Article

Parameterization of Eddy Mass Transport in the Arctic Seas based on the Sensitivity Analysis of Large-scale Flows

Gennady Platov [1,*], Dina Iakshina [1] and Elena Golubeva [1]

[1] Institute of Computational Mathematics and Mathematical Geophysics, SB RAS, Novosibirsk, 630090, Russia; platov.g@gmail.com (G.P.); iakshina.dina@gmail.com (D.I.); e.golubeva.nsk@gmail.com (E.G.)
* Correspondence: platov.g@gmail.com

Abstract: The characteristics of the eddy mass transport are estimated depending on the values of the parameters of a large-scale flow that forms under the conditions of the shelf seas in the Arctic. For this, the results of numerical simulation of the Kara Sea with a horizontal resolution permitting the development of mesoscale eddies are used. The parameters resulting from numerical experiment are considered as a statistical sample and are analyzed using methods of sensitivity study and clustering of sample elements. Functional dependencies are obtained that are closest to the simulated distributions of quantities. These expressions make it possible, within the framework of large-scale models, to evaluate the characteristics of the cross-isobatic eddy mass transport in the diffusion approximation with a counter-gradient flux. Numerical experiments using the SibCIOM model showed that areas along the Fram branch of the Atlantic waters trajectory in the Arctic as well as the shelf of the East Siberian and Laptev seas with adjacent deep water areas are most sensitive to proposed parameterization of eddy exchanges. Accounting for counter-gradient eddy fluxes turned out to be less important.

Keywords: eddy mass transport; subgrid-scale processes; parametrization; sensitivity study; clustering

1. Introduction

One of the most important tasks in large-scale modeling of the oceans in the framework of climate models is an adequate description of subgrid-scale processes, that is, processes that, within the framework of the accepted horizontal and vertical resolution of the model, as well as due to a number of simplifying assumptions, cannot be explicitly described by numerical solution of the relevant set of differential equations. In such cases, one resorts to a parametric description of the large-scale consequences of such mesoscale processes.

Among these processes is the eddy transport of scalar quantities. The scales of eddy formations in the ocean are varying in a fairly wide range, and not all of them can be properly described. However, the result of the action of such mesoscale (and submesoscale) eddies, namely the exchange of properties of the waters involved in the movement, requires the search for additional possibilities for parameterization of these processes.

The most common method is the parametrization of eddy fluxes using the diffusion approximation, when large-scale diffusion fluxes are enhanced by a specially chosen diffusion coefficient. A uniform increase in the diffusion coefficient leads to a smoothing of the thermodynamic characteristics of the ocean, including cases of the absence or weak eddy motion. A more differentiated approach is associated with the introduction of the diffusion coefficient being variable in space and time. However, the question arises of how to recognize the presence or absence of an eddy activity with only large-scale flow characteristics in hands. For example, in his approach, Smagorinsky [1] uses the presence of sheer and divergence in the velocity field to calculate the viscosity coefficient associated with mesoscale eddies.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 December 2022 doi:10.20944/preprints202212.0226.v1

Eddy diffusion and viscosity models are introduced into the coarse models to simulate unresolved eddy-driven motions. This mechanism is often represented by some functional statements [3,4] that depend on the resolved flow properties.

The most used parametrization of the mesoscale effect is based on the eddy advection scheme [5–7], which consists in simulating the effect of baroclinic instability by flattening the isopycnal surfaces, which transfers an available potential energy toward the eddy kinetic energy of subgrid-scale motions. However, such a scheme reduces the total energy, since it does not take into account the reverse transfer of kinetic energy to large scales [8]. Usually, the development of parameterization schemes and the evaluation of their parameters occur regardless of the climate models in which they are eventually to be included. They are tested in field experiments and the test area is relatively small compared to the area covered by large-scale climate models. In addition, parametrizations usually contain parameters that are uncertain, that is, there is parametric and structural uncertainty [9].

After parameterization schemes are developed and included into a climate model, modelers tune the parameters to make model adequately simulate the known physical processes and/or the observations of them [10]. Recently, to tune parameters modelers use data-driven algorithms. This is due to big data accumulation and the rapid development of methods for processing them, including data assimilation methods [11,12], as well as machine learning methods [13]. A whole class of data-driven parameterizations has emerged [14,15], rather than using idealized theories. Statistical methods are also traditionally used to manage and analyze data. They can be used to integrate high-resolution targeted local modeling into a large-scale climate model, systematically learning from the results of the local model and quantifying uncertain parameters of large-scale modelling.

This article proposes to use the results of regional modeling based on a model that resolves mesoscale eddies and well-known statistical approaches to analyze the sensitivity of eddy fluxes in relation to the characteristics of large-scale motion. Using this approach, an attempt will be made to obtain a functional dependence of the eddy transport characteristics on large-scale ocean thermodynamical characteristics using the Kara Sea shelf model in the Arctic as an example. In recent studies a similar approach was used in [16] and [17,18] but the eddy-diffusivity/mass-flux approach (EDMF) was used to parameterize convection and planetary boundary layer in atmospheric models.

Finally, the obtained expressions for the parametrization dependences were used in the framework of large-scale modeling of processes in the Arctic and the North Atlantic using the coupled ice-ocean SibCIOM model. With its help, it was possible to evaluate the effectiveness of the developed parametrization and identify the main trends in the simulated state of the ocean, associated with the inclusion of cross-adiabatic eddy flows in such large-scale models of the Arctic region.

2. Materials and Methods

2.1. The model of the Kara Sea

SibPOM sigma coordinate shelf model [19,20], which is a modification of the Princeton Ocean Model (POM) [2], was used as such model. It includes the parameterization of vertical turbulent processes and the correction of the horizontal pressure gradient error due to sigma coordinate [21]. The simulation area is shown in Figure 1. The quasi-regular grid of the region is constructed on the basis of a rotated spherical coordinate system with the poles selected so that the new equator is the central axis of the Kara Sea, while the horizontal resolution, which, according to [22], allows reproducing large mesoscale eddies.

The Kara Sea model is nested in Arctic and North Atlantic coupled ice-ocean model SibCIOM [23]. The model embedding scheme is given earlier in [19]. The main idea is that the Laplace operator, which describes the thermal conductivity and diffusion of the salt in fine resolution model, was applied only to deviations of temperature and salinity from their large-scale distributions. Feedback in this case was not taken into account.

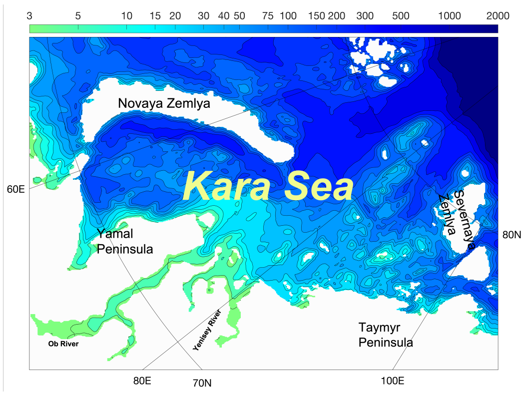


Figure 1. Model domain and its topography

The results discussed below refer to a numerical experiment covering the period from September 2006 to September 2008, the results of which were presented in our previous paper [20]. From this experiment, only 2007 was taken into account in our analysis below. Thus, we expected to incorporate the extremal features of this year when considering cross-isobatic transport.

2.2. Parametrization of cross-isobatic transport

The horizontal transport of density fluctuations by velocity fluctuations is described by the terms in the mass conservation equation of the form

$$-\frac{\partial}{\partial x}\overline{\rho' u'} - \frac{\partial}{\partial y}\overline{\rho' v'} = -\vec{\nabla} \cdot \overline{\rho' \vec{u}'}, \tag{1}$$

where $\vec{u}'$ is the horizontal velocity fluctuation vector with components $(u', v') = (u, v) - (\bar{u}, \bar{v})$, and $\bar{u}, \bar{v}$ are the regular components of the horizontal current velocity, that is, the averaged velocity components over a certain characteristic period of time T. Accordingly, the operator $\overline{(\cdot)}$ is the result of averaging the value over this time-period. Thus, the mass flux $\overline{\rho' \vec{u}'}$ in the direction of some vector $\vec{n}$ is defined as follows

$$-\overline{\rho' u'_n} = -\overline{\rho' \left(\vec{u}' \cdot \vec{n}\right)} = -\left(\left(\overline{\rho \vec{u}} - \bar{\rho}\bar{\vec{u}}\right) \cdot \vec{n}\right). \tag{2}$$

In numerical experiment with a detailed resolution the time-averaged values of both components $\bar{u}$ and $\bar{v}$ along with the value of the density $\bar{\rho}$ at each point of the grid area, as well as the time-averaged values of the flux componens $\overline{\rho u}$ and $\overline{\rho v}$, were stored for each 12 hour periods. Thus, any averaging period can be chosen with a resolution of 12 hours, that is, from 12 hours to several days. The time scale T characterizing the time of the influence of the mesoscale eddy on the state of the ocean at a certain point is about 10 days [24]. Therefore, hereafter we will consider T =10 days (averaging from the 1st to the 10th day of the month from 11 to 20 and from 21 to the end of the month), that is, about 20 12-hour records will be used for the averaging operation $\overline{(\cdot)}$. The most interesting direction of mass transport carried out by mesoscale eddies is the direction perpendicular to the geostrophic flow, that is, locally this direction coincides with the direction of ocean depth growth $\vec{n} = \vec{\nabla} H / \left|\vec{\nabla} H\right|$. Within the framework of the diffusion approximation, such an eddy flux is parameterized using large-scale characteristics in the form [25–27]

$$-\left(\overline{\rho' \vec{u}'} \cdot \vec{n}\right) = K\left(\frac{\partial \bar{\rho}}{\partial \vec{n}} + \gamma\right) = K\frac{\partial \bar{\rho}}{\partial \vec{n}} + q \tag{3}$$

using the eddy diffusion coefficient K and the counter-gradient γ due to which the counter-gradient flux $q = K\gamma$ is formed. Such a flux is formed in the absence of a change in density in the direction of the vector $\vec{n}$ due to an ordered vortex structure. To calculate the large-scale horizontal averaging, we will use a scale of Δ =50 km, since this approximately corresponds to the grid spacing of the large-scale model. The horizontal averaging of a certain value ϕ will be denoted by angle brackets $\langle\phi\rangle$ and mean the value at the point (x_0, y_0, z_0), which is obtained by averaging ϕ over all points of the model with a detailed resolution located in the square $\{-\Delta < x - x_0 < \Delta, -\Delta < y - y_0 < \Delta, z = z_0\}$. Thus, the values K and q can be found using the least squares method as linear regression coefficients

$$K\langle\frac{\partial\bar{\rho}}{\partial\vec{n}}\rangle + q = -\langle\overline{\rho'\vec{u'}}\cdot\vec{n}\rangle, \tag{4}$$

that is

$$K = -\frac{\langle\left(\overline{\rho'\vec{u'}}\cdot\vec{n}\right)\frac{\partial\bar{\rho}}{\partial\vec{n}}\rangle - \langle\overline{\rho'\vec{u'}}\cdot\vec{n}\rangle\langle\frac{\partial\bar{\rho}}{\partial\vec{n}}\rangle}{\langle\left(\frac{\partial\bar{\rho}}{\partial\vec{n}}\right)^2\rangle - \langle\frac{\partial\bar{\rho}}{\partial\vec{n}}\rangle^2}, \tag{5}$$

$$q = -\langle\left(\overline{\rho'\vec{u'}}\cdot\vec{n}\right)\rangle - K\langle\frac{\partial\bar{\rho}}{\partial\vec{n}}\rangle. \tag{6}$$

2.3. Large-scale representation of mesoscale characteristics

Any value ϕ can be represented as its linearization in the vicinity of a point (x_0, y_0, z_0) in horizontal coordinates $(x', y') = (x - x_0, y - y_0)$ in the form

$$\phi(x', y') \approx Ax' + By' + C = \langle\frac{\partial\phi(x_0, y_0)}{\partial x}\rangle x' + \langle\frac{\partial\phi(x_0, y_0)}{\partial y}\rangle y' + \langle\phi(x_0, y_0)\rangle, \tag{7}$$

where the coefficients A, B, and C can also be found by the least squares method, i.e.

$$A = \langle\frac{\partial\phi(x_0, y_0)}{\partial x}\rangle = \frac{\Lambda(\phi, x')\Lambda(y', y') - \Lambda(\phi, y')\Lambda(x', y')}{D}, \tag{8}$$

$$B = \langle\frac{\partial\phi(x_0, y_0)}{\partial y}\rangle = \frac{\Lambda(\phi, y')\Lambda(x', x') - \Lambda(\phi, x')\Lambda(x', y')}{D} \tag{9}$$

$$C = \langle\phi(x_0, y_0)\rangle = \langle\phi\rangle - A\langle x'\rangle - B\langle y'\rangle, \tag{10}$$

$$D = \Lambda(x', x')\Lambda(y', y') - \Lambda(x', y')^2, \tag{11}$$

where the following operator is introduced as a notation

$$\Lambda(\varphi, \psi) = \langle\varphi\psi\rangle - \langle\varphi\rangle\langle\psi\rangle. \tag{12}$$

As before, the designation $\langle\cdot\rangle$ here means such a value of the characteristic at the point (x_0, y_0, z_0), which is obtained by its averaging over all points of the model with a detailed resolution located in the square $\{-\Delta < x - x_0 < \Delta, -\Delta < y - y_0 < \Delta, z = z_0\}$.

