Genesis and Decay of Mesoscale Baroclinic Eddies in the Seasonally Ice-Covered Interior Arctic Ocean

GIANLUCA MENEGHELLO,a JOHN MARSHALL,a CAMILLE LIQUE,b PÅL ERIK ISACHSEN,c,d EDWARD DODDRIDGE,a,e JEAN-MICHEL CAMPINA, HEATHER REGAN,b AND CLAUDE TALANDIERb

a Massachusetts Institute of Technology, Cambridge, Massachusetts
b Univ. Brest, CNRS, IRD, Ifremer, Laboratoire d’Oceanographie Physique et Spatiale, IUEM, Brest, France
c University of Oslo, Oslo, Norway
d Norwegian Meteorological Institute, Oslo, Norway
e Institute for Marine and Antarctic Studies, University of Tasmania, Hobart, Tasmania, Australia

ABSTRACT: Observations of ocean currents in the Arctic interior show a curious, and hitherto unexplained, vertical and temporal distribution of mesoscale activity. A marked seasonal cycle is found close to the surface: strong eddy activity during summer, observed from both satellites and moorings, is followed by very quiet winters. In contrast, subsurface eddies persist all year long within the deeper halocline and below. Informed by baroclinic instability analysis, we explore the origin and evolution of mesoscale eddies in the seasonally ice-covered interior Arctic Ocean. We find that the surface seasonal cycle is controlled by friction with sea ice, dissipating existing eddies and preventing the growth of new ones. In contrast, subsurface eddies, enabled by interior potential vorticity gradients and shielded by a strong stratification at a depth of approximately 50 m, can grow independently of the presence of sea ice. A high-resolution pan-Arctic ocean model confirms that the interior Arctic basin is baroclinically unstable all year long at depth. We address possible implications for the transport of water masses between the margins and the interior of the Arctic basin, and for climate models’ ability to capture the fundamental difference in mesoscale activity between ice-covered and ice-free regions.

KEYWORDS: Arctic; Baroclinic flows; Eddies; Ocean circulation; In situ oceanic observations; Quasigeostrophic models

1. Introduction

Mesoscale eddies are an ubiquitous feature of the Arctic Ocean (Timmermans and Marshall 2020). Recent satellite observations (Kozlov et al. 2019) have identified an active surface eddy field in ice-free regions and in the marginal ice zone, while observations of temperature and salinity profiles from Ice-Tethered Profilers (ITP; Toole et al. 2011) and the Beaufort Gyre Exploration Project moorings (https://www.whoi.edu/beaufortgyre) have begun to characterize the eddy field below the surface (Zhao et al. 2014, 2018). At the same time, the presence of sea ice and the difficulties of accessing this remote region have led to a rather patchy description of the mesoscale field which lacks the detailed quantification available at the surface of other oceans (e.g., Faghmous et al. 2015). Satellite observations cannot provide information within the ice covered regions, or about the eddy structure at depth. In situ observations, while improving, still remain sparse in both space and time. This lack of observability, together with the challenge of producing accurate numerical simulation in the Arctic—owing, for example, to the small deformation radius (Nurser and Bacon 2014) and the need to represent the complexity of ocean–sea ice interaction—have limited our ability to understand the nature and role of the mesoscale in the general circulation of the Arctic.

Numerical modeling and satellite observations point to an important role played by eddies in the transport of heat and freshwater into or out of the Arctic interior (Watanabe 2011; Spall 2013) and the equilibration of the large-scale current (Davis et al. 2014; Manucharyan et al. 2016; Manucharyan and Spall 2016; Meneghello et al. 2018b; Spall 2020). However, eddy flux estimates from observations (Timmermans et al. 2012; Meneghello et al. 2018b), and results from simulations at high resolution (Regan et al. 2020), suggest a lack of energy at the mesoscale in the Arctic interior compared to characteristic midlatitude ocean dynamics (Wunsch 2002). Indeed, the usual ingredients generating mesoscale ocean turbulence may be inefficient in the Arctic, owing to the presence of the ice–ocean governor diminishing the rates of Ekman pumping (Meneghello et al. 2017, 2018a, 2020; Doddridge et al. 2019; Dewey et al. 2018; Zhong et al. 2018) and hence the creation of potential energy stored in sloping isopycnals.

Another possible explanation for the lack of energy is for mesoscale activity to be fundamentally different in ice-free and ice-covered oceans. We still do not understand the details of the eddy generation process in the Arctic (Zhao et al. 2014), and the eddies’ origin itself remains a subject of debate: are they generated only within the relatively fast and unstable currents along the Arctic margins, from where they can be advected into the central Arctic (see, e.g., Spall et al. 2008, and references therein), or is the entire Arctic baroclinically unstable?
In this work we address the dynamical mechanisms behind the origin and seasonal evolution of mesoscale eddies in the seasonally ice-covered Arctic interior, and discuss their sources and their role for the large-scale Arctic circulation. We start in section 2 by presenting observations of kinetic energy in the Canada basin over the seasonal cycle, and reviewing previous attempts to explain their peculiarities. In section 3 we interpret those observations by exploring the combined effect of stratification and friction at the ice–ocean interface on the development of linear baroclinic instability. In section 4 we use a high-resolution, pan-Arctic numerical model to explore if the predictions from linear theory extend to a well-developed eddy field over the entire Arctic. In section 5 we address the possible origin of baroclinic instability. We conclude in section 6 by discussing implications for the large-scale circulation and the future evolution of the Arctic.

