Dear Editor and Reviewer,

We would like to thank you for accurately reading and commenting the manuscript, and suggesting how to improve it. Answers to your comments are given in details hereafter. We hope that you will find them satisfactory. All authors agree with the modifications made to the manuscript.
Reviewer comments are in bold, and are followed by our response (in blue) that includes changes and/or additions to the text.

For the authors,
Dorotea Iovino

Answer to **Referee #1**

General comments
This paper presents a summary of the basic performance for a _ten year run of a new 1/16th global model, which is to be employed for operational ocean forecasts.
Some performance comparisons are made with a similar 1/4 degree configuration of the model.
Overall I thought it was a well written paper which presented some interesting results.
Most of the analyses are fairly standard, but to me this is as one would expect for a system definition paper.
My key comments are the paper could make better use of the 1/4 degree parallel experiments they use and describe in places. For example:
*** It is key to emphasize that resolution is not a panacea – there are many atmospheric forcing set errors, and errors due to limited physics, i.e. parameterisation, e.g. mixing in particular. There may also be error cancellation, e.g. a bias might get worse due to removing a cancelling error, e.g. with ocean forcing set errors, by increasing resolution.**

We do agree that high horizontal resolution is not a panacea for all ocean modelling problems and it does not guarantee that all undesirable characteristics of our models are ameliorated. In fact, a number of prominent biases and model errors persist, or even worsen, despite increases in model resolution. The increase in horizontal and vertical resolutions needs to be accompanied by improved modelling of relevant physical processes at the appropriate scale. Model parameterizations that have been developed for coarse resolutions mat not be ideal for considerably finer spatial-scales and may need to be revised, requiring model developments. However, the finer resolution remains one possible way in which model capabilities can be enhanced, thanks to the explicit solution of eddies. In this paper, we want to document if and how an "eddy-resolving" ocean model resolution impacts the simulation of large-scale ocean variability with respect to an eddy-permitting configuration. We do believe that our GLOB16 results, in accordance to the current literature, provide compelling evidence that an eddying ocean component can significantly impact the simulation of the large-scale ocean dynamics. We modified the Introduction to account for this comment: "Although the increase of resolution does not necessarily lead *per se* to an improved representation of the ocean general circulation, the aim of this work is to evaluate the effect of the explicit solution of eddy dynamics at low- and mid- latitudes on the large-scale dynamics of a high-resolution global ocean model, compared to a coarser resolution configuration".

*** I would ideally prefer to see many more comparison sub-plots for existing figures to identify differences of the 1/12th compared to the 1/4 twin run so we can see how many of the features you describe are common to both 1/4 and 1/12 and how many are really down to**
We agree with the referee that a more detailed comparison between GLOB16 and GLOB4 can definitely improve the manuscript. To present a more complete comparison with GLOB4, we modified subplots and/or add text in any relevant part of section 3, trying not to alter too much the manuscript structure or increase its length.

*** The reader ideally needs to know exactly which parameters and config settings are different for the 1/4 degree run compared to the 1/12th run. For example, is it the same version of NEMO, you mention it has level vertical levels so some differences could be due to vertical rather than spatial resolution. Also what coefficient for isopycnal mixing on tracers do you use at 1/4 (as the 1/4 simulation and biases are quite sensitive to this)?**
We added the key differences in parameters and configuration settings at the end of section 2.

Minor comments:

Line 42-43 – I think it may depend on the definition but is there not a factor of pi between the Rossby radius (based on wave number) and eddy scale? This makes a difference as one only needs two grid cells per Rossby radius to get 6 cells per eddy. This would be worth clarifying?

The direct relation between the Rossby radius, R, and the baroclinic Rossby wavelength, λ, is $R = \lambda/2\pi$. The grid spacing Δ has to satisfy $\Delta < R$, and then $2\pi\Delta < \lambda$, following the traditional criteria for resolving a wave on a discrete grid. Nevertheless, Hallberg (2013) considered a $2\Delta < R$ criteria.

In order to clarify this, we modified the text in section 1 and added the following lines:

"Hallberg (2013) showed the model horizontal resolution required to resolve the first baroclinic deformation radius with two grid points, based on a Mercator grid. From his analysis, 1/4° Mercator spacing is insufficient to resolve mesoscale eddies that have a typical scale of 50 km at mid-latitudes."

Line 53 – should you mention 'often (wrongly) termed eddy resolving' in view of the fact you go on to state they are not eddy resolving at high latitudes as shown by Hallberg (2013)?