2.4. Large-scale flow characteristics

The purpose of our analysis is to single out among the large-scale characteristics those on which the eddy mass transport in the cross-isobatic direction depends to a greater extent. That is, taking into account the parametrization dependence (3), we aim to find out on which large-scale characteristics of the flow the values K and q depend to the greatest extent. The following geographic and physical characteristics were considered:

1. Ocean depth $H(x, y)$;
2. Local depth z, which in terms of the topography following σ-coordinate system follows from the relation $z(x, y, \sigma) = \sigma H(x, y)$;
3. The value of the bottom slope $s = \left|\vec{\nabla}H\right|$ or $s = \left(\vec{\nabla}H \cdot \vec{n}\right)$, where $\vec{n} = \vec{\nabla}H/\left|\vec{\nabla}H\right|$;

 doi:10.20944/preprints202212.0226.v1

4. Component of the density gradient in the direction of the bottom slope $\partial\rho(x,y,\sigma)/\partial\vec{n} = \left(\vec{\nabla}\rho\cdot\vec{n}\right)$;
5. The density gradient component in the direction along the isobath, that is, along the vector $\vec{m} = (n_y, -n_x)$, where $(n_x, n_y) = \vec{n}$, that is $\partial\rho(x,y,\sigma)/\partial\vec{m} = \left(\vec{\nabla}\rho\cdot\vec{m}\right)$;
6. Component of the bottom density gradient in the direction of the slope $\partial\rho(x,y,\sigma=-1)/\partial\vec{n} = \left(\vec{\nabla}\rho\cdot\vec{n}\right)\Big|_H$;
7. Component of the bottom density gradient in the direction along the isobath $\partial\rho(x,y,\sigma=-1)/\partial\vec{m} = \left(\vec{\nabla}\rho\cdot\vec{m}\right)\Big|_H$;
8. The speed of the current in the direction of the slope $U = (\vec{u}\cdot\vec{n})$;
9. Flow velocity along the isobath $V = (\vec{u}\cdot\vec{m})$;
10. Divergence of the velocity component U in the direction of the slope $\partial U/\partial\vec{n}$;
11. Shift of the velocity component U in the direction of the isobath $\partial U/\partial\vec{m}$;
12. Shift of the velocity component V in the direction of the slope $\partial V/\partial\vec{n}$;
13. Divergence of the velocity component V in the direction of the isobath $\partial V/\partial\vec{m}$;
14. Vertical component of the density gradient $\partial\rho(x,y,\sigma)/\partial z = \partial\rho(x,y,\sigma)/(H(x,y)\partial\sigma)$;
15. Geographic latitude $\theta(x,y)$;
16. Geographic longitude $\lambda(x,y)$.

For each of these values at each grid point of the high spatial resolution model, the corresponding large-scale value can be found by using $\overline{(\cdot)}$ and $\langle\cdot\rangle$ operators described previously. Based on the results of modeling the Kara Sea using the SibPOM [20,28] model, about 78 million records of these values and the corresponding values of K and q were obtained. However, later on, the seven most independent values were selected from 16 values (see below).

2.5. Independence of large-scale characteristics

It will be further assumed that K and q are functions of several large-scale characteristics. The sensitivity analysis of K and q in relation to these characteristics provides for their statistical independence of each other. That is, 16 selected characteristics will be considered as 16 independent variables on which the value of these functions depends. An analysis of the 78 million records mentioned earlier showed that, based on the Fisher test, none of these variables is independent. Even latitude θ and longitude λ are not statistically independent because, for example, due to the shape of the basin, latitude cannot take on certain values at some fixed longitude. The same applies to local depth z. However, it was possible to rule out the characteristics that are most dependent on others.

1. The depth of the ocean $H(x,y)$ turned out to be strongly related to the values $\partial\rho/\partial\vec{n}$, $\partial\rho/\partial z$ and θ. Being, in principle, an independent value, and it can hardly be assumed that the depth of the ocean depends on $\partial\rho/\partial\vec{n}$ or $\partial\rho/\partial z$, rather, vice versa. Nevertheless, it was neglected, since it is better to have a connection with physical state characteristics such as $\partial\rho/\partial\vec{n}$ or $\partial\rho/\partial z$, and not with geographical ones, especially since the former are strongly depend on the latter. Thus, rejecting the ocean depth as an independent variable, it is assumed that the dependence of the values K and q on it can be replaced by the dependence on $\partial\rho/\partial\vec{n}$ and $\partial\rho/\partial z$ having a close relationship with the ocean depth.
2. Local depth z, just as an ocean depth H, is in principle an independent variable. However, it turned out to be strongly related to physical values $\partial\rho/\partial\vec{n}$ and $\partial\rho/\partial z$, therefore, it was also neglected, since instead of a geographical location it is better to deal with the physical characteristics associated with it.
3. The component of the density gradient in the direction of the slope $\partial\rho/\partial\vec{n}$ turned out to be strongly dependent on the value of this characteristic near the bottom $\partial\rho/\partial\vec{n}|_H$. Bearing in mind that, according to [29], the latter is an important characteristic for the formation of cascading, when the bottom density decreases with increasing depth,

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 December 2022 doi:10.20944/preprints202212.0226.v1

the first value was neglected, since it is largely explained by the second, which has a clearer physical meaning.

4. Similarly, the near-bottom density gradient component along the isobath $\partial\rho/\partial\vec{m}|_H$ turned out to be strongly dependent on the local value of this gradient $\partial\rho/\partial\vec{m}$, so this value was neglected in favor of the second one.
5. The flow velocity in the direction of the slope U turned out to be dependent on the density gradient in this direction $\partial\rho/\partial\vec{n}$ and on the magnitude of the velocity along the isobaths V and its variability along the slope $\partial V/\partial\vec{n}$.
6. The flow velocity in the direction of the isobaths V depends on the density gradient in the direction of the slope $\partial\rho/\partial\vec{n}$, on the steepness of that slope s, and is closely related to the V variability along the slope $\partial V/\partial\vec{n}$.
7. The divergence of the velocity component U in the direction of the slope $\partial U/\partial\vec{n}$ is related to the divergence of the flow along the isobaths $\partial V/\partial\vec{m}$. This relationship is based on the continuity equation.
8. The dependence on latitude θ and longitude λ was also neglected, since the goal is to be tied to physical processes, and not to a specific geographical location.

As a result, the following seven values in the large-scale approximation were considered as variables on which the values K and q can depend: s, $\partial\rho/\partial\vec{n}|_H$, $\partial\rho/\partial\vec{m}$, $\partial U/\partial\vec{m}$, $\partial V/\partial\vec{n}$, $\partial V/\partial\vec{m}$ and $\partial\rho/\partial z$.

2.6. Clustering

The total sample, built on the results of a fine resolution simulation, contains elements consisting of a set of parameters characterizing the large-scale motion

$$\left(s, \left.\frac{\partial\rho}{\partial\vec{n}}\right|_H, \frac{\partial\rho}{\partial\vec{m}}, \frac{\partial U}{\partial\vec{m}}, \frac{\partial V}{\partial\vec{n}}, \frac{\partial V}{\partial\vec{m}}, \frac{\partial\rho}{\partial z}\right)$$

and parameters describing the integral effect of mesoscale pulsations on the large-scale motion (K, q). Since the nature of mesoscale movements can be completely different and refer to completely unrelated physical mechanisms, it makes sense to divide in the way it was done in [30,31] the entire sample into clusters, that is, into groups of the most closely related sample elements.

There are a number of approaches related to the choice of the criterion for the tightness of the connection between elements. In this study, the so-called k-means method was used [32], in which belonging to a cluster is determined by the fact that the distance to its center is minimal among the centers of all clusters. The clustering procedure is iterative, after determining the belonging of elements to clusters, the center of each cluster is redefined in accordance with which elements are included in it. The iterations stop as soon as the composition of the clusters becomes unchanged.

An important issue in the implementation of the k-means method is the choice of the number of clusters k and the initial position of their centers. The choice of the number of clusters is based on the fact that the source of mesoscale motions can be barotropic or baroclinic instabilities in the region of jet streams or near density fronts, as well as in regions of intense convective and wind mixing. Having considered the values k from 2 to 8 as options, it was decided to stop at the value of $k = 3$, due to the fact that the resulting clusters in this case are more cohesive according to the relevant criteria, and also have a clear geographical localization, indicating a certain nature associated with this cluster physical processes (see below).

The values of the parameters characterizing the large-scale movement is a vector $(x_1, x_2, \ldots, x_N)$ in N-dimensional space, where N is equal to the number of these parameters (in our case, $N = 7$). Since the parameters are heterogeneous in nature, normalization

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 December 2022 doi:10.20944/preprints202212.0226.v1

is used to achieve their equivalence, that is, instead of a vector $(x_1, x_2, \ldots, x_N)$, modified vector $(X_1, X_2, \ldots, X_N)$ is used, where

$$X_i = \frac{x_i - \overline{x_i}}{\sigma_i}, \;\; i = 1, \ldots, N \tag{13}$$

and

$$\overline{x_i} = \frac{1}{M}\sum_{j=1}^{M} x_{i,j}, \;\; \sigma_i = \sqrt{\frac{\sum_{j=1}^{M}\left(x_{i,j} - \overline{x_i}\right)^2}{M}}, \tag{14}$$

M is the length of the sample, $x_{i,j}$ is the value of the i-th parameter in the j-th element of the sample. After normalization, the coordinates of the center of the l-th cluster are defined as

$$\overrightarrow{R_l} = (R_{1,l}, R_{2,l}, \ldots, R_{N,l}), \;\; R_{i,l} = \frac{1}{M_l}\sum_{j \in S_l} X_{i,j}, \tag{15}$$

where S_l is the set of sample elements belonging to the l-th cluster, M_l is the number of these elements, $X_{i,j}$ is the value of the normalized i-th parameter in the j-th sample element. After finding the centers of clusters, the belonging of the sample elements to clusters is redefined as

$$\overrightarrow{X_j} = \left(X_{1,j}, X_{2,j}, \ldots, X_{N,j}\right) \in S_l \tag{16}$$

if

$$l: \;\; r_{j,l} = \min_{p=1}^{k} r_{j,p}, \;\; \text{where } r_{j,p} = \sqrt{\sum_{i=1}^{N}\left(X_{i,j} - R_{i,p}\right)^2}. \tag{17}$$

The initial position of cluster centers is determined by following the k-means++ [33] algorithm:

1. The center of the first cluster is determined randomly.
2. The center of the second cluster is determined randomly with a probability proportional to the distance to the center of the first cluster.
3. The center of the next i-th cluster is also determined randomly with a probability proportional to the minimum among the distances to the known cluster centers.

Conventional indices are used to determine the quality of clustering and search for the most appropriate partition.

2.6.1. Davies-Bouldin Index

The index is calculated using the formula [34]:

$$DB = \frac{1}{k}\sum_{l=1}^{k} \max_{m \neq l}\left(\frac{\sigma(l) + \sigma(m)}{\left|\overrightarrow{R_l} - \overrightarrow{R_m}\right|}\right), \tag{18}$$

where k, as before, is equal to the number of clusters, $\left|\overrightarrow{R_l} - \overrightarrow{R_m}\right|$ is the distance between l-th and m-th cluster centers,

$$\sigma(n) = \sqrt{\frac{1}{M_n}\sum_{j \in S_n}\left|\overrightarrow{X_j} - \overrightarrow{R_n}\right|^2}, \;\; n = l, m, \tag{19}$$

where the notation introduced earlier is used, and the operation $|\cdot|$ determines the distance by the formula

$$\left|\overrightarrow{X_1} - \overrightarrow{X_2}\right| = \sqrt{\sum_{i=1}^{N}(X_{i,1} - X_{i,2})^2}. \tag{20}$$

The numerator of each term in the expression (18) for the index DB depends on the distances within the clusters, and the denominator is equal to the distance between their centers. Thus, the smaller the distances within a cluster compared to the distances to other clusters, the better clustering is considered. Therefore, the clustering with the smallest index is the most preferable.

2.6.2. Dunn index

This criterion is based on the index calculated by the formula [35]:

$$D = \frac{\min\limits_{1 \le l < m \le k} \left| \overrightarrow{R_l} - \overrightarrow{R_m} \right|}{\max\limits_{1 \le n \le k} \sigma(n)}, \tag{21}$$

that is, on the contrary, the numerator depends on the distance between the cluster centers, and the denominator depends on the distances within the clusters. Therefore, clustering with the highest index value is considered more preferable in this case.

2.6.3. Silhouette coefficients

For each sample element i, two values are calculated [36]:

$$a(i) = \frac{1}{M_l - 1} \sum_{j \in S_l, j \ne i} \left| \overrightarrow{X_i} - \overrightarrow{X_j} \right|, \tag{22}$$

$$b(i) = \min_{m \ne l} \frac{1}{M_m} \sum_{j \in S_m} \left| \overrightarrow{X_i} - \overrightarrow{X_j} \right|, \tag{23}$$

where S_l is the set of elements of the l-th cluster to which this element belongs, $a(i)$ is the average distance from this element to the remaining elements of this cluster, $b(i)$ is the minimum of the average distances to elements of other clusters. The silhouette coefficient of a given sample element is determined using these values as follows:

$$s(i) = \frac{b(i) - a(i)}{\max[b(i), a(i)]}, \tag{24}$$

so that the value of this coefficient is $-1 \le s(i) \le 1$, and if the distances within a cluster are negligibly small compared to the distances between clusters, then $s(i) \to 1$, and if vice versa, then $s(i) \to -1$. That is, in this case, the highest values of the silhouette coefficients are more preferable. In our analysis, we consider the mean values of the silhouette coefficients for each cluster $S(l) = \frac{1}{M_l} \sum_{i \in S_l} s(i)$, and the silhouette coefficient for the entire clustering $\bar{S}^k = \frac{1}{M} \sum_{i=1}^{M} s(i)$ [37].

To speed up the calculation of these coefficients, randomly selected M_0 elements of each cluster were used, that is, it was assumed that $S(l) \approx S^*(l) = \frac{1}{M_0} \sum_{i=i_1}^{i_{M_0}} s(i)$, where $i_1, i_2, \ldots, i_{M_0} \in S_l$ are randomly selected elements of the l-th cluster. Assuming that $S^*(l)$ is an estimate for $S(l)$, the estimate for the coefficient $\bar{S}^k$ can be obtained from the formula $\bar{S}^k \approx \frac{1}{M} \sum_{l=1}^{k} M_l \cdot S^*(l)$.

2.7. Analysis of dependencies on the selected parameters

The values of eddy diffusion coefficient K and counter-gradient flux q resulting from fine resolution simulations vary in a fairly wide range. The variability of the K value is several orders of magnitude. Therefore, in order to narrow the range, the logarithm of this value was considered, made dimensionless with the help of some characteristic value K_0, thus, instead of the eddy diffusion coefficient, the value $\log \frac{K}{K_0}$ was considered, while only the values of $K > 1$ m^2/s were taken into account. The variability of the counter-gradient eddy flux q is also several orders of magnitude and, at the same time, has both positive

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 December 2022 doi:10.20944/preprints202212.0226.v1

and negative values in a comparable proportion. In addition, it takes most of the values in a fairly narrow range. Therefore, to isolate the region of the most frequent values of the characteristic, instead of the eddy flux itself, the following function was considered $\tanh\left((q-\bar{q})/\sigma(q)\right)$ where $\bar{q}$ is the sample mean value of the flux, and $\sigma(q)$ is its standard deviation. The use of hyperbolic tangent narrows the range of values to the segment $[-1, 1]$, while the values in the central part, represented by the overwhelming number of sample elements, experience only a linear normalization transformation.