### 2. Peculiarities of the Arctic mesoscale activity

Figure 1 shows a seasonal climatology of the vertical distribution of kinetic energy in the central Canada basin. It is obtained by combining multiple years of current observations from a mooring located at 75°N, 150°W. (The full time series over 2003–18, showing the availability of observations over time, is shown in Fig. A1 in the appendix.) Clearly visible in Fig. 1 are two layers of enhanced kinetic energy, bounded by the two peaks in the background stratification $N^2 = -(g/p)(\partial p/\partial z)$ (left panel), where $p$ is potential density and $g$ is the gravitational constant. The surface layer extends down to approximately 50-m depth and is characterized by a marked seasonal cycle, apparently anticorrelated with ice draft (top panel). The subsurface layer, extending between approximately 50- and 250-m depth, is characterized by a more homogeneous kinetic energy level throughout the year. The low kinetic energy patches in the subsurface layer can be associated with the limited number of eddies contained in the time series (less than 40) rather than actual variability.

Subsurface intensified mesoscale activity in the Arctic was first reported in the 1970s by Newton (1973), Newton et al. (1974), and Hunkins (1974). Almost a decade later, Hunkins (1981) and Manley and Hunkins (1985) investigated its origin by studying the linear stability of the Arctic mean baroclinic flow in the presence of ice–ocean friction. Their analysis assumed velocity and stratification profiles exponentially decaying with depth over the same length scale, and concluded that conditions in the central Arctic Ocean are not favorable for baroclinic instability to develop, due to the friction between eddies and sea ice. It was then suggested that eddies must originate elsewhere, for example in the more unstable coastal regions, and propagate into the interior afterward. Additionally, Ou and Gordon (1986) used a simple linear model to show how an eddy’s surface flow would be spun down by frictional effects on a time scale of days once below sea ice, while the subsurface flow, shielded from sea ice by a strong stratification, would be characterized by a much longer spin-down time scale of a year or more (see section 3c below).

More recently, the vertical distribution of kinetic energy in the Arctic has been addressed by Zhao et al. (2018), who...
partitioned the observed kinetic energy onto the neutral modes of the observed stratification, i.e., the eigenvectors of the linearized, quasigeostrophic equations under the assumption of a flat bottom, resting ocean (see also Pedlosky 1982; Wunsch 1997). Their analysis showed how a large part of the eddy kinetic energy is contained in the second baroclinic mode, characterized by a subsurface maximum, in agreement with Fig. 1. Yet, they did not explain the dynamical origin of this distribution. In the next section, we will address the origin of the peculiar vertical structure of Arctic kinetic energy.

3. Baroclinic instability and eddy dissipation in a seasonally ice covered ocean

The available potential energy stored in the sloping isopycnals visible in Fig. 2 provides a ready source of energy for baroclinic instability. But, in order for it to be extracted, at least one of three conditions must be present (Pedlosky 1982, section 7.3): (i) same-sign horizontal buoyancy gradients at the surface and at the bottom of the ocean, as described by Eady (1949), (ii) same-sign, horizontal surface buoyancy gradients and interior potential vorticity gradients (or opposing signs at the bottom), as described by Charney (1947), or (iii) opposing horizontal potential vorticity gradients in the interior, as discussed by Phillips (1954).

However, along with energy sources, there are many energy sinks, associated with, for example, frictionally driven Ekman layers at the surface and the bottom (Charney and Eliassen 1949; Barcilon 1964; Williams and Robinson 1974). As discussed by Hunkins (1981), ice-driven Ekman layers at the surface can strongly damp the growth of baroclinic instability. Yet, it must be remarked that Hunkins (1981) only considered the first of the above conditions, namely, the presence of buoyancy gradients at the boundaries. We shall find in this paper that the presence of horizontal potential vorticity gradients in the interior enables the growth of independent subsurface eddies.

Our analysis starts from the linearized quasigeostrophic potential vorticity (PV) equation (Smith 2007; Tulloch et al. 2011)

\[
\frac{Dq}{Dt} = -\mathbf{u} \cdot \nabla Q, \tag{1a}
\]

\[
q = \nabla^2 \psi + \frac{\partial}{\partial z} \left( f_0^2 \frac{\partial \psi}{N^2 \partial z} \right), \tag{1b}
\]

where \(q\) is the perturbation PV, \(f_0\) is the Coriolis parameter, \(N^2\) is the stratification, and \(\psi = p/(\rho_0 f_0)\) is the perturbation streamfunction, with \(p\) being pressure and \(\rho_0\) a reference density, so that \(\mathbf{u} = \mathbf{k} \times \nabla \psi\), where \(\mathbf{k}\) is the unit vector pointing upward. The total derivative \(D/Dt = (\partial/\partial t + \mathbf{U} \cdot \nabla)\) is based on the background velocity \(\mathbf{U}\), and \(\nabla Q\) is the background PV gradient.

Boundary conditions for (1) are provided at the surface and at the ocean floor by imposing the quasigeostrophic density equation

\[
\frac{D}{Dt} \left( \frac{\partial \psi}{\partial z} \right) = -\frac{\partial \mathbf{U}}{\partial z} \cdot \nabla \psi \pm \frac{N^2}{f_0^2} \frac{1}{\partial z} \nabla^2 \psi. \tag{2}
\]

Here, the vertical velocity at the boundaries is not zero (as commonly done to model a free slip condition) but is set to
the Ekman pumping $w_E = \pm (d/2)\nabla^2 \psi$ where $d$ is the Ekman layer depth and $\nabla^2 \psi$ is the interior relative geostrophic vorticity. The sign is positive at the ocean floor and negative at the surface.