We apologise for pointing out that we did find this specific text neither in line 53 nor elsewhere in the manuscript. As reported by emails, we did not detect a direct correspondence between some line numbers indicated by the Referee's comments and the line numbers reported in the pdf file (available online http://www.geosci-model-dev-discuss.net/gmd-2015-268/gmd-2015-268.pdf). However, following this comment, we highlighted in the introduction that GLOB16 and configurations with comparable horizontal resolution do not resolve mesoscale globally, and are only eddy-permitting north of a certain latitude and over shelf regions. The test now reads "Resolving mesoscale eddy variability remains anyway elusive at higher latitudes. For example, in the Arctic Ocean where the first Rossby radius decreases down to few kilometres on the continental shelf or in weakly stratified regions, typical eddy-resolving resolution does only permit eddies at best (Nurser and Bacon 2014)."

Line 71-74 – This is a long sentence. Also should you qualify this statement in view of your paragraph above, i.e. state we are now able to at least resolve eddies mostly equatorward of 50-60N/S BUT we don't resolve high latitude eddies or sub mesoscale or associated energy cascade anywhere. Furthermore, results are sensitive to grid scale closure, particularly viscosity, as you state in your eddy kinetic energy section?

In accordance with the Referee's suggestion, text has been added to specify where the 1/16° resolution is not eddy-resolving: "In this context, we developed a global eddying configuration, where eddying means that the numerical simulation is eddy-resolving in most deep ocean regions equatorward of 60°, while it is mostly only eddy-permitting at higher latitude."

Line 88 – As stated above should it be described as a step forward, particularly for mid latitudes where this resolution resolves eddies but 1/4 doesn't?

We do confirm that, as stated in the abstract, GLOB16 configuration represents a step forward in the CMCC global ocean modelling to resolve eddies in the ocean at mid-latitudes. At line 88, we want to emphasise also that this configuration is, for our modelling group, an accomplishment that open the way toward a new, operational short-term ocean forecast system. We did slightly modify the sentence in "…is a foothold that opens the way for the development of a new, operational short-term ocean forecast system…"

Line 123 – what about connection from Marmara Sea to Aegean - Dardanelles Strait is very important for seasonal freshwater input to Northern Aegean.

We did not mention the Dardanelles Strait into the bathymetry description (section 2.2) because no specific modification was applied there. After interpolating GEBCO dataset on the GLOB16 grid, the Dardanelles strait resulted open and 3-grid-point wide in its narrowest area. We reasoned that additional hand editing was unnecessary in that location. On the other hand, the Bosphorus strait (with a maximum width of ~3.5 km) was close and we had to modify the bathymetry to connect the Marmara Sea and the Black Sea.

Line 127 – Comment only – I believe some 1/12th NEMO configurations use partial slip in Labrador Sea to generate more eddies to help re-stratification after convection?
This comment refers to ORCA12 configuration. We are aware of a set of sensitivity tests performed by the DRAKKAR group to study the impact of lateral boundary condition on the mixed layer depth (MLD) and eddy kinetic energy (EKE) in the Labrador Sea, the Mediterranean Sea and other location where the boundary current dynamics is relevant. Free slip, partial slip and "variable (local partial)" slip were considered. In the Labrador Sea, for example, their results showed that local partial slip helps to increase EKE with its maximum well located between 60-62°N, and also reduce MLD. They suggested that the best choice is a combination of free slip applied everywhere except few limited patches with no-slip. Those locations have to be identified in order to build a sort of mask. We do agree that this approach might help to improve the GLOB16 results in some areas. A set of improvements is already planned for this configuration, some of them already employed. The lateral boundary conditions are for sure on the list.