Following the Sobol' method [38,47], we represent the dependence of the value $Y = \log\frac{K}{K_0}$ or $Y = \tanh\left((q-\bar{q})/\sigma(q)\right)$ on the N parameters of the large-scale model $(Z_1, Z_2, \ldots, Z_N)$ in the form

$$Y = f_0 + \sum_i f_i + \sum_{i<j} f_{ij} + \ldots + \sum_{\substack{i_1<\ldots<i_m \\ 2<m<N}} f_{i_1\ldots i_m} + f_{1\ldots N}, \tag{25}$$

where

$$\begin{aligned} f_0 &= E(Y), \\ f_i &= f_i(Z_i) = E(Y|Z_i) - f_0, \\ f_{ij} &= f_{ij}(Z_i, Z_j) = E(Y|Z_i, Z_j) - f_0 - f_i - f_j, \\ &\ldots \\ f_{i_1\ldots i_m} &= E(Y|Z_{i_1}, \ldots, Z_{i_m}) - f_0 - \sum_{i\in[i_1,\ldots,i_m]} f_i - \sum_{\substack{j_1<\ldots<j_k \\ 1<k<m \\ j_1,\ldots,j_k\in[i_1,\ldots,i_m]}} f_{j_1\ldots j_k}. \end{aligned} \tag{26}$$

Here E denotes the mathematical expectation of the value obtained as a result of the statistical evaluation, and $Y|Z_1, \ldots$ denotes the Y value at fixed values of the $Z_1, \ldots$ values. According to [38], the values Z_i are assumed to be uniformly distributed over the interval [0,1]. Therefore, we will divide the entire range of X_i changes into L boxes with an equal number of elements in each of them. Thus, as Z_i values, we can consider values that are discrete on a segment $[0, L]/L$ and obtained as a result of the distribution of X_i elements over L successive intervals or boxes each containing M/L of the sample elements.

3. Results

3.1. Clustering results

Due to the fact that the process of selecting clusters is random at the stage of initial separation, the result may not be optimal. Therefore, for each k value from 2 to 8, 25 clusterings were carried out, among which the variant with the optimal values of the above indices was selected. Naturally, under such conditions, we are still not guaranteed that the partition will be optimal, but the chance of optimality increases significantly. From the analysis of the resulting partitions, it follows that the process of selecting clusters has about 2-3 limit states, so choosing the most optimal one is not a problem.

Table 1 presents the results of the performed clusterings with the number of clusters identified from $k = 2$ to 5, and Figures 2 and 3 show the geographical location of the clusters and their depth distribution. The table shows that in the case of selecting 2 clusters, their size (the number of elements in them) differs by an order of magnitude. A large cluster is presented, covering 91% of all elements, and a small one, including 9% of the elements. Geographically, the former is distributed throughout the entire Kara Sea basin over the entire range of depths. The second is localized in the coastal part in the places where the waters of Siberian rivers spread, and in depth it is located in the upper layer within a few tens of meters. The latter, in our opinion, indicates that the elements of this cluster correspond to a set of parameters that describe the specifics of the distribution of the river plume and the development of the salinity front in the upper layer of the sea.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 December 2022 doi:10.20944/preprints202212.0226.v1

Table 1. The results of clustering the selected parameters of large-scale movement for the number of clusters from 2 to 5.

Number of clusters	No.	Cluster percentage	DB	D	$\bar{S}^k$ %	$S^*(l)$ %
2	1	91	1.62	0.953	78	95
	2	9				-89
3	1 (A)	88	1.37	0.958	74	96
	2 (B)	9				-90
	3 (C)	3				-94
4	1	81	1.70	0.526	57	91
	2	13				-84
	3	3				-93
	4	3				-95
5	1	51	1.73	0.376	3	46
	2	35				-22
	3	9				-81
	4	3				-89
	5	3				-93

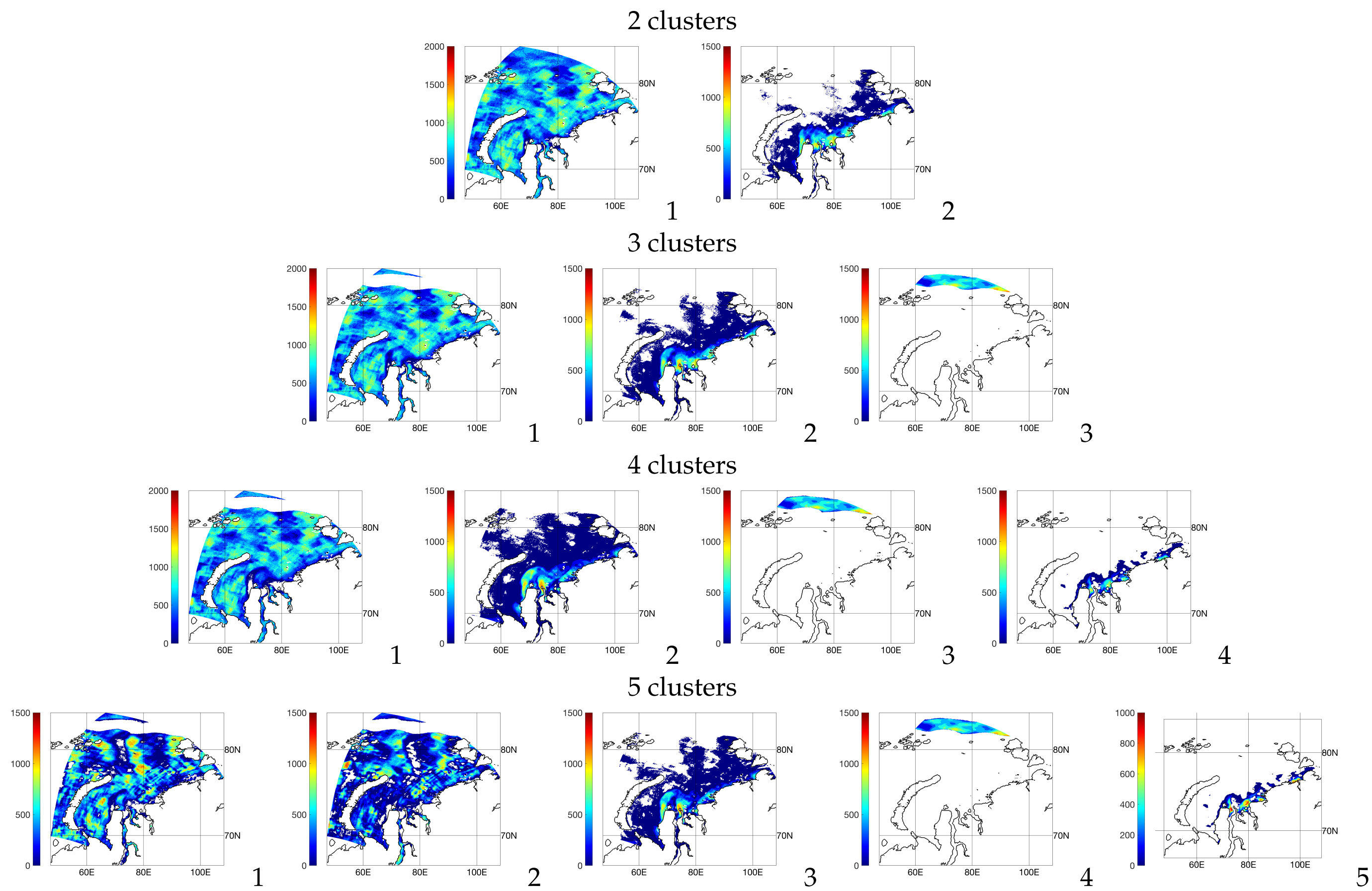


Figure 2. Geographical localization of clusters when split into 2, 3, 4 and 5 clusters: to the right of the panels is the cluster number in this split in descending order of the number of elements. Color represents number of elements per model grid cell.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 December 2022 doi:10.20944/preprints202212.0226.v1

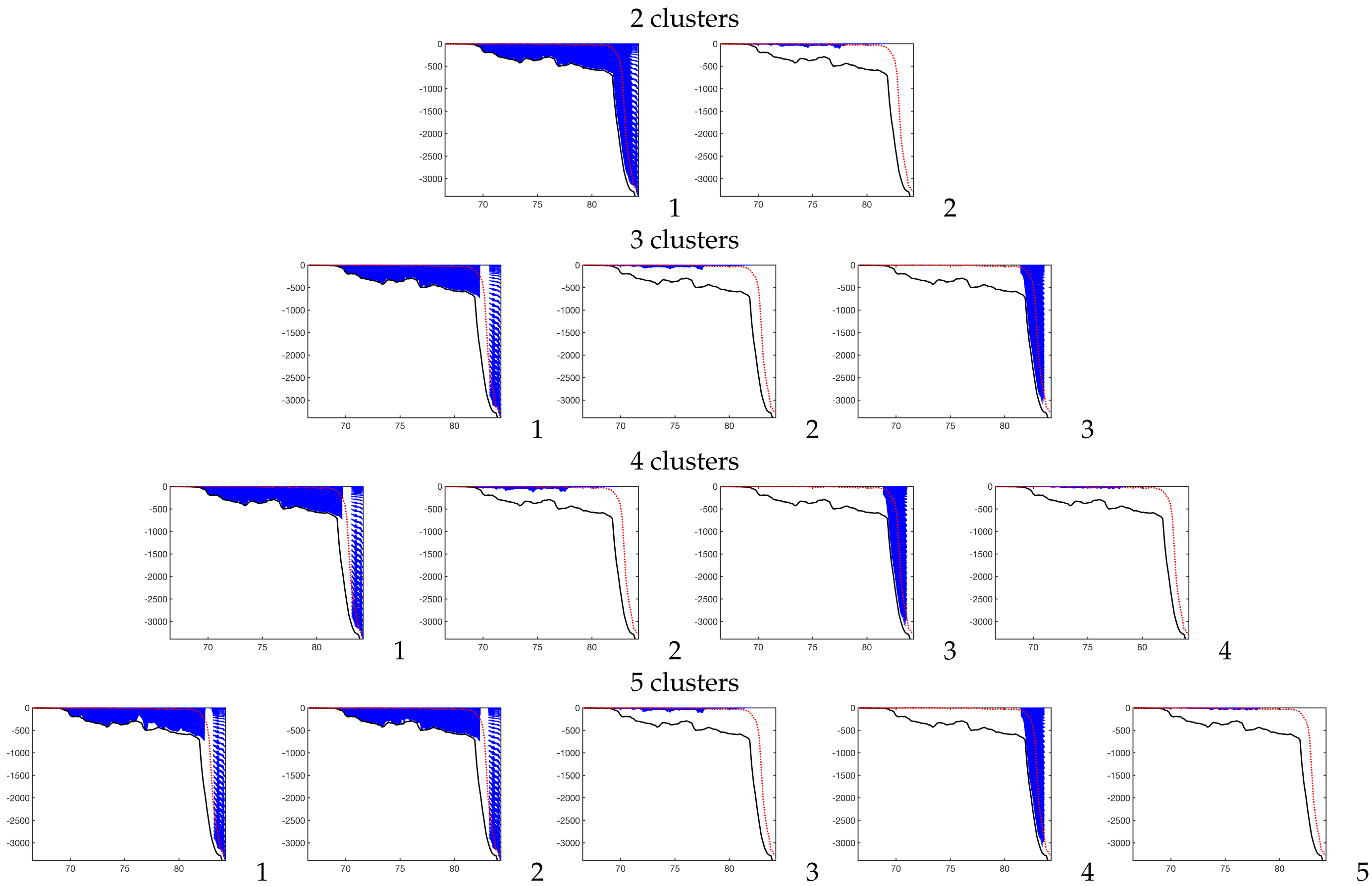


Figure 3. Localization of clusters in depth (vertical axis) depending on latitude (horizontal axis) when divided into 2, 3, 4 and 5 clusters: the labels are similar to Figure 2, the solid black line shows the maximum depth for different latitudes of the Kara Sea, the dotted red line shows the minimum depth.

With an increase in the number of clusters to three, the previously mentioned cluster with river influence remains practically unchanged and makes up the same 9% of the elements. For brevity, we will denote this cluster as B. An even smaller cluster is separated from the large cluster, covering only 3%. We will denote it with the symbol C. Thus, the share of a large cluster has decreased to 88%. This cluster will be denoted by the symbol A. Cluster C turned out to be geographically localized in a narrow strip of the steepest slope at the boundary between the shelf and the deep ocean. Thus, this cluster contains elements with parameters that describe the specifics of mesoscale movements in the region of a steep shelf slope. This partition has the smallest among the considered Davies-Bouldin index and the largest Dunn index, which indicates its optimality according to these criteria. Its silhouette coefficient turned out to be somewhat smaller compared to splitting into two clusters (0.74 instead of 0.78), but the silhouette coefficient of the largest cluster slightly increased from 0.95 to 0.96. This partition will be further considered as the main one.

An increase in the number of clusters to four and five leads to the formation of an intermediate cluster associated with the upper layer no deeper than 100 m due to a decrease in the proportion of the large cluster A and the proportion of the river cluster B, the latter becomes the smallest as a result and covers the areas immediately adjacent to the river mouths. Our interpretation of the intermediate cluster is to identify areas of convective and wind mixing. In addition, a large cluster also splits into two, while part of the intermediate cluster is captured. As a result, a reduced version is formed from its remaining part.

The introduced designations of clusters for reference are presented in Table 1.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 December 2022 doi:10.20944/preprints202212.0226.v1

3.2. Parametrization dependencies. Cluster A

We restrict our analysis of the resulting partition to the consideration of the largest obtained cluster A, leaving consideration of other clusters and other partitions for future research.

The dependence of the eddy diffusion coefficient K on large-scale parameters was estimated on the basis of dividing the variability ranges of each of the parameters into $L = 5$ successive intervals with an equal number of sample elements in them. Dividing by more intervals results in a lot of computation when evaluating sensitivity, because the number of boxes in N-dimensional hyperspace is L^N. In the considered case $N = 7$ their number is $5^7 = 78125$.