At the scales at which the background PV $Q$ is computed, characterized by horizontal length scales $L \approx 100$ km and vertical length scales $H \approx 100$ m, the relative vorticity term $\nabla^2 \Psi \approx \Psi/L^2$ (where $\Psi$ is the background streamfunction) is at least one order of magnitude smaller than the stratification term $(\partial \psi/\partial z)((f_0^2/N^2) \partial \Psi/\partial z) \approx (f_0^2/N^2)\Psi/H^2$ and can be neglected in the computation of the background PV gradient. It then reduces to

$$\nabla Q = \nabla \frac{\partial}{\partial z} \left( \frac{f_0^2}{N^2} \frac{\partial \Psi}{\partial z} \right) + \beta_0 \hat{j},$$

where $\beta_0$ is the gradient of the Coriolis parameter, and $\hat{j}$ is the unit vector pointing northward.

We recall how the dynamics described by (1) and (2) can also be interpreted in terms of the organization of isopycnal slopes, shown in Fig. 2. Of particular interest are the velocity shear and PV gradient on the right-hand side of (2) and (1a), as they represent the processes enabling the instability. The vertical velocity shear $\partial U/\partial z$ on the right-hand side of (2) can be related to the large-scale density field by the thermal wind relation

$$\frac{\partial U}{\partial z} = \frac{g}{\rho_0} \frac{\partial \rho}{\partial z} \times \nabla z_p = -\frac{N^2}{f_0} \hat{k} \times \nabla z_p,$$

where $\nabla z_p$ is the isopycnal slope. An important implication of (4), to be used later, is that a combination of strong stratification $N^2$ and isopycnal slope $\nabla z_p$ can result in a large vertical velocity shear $\partial U/\partial z$. Similarly, the first term in the background PV gradient (3) can be rewritten as

$$\nabla \frac{\partial}{\partial z} \left( \frac{f_0^2}{N^2} \frac{\partial \Psi}{\partial z} \right) = -\frac{\partial}{\partial z} \left( \frac{f_0^2}{N^2} \hat{k} \times \frac{\partial U}{\partial z} \right) = -f_0 \frac{\partial}{\partial z} \nabla z_p,$$

where $U = \hat{k} \times \nabla \Psi$ and (4) have been used for the first and second equalities, respectively. Equation (5) implies that the presence of horizontal PV gradients in the interior is a function of the isopycnal slope variation with depth rather than of the isopycnal slope itself. Otherwise said, $\nabla Q$ is zero on the right-hand side of (1a) if the isopycnal slope is constant with depth, and the condition (iii) mentioned above is not satisfied, as is the case in the Eady (1949) and Hunkins (1981) models.

The specific solution of (1) with boundary conditions (2) depends on the background stratification $N^2$ and velocity profile $U$—or, equivalently, the isopycnal slope $\nabla z_p$ through Eq. (4). As we will show, the choice of stratification has important consequences for our ability to explain the observations presented in Fig. 1.

Our reference stratification $N^2$ and velocity profile $U$ are motivated by Arctic observations and shown as solid black lines in Figs. 3a and 3b. While idealized, they are representative of the Canadian basin stratification, shown in Fig. 2. Specifically, our reference stratification profile is a B-spline mimicking the two-peak structure characteristic of the mooring location at 75°N, 150°W. Our reference velocity profile has a maximum surface velocity of 2 cm s$^{-1}$, and decays exponentially with depth on a length scale of 200 m. This is the same velocity profile used by Hunkins (1981) in his baroclinic
instability analysis, and compares well with the currents’ speed profile computed from the World Ocean Atlas (WOA) climatology using the thermal wind relation (4) and assuming no flow at the ocean floor (thick gray line in Fig. 3b). Choosing a more complex velocity profile, including the inflection points visible in the WOA-derived velocity at approximately 25, 150, and 250 m, would introduce additional instabilities that do not provide any further insight to the physics discussed in this work. Accordingly, they are omitted from our discussion. The only difference between our analysis and the one performed by Hunkins (1981) is the prescribed vertical stratification.

Also shown in Fig. 3c is the quasigeostrophic PV gradient, computed from the reference stratification $N^2$ and velocity profile $U$. The PV gradient reversal within the halocline is a consequence of our choice of a realistic stratification, and will play a central role in facilitating the growth of subsurface eddies through the right-hand side of (1a). The quasigeostrophic PV gradient obtained directly from the WOA profiles exhibits a similar vertical structure, although more noisy (not shown).

Once the stratification and velocity profiles have been chosen, the growth rate, phase speed and vertical structure of the perturbations can be obtained by further assuming a separation of variables in the form

$$\psi = \hat{\psi}(z)e^{i(kx-\omega t)},$$

where $\omega$ is the complex frequency and $k$ the real wavenumber, and numerically solving the generalized eigenvalue problem obtained from (1) with boundary conditions (2). We discretize the 3800-m vertical domain with 1000 points, and a grid spacing ranging from 0.5 m at the surface to 7 m at depth. We refer the reader to Smith (2007) and Trodahl and Isachsen (2018) for details on the numerical solution of the eigenvalue problem.

