21, 413–431.
- Khairoutdinov, M., Randall, D., 2001. A cloud resolving model as a cloud parameterization in the NCAR Community Climate System Model. Geophysical Research Letters 28 (18), 3617–3620.
- Khairoutdinov, M., Randall, D., DeMott, C., 2005. Simulations of the atmospheric general circulation using a cloud-resolving model. Journal of the Atmospheric Sciences 62, 2136–2154.

- Klinger, B. A., Marshall, J., Send, U., 1996. Representation of convective plumes by vertical adjustment. *Journal of Geophysical Research* 101 (C8), 18,175–18,182.
- Kraus, E., Turner, J., 1967. A one-dimensional model of the seasonal thermocline II: the general theory and its consequences. *Tellus* 19, 98–105.
- Large, W., Danabasoglu, G., Doney, S., McWilliams, J., 1997. Sensitivity to surface forcing and boundary layer mixing in a global ocean model: Annual-mean climatology. *Journal of Physical Oceanography* 27, 2418–2447.
- Large, W., McWilliams, J., Doney, S., 1994. Oceanic vertical mixing: A review and a model with nonlocal boundary layer parameterization. *Reviews of Geophysics* 32, 363–403.
- Lin, J. W.-B., Neelin, J. D., 2003. Toward stochastic deep convective parameterization in general circulation models. *Geophysical Research Letters* 30, 1162.
- Majda, A., 2007. Multiscale models with moisture and systematic strategies for superparameterization. *Journal of the Atmospheric Sciences* 64, 2726–2734.
- Marshall, J., Adcroft, A., Hill, C., Perelman, L., Heisey, C., 1997a. A finite-volume, incompressible Navier Stokes model for studies of the ocean on parallel computers. *Journal of Geophysical Research* 102, 5753–5766.
- Marshall, J., Hill, C., Perelman, L., Adcroft, A., 1997b. Hydrostatic, quasi-hydrostatic, and nonhydrostatic ocean modeling. *Journal of Geophysical Research* 102, 5733–5752.
- Marshall, J., Jones, H., Hill, C., 1998. Efficient ocean modeling using non-hydrostatic algorithms. *Journal of Marine Systems* 18, 115–134.
- Marshall, J., Schott, F., 1999. Open ocean deep convection: observations, models and theory. *Reviews of Geophysics* 37, 1–64.
- Mellor, G. L., Yamada, T., 1974. A hierarchy of turbulent closure models for planetary boundary layers. *Journal of the Atmospheric Sciences* 31 (4), 1791–1806.
- Mellor, G. L., Yamada, T., 1982. Development of a turbulence closure model for geophysical fluid problems. *Reviews of Geophysics* 20 (4), 851–875.
- Molod, A., 2009. Running GCM physics and dynamics on different grids: algorithm and tests. *Tellus A* 61A (3), 361–395.
- Molod, A., Salmun, H., Waugh, D., 2004. The impact on a GCM climate of an extended mosaic technique for the land-atmosphere. *Journal of Climate* 17 (20), 3877–3891.
- Müller, P., Holloway, G., Henyey, F., Pomphrey, N., 1986. Nonlinear interactions among gravity waves. *Reviews of Geophysics* 24 (4), 493–536.
- Nguyen, A., Menemenlis, D., Kwok, R., 2009. Improved modeling of the arctic halocline with a sub-grid-scale brine rejection parameterization. *Journal of Geophysical Research* 114, C11014.
- Nurser, A. J. G., 1996. A review of models and observations of the oceanic mixed layer. Internal document No 14, Southampton Oceanography Centre, Southampton, UK, 247 pp.
- Price, J., Weller, R., Pinkel, R., 1986. Diurnal cycling: Observations and models of the upper ocean response to diurnal heating. *Journal of Geophysical Research-Oceans* 91 (C7), 8411–8427.
- Suarez, M., Trayanov, A., Hill, C., Schopf, P., Vikhliav, Y., 2007. MAPL: a high-level programming paradigm to support more rapid and robust encoding of hierarchical trees of interacting high-performance components. In: *CompFrame '07: Proceedings of the 2007 symposium on Component and framework technology in high-performance and scientific computing*. ACM, New York, NY, USA, pp. 11–20.
- Tao, W., Chern, J., Atlas, R., Randall, D., 2009. A multiscale modeling system: Developments, applications, and critical issues. *Bull Amer Met Soc*, 515–534.
- Wang, L., Ayala, O., Kasprzak, S., Grabowski, W., 2005. Collision efficiency of hydrodynamically interacting cloud droplets in turbulent atmosphere. *Journal of the Atmospheric Sciences* 62, 2433–2450.
- Wyant, M., Khairoutdinov, M., Bretherton, C., 2006. Climate sensitivity and cloud response of a GCM with a superparame-

terization. Geophys. Res. Lett 33, 4.