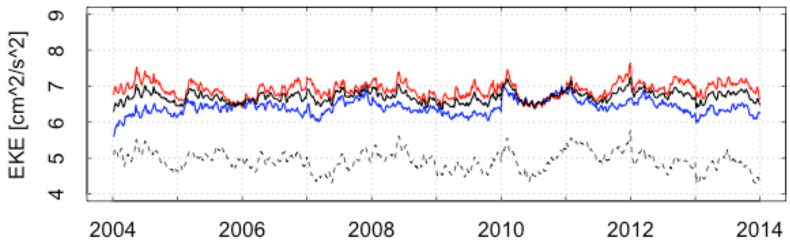
Line 141 – Comment only – important to note that uncorrected ERAI interim fields, e.g. radiation fields due to cloud errors, will have large errors which would be expected to impact on or even dominate near surface biases.
We do agree with the referee that errors in the atmospheric forcing may largely affect surface biases. A set of corrections can be applied to ECMWF ERAinterim variables to reduce global and regional biases. For example, in developing the DRAKKAR Forcing Set, Brodeau et al. (2014) corrected ERAinterim winds (weaker than QuikSCAT in the inter-tropical band between 40°S and 40°N), modified air temperature and humidity in the polar regions, reduced the shortwave radiation and increased the longwave radiation, and corrected the precipitation field in the western tropical Atlantic and Pacific oceans. How those corrections can improve the forced simulation has been addressed. Nevertheless, we decided not to follow this approach, but rather using some surface restoration (see reply to next point). Our decision may be justified by a number of arguments. Our ocean-modelling group generally uses the uncorrected-ERAinterim atmospheric reanalysis, conducting an analysis of the effect of the quality of the atmospheric forcing on the high-resolution system was beyond the scope of this study, and this GLOB16 simulation has been used as bedrock for the CMCC global ocean forecast system that uses global ECMWF operational system and forecast as it is released. Note that applying corrections to operational analyses/forecasts is rather dangerous, due to possible unexpected changes in quality of the real-time atmospheric forcing.

Line 158 – Why do you need to use SST restoration? This will mask model errors in near surface fields and there is already inherent relaxation back to air temperature in the forcing set?
The SST relaxation is implemented for a twofold reason: to limit the propagation of the atmospheric forcing biases onto the upper ocean and to compare the results with the twin experiment GLOB4 at lower resolution. The value of the relaxation time-scale was set equal to that used in the 1/4° reanalysis and simulation system (Storto et al., 2016; Haid et al. submitted to The Cryosphere).
However, we acknowledge that, due to the SST relaxation, the verification of the sea surface temperature may be misleading and driven by the relaxation dataset and time-scale. Thus, in the revised version of the manuscript, we focused our analysis (in section 3.1) on the subsurface biases.

Storto, A., S. Masina, and A. Navarra (2016), Evaluation of the CMCC eddy-permitting global ocean physical reanalysis system (C-GLORS, 1982–2012) and its assimilation components. Q.J.R. Meteorol. Soc., 142: 738–758. doi: 10.1002/qj.2673

Haid, V., D. Iovino, and S. Masina. Impacts of Antarctic runoff changes on the Southern Ocean sea ice in an eddy-permitting sea ice-ocean model. Submitted to The Cryosphere

Line 194 – I wonder why you didn't remove the seasonal cycle of MKE from EKE as otherwise the seasonal cycle of flows will be included in EKE estimates?
Line 194 falls within section 2.7 on the "Output and analysis strategy". We assumed the referee's comment is about the calculation of EKE in section 3. As suggested, the seasonal cycle was estimated and then removed from the time series (Plot 1).



Plot 1. (Fig. 1b in the manuscript) Time variations of volume-averaged EKE (in $cm^2 s^{-2}$), where the black line represents the global basin-mean value and the red (blue) the contribution of the Southern (Northern) Hemisphere in GLOB16. Thin-dashed line represents the basin-mean EKE in GLOB4.

Line 215 – Can one really say much about SST biases from a ten year run with SST relaxation?
We do agree with the referee. Due to the length of the simulation and the temperature restoring that we applied at the sea surface, the analysis of the SST biases does not add much to our study. We removed the plots showing the surface biases (Fig 2a, b) and accordingly correct the text at the beginning of section 3.1 as follow "The mean fields of modelled potential temperature and salinity are here validated against the mean of EN3 climatology (the UK Met Office Hadley Centre observational dataset, Ingleby and Huddleston 2007), both averaged over the same period 2009-2013. As expected, due to the temperature and salinity restoring applied at the ocean surface, the global mean SST and SSS biases are small (-0.06 for SST and -0.04 for SSS). There are weak cold biases in the tropics, extending over much of the subtropical band, with the largest SST biases (~1 °C warmer) collocated with positive SSS error (0.5–1.5 psu) over the western boundary currents in the Atlantic and North Pacific oceans (not shown). The overall pattern of surface biases is similar between the two models."

Line 215 – Should you show equivalent plots for figs 2 for 1/4 degree or state that they are indistinguishable from the 1/12th if it is? The difference with the 1/4 degree model is surely a key result?
We followed this comment and we added two subplots (e, f) in Figure 2 showing the zonal mean temperature and salinity differences between GLOB4 and observations. Section 3.1 includes now the following text "Although the overall biases are similar between the two model configurations in many latitude bands, there are some relevant differences (Fig. 2e, f). For instance, the Southern Ocean is generally warmer in GLOB4, with a larger positive salinity bias at ~400 m depth around 50°S. Both models are warm and saline in the above depth range in the northern mid and high latitudes, but the biases differ in magnitude and locations, highlighting the difference in path of the western boundary currents. Both models are warmer than observations in the Artic Ocean: the largest warming is confined in the upper 200 m depth in GLOB16, while the maximum, with a similar value, is located between 300 and 500 m depth in GLOB4."