Cluster A covers 88% of all sample elements. The strongest dependence of the coefficient of eddy diffusion K in the representation of eddy mass transport (3) was revealed on the value of the vertical derivative of the density, which is related to the Brunt-Väisälä frequency N_B by the relation

$$\frac{\partial \rho}{\partial z} = -\frac{\rho}{g} N_B^2. \tag{27}$$

This value is involved in explaining 69% of the variability in the value $\log \frac{K}{K_0}$ in this cluster.

The second most important value is the rate of density change along the direction of the bottom slope, which is responsible for 47% of the variability. The third value, which explains 39% of the variability, is the rate of the bottom slope. It is from these values that the individual dependence of the value $\log \frac{K}{K_0}$ on one-dimensional functions in representation (25) is most pronounced: $\frac{\partial \rho}{\partial z}$ – 33.4%, $\left.\frac{\partial \rho}{\partial \vec{n}}\right|_H$ – 7.6%, and s – 1.7%. The distributions of the cluster elements against these values are presented in Figure 4 using histograms in terms of their normalized deviations from the mean. The role of other parameters in explaining the dependence is not negligible and varies from 24 to 33%. Among the two-dimensional functions, the most pronounced dependence on pairs of variables is $\left(\frac{\partial \rho}{\partial z}, \left.\frac{\partial \rho}{\partial \vec{n}}\right|_H\right)$ – 3% and $\left(\left.\frac{\partial \rho}{\partial \vec{n}}\right|_H, s\right)$ – 1.3%. For multivariate functions, it can be noted that any combination of six variables explains from 0.8 to 1.6%, totaling about 8%, and the term with a function of all variables explains 2.8% of the variability. The analysis of such multidimensional dependences requires extensive theoretical substantiation, and in our analysis we will restrict ourselves mainly to one-dimensional dependences.

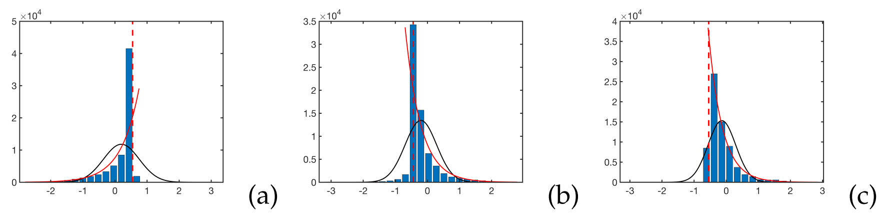
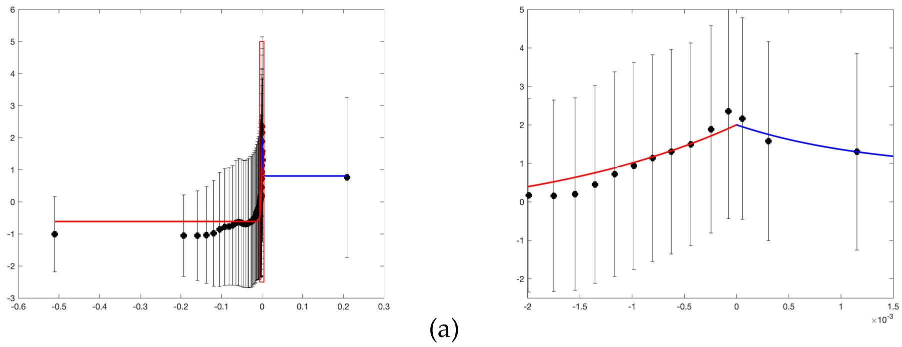


Figure 4. Histograms of the distribution of the number of cluster elements depending on the values of parameters a) $\frac{\partial \rho}{\partial z}$, b) $\left.\frac{\partial \rho}{\partial \vec{n}}\right|_H$ and c) s relative to their averages. The horizontal intervals show the deviation from the mean in units of standard deviation, the vertical axes shows the number of sample elements in thousands. The red dotted line shows the shift of the mean relative to the zero value of the parameters. For comparison, the black solid line shows the corresponding normal distribution, and the red solid line shows the exponential distribution.

Having obtained an estimate of the sensitivity of the eddy flux characteristics with respect to the selected set of parameters, we can now refine the dependence on them by increasing the number of intervals in the discretization of variables. The dependence on the vertical component of the density gradient $\frac{\partial \rho}{\partial z}$ is shown in Figure 5 for the case of dividing the range of variability into $L = 50$ intervals.

 doi:10.20944/preprints202212.0226.v1



(a) (b)

Figure 5. One-dimensional dependence of the value $\log \frac{K}{K_0}$ on the parameter $\frac{\partial\rho}{\partial z}$ in cluster A: (a) deviation of average values for each interval from the average for the cluster when divided into 50 intervals with an equal number of cluster elements (horizontal value in kg/m^4); (b) an enlarged view of the group of dots in a rectangle shown in (a). Panels contain the closest curves for negative and positive values of the argument (red and blue solid lines). The vertical bars show the standard deviation for each interval.

The histogram of the distribution of sample elements depending on $\frac{\partial\rho}{\partial z}$ (Figure 4a) has the form of an exponential distribution with only a small number of elements that go into the region of positive values (unstable stratification), and the highest concentration of elements is observed in the region of zero value. In this case, it makes sense to look for dependence in the form of an exponent for values $\frac{\partial\rho}{\partial z}$ less than zero and for values greater than zero. This gives a tendency towards some extreme value in the case of neutral stratification, and a limiting value when $\frac{\partial\rho}{\partial z}$ tending to infinity. The value K_0 is defined so that $\log K_0$ is equal to the average value of the value $\log K$ in the given cluster. In the case under consideration, $\log K_0 = 5.98$, which corresponds to $K_0 = 394$ m^2/s. Figure 5 shows the exponential curves with the help of red and blue lines, which most closely describe the $\log \frac{K}{K_0}$ behavior in the region of small absolute values of $\frac{\partial\rho}{\partial z}$ and somewhat worse in the case of large values:

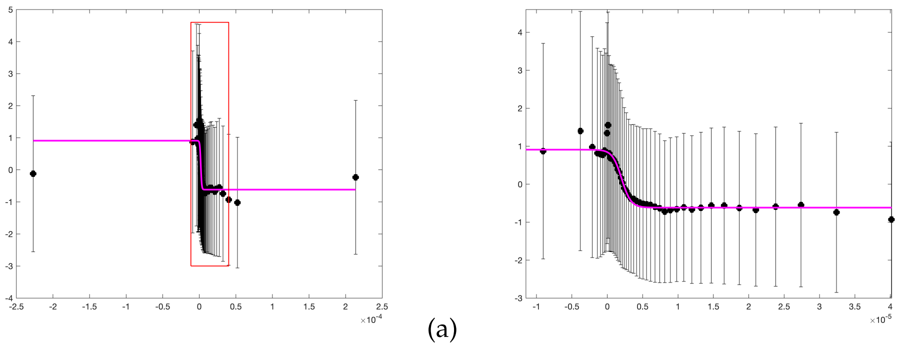
$$\left(\log \frac{K}{K_0}\right)_{\frac{\partial\rho}{\partial z}} = \begin{cases} 2.614 \cdot \exp\left(\dfrac{\frac{\partial\rho}{\partial z}}{2.093 \cdot 10^{-3}}\right) - 0.611, & \text{if } \dfrac{\partial\rho}{\partial z} < 0 \\ 1.195 \cdot \exp\left(-\dfrac{\frac{\partial\rho}{\partial z}}{1.302 \cdot 10^{-3}}\right) + 0.808, & \text{if } \dfrac{\partial\rho}{\partial z} \geq 0 \end{cases}, \qquad (28)$$

where the assumed $\frac{\partial\rho}{\partial z}$ dimension is expressed in kg/m^4. These dependencies have a standard deviation for the range of negative values of the argument $\sigma = 0.262$, and for positive values $\sigma = 0.138$. This means that the coefficient K is determined up to a factor of $e^{0.262} = 1.3$ in the first case and $e^{0.138} = 1.15$ in the second. Alternatively, one can also use tabular values corresponding to the depicted points in Figure 5.

The dependence on the parameter $\left.\frac{\partial\rho}{\partial \vec{n}}\right|_H$ is shown in Figure 6. It is interesting that the range of this parameter also includes negative values (see also Figure 4b), that is, the bottom density decreases along the sloping bottom. This is a characteristic condition for the formation of cascading [29]. As we have shown earlier in [28], the movement of dense waters along the sloping bottom is accompanied by the active generation of mesoscale eddies due to the released potential energy.

We searched for the curve closest to the given arrangement of points in Figure 6 in the form of a hyperbolic tangent by adjusting its slope, the position of the center of symmetry,

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 December 2022 doi:10.20944/preprints202212.0226.v1



(a) (b)

Figure 6. Same as Figure 5, but depending on the derivative of the bottom density in the direction of the slope $\left.\frac{\partial\rho}{\partial\vec{n}}\right|_H$. Panels contain the closest functional dependence curve.

and reaching a plateau for the parameter values to the right and left of the center. The following curve turned out to be optimal in this class

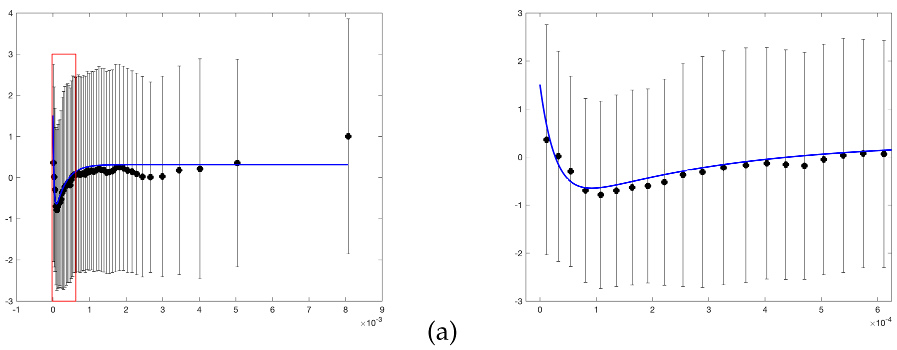
$$\left(\log\frac{K}{K_0}\right)_{\left.\frac{\partial\rho}{\partial\vec{n}}\right|_H} = 0.145 - 0.763\cdot\tanh\left(\frac{\left.\frac{\partial\rho}{\partial\vec{n}}\right|_H - 2.093\cdot 10^{-6}}{1.306\cdot 10^{-6}}\right), \tag{29}$$

Here, as before, the dimension of $\left.\frac{\partial\rho}{\partial\vec{n}}\right|_H$ is expressed in $\mathrm{kg/m^4}$. The value of the standard deviation from the averages over the intervals is $\sigma = 0.231$, that is, taking into account the logarithmic dependence, the value K is determined with an accuracy of up to a factor or divisor of $e^{0.231} = 1.26$.

The one-dimensional dependence on the bottom slope s explains only 1.7% of the variability in $\log K$. The dependence on this parameter is shown in Figure 7. We looked for the closest fitness to the location of the points in the form of a linear combination of two exponentials: the first with a slow decay to provide a general dependence, and the second with a fast decay to ensure growth near small slope values. In this representation, the following curve turned out to be optimal in the sense of the smallest standard deviation

$$\left(\log\frac{K}{K_0}\right)_s = 0.317 - 1.474\cdot\exp\left(-\frac{s\cdot 10^4}{2.878}\right) + 2.666\cdot\exp\left(-\frac{s\cdot 10^4}{0.285}\right), \tag{30}$$

The value is dimensionless and represents the increase in depth when moving along the slope per unit length (for example, m/m). The standard error in this formula is $\sigma = 0.177$, which gives the value a multiplier or divisor of $e^{0.177} = 1.19$.



(a) (b)

Figure 7. Same as Figure 5, but depending on the slope of the bottom s.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 December 2022 doi:10.20944/preprints202212.0226.v1

The dependence of the characteristic $\tanh \frac{q-\bar{q}}{\sigma(q)}$ in representation (25) turned out to be not so pronounced. The mean value and standard deviation in cluster A for the counter-gradient flux turned out to be $\bar{q} = -0.0467$ kg/(m^2·s) and $\sigma(q) = 473$ kg/(m^2·s). The same parameters as before, but in a different order $\left.\frac{\partial\rho}{\partial\vec{n}}\right|_H$, s and $\frac{\partial\rho}{\partial z}$ turned out to be most useful in explaining the variability of this value at the level of 69%, 65% and 53%, respectively. However, the one-dimensional dependence on these parameters explains only 7.6%, 3.4%, and 2.2% of the variability. The largest contribution up to 20%, as it turned out, is given in the aggregate by functions of five variables, while one-dimensional ones give only 16%, and two-dimensional ones only 12%. The total proportion of explained variability was 98.7%.

To get a more detailed picture of the dependence on these three parameters, we reduced the number of variables considered to $N = 3$, while increasing the number of intervals to $L = 50$, so that the total number of boxes became $50^3 = 125000$. As a result, the total part of explained variability decreased to 78.9%, but the dependence pattern became clearer. The parameter $\left.\frac{\partial\rho}{\partial\vec{n}}\right|_H$ is involved in explaining 92% of the variability of the value $\tanh \frac{q-\bar{q}}{\sigma(q)}$, the parameter $\frac{\partial\rho}{\partial z}$ – 78%, and s – 63%. The one-dimensional dependence on $\left.\frac{\partial\rho}{\partial\vec{n}}\right|_H$ explains 14% of the variability of this value, the $\frac{\partial\rho}{\partial z}$ contribution is 4%, the s contribution is less than 1%. Among the two-dimensional dependences, the maximum contribution is made by the dependence on the pair $\left(\left.\frac{\partial\rho}{\partial\vec{n}}\right|_H, \frac{\partial\rho}{\partial z}\right)$ – 19%. The rest give $\left(\left.\frac{\partial\rho}{\partial\vec{n}}\right|_H, s\right)$ – 8%, $\left(\frac{\partial\rho}{\partial z}, s\right)$ – 3%. The three-dimensional dependence on all three parameters explains 51% of the variability.