One might wonder whether the quasigeostrophic approximation is appropriate for our problem. The large background Richardson number $Ri = N^2/(\partial U/\partial z)^2$ (Fig. 3d) confirms that this is indeed the case. As discussed by Stone (1971) and Thomsen et al. (2014), the quasigeostrophic approximation is valid for $Ri > 1$, with ageostrophic and nonhydrostatic effects being relevant only for $Ri = \mathcal{O}(1)$ or lower. The combination of our reference stratification, reaching a minimum value of $10^{-7}$ s$^{-2}$ at the surface, and velocity shear, reaching a maximum value of $10^{-5}$ s$^{-1}$ at the surface (see Fig. 3), results in a minimum Richardson number $Ri = 10^3$. Even a two order of magnitude lower stratification of $N^2 = 10^{-7}$ s$^{-2}$ would result in a $Ri = 10$.

It is also important to note that Eq. (2) implicitly assumes that the predominant role of friction is to produce an Ekman pumping—the last term on the right-hand side. That is, we are considering an Ekman and mixed layer of infinitesimal thickness (Williams and Robinson 1974; Pedlosky 1982, section 4.6). This assumption does not affect the mesoscale baroclinic instabilities that are the main focus of this work.

In addition to the quasigeostrophic approximation implicit in (1) and (2), and the choice of stratification and velocity profiles shown in Fig. 3, we make the following assumptions:

- we neglect the effect of bathymetry;
- we set the Coriolis parameter $f_0$ to $1.4 \times 10^{-4}$ s$^{-1}$ and its gradient $\beta_0$, very small in the Arctic, to zero;
- we align the reference system with the flow, so that $U = (U, 0)$; and
- friction for all computations is set to zero ($d = 0$ m) at the ocean floor in (2).

The results presented below do not change appreciably for $d$ in the range 0–100 m at the ocean floor, or for values of $\beta_0$ different from zero but characteristic of the Arctic Ocean.

Results of our linear instability analysis are summarized in Fig. 4. They will be discussed next by first considering the effect
of stratification in the absence of surface friction, and then adding the effect of friction against sea ice.

a. Effect of stratification

We start by analyzing the effect of stratification by considering the ice-free, frictionless case [corresponding to $d = 0$ m in (2)]. The growth rates of all unstable modes are plotted as a function of the wavenumber in Fig. 4a (thick lines). Three different unstable branches can be identified, marked with blue, green and red. For each branch, the amplitude $|\phi|$ of the mode corresponding to the fastest-growing wavenumber is plotted as a function of depth with thick lines and matching colors in Fig. 4b: a surface-intensified mode (blue), a halocline-intensified mode (green), and a deep-intensified mode (red) can be identified, the latter decaying monotonically with depth down to the ocean floor at 3800 m (not shown).

The surface mode (blue) grows on a time scale of order 10 days, and is characterized by a phase speed of order 2 cm s$^{-1}$ ($\sim$2 km per day, not shown) and a horizontal length scale of order 100 m. Its signature is concentrated between the surface and the shallower peak in stratification located at 50-m depth (see inset in Fig. 4b). The halocline mode (green) is characterized by a slower time scale of order 2 months, a phase speed of order 1 cm s$^{-1}$ ($\sim$1 km per day) and a larger horizontal length scale of order 10 km. Its signature reaches its maximum in the halocline between the two peaks in stratification at 50 and 240 m, but, in the absence of friction, its imprint is still visible at the surface. The deep mode (red) is characterized by similar time and length scales, yet it has a much slower phase speed of order 1 mm s$^{-1}$ ($\sim$100 m day$^{-1}$), its maximum lies below the deepest peak in stratification, and decays across the halocline. As shown by Zhao et al. (2018) when making the assumption of an ocean at rest to determine the neutral (stable) modes, the shape of the stratification profile is largely responsible the vertical structure of the different modes. By relaxing this assumption, we have here identified growth rate, phase speed and vertical structure of the unstable modes.

The vertical structure shown in Fig. 4b is remarkably consistent with the picture of the kinetic energy field shown in Fig. 1: independent perturbations grow in the surface layer and within the halocline. Deeper eddies, laying below the 240-m peak and corresponding to the deep mode (red), are not shown in the climatology of Fig. 1 but can be seen in Fig. A1 in the appendix (see, e.g., in October 2012, and February and July 2013) and have been observed in other mooring- and ITP-based observations (see, e.g., Fig. 3 in Zhao and Timmermans 2015). It is important to note that the three modes grow independently, at different depths, and with different characteristic growth rates and horizontal length scales.

Note that while the shallower stratification peak at around 50-m depth extends across the whole Arctic Basin, all the way from North America shelf (left side of Fig. 2) to Fram Strait (right side), the deeper peak identifying the bottom of the halocline in the Canada basin is not present in the Eurasian side of the Arctic. Repeating the same analysis maintaining only the shallowest stratification results in only one subsurface mode extending from the bottom to 50-m depth and characterized by growth rate and length scale similar to the deep (red) mode, but the general structure of the problem remains unchanged (not shown). The effect of the variability in the surface stratification will be discussed in the next section.