Line 236 – Will the deep (>1000m) ocean really have equilibrated in a ten year run? I am guessing you are probably looking mainly as biases due to isopycnal heave which occurs reasonably fast?

We do agree with the Referee that the integration time required to "spin up" an ocean model from an initial state of rest to a near equilibrium state is much longer for the assessment of the deep ocean. Unfortunately, obtaining equilibrium at eddying resolution is not practical. In 10-year integration, GLOB16 is still in its adjustment phase, especially in the deep ocean. This short simulation is not appropriate for studying the long-term evolution of deep-water masses, as we assert in section 2.6, "The model run for 11 years through the end of 2013, which appears to be a sufficient amount of time for the near-surface velocity field to adjust to the initial density field and for mesoscale processes in the upper ocean to have reached a quasi-equilibrium, while the deep ocean takes much longer to reach steady state." In the paper, we are limited to describe the state of the deep ocean hydrography and overturning circulation after 10 years of integration.

Line 253/296 – More discussion or plots for GLOB4 on figs 3 and 4 would be useful?
Figure 3 includes the time series of AMOC at 26.5°N as reproduced by GLOB4, together with GLOB16 output and RAPID estimates. These results are discussed in section 3.2. We decided not to add plots of the MOC stream functions in the Atlantic, Indo-Pacific and Southern oceans for GLOB4, but we added descriptions of the differences in the overturning circulation between the two configurations for each basin. Both subplots in Figure 4 included GLOB4 results (black lines) as the time-mean Atlantic MHT and its 10-y time series at 26.5°N, also in the previous version of the manuscript. A more detailed description has been added.

Line 263 – Surely one can not say much about AABW in a ten year run?
We do agree that in GLOB16 we cannot evaluate the performances of the model to reproduce the deep ocean circulation in a 10-year run. We limit our purpose to providing estimate of the mean model transport in the bottom layer in such a short simulation.

Line 317 – Implied heat transports assume equilibrium? How large are you heat content tendencies with a short ten year run that may well still be drifting?
One caveat of this study is that the simulations lasted only 10 years.
Comparing actual ocean heat transports with those implied by surface fluxes gives an indication of the volume-averaged drift in temperature. The time evolution of the volume-averaged ocean temperature is shown in Fig 1c to demonstrate the extent to which a quasi-steady state has been reached at the end of the 10-year integration. The potential temperature is seen to decrease by ~0.005°C over 10 years. This drift (related to discrepancies between the actual heat transports by the ocean and the heat transport implied by the surface fluxes) is small enough to assume a close agreement between implied and actual meridional heat transports. That suggests that the GLOB16 ocean is close to quasi-equilibrium.
Then, the time series of the MHT in the Atlantic Ocean at 26°N (at other selected latitudes in the Atlantic Ocean as well – not shown in the paper) does not present any particular trend over the 10-year time window.

Line 430 – I did not think the location of mixed layer maxima agreed so well with the observations in Southern Ocean?
Line 430 is in section 3.3 and concerns the transport in the Mozambique Channel. We think that this comment refers to section 3.4 where locations of mixed layer maxima were mentioned in lines 476-477 (following line numbers in previous section). We deleted this sentence and we kept line 489-490 in which we state that maxima in the Southern Ocean Pacific sector are "not exactly collocated with the observed ones".

Line 430 – How does GLOB4 look in spatial plots? If it is very similar is it worth stating this?
Figure 6 includes now a subplot showing GLOB4 mixed later depth averaged in the Northern (Southern) in March (September) 2009-2013.

Line 439 – I wonder if a different temperature based criterion might make the model mixed depths appear better? For example, if there is salinity compensation then density criterion can be rather sensitive to salinity errors.
We applied different criteria to our model output to compute the mixed layer depth. The differences are mainly local and we did not find any general improvement or worsening using different MLD calculations in both hemispheres.

Line 550 – I would emphasize the point about viscosity sensitivity more and include it in introduction. There is often an optimum viscosity level for EKE (and associated MKE) as too little enhances grid scale noise which damps eddies and too much obviously damps eddies.
A clear weakness of this first GLOB16 experiment is its ability in reaching the observed magnitude of EKE, especially in the Southern Ocean. This behaviour is most likely related to the coefficient chosen for lateral