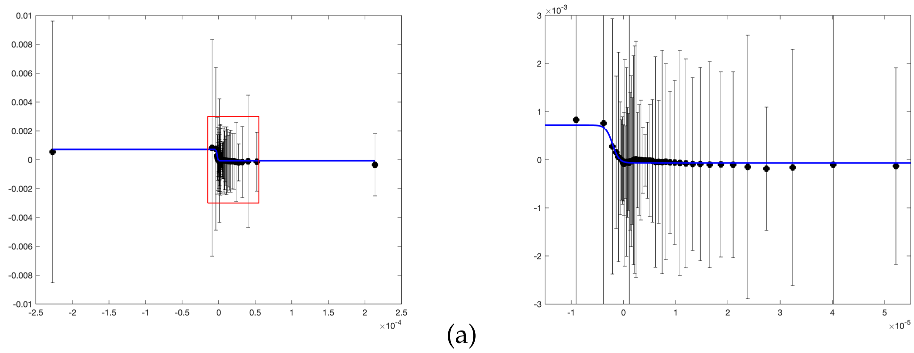
Next, we aim to consider one-dimensional functions of the parameters $\left.\frac{\partial\rho}{\partial\vec{n}}\right|_H$ and $\frac{\partial\rho}{\partial z}$, and neglect the dependance on s. Figure 8 shows a one-dimensional dependence of $\tanh \frac{q-\bar{q}}{\sigma(q)}$ on the parameter $\left.\frac{\partial\rho}{\partial\vec{n}}\right|_H$ expressed in units (kg/m^4). The closest functional dependence was sought in the form of a hyperbolic tangent, which resulted in the expression

$$\left(\tanh \frac{q-\bar{q}}{\sigma(q)} - f_0\right)_{\left.\frac{\partial\rho}{\partial\vec{n}}\right|_H} = \\ = 10^{-4} \cdot \left[3.269 - 3.930 \cdot \tanh\left(\frac{10^6 \cdot \left.\frac{\partial\rho}{\partial\vec{n}}\right|_H - 2.093}{1.269}\right)\right], \quad (31)$$

where $f_0 = 1.058 \cdot 10^{-4}$ is equal to the average value of $\tanh \frac{q-\bar{q}}{\sigma(q)}$ in the cluster A. The standard deviation of graph points from this dependence is $\sigma = 6.68 \cdot 10^{-5}$. This leads to an error in determining the value q of the order of ± 0.03 kg/(m^2·s), despite the fact that the maximum value within this dependence will be 0.34, and the minimum -0.03 kg/(m^2·s). The latter turns out to be at the error level, so the presence of negative values of counter-gradient fluxes remains questionable. A positive mass flux, as can be seen from Figure 8, takes place in the presence of a negative value of the density derivative in the direction of the slope, that is, under conditions of cascading. Based on the distribution of the value $\left.\frac{\partial\rho}{\partial\vec{n}}\right|_H$ presented in Figure 4b, this does not happen often.

The one-dimensional dependence of the value $\tanh \frac{q-\bar{q}}{\sigma(q)}$ on the parameter $\frac{\partial\rho}{\partial z}$ (kg/m^4) in cluster A in the region of negative values of the parameter resembles a log-normal

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 December 2022 doi:10.20944/preprints202212.0226.v1



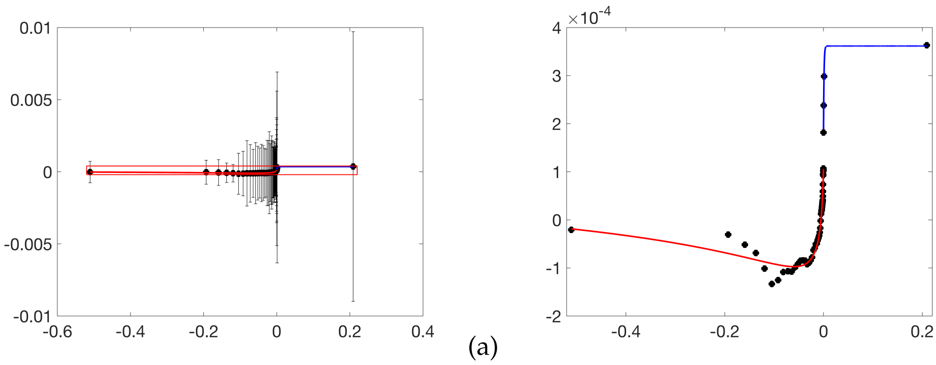
(a) (b)

Figure 8. One-dimensional dependence of tanh $\frac{q-\bar{q}}{\sigma(q)}$ on the parameter $\left.\frac{\partial \rho}{\partial \vec{n}}\right|_H$ (kg/m4) in cluster A: (a) deviation of average values for each interval from the average for the cluster when divided into 50 intervals with an equal number of cluster elements; (b) an enlarged view of the group of dots in a red box shown in (a). The blue curve represents the closest functional relationship. The vertical bars show the standard deviation for each box.

distribution and exponential one in the region of its positive values (Figure 9). The best dependence in this class of functions is given by the following expression

$$\left(\tanh\frac{q-\bar{q}}{\sigma(q)} - f_0\right)_{\frac{\partial\rho}{\partial z}} = \\ = 10^{-4}\cdot\begin{cases} \dfrac{1.298}{\dfrac{\partial\rho}{\partial z}}\cdot\exp\left(-\dfrac{\left(\log\left|\dfrac{\partial\rho}{\partial z}\right| - 1.901\right)^2}{9.420}\right) + 1.072, & \text{if } \dfrac{\partial\rho}{\partial z} < 0 \\ 3.615 - 1.833\cdot\exp\left(-\dfrac{\dfrac{\partial\rho}{\partial z}}{9.902\cdot 10^{-4}}\right), & \text{if } \dfrac{\partial\rho}{\partial z} \geq 0 \end{cases}, \quad (32)$$

Standard deviation in first case $\sigma = 1.33\cdot 10^{-5}$, for the second $\sigma = 0.72\cdot 10^{-5}$.



(a) (b)

Figure 9. Same as Figure 8, but for the parameter $\frac{\partial\rho}{\partial z}$.

3.3. Practical use

The system of equations of numerical models usually proceeds from the Boussinesq approximation, which reduces the mass conservation equation to a continuity equation, and therefore the mass flux is not explicitly taken into account. However, it can be calculated if

 doi:10.20944/preprints202212.0226.v1

we assume that the heat and salt fluxes can be represented in the equations for temperature T and salinity S in a form similar to (3), i.e.

$$\begin{cases} -\left(\overline{T'\vec{u}'}\cdot\vec{n}\right) &= K_T\dfrac{\partial\bar{T}}{\partial\vec{n}} + q_T \\ -\left(\overline{S'\vec{u}'}\cdot\vec{n}\right) &= K_S\dfrac{\partial\bar{S}}{\partial\vec{n}} + q_S \end{cases} \tag{33}$$

If the equation of state is presented in a linearized form, then the density change can be written as $\delta\rho = \alpha\delta T + \beta\delta S$, where $\alpha = \frac{\partial\rho}{\partial T}$ and $\beta = \frac{\partial\rho}{\partial S}$. Then, multiplying the first row in (33) by α, and the second by β and after adding them and comparing with (3), we get

$$K\frac{\partial\bar{\rho}}{\partial\vec{n}} + q = \left(\alpha K_T\frac{\partial\bar{T}}{\partial\vec{n}} + \beta K_S\frac{\partial\bar{S}}{\partial\vec{n}}\right) + (\alpha q_T + \beta q_S). \tag{34}$$

To ensure that the first term on the left is equal to the expression in the first bracket on the right, we set $K_T = K_S = K$. For counter-gradient fluxes, then we get

$$q = \alpha q_T + \beta q_S. \tag{35}$$

We assume that $q_T = \mu_T q$ and $q_S = \mu_S q$, where μ_T and μ_S are some constants. Then after substitution we obtain for them the following expression

$$\alpha\mu_T + \beta\mu_S = 1. \tag{36}$$

First we assume that $\frac{\mu_T}{\mu_S} = \frac{\alpha}{\beta}$. It means that the more sensitive the density is to changes in a variable, the greater the eddy flux of that variable. In this case, we get

$$\begin{cases} \mu_T &= \dfrac{\alpha}{\alpha^2+\beta^2} \\ \mu_S &= \dfrac{\beta}{\alpha^2+\beta^2} \end{cases} \tag{37}$$

Assuming the opposite that $\frac{\mu_T}{\mu_S} = \frac{\beta}{\alpha}$, that is, the more sensitive the density to changes in the variable, the less eddy flux of this variable will be needed for mass transfer, we get

$$\begin{cases} \mu_T &= \dfrac{1}{2\alpha} \\ \mu_S &= \dfrac{1}{2\beta} \end{cases} \tag{38}$$

In general

$$\begin{cases} \mu_T &= \dfrac{\alpha(1-p)}{\alpha^2+\beta^2} + \dfrac{p}{2\alpha} \\ \mu_S &= \dfrac{\beta(1-p)}{\alpha^2+\beta^2} + \dfrac{p}{2\beta}, \end{cases} \tag{39}$$

where p is a parameter which could be equal to any value from $(-\infty, +\infty)$ but giving (37) in case $p = 0$ and (38) in case $p = 1$. When p is outside $[0, 1]$ interval then the terms in (39) will have opposite signes. Since the expressions (33) give the values of the eddy flux modulus in the direction of the bottom slope $\vec{n}$, then finally for the heat fluxes in the direction of the model coordinates we obtain a two-component vector

$$\left(K_T\frac{\partial\bar{T}}{\partial\vec{n}} + q_T\right)\vec{n} = K_T\begin{pmatrix} n_x \\ n_y \end{pmatrix}\left(n_x\frac{\partial\bar{T}}{\partial x} + n_y\frac{\partial\bar{T}}{\partial y}\right) + q_T\begin{pmatrix} n_x \\ n_y \end{pmatrix}. \tag{40}$$

The same is for salinity flux. The first term in (40) could be written as

$$K_T R\cdot\vec{\nabla}\bar{T}, \tag{41}$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 December 2022 doi:10.20944/preprints202212.0226.v1

where

$$R = \begin{pmatrix} n_x^2 & n_x n_y \\ n_x n_y & n_y^2 \end{pmatrix}. \quad (42)$$

The diagonal elements of R are always positive (not negative) but off-diagonal elements could be both positive or negative depending on vector $\vec{n}$ direction.

The next series of numerical experiments is to use the obtained parametrizations and the proposed way of including them in a large-scale model. As such a model, we used the coupled ocean-ice model SibCIOM, the computational domain of which includes the Atlantic Ocean north of 20S and the Arctic Ocean, whose boundary with the Pacific Ocean in the Bering Strait is considered to be the boundary of the domain. The model is described in more detail in [39]. The horizontal resolution of the model is 0.5° in the Atlantic Ocean and is variable from 10 to 25 km in the Arctic. The application of the obtained parametrizations in numerical experiments was extended beyond the Arctic to the entire domain, including not only middle latitudes, but also the subtropics and tropics of the Atlantic Ocean. However, only the results of numerical experiments related to the Arctic region are presented below.

3.4. Diffusion coefficient test

In the first experiment B, we set the eddy diffusion coefficient equal to the sum of the coefficients obtained using the expressions (28), (29) and (30) along with the equation $K_T = K_S = K$ and (40). The experiment is a restart from the fields of January 1, 2000, obtained during the experiment from 1948 to 2020 using the state of the lower atmosphere and radiation fluxes from the NCEP/NCAR reanalysis data as a forcing (see details in [39]).

Since the proposed parametrization is designed to take into account additional eddy mass fluxes, and associated fluxes of heat and salts, we consider the integral difference of these values at latitudes above 65N latitude for two experiments: A - without the inclusion of the proposed parameterization and B - with the inclusion of this parameterization in the variant proposed above, that is, without taking into account the counter-gradient. Figure 10a shows the timeseries of the difference in mass, heat content and salinity in terms of the increment in the mass of water in the Arctic. More precisely, we consider the change in time of the difference in the mass of water between two experiments located north of latitude 65N. For reference, the total mass of water in this region according to the model grid is $1.77 \cdot 10^{19}$ kg. The heat content of water decreases until about 2007, which in terms of mass means an increase (red curve in the graph), after which it reaches a certain quasi-constant level. But after 2013 it continues growing. On average, the change in the heat content of one cubic meter of water decreased by 0.25 J by the end of the period, which is equivalent to a decrease in temperature by $6 \cdot 10^{-8}$ °C. The salinity has been decreasing (magenta curve) during all this time, and since 2004 the rate of decrease has been approximately constant but in 2013 the rate of decrease has been growing significantly. This leads to a decrease in the integral mass (black curve). As a result, the fluxes of heat and salts act in different directions, but the change in salinity is dominant and therefore we get a general reduction in the mass of water in the Arctic. The maximum difference between the two experiments was $7.23 \cdot 10^{14}$ kg, which is approximately 0.00004 of the total mass, or in terms of water density, about 42 g per cubic meter. Figure 10b shows the vertical distribution of mass change due to heat and salt fluxes. In the upper 30 m layer, the salt content (magenta curve) is on average higher in experiment B than in experiment A, while in the layer from 30 to 600 m the situation is reversed. As for the heat content (red curve), it decreases in both cases, but the changes are several times smaller in terms of mass change. As a result, the average change in the mass of the layers (black curve) is only slightly greater than as a result of the action of salt fluxes. Thus, the general trend of water mass reduction in the Arctic during experiment B is achieved by increasing the difference between the mass inflow in the 30-m layer and its decrease in the 30-600 m layer. The Figure 11 shows the deviations of the mass of water, its heat content and the mass of salt per square meter of the basin area, obtained as a result of vertical integration from the surface to the ocean floor. The growth of salinity

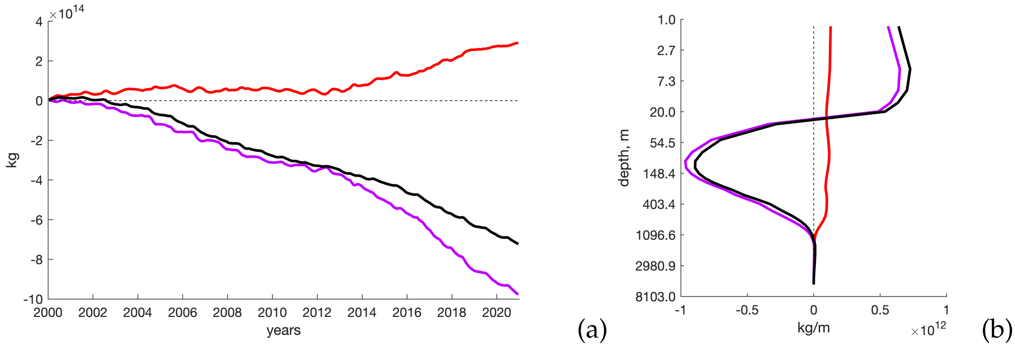


Figure 10. Difference in mass (black curves), heat (red curves) and salt (magenta curves) content in terms of the increment in the mass of water in the Arctic (above 65N): (a) timeseries of the whole content, (b) vertical distribution avaraged over time in terms of incremental values in experiment B with respect to experiment A.

in the area of the East Siberian and Chukchi Seas, off the shelf break in Barents and Kara Seas and also in the eastern part of the Beaufort Sea leads to subsequent growth in the mass per unit area in this regions. Heat content changes play a minor role but its reduction at the very shelf brake make the whole region in the vicinity of Barents and Kara Seas shelf break to be of positive mass change. Salinity content reduces substantially in the north part of Laptev Sea and farther off to the east, making mass tendency similar to it despite of heat content acting in opposite.