We have shown that different perturbations can grow independently at different depths, but we have not yet explained the strong seasonality of the surface eddies. To understand the origin of this seasonality, we next introduce friction at the surface.

b. Effect of friction on the growth of perturbations

To explain the relationship between ice draft and surface eddy activity suggested by Fig. 1, we now introduce the frictional effects due to the presence of the ice cover. In our model, the surface Ekman layer depth $d$ is a proxy for the ice’s ability to sustain internal stress, itself showing a very large variability over the seasonal cycle. In the limit of negligible friction just discussed ($d = 0$ m, which can be interpreted as representing the ice-free summer or ice in free drift) we recover the free-slip boundary condition characterizing baroclinic instability analysis in open oceans. As sea ice friction and Ekman layer depth increases during autumn and winter, the growth rate should decrease due to the dissipation of vorticity by the second term on the right-hand side of (2), as shown by Williams and Robinson (1974) and Hunkins (1981).

Observational estimates of Ekman layer depth in the Arctic suggest a median value of order $d \approx 11$ m (see, e.g., Fig. 5 of Cole et al. 2014) or, equivalently, a vertical diffusivity $\nu_E = d^2 f/2$ of almost $10^{-2}$ m$^2$ s$^{-1}$. We then analyze the effect of friction by varying the Ekman layer depth between 0 and 12 m. The obtained growth rates and vertical structure of the perturbations are shown for selected values of $d$ in Fig. 4 in shades of blue, green and red, and the dependence of the growth rate and the fastest growing wavenumber on $d$ is summarized in Fig. 5. Even for very small Ekman layer depths, friction has a
strong effect on both the growth rate and length scale of the surface mode (blue). An Ekman layer depth of 2 \( (\nu_E \approx 3 \times 10^{-4} \text{m}^2\text{s}^{-1}) \) is enough to reduce the growth rate by more than an order of magnitude to below the level of the subsurface modes, with higher friction damping the perturbation growth even more (Fig. 5a, blue). At the same time, the fastest growing horizontal length scale increases from 100 m for the ice free case \( (d = 0 \text{ m}) \) to almost 1 km for \( d = 10 \text{ m} \) \( (\nu_E \approx 7 \times 10^{-3} \text{m}^2\text{s}^{-1}) \) (Fig. 5b, blue). The vertical structure of the surface mode shows the effect of dissipation. Increasing friction drives the mode toward zero at the surface, and a subsurface peak in the streamfunction amplitude is developed, but is still contained in the surface layer, as can be seen in the inset of Fig. 4b.

The strong effect of friction on the surface mode must be contrasted with its effect on the halocline (green) and the deep (red) modes. The only visible effect is a reduction in the surface amplitude of the halocline modes (Fig. 4b, green lines), while the bulk of the perturbations, lying within the halocline below the stratification peak at 50 m remains unchanged. More importantly, the growth rate and fastest growing length scale are unaffected by increased friction, as seen in Figs. 4a and 5.

It is noteworthy how the surface mode’s growth rate is of the same order as that of the halocline and deep modes one for an Ekman layer of order 1 m, while being 10 times larger in the ice free case \( (d = 0 \text{ m}) \) and 10 times smaller for \( d = 10 \text{ m} \). As suggested by the climatology of kinetic energy (Fig. 1), the growth of perturbations in the surface layer is strongly affected by interaction with sea ice, with even small friction reducing growth rates by an order of magnitude or more, as suggested by Fig. 4.

Strong stratification is required to protect the halocline mode from the effect of friction against the ice cover. To understand how regional and seasonal variability in the stratification affects our results, we repeat our baroclinic instability computations, this time varying the intensity of the upper peak from \( N^2 = 10^{-3} \text{s}^{-2} \), a value characteristic of the Canadian basin, to \( N^2 = 10^{-4} \text{s}^{-2} \), a lower limit for the stratification in the Eurasian basin. While varying the stratification, we keep friction constant by imposing an Ekman layer depth \( d = 5 \text{ m} \). As can be seen in Fig. 6, both the growth rate and wavenumber of the halocline mode (green) are affected once the stratification drops below approximately \( N^2 = 6 \times 10^{-4} \text{s}^{-2} \). For decreasing stratification, the model’s halocline is only partially shielded from the ice above, but, for the given Ekman layer thickness of \( d = 5 \text{ m} \), a very low stratification \( N^2 = 2 \times 10^{-4} \text{s}^{-2} \) has to be reached before the halocline mode growth rate matches the surface mode one.

c. Effect of friction on preexisting eddies

How about preexisting eddies generated, e.g., in ice free regions during summer? Their spindown time scale can be estimated by energetic considerations as (Pedlosky 1982, section 4.3)

\[
T_s = \frac{K}{W} = \frac{H}{d} f^{-1},
\]

where \( K \) is the kinetic energy of the eddy, \( W \) is the power dissipated by friction within the Ekman layer, \( H \) is the depth of the eddy, and \( d \) is the Ekman layer depth. If we consider a vertical scale \( H \approx 50 \text{ m} \), characteristic of the surface layer, and an Ekman layer depth of order 1 m \( (\nu_E \approx 10^{-4} \text{m}^2\text{s}^{-1}) \), the resulting time scale is about 4 days, with larger Ekman layer depths resulting in even faster dissipation times. This short time scale precludes the possibility of eddies traveling long distances within the surface layer while in contact with the ice cover. Indeed, any eddy with an even relatively high propagation speed of 5 cm s\(^{-1}\) would travel less than 20 km in 4 days before being dissipated by an Ekman layer that is only 1 m deep.