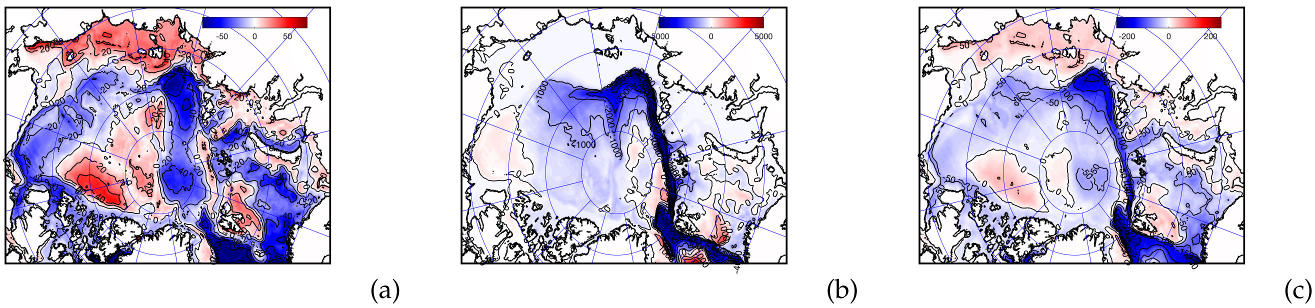


Figure 11. Annually averaged deviations of: (a) mass of water (kg/m^2), (b) its heat content (MJ/m^2) and (c) mass of salt (kg/m^2) per square meter of the Arctic area (above 65N), obtained as a result of vertical integration from the surface to the ocean floor in terms of incremental values in experiment B with respect to experiment A.

Seasonal changes are strongest in the upper 30-meter layer (Figure 12), but even here it can be noted that the seasonal differences of all values in the period of their positive values from April to June from the period of their negative values in the period from September to November are not so large (Figure 13). In general, it can be noted that the changes associated with the introduction of parameterization in experiment B do not exceed 2-5% of the total seasonal variability and basically enhance them, that is, they work towards increasing the seasonal variability of the mass, heat and salt content of water in the Arctic.

In the upper 30-meter layer, as a whole, one can note (regardless of the season) an increase in the salt content in the shelf areas of the East Siberian and Chukchi Seas in the southeastern part of the Laptev Sea and in the shallow waters of the Kara Sea, where the influence of the river runoff of the Lena, Ob and Yenisei rivers is strongest. In our analysis, we did not consider the role of cluster B associated with river plumes. Therefore, we can assume that the selection of the last two areas as areas with the most important change in salinity is not entirely fair. The strongest manifestation of positive trends in salinity, as can be seen from Figure 12a, is observed in the first 5-6 years of integration after the introduction of the proposed parameterization. Further, in years 7-9, the role of salinity decreases to almost zero, after which it rises again. The greatest positive deviation of

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 December 2022 doi:10.20944/preprints202212.0226.v1

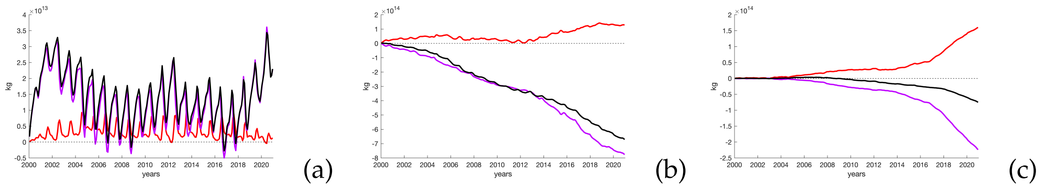


Figure 12. Timeseries of difference in mass (black curves), heat (red curves) and salt (magenta curves) content in terms of the increment in the mass of water in the Arctic (above 65N): (a) upper 30-meter layer, (b) 30-600 m layer, (c) layer from 600 m to bottom in terms of incremental values in experiment B with respect to experiment A.

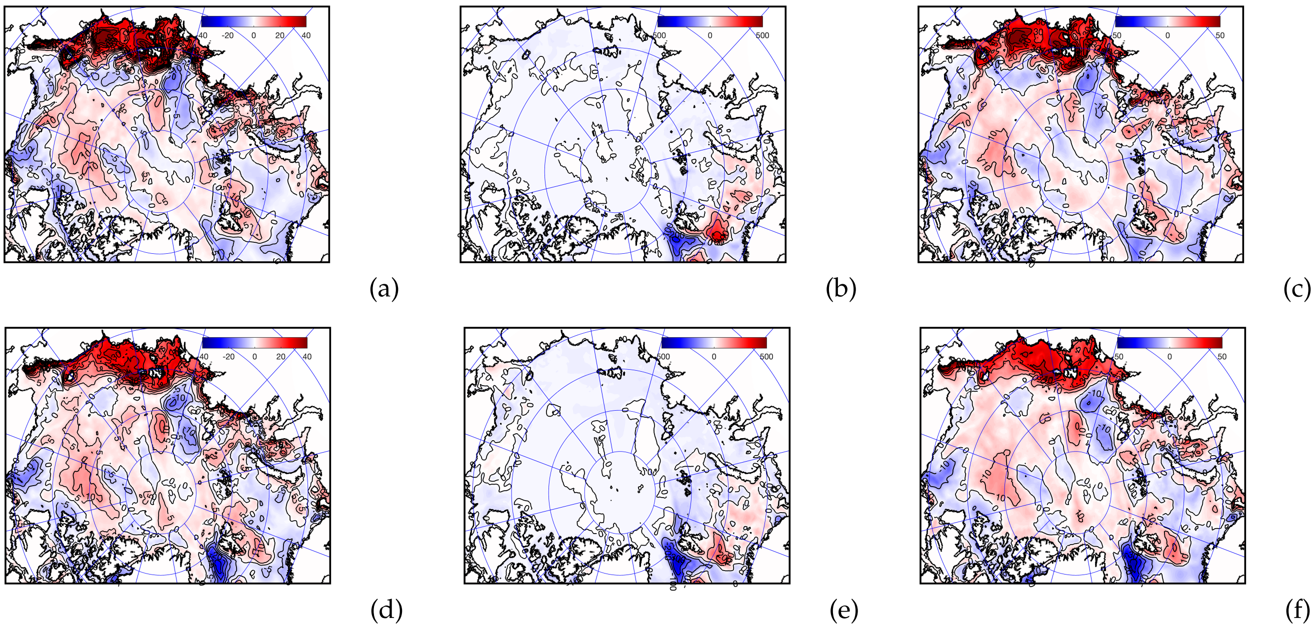


Figure 13. Deviations of: (a,d) mass of water (kg/m^2), (b,e) its heat content (MJ/m^2) and (c,f) mass of salt (kg/m^2) per square meter of the Arctic area (above 65N), obtained as a result of vertical integration from the surface to the 30-meter depth and averaged over April-June (a,b,c) and September-November (d,e,f) periods in terms of incremental values in experiment B with respect to experiment A.

salinity is in the winter-spring period (the period of ice growth), and the negative one is at the end of summer (the period of thawing). Thus, we again note an increase in seasonal changes in salinity. The role of temperature changes, as can be seen from the Figure 13b,e, is not so significant and more important in the seas of the North Atlantic and partly in the Barents Sea, but has almost no effect on changes in water mass.

Deeper layers show less seasonal variability and significant trends in heat and salt content (Figure 12b,c). The salt content in the 30-600 m layer falls almost linearly, which provides a corresponding trend in the mass of this layer. At the same time, the temperature drop of this layer works in the opposite direction to increase the mass, but the changes themselves are insufficient to withstand changes in salinity. The largest decrease in mass is observed in the area of the Barents Sea, the Amundsen Basin, in the north of the Laptev, East Siberian and Chukchi Seas off the shelf slope, as well as along the coast of the Beaufort Sea (Figure 14a). In all of the above areas, there is also a drop in the salt content in the layer (Figure 14c). An increase in mass can be seen only near the islands of the Canadian Archipelago, along the Lomonosov Ridge and off the shelf slope of the Barents and Kara Seas. In the latter case, an increase in mass occurs not only due to an increase in salinity, but also due to a decrease in the temperature of the layer (Figure 14b). In the deepest layer from 600 m to the bottom, changes become noticeable only after 5 years (Figure 12c), and the next 5 years, changes in the content of heat and salt affect the water mass in different

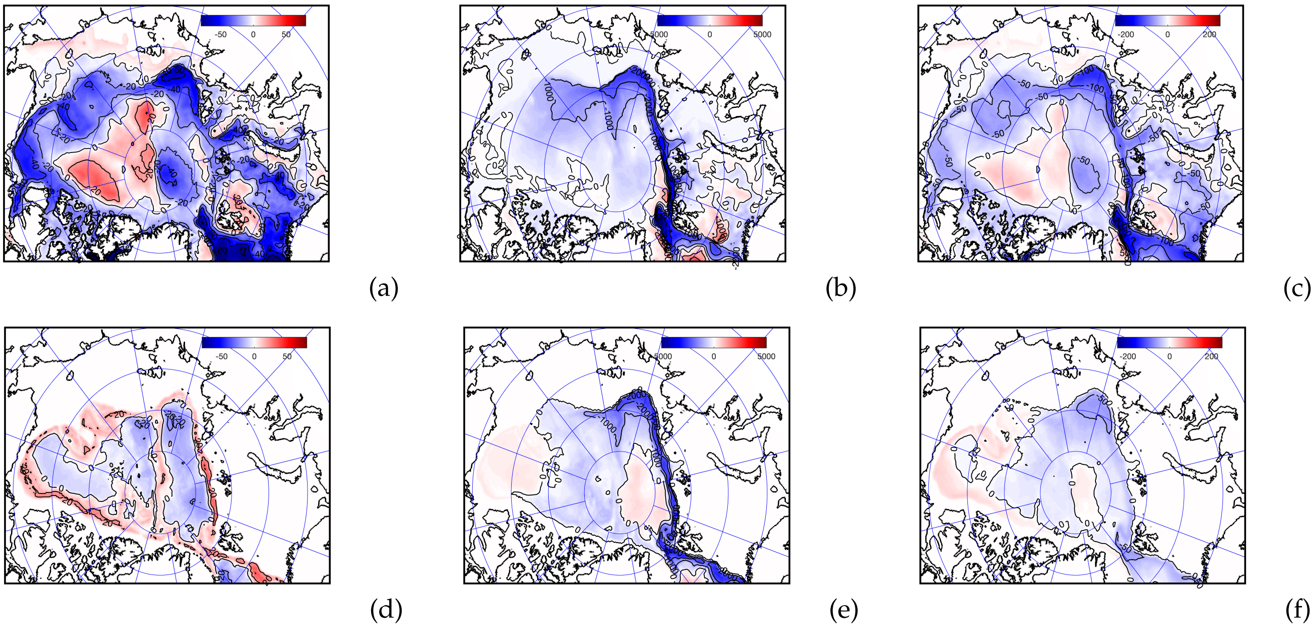


Figure 14. Annually averaged deviations of: (a,d) mass of water (kg/m^2), (b,e) its heat content (MJ/m^2) and (c,f) mass of salt (kg/m^2) per square meter of the Arctic area (above 65N), obtained as a result of vertical integration from 30 to 600 m depth (a,b,c) and from 600 m to the ocean floor in terms of incremental values in experiment B with respect to experiment A.

directions and almost completely compensate each other. Only after 10 years, the influence of salinity becomes dominant and its fall causes a decrease in the mass of water in the layer. At the same time, the decrease in mass, according to the Figure 14d, occurs in the central part of all basins, where the bottom depth is maximum, and the increase occurs in shallower areas of the ridges and shelf slopes bordering these basins.

3.5. Counter-gradient tests

The next two experiments are related to the introduction of the counter-gradient parametrization based on the derived expressions for the counter-gradient mass flux, as well as using the equations (35,36,39) under the assumption that the counter-gradient fluxes of heat and salts are expressed as $q_T = \mu_T q$ and $q_S = \mu_S q$ with μ_T and μ_S dependent on two linearization coefficients derived from equation of state $\alpha = \frac{\partial \rho}{\partial T}$ and $\beta = \frac{\partial \rho}{\partial S}$.

In the first C1 experiment, we assumed the p parameter in the equation (39) to be equal to zero, which corresponds to the situation when the heat and salt fluxes are taken in proportion to the contribution of temperature and salinity variations to density variations in accordance with the equation of state. In this case, the counter-gradient salt flux turns out to be co-directed with the counter-gradient mass flux, and the heat flux is opposite to them. The coefficient $\beta\mu_S$ in the equation (36) is approximately 16 times greater than the value of the coefficient $\alpha\mu_T$ and is approximately equal to 0.94, while the coefficient $\alpha\mu_T$ is approximately 0.06.