In contrast, as initially suggested by Ou and Gordon (1986), halocline and deep eddies are shielded from the ice by the strong stratification at 50 m depth. This situation is exemplified in Fig. 7, where we show mooring-based observations of a subsurface, anticyclonic eddy during the winter of 2017 (see also Fig. A1 for other similar cases). The eddy appears unaffected by the presence of thick sea ice (top panel) during its 1-month-long transit. The eddy induced isopycnal displacement (gray lines), coupled with the strong stratification (Fig. 7b), results in an inviscid thermal-wind shear [Eq. (4)] reducing the

---

**Fig. 6.** (a) Variability in the peak value of stratification affects the (b) growth rate and (c) wavenumber of the fastest growing perturbation. The Ekman layer depth is fixed at 5 m. Colors as in Fig. 4.
current intensity across the peak in stratification. Frictional dissipation within the Ekman layer is therefore reduced, and the eddy can move undisturbed under ice.

4. Pan-Arctic model

We have shown how a realistic stratification profile characteristic of the Canadian Arctic, combined with friction at the ice–ocean interface, results in an approximately 50-m-deep surface layer that is only seasonally baroclinically unstable. Below that, conditions favorable to the development of baroclinic instability can be found all year, and the mooring observation suggests that resulting eddies are not dissipated and can survive all year long. Using a pan-Arctic, high-resolution model we now explore if results from our linear analysis are broadly applicable to the entire ice-covered Arctic Ocean (despite the regional differences in stratification and sea ice conditions). This also allows us to verify if the results from our simple model can be extended to a well-developed eddy field where nonlinearities may become important.

Here we use results from a simulation based on the CREG12 configuration (Dupont et al. 2015), encompassing the Arctic and parts of the North Atlantic. It is based on the NEMO (Madec and NEMO Team 2016) and LIM3 (Roussset et al. 2015) numerical models for the ocean and sea ice components, respectively. LIM3 uses an EVP (elastic-viscous-plastic) rheology (Hunke and Dukowicz 1997). The ocean model has a high vertical (75 levels) and horizontal (3–4 km) resolution in the Arctic Ocean. Wavenumbers up to approximately $10^{-3}$–$10^{-4}$ m$^{-1}$ in Fig. 4a are then resolved (Soufflet et al. 2016), including a large part of the unstable wavenumber range for subsurface eddies (red and green lines) and the lower wavenumber range for surface instabilities (blue lines). Surface instabilities generating small submesoscale eddies as those observed by Zhao et al. (2014) are not represented in this model. See also Fig. 2 of Dupont et al. (2015) for details on the eddy resolving abilities of the model. The simulation used in the present study runs from 1979 to 2014. The atmospheric forcing is the Drakkar forcing set 5.2 (which is an updated version of the fields described in Brodeau et al. 2010). More details about the numerical design as well as representation of the mesoscale activity in the Arctic basin can be found in Regan et al. (2020).

The results of our analysis are shown in Fig. 8, where we present maps of 5-day averaged relative vorticity at 17 m (top) and 147 m (bottom) depth for the summer (left) and winter (right) of 2012, the year of the lowest sea ice minimum recorded. An animation of the vorticity field shown in Fig. 8 is provided in the online supplemental material. The ice concentration is largely below 80% (white contour) during summer and an active eddy field is present both at the surface and at depth. In contrast, the surface vorticity intensity is much lower during winter, when the ice concentration is larger than 80% over most of the basin. Note how the 80% ice concentration, broadly corresponding to the onset of internal stresses in the ice, marks the demarcation line between the active eddy field to the south and an eddy “desert” to the north. As suggested by our previous analysis, the subsurface vorticity (bottom panels) field is broadly unaffected by the seasonal cycle.

The evolution of the relative vorticity with the changing ice cover is summarized in the bottom panel of Fig. 8, where we plot the evolution of the vorticity field’s standard deviation computed over the orange area in the inset map, covering a region where the time-averaged ice concentrations is larger than 80% and the ocean depth is larger than 300 m. The contrast between the surface layer and in the halocline is stark. A strong seasonal cycle is seen in the surface layer (blue line), anticorrelated to the average sea ice concentration (black line, inverted scale), while the relative vorticity is roughly constant over the year in the deeper layer (orange line).

Our high-resolution model thus suggests that the prediction of our linear baroclinic instability analysis extends qualitatively...
FIG. 8. Relative vorticity (5-day average, normalized by $f_0$) at (top) 17- and (middle) 147-m depth for the 2012 (left) summer minimum and (right) winter maximum ice extension; the vorticity color scale is logarithmic between $10^{-3}$ and $10^{-2}$ to accommodate the larger vorticity east of Greenland, and linear for values closer to zero. The white line marks the 80% ice concentration contour. Light gray corresponds to shelf regions. The model is not eddy resolving in the Barents and Chukchi Seas [see Fig. 2 of Dupont et al. (2015) for details]. A time resolved animation is provided in online the supplemental material. (bottom) The time evolution of the vorticity field’s standard deviation at 17 (blue) and 147 m (orange) over the region marked in orange in the inset map. The black line is ice concentration (inverted scale).
5. Sources of baroclinic instability and implications for the large-scale circulation

As mentioned at the beginning of section 3, the development of baroclinic instability relies on the presence of a combination of density gradients at the boundaries and/or PV gradients in the interior (Eady 1949; Charney 1947; Phillips 1954). At the same time, baroclinic eddies tend to homogenize both fields. How can density and PV gradients be maintained? We discuss two possible mechanisms with the help of Fig. 9 and the schematic in Fig. 10.

In the surface layer, conditions favorable to baroclinic instability are provided by the large-scale density and PV gradients visible in Fig. 9 (black line and filled contours, respectively). These are produced by wind-driven Ekman transport accumulating freshwater in the Canada Basin and inflating isopycnals.

At depth, baroclinic instability can be enabled by the subsurface PV gradients marked by red and blue contours in Fig. 9. Sources of PV can be found at the outcropping of isopycnals and in the cyclonic, topographically constrained circulation of Atlantic and Pacific waters on the margin of the Arctic basin (Aksenov et al. 2011; Rudels 2012; Stabeno and McCabe 2020). As an example, Spall et al. (2008) has discussed how the Alaskan shelf-break current is a source of low PV. Its effect can be seen in Fig. 9 by the change in sign of the horizontal gradient of PV with depth occurring next to the North America coast at around 100-m depth (see the two leftmost arrows). Crucially, Fig. 9 shows how similar reversals of PV gradient with depth are present almost everywhere in the Arctic basin.

This picture is summarized by the schematic in Fig. 10. Ekman transport at the surface (red arrows) is compensated by surface intensified eddy fluxes (cyan arrows in the surface layer). In subsurface layers, PV gradients are instated by sources of PV represented by, e.g., the shelf currents (blue half circle), and are then smoothed by the resulting eddy fluxes (cyan arrows). Note that the presence of sustained PV gradients additionally implies a downgradient, residual volume transport along isopycnals. The resulting residual overturning circulation might be readily closed by diapycnal mixing across the dashed isopycnal (orange arrow), representing the weakly stratified region between peaks in stratification, visible in Fig. 2.

It is important to note that, in the absence of these PV sources bending isopycnals (as suggested by the dashed line in Fig. 10), no subsurface eddies would develop. Should all halocline isopycnals have the same slope the interior PV gradient would be zero, as discussed when we introduced Eq. (5). This would be the case if the slope of the isopycnals were solely set by Ekman pumping at the surface.

From an energetic perspective, the available potential energy for the surface instability is provided by the relatively large slope of the isopycnals shown as a solid lines in Fig. 10. In contrast, the available potential energy for the halocline layer is given only by the relative slope between the dashed and solid isopycnals. Once isopycnals are aligned, no interior PV gradient (and no potential energy) is available to the subsurface mode, which cannot grow as was the case in the analysis of Hunkins (1981).

6. Discussion and conclusions

Guided by observations (section 2), and informed by results of a linear stability analysis (section 3) and the output from a
high-resolution numerical model (section 4), we have addressed the peculiar vertical distribution of mesoscale eddy kinetic energy in the Arctic, and discussed its origins in the large-scale Arctic circulation (section 5).

As shown by the climatology in Fig. 1, the Arctic mesoscale activity is characterized by a marked seasonal cycle close to the surface: the vigorous summer eddy activity, captured by both remote and in situ observations, is followed by very quiet winters. In contrast, subsurface intensified eddies persist all year long below the strong peak in stratification at approximately 50-m depth.

Our linear baroclinic instability analysis, based on the stratification and velocity profiles shown in Fig. 3 and summarized in Fig. 4, combines the effect of friction and stratification to show that the Arctic is indeed only seasonally baroclinically unstable at the surface, but that independent instabilities can develop throughout the year at depth. Subsurface instabilities are enabled by interior PV gradients (Fig. 9) and are shielded from the ice by a strong stratification at approximately 50-m depth (Fig. 2). Both PV gradients and stratification extend across the entire ice-covered Arctic. Our results contrast with previous analysis by Hunkins (1981) and Manley and Hunkins (1985). Their stratification and velocity profiles shared the same vertical structure—an exponential decay with depth over the same length scale—thus preventing the presence of interior PV gradients. As a consequence, it was suggested that frictional dissipation against the ice cover would prevent the local development of baroclinic eddies across the entire water column, and that the origin of eddies in the Arctic interior could be found only in the more unstable coastal regions. We now know that conditions enabling subsurface baroclinic instability—namely, the peak in stratification visible in Fig. 2 and the change in interior PV gradient visible in Fig. 9—are present across the entire Arctic basin. Our numerical model results, summarized in Fig. 8 as well as in the animation provided in the supplemental material, suggests that the separation between the seasonality in the surface layer and the more persistent mesoscale activity in the subsurface layer extends beyond linear theory and to the entire seasonally ice-covered Arctic.

We discussed in section 5 how density and PV gradients in the surface layer are most likely sustained by the combined effect of (i) density variations between the saltier Atlantic waters and the fresher Canada basin and (ii) Ekman pumping, sloping isopycnals. At depth, interior potential vorticity gradients might be sustained by the topographically constrained circulation of Atlantic and Pacific waters (Nøst and Isachsen 2003), as discussed by, e.g., Spall et al. (2008) in the case of the Alaskan shelf current.