Figure 15a shows the timeseries of mass increments due to changes in water temperature and salinity. It can be seen that, compared with experiment B, the mass increment is more than two orders of magnitude smaller and mostly becomes noticeable about 17 years after the start of the experiment. As expected, the contribution of salinity changes has a more significant effect on mass changes and is mostly negative within first 17 years and becomes positive after. An decrease in density due to an increase in temperature counteracts this increment in mass but is not dominating. Figure 15b shows that averaged over the basin density drop is most pronounced in the upper 20 m layer but substantial growth happen in deeper 20-50 m layer and less distinct but more extended in 50-400 m layer. In both cases the salinity contribution is dominating.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 December 2022 doi:10.20944/preprints202212.0226.v1

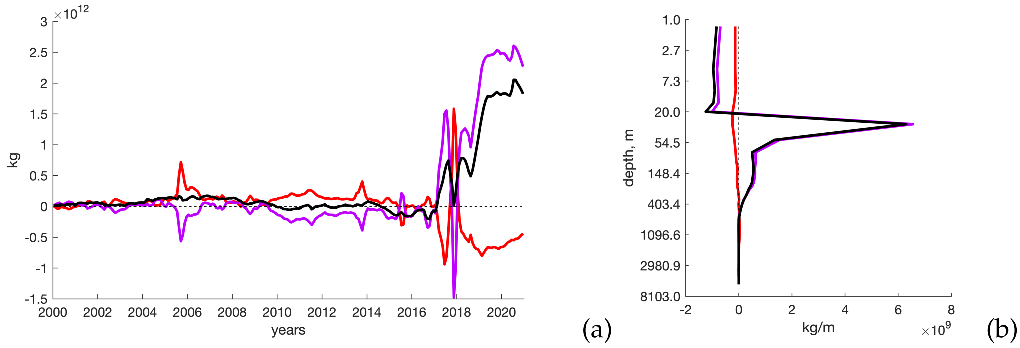


Figure 15. Difference in mass (black curves), heat (red curves) and salt (magenta curves) content in terms of the increment in the mass of water (kg) in the Arctic (above 65N): (a) timeseries of the whole content, (b) vertical distribution avaraged over time in terms of incremental values in experiment C1 with respect to experiment B.

According to the Figure 16, the greatest changes in salinity occur on the shelf and its vicinity in the Laptev Sea and the East Siberian Sea. Moreover, salinity increases in the East Siberian Sea extending positive trend toward Chukchi Sea and decreases in the Laptev Sea and less significantly in the rest of Arctic. Obviously, the reason for this is the accelerated eddy counter-gradient fluxes of fresh waters of the Lena, Olenek and Khataga rivers towards the open ocean and the subsequent deficit of these waters in the East Siberian Sea. As a result, a similar picture develops in the field of changes in the water mass. The rise in temperature in the Laptev Sea and its fall in the East Siberian Sea act in concert and also contribute to mass trends in these areas. However, in general, the opposite positive effect of temperature is manifested in the vicinity of the Fram Strait, where Atlantic waters intrude into the polar Arctic. According to the vertical distribution of temperature changes, the strongest temperature growth occurs precisely in the Atlantic water layer (Figure 15b does not show it clearly because of its small contribution to the mass distribution). Since the difference between experiment C1 and experiment B lies in taking into account counter-gradient fluxes, it is the ordered eddy motions in the area of the shelf slope that have the effect on the noticed changes in heat content of the Atlantic water layer in this area.

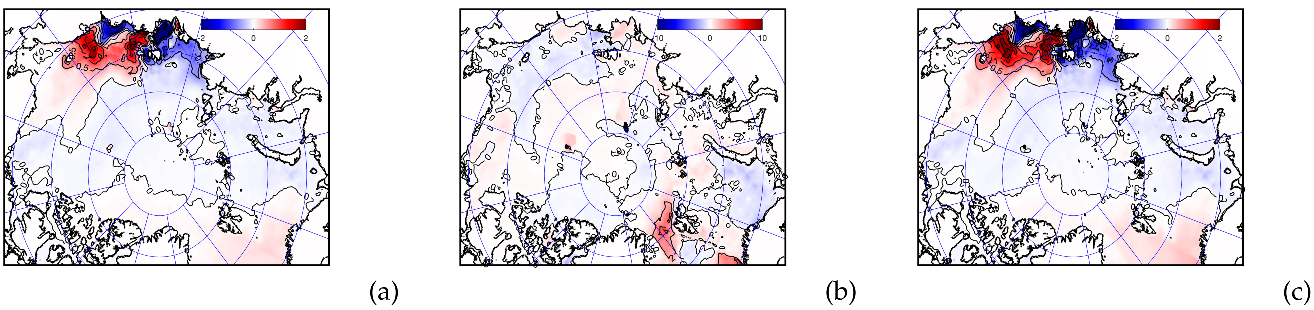


Figure 16. Annually averaged deviations of: (a) mass of water (kg/m^2), (b) its heat content (MJ/m^2) and (c) mass of salt (kg/m^2) per square meter of the Arctic area (above 65N), obtained as a result of vertical integration from the surface to the ocean floor in terms of incremental values in experiment C1 with respect to experiment B.

In the second experiment C2, we set the parameter p in the equation (39) to be equal to one, which corresponds to the situation when the heat and salt fluxes are taken inversely proportional to the contribution of temperature and salinity variations to density variations in accordance with the equation of state. As in experiment C1, the counter-gradient salt flux turns out to be co-directed with the counter-gradient mass flux, and the heat flux is opposite to it. The coefficient $\beta\mu_S$ in the equation (36) is equal to the coefficient $\alpha\mu_T$ and, accordingly, both are equal to 0.5. Thus, the contribution of the counter-gradient heat flux becomes more significant relative to the counter-gradient salinity flux than in the C1

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 December 2022 doi:10.20944/preprints202212.0226.v1

experiment. Nevertheless the resulting mass increment is lower than in C1 experiment and even three orders of magnitude smaller than in B experiment.

Figure 17a shows that first ten years both salinity and temperature contribution to water mass was growing being negative. Right after this period density change rate due to salinity started growing so that in 2019-2020 it became positive. In the same time temperature contribution became positive but after 2017 was mostly negative. Vertically salinity tendencies decreased water density in upper 400 m layer, but temperature tendencies worked opposite in upper 150 m but supported them in 150-1000 m layer.

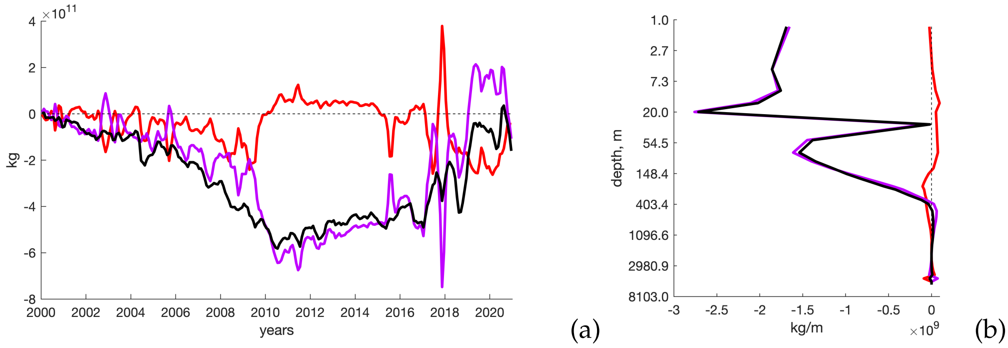


Figure 17. Difference in mass (black curves), heat (red curves) and salt (magenta curves) content in terms of the increment in the mass of water (kg) in the Arctic (above 65N): (a) timeseries of the whole content, (b) vertical distribution avaraged over time in terms of incremental values in experiment C2 with respect to experiment B.

The most influenced by counter-gradient fluxes area is still in the Laptev and East Siberian seas but the anomaly distribution is more complicated than in C1 experiment. We still have a heat content growth in the vicinity of Fram Strait in the upper layer but also there is a noticeable area of counter-gradient fluxes in the vicinity of the Lomonosov ridge in central Arctic where Atlantic water current following Laptev Sea shelf break turns northward along the ridge.

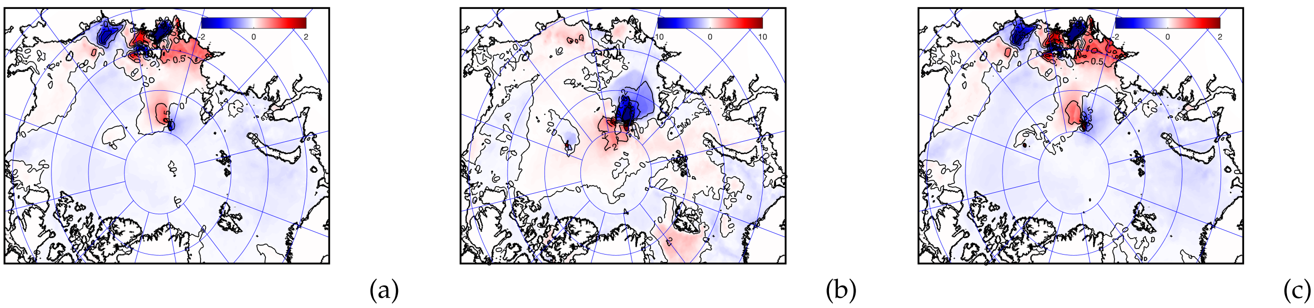


Figure 18. Annually averaged deviations of: (a) mass of water (kg/m^2), (b) its heat content (MJ/m^2) and (c) mass of salt (kg/m^2) per square meter of the Arctic area (above 65N), obtained as a result of vertical integration from the surface to the ocean floor in terms of incremental values in experiment C2 with respect to experiment B.

4. Discussion and Conclusion

As a result of the analysis performed, some parametrization dependences of the characteristics of eddy fluxes on large-scale thermodynamic characteristics of the Arctic shelf zone in the Kara Sea were obtained. The resulting expressions for diffusion coefficient K and counter-gradient flux q can be directly used in a large-scale oceanic model instead of the available diffusive fluxes if the eddy ones exceed them. For example, in the current version of the World Ocean model built using the SibCIOM model mentioned above, the diffusion coefficient value is equal to 100 m^2/s. Considering that all the above dependencies were obtained at the value $K_0 = 394$ m^2/s, we can assume that on the previous plots

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 December 2022 doi:10.20944/preprints202212.0226.v1

(Figures 5-7) this corresponds to the level $\log \frac{100}{394} \approx -1.37$. As can be seen, this level is below the minimum on all graphs and, therefore, eddy fluxes will be dominant. Since there is no counter-gradient flux in this model, the resulting value of the flux q is unconditionally applicable.

It should also be noted that in our analysis we considered only one-dimensional dependencies, which were significant, but not dominant. Therefore, in the development of this approach, more attention should be paid to multidimensional dependencies based on various (non-additive) combinations of large-scale parameters. For example, eddy fluxes turned out to be the most sensitive with respect to the parameters $\frac{\partial\rho}{\partial z}$, $\left.\frac{\partial\rho}{\partial\vec{n}}\right|_H$ and s, from which it is possible to construct a dimensionless combination equal to the angle θ between the bottom surface and the isopycnal surface

$$\theta(\frac{\partial\rho}{\partial z}, \frac{\partial\rho}{\partial\vec{n}}, s) = \arccos\frac{\left(\vec{\nabla}\rho\cdot\vec{\nabla}H\right)}{\left|\vec{\nabla}\rho\right|\cdot\left|\vec{\nabla}H\right|},$$

where $\vec{\nabla}\rho = \left(\frac{\partial\rho}{\partial\vec{n}}, \frac{\partial\rho}{\partial z}\right)$, and $s = \left|\vec{\nabla}H\right|$. As we can see, on the one hand, the value of this angle depends on all three parameters, and on the other hand, the angle itself is a key parameter for the formation and intensification of cascading, so the eddy flux has many chances to be sensitive to this value. The study of such combinations is planned in future work within the framework of this approach.

Our conclusion about the dependence of eddy parameters on density gradients is closely related to the popular approach to the parametrization of mesoscale motions in the form of isopycnal diffusion. In this approach either the coordinate system is rotated along isopycnal surfaces, or an isopycnal diffusion tensor is composed to consider the action of eddy fluxes along the surfaces of constant density. Unlike this method, we use the obtained coefficients for horizontal fluxes, not isopycnal, but we note that the values of both coefficients, gradient and counter-gradient, are calculated from the density slope.

These methods used to parameterize eddy motions in numerical models using diffusion operators have existed for a long time, but there is still no complete understanding of the required values of the diffusion coefficients. There are many approaches to quantify unknown diffusion parameters, both horizontal/vertical and isopycnal. The possibility of using the eddy-permitting model output, which we are implementing here, arose not so long ago with the development of computer technology, which made it possible eddy-resolving simulations.

The earlier estimations of horizontal diffusion coefficients were based on numerical and theoretical models [40–42], or observations [43]. The horizontal diffusion coefficients estimated in these works vary depending on the location, ranging from almost zero to about 10^4 m^2/s. Based on an ensemble of tracers from subsurface lagrangian drifters, [44] estimated an isopycnal diffusion of 800±200 m^2/s using the lagrangian dispersion method [45]. In [46] the contribution of eddy kinetic energy was evaluated using a very detailed (1 km) model over the entire Arctic region. They found that the largest contribution from mesoscale eddies comes from continental slope regions along the main currents in autumn season. In November the kinetic energy of eddies averages about $10^{0.2}$ m^3/s^2 along the coast of Alaska and $10^{0.05}$ m^3/s^2 in the Laptev Sea.

In future we plan a quantitative assessment of the total contribution of mesoscale movements using an eddy-resolving model and comparison with those obtained as a result of a large-scale simulation with parametrization of isopycnal diffusion.

It is also worth noting that here we were unable to trace the relationship between eddy fluxes and the value of the Coriolis parameter (i.e., geographic latitude), largely due to the fact that its variability is small within the Kara Sea. To take it into account, it is necessary to consider several seas or even a series of them located at different latitudes.

In addition, the ratios obtained should be refined by considering other Arctic seas and in other periods of time, which, perhaps, will make it possible to get rid of the specific features of the Kara Sea and the period of 2007.

Nevertheless, the result obtained is hopefully important and requires its further approbation within the framework of a large-scale modelling of the Arctic.