The presence of subsurface eddies has an important consequence on the Arctic large-scale circulation. As discussed by Nøst and Isachsen (2003) and Isachsen et al. (2003), PV conservation dictates that, in the absence of eddy fluxes, no transport across isobaths is allowed with the exception of the frictionally driven surface and bottom Ekman layers. Subsurface baroclinic instability provides an alternative mid-depth pathway for communication between the Arctic’s margins and its interior, providing an explanation for, e.g., the presence of a Pacific Water temperature maximum in the Canadian halocline (Timmermans et al. 2017; Spall et al. 2018) and the presence of Atlantic Waters in the central Canadian basin (Spall 2013). We stress that this could not be possible if subsurface eddies were not shielded from sea ice by the stratification.

Our analysis provides indications on how future changes in sea ice conditions and their seasonality will impact the Arctic Ocean mesoscale activity. In current conditions, the winter ice cover is strong and rough enough to prevent the growth of baroclinic instabilities close to the surface, thus blocking the main mechanism driving lateral mixing in the surface layer. The ice is therefore effectively shielded from intrusion of warm water from the side (Ramudu et al. 2018; Horvat et al. 2016). This would not be the case with thinner or smoother ice: less friction would result in a more active surface eddy field, even below ice. Consequences on the transport of heat below the ice, and on the regeneration of the ice cover the following winter, are to be expected. Changes in stratification, observed in the past decade (e.g., Cole and Stadler 2019), are equally important and affect the vertical structure of the eddies and the interaction of halocline and deep eddies with the ice. As summarized by Fig. 6, an erosion of the shallow peak in stratification, driven, e.g., by an increase of momentum transfer from the atmosphere associated with a reduced ice cover, would result in the halocline eddies interacting with the ice itself.

We stress that our analysis is based on idealized, but representative, profiles of Arctic stratification and currents. As such, it presents a qualitative picture of the dynamics associated with the interplay of stratification and ice friction in the Arctic. It now remains to render this discussion quantitative. More studies will be required to analyze the impact of sea ice state on ice–ocean friction, to include a more accurate representation of the Ekman layer, and to analyze the nonlinear saturation of perturbations. Regional and seasonal variability of stratification, mixed layer properties, and Pacific and Atlantic layer circulation should be taken into account.
Fig. A1. Observed kinetic energy at mooring A, 2003–18. Note that the vertical scale is linear until 50 m, and logarithmic thereafter.
Submesoscale ageostrophic instabilities have not been addressed here and require dedicated studies as done, for example, by Mensa and Timmermans (2017) and Manucharyan and Thompson (2017) for the Arctic, and Swart et al. (2020) and Biddle and Swart (2020) for Antarctica. Finally, we have neglected the role of a sloping bottom on the development of instability. Topographic suppression of unstable growth can be expected in the Arctic (see, e.g., Trodahl and Isachsen 2018; Manucharyan and Isachsen 2019). Our calculations here, based on realistic density profiles, suggest that impacts may be primarily on deep modes below the halocline, but a dedicated study of this is also warranted.

From the modeling perspective, it is important to note that our numerical model has a very high horizontal resolution of 3–4 km. The processes investigated here are thus largely parameterized in state-of-the-art climate models, characterized by grid size of order 100 km, and this will be the case for the foreseeable future. Current parameterizations for mesoscale turbulence are completely unaware of the effect of either ice–ocean friction or vertical variation of stratification on eddy fluxes. This will need to be addressed to improve the representation of the Arctic ocean, one of the largest sources of bias in climate models.

From the observational perspective, the surface layer is where a large part of the eddy activity and most of the variability take place, but is not sampled by current moorings—which cannot extend to the surface because of the presence of ice—and only partially so by ITPs, which are limited to depths larger than 7 m. This shallow region is also the location where we expect the largest impact from changes in sea ice conditions. More innovative in situ observations should provide very valuable information to further test our hypothesis, and inform the development and validation of models and parameterizations.

Acknowledgments. G.M. J.M., E.D., and J-M. C. thankfully acknowledge support from the National Science Foundation Division of Polar Programs under Awards 1603557 and 1543366. C.L., H.R., and C.T. thankfully acknowledge support from the French ANR through the project ImMEDIAT under Award ANR-18-CE01-0010. P.E.I. was partially funded by the KeyClim project, funded by the Research Council of Norway (Award 295046). Our collaboration has been made possible by funding from the MIT-France MISTI Global Seed Funds for the project “The Arctic in a Warming World.” The Pan-Arctic simulation was performed using HPC resources from GENCI-CINES, Award 2018-A0050107420.

Data availability statement. Mooring data were collected and made available by the Beaufort Gyre Exploration Program based at the Woods Hole Oceanographic Institution (https://www.whoi.edu/beaufortgyre) in collaboration with researchers from Fisheries and Oceans Canada at the Institute of Ocean Sciences. World Ocean Atlas climatologies are available from the National Ocean Data Center: https://www.nodc.noaa.gov/OC5/woa18/, see Locarnini et al. (2013); Zweng et al. (2018). Ice data used in Fig. 2 are available at https://data.seaiceportal.de/data/cs2smos/version1.3/n/, see Ricker et al. (2017). The outputs from the high-resolution Pan-Arctic model used in this paper are available from Camille Lique.

APPENDIX

Figure A1 shows the time series of mooring-observed kinetic energy from which the climatology in Fig. 1 is obtained.

REFERENCES


Faghmous, J. H., I. Frenger, Y. Yao, R. Warmka, A. Lindell, and V. Kumar, 2015: A daily global mesoscale ocean eddy dataset.