Our numerical experiments carried out using the SibCIOM model showed that the results are most sensitive with respect to the parametrization of the eddy diffusion coefficient. The greatest differences from the experiment without the proposed parameterization are achieved in areas along the Fram branch of the Atlantic waters trajectory in the Arctic. Moreover, the manifestation is most pronounced in the fields of final salinity and temperature, while the density field turned out to be less sensitive. Another area where parametrization of eddy exchange turned out to be important is the shelf of the East Siberian and Laptev seas and adjacent deep water areas. Here, due to the increase in cross-isobatic exchanges, the salinity of these seas has noticeably increased and the amount of salt has decreased in the adjacent regions of the Arctic in 30-600 m layer. In this regard, it can be noted that for a more accurate description of the processes in these regions, it is also necessary to take into account the elements of the sample from cluster B (eddy structures at the river plume boundary) and cluster C (eddy structures in areas of a sharp bottom slope). In this work, we have left the features associated with these regions aside and we cannot yet say to what extent the identified dependencies correlate with the statistical distributions in these specific clusters.

Accounting for counter-gradient eddy fluxes turned out to be less important, and the corresponding response differs from the response to the introduction of eddy diffusion by 2–3 orders of magnitude. Considering also the fact that, based on the statistical analysis, the inequality of the counter-gradient flux to zero is not significant, we can conclude that they can be neglected in large-scale models. Although we could still notice that area where counter-gradient parametrization turned out to be most valuable is again the shelf of the East Siberian and Laptev seas and adjacent deep water areas.

Author Contributions: Conceptualization, methodology, formal analysis, investigation, funding acquisition, G.P.; validation, G.P. and E.G., writing—original draft preparation, G.P. and D.I.; writing—review and editing, D.I.. All authors have read and agreed to the published version of the manuscript.

Funding: The study was carried out within the framework of the CRiceS project under the Horizon-2020 cooperation program (GA No 101003826) with the financial support of the Russian Federation represented by the Ministry of Science and Higher Education of Russia (contract No. 075-15-2021-947).

Institutional Review Board Statement: Not applicable

Informed Consent Statement: Not applicable

Data Availability Statement: Not applicable

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Smagorinsky, J. General circulation experiments with the primitive equations I: The basic experiment. *Monthly Weather Review* **1963**, *91*, 99–164. doi: 10.1175/1520-0493(1963)091<0099:GCEWTP>2.3.CO;2
2. Blumberg, A.F.; Mellor, G.L. A Description of a Three-Dimensional Coastal Ocean Circulation Model. In *Three-Dimensional Coastal Ocean Models*; Heaps, N. S. Eds.; *Coastal and Estuarine Sciences* **1987**, *4*, American Geophysical Union, Washington, D.C., 1–16.
3. Griffies, S.M.; Hallberg, R.W. Biharmonic friction with a Smagorinsky-like viscosity for use in large-scale eddy-permitting ocean models. *Monthly Weather Review* **2000**, *128*, 2935–2946.
4. Bachman, S.D.; Fox-Kemper, B.; Pearson, B. A scale-aware subgrid model for quasi-geostrophic turbulence. *Journal of Geophysical Research: Oceans* **2017**, *122*, 1529–1554.
5. Gent, P.R.; McWilliams, J.C. Isopycnal mixing in ocean circulation models. *Journal of Physical Oceanography* **1990**, *20*, 150–155. doi: 10.1175/1520-0485(1990)020<0150:IMIOCM>2.0.CO;2
6. Gent, P.R.; Willebrand, J.; McDougall, T.J.; McWilliams, J.C. Parameterising eddy-induced tracer transports in ocean circulation models. *Journal of Physical Oceanography* **1995**, *25*, 463–474. doi: 10.1175/1520-0485(1995)025<0463:PEITTI>2.0.CO;2

7. Griffies, S.M. The Gent-McWilliams skew flux. *Journal of Physical Oceanography* **1998**, *28*, 831–841. doi: 10.1175/1520-0485(1998)028<0831:TGMSF>2.0.CO;2
8. Bachman, S.D. The GM+E closure: A framework for coupling backscatter with the Gent and McWilliams parameterization. *Ocean Modelling* **2019**, *136*, 85–106. doi: 10.1016/j.ocemod.2019.02.006
9. Draper, D. Assessment and propagation of model uncertainty. *Journal of the Royal Statistical Society Series B* **1995**, *57*, 45–97. Available online: http://www.jstor.org/stable/2346087 (accessed on 12 Nov. 2022).
10. Hourdin, F.; Mauritsen, T.; Gettelman, A.; Golaz, J. -C.; Balaji, V.; Duan, Q.; Folini, D.; Ji, D.; Klocke, D.; Qian, Y.; Rauser, F.; Rio, C.; Tomassini, L.; Watanabe, M.; Williamson, D. The art and science of climate model tuning. *Bulletin Of The American Meteorological Society* **2017**, *98*, 589–602. doi:10.1175/BAMS-D-15-00135.1
11. Heemink A.W.; Segers A.J. Modeling and prediction of environmental data in space and time using Kalman filtering. *Stochastic Environmental Res. and Risk Assessment* **2002**, *16*, 225–240. doi:10.1007/s00477-002-0097-1
12. Klimova, E.G.; Platov, G.A.; Kilanova, N.V. Development of an environmental data assimilation system based on the ensemble Kalman filter. *Computational Technologies* **2014**, *19(3)*, 27–37. (in Russian)
13. Ling, J.; Kurzawski, A.; Templeton, J. Reynolds averaged turbulence modelling using deep neural networks with embedded invariance. *Journal of Fluid Mechanics* **2016**, *807*, 155–166. doi:10.1017/jfm.2016.615
14. Ling, J.; Ruiz, A.; Lacaze, G.; Oefelein, J. Uncertainty analysis and data-driven model advances for a jet-in-crossflow. In Procedings of ASME Turbo Expo 2016, Seoul, South Korea, 13-17 June 2016.
15. Schneider, T.; Lan, S.; Stuart, A.; Teixeira, J. Earth system modeling 2.0: A blueprint for models that learn from observations and targeted high-resolution simulations. *Geophysical Research Letters* **2017**, *44*, 12,396–12,417. doi: 10.1002/2017GL076101
16. Kurowski, M.J.; Thrastarson, H.T.; Suselj, K.; Teixeira, J. Towards Unifying the Planetary Boundary Layer and Shallow Convection in CAM5 with the Eddy-Diffusivity/Mass-Flux Approach. *Atmosphere* **2019**, *10*, 484. doi: 10.3390/atmos10090484
17. Zhang, J.A.; Kalina, E.A.; Biswas, M.K.; Rogers, R.F.; Zhu, P.; Marks, F.D. A Review and Evaluation of Planetary Boundary Layer Parameterizations in Hurricane Weather Research and Forecasting Model Using Idealized Simulations and Observations. Atmosphere 2020, 11, 1091. doi: 10.3390/atmos11101091
18. Kalina, E.A.; Biswas, M.K.; Zhang, J.A.; Newman, K.M. Sensitivity of an Idealized Tropical Cyclone to the Configuration of the Global Forecast System–Eddy Diffusivity Mass Flux Planetary Boundary Layer Scheme. Atmosphere 2021, 12, 284. doi: 10.3390/atmos12020284
19. Platov, G.A. Numerical modeling of the Arctic Ocean deepwater formation: Part II. Results of regional and global experiments. *Izvestiya, Atmospheric and Oceanic Physics* **2011**, *47(3)*, 377–392. doi: 10.1134/S0001433811020083
20. Platov, G.A.; Golubeva, E.N. Interaction of dense shelf waters of the Barents and Kara seas with the eddy structures. *Physical Oceanography* **2019**, *26(6)*, 484–503. doi: 10.22449/1573-160X-2019-6-484-503
21. Platov, G.A.; Middleton, J.F. Notes on Pressure Gradient Correction. *Bulletin of the Novosibirsk Computing Center***2001**, *7*, 43–58. Available online: https://nccbulletin.ru/files/article/platov_3.pdf (accessed on 12 Nov. 2022).
22. Atadzhanova, O.A.; Zimin, A.V.; Romanenkov, D.A.; Kozlov, I.E. Satellite Radar Observations of Small Eddies in the White, Barents and Kara Seas. *Physical Oceanography* **2017**, *2*, 75–83. doi: 10.22449/1573-160X-2017-2-75-83
23. Golubeva, E.N.; Platov, G.A. Numerical Modeling of the Arctic Ocean Ice System Response to Variations in the Atmospheric Circulation from 1948 to 2007. *Izvestiya, Atmospheric and Oceanic Physics* **2009**, *45*, 137–151. doi: 10.1134/S0001433809010095
24. Pnyushkov, A.; Polyakov, I. V.; Padman, L.; Nguyen, A. T. Structure and dynamics of mesoscale eddies over the Laptev Sea continental slope in the Arctic Ocean. *Ocean Sci.* **2018**, *14*, 1329–1347. doi: 10.5194/os-14-1329-2018
25. Deardorff,J. W.; Willis, G. E. Turbulence within a baroclinic laboratory mixed layer above a sloping surface. *J. Atmos. Sci* **1987**, *44*, 772–778.
26. Lykossov, V. N. K-theory of atmospheric turbulent planetary boundary layer and the Boussinesq's generalized hypothesis.*Sov. J. Numer. Anal. Math. Modeling* **1990**, *3*, 221–240.
27. Lykossov, V. N.; Platov, G. A. A numerical model of interaction between atmospheric and oceanic boundary layers. *Rus. J. Numer. Anal. Math. Modeling* **1992**, **7**, 419–440. doi: 10.1515/rnam.1992.7.5.419
28. Platov, G.; Golubeva, E. Characteristics of mesoscale eddies of Arctic marginal seas: results of numerical modeling. *IOP Conf. Ser.: Earth Environ. Sci* **2020**, *611*, 012009. doi: 10.1088/1755-1315/611/1/012009
29. Ivanov, V.V.; Shapiro, G.I.; Huthnance, J.M.; Aleynik, D.L.; Golovin, P.N. Cascades of Dense Water around the World Ocean. *Progress in Oceanography* **2004**, *60*, 47–98. doi: 10.1016/j.pocean.2003.12.002
30. Walter, J.; Chesnaux, R.; Gaboury, D.; Cloutier, V. Subsampling of Regional-Scale Database for improving Multivariate Analysis Interpretation of Groundwater Chemical Evolution and Ion Sources. *Geosciences* **2019**, *9*, 139. doi: 10.3390/geosciences9030139
31. Wang, X.; Wang, S.; Zhang, S.; Gu, C.; Tanvir, A.; Zhang, R.; Zhou, B. Clustering Analysis on Drivers of O3 Diurnal Pattern and Interactions with Nighttime NO3 and HONO. *Atmosphere* **2022**, *13*, 351. doi: 10.3390/atmos13020351
32. Lloyd, S. P. Least squares quantization in PCM. *IEEE Transactions on Information Theory* **1982**, **28**, 129–137. doi: 10.1109/TIT.1982.1056489
33. Arthur, D.; Vassilvitskii, S. k-means++: the advantages of careful seeding. In Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms, Society for Industrial and Applied Mathematics Philadelphia, PA, USA, 2007; 1027–1035.
34. Davies, D.L.; Bouldin, D.W. A Cluster Separation Measure. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **1979**, *1*, 224–227. doi:10.1109/TPAMI.1979.4766909
35. Dunn, J. Well separated clusters and optimal fuzzy partitions. *Journal of Cybernetics* **1974**, *4*, 95–104. doi:10.1080/01969727408546059

36. Rousseeuw, P.J. Silhouettes: a Graphical Aid to the Interpretation and Validation of Cluster Analysis. *Computational and Applied Mathematics* **1987**, *20*, 53–65. doi:10.1016/0377-0427(87)90125-7
37. Kaufman, L.; Rousseeuw, P.J. *Finding groups in data: An introduction to cluster analysis*; Publisher: Hoboken, NJ: Wiley-Interscience, USA, 1990; 87 pp. doi:10.1002/9780470316801
38. Sobol, I.M. Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Math. Comput. Simulat.* **2001**, *55*, 271–280, doi:10.1016/S0378-4754(00)00270-6
39. Golubeva, E.; Kraineva, M.; Platov, G.; Iakshina, D.; Tarkhanova, M. Marine Heatwaves in Siberian Arctic Seas and Adjacent Region. *Remote Sens.* **2021**, *13*, 4436. doi: 10.3390/rs13214436
40. Ferreira, D.; Marshall, J.; Heimbach, P. Estimating eddy stresses by fitting dynamics to observations using a residual-mean ocean circulation model and its adjoint. *J. Phys. Oceanogr.* **2005**, *35*, 1891–1910. doi: 10.1175/JPO2785.1
41. Griesel, A.; GIlle, S. T.; Sprintall, J.; McClean, J. L.; La Casce, J. H.; Maltrud, M. E. Isopycnal diffusivities in the Antarctic Circumpolar Current inferred from Lagrangian floats in an eddying model. *J. Geophys. Res.* **2010**, *115*, C06006. doi:10.1029/2009JC005821
42. Vollmer, L.; Eden, C. A global map of meso-scale eddy diffusivities based on linear stability analysis. *Ocean Modell.* **2013**, *72*, 198–209. doi: 10.1016/j.ocemod.2013.09.006
43. Funk, A.; Brandt, P.; Fischer, T. Eddy diffusivities estimated from observations in the Labrador Sea. *J. Geophys. Res.* **2009**, *114*, C04001. doi:10.1029/2008JC005098
44. La Casce, J.; Ferrari, R.; Marshall, J.; Tulloch, R.; Balwada, D.; Speer, K. Float-derived isopycnal diffusivities in the DIMES experiment. *Journal of Physical Oceanography* **2014**, *44*, 764–780. doi: 10.1175/JPO-D-13-0175.1
45. Garrett, C. On the initial streakness of a dispersing tracer in two-and three-dimensional turbulence. *Dynamics of Atmospheres and Oceans* **1983**, *7*, 265–277. doi: 10.1016/0377-0265(83)90008-8
46. Wang, Q.; Koldunov, N. V.; Danilov, S.; Sidorenko, D.; Wekerle, C.; Scholz, P.; Bashmachnikov, I. L.; Jung, T. Eddy kinetic energy in the Arctic Ocean from a global simulation with a 1-km Arctic. *Geophysical Research Letters* **2020**, *47*, e2020GL088550. doi: 10.1029/2020GL088550
47. Wan, H.; Xia, J.; Zhang, L.; She, D.; Xiao, Y.; Zou, L. Sensitivity and Interaction Analysis Based on Sobol' Method and Its Application in a Distributed Flood Forecasting Model. Water 2015, 7, 2924-2951. https://doi.org/10.3390/w7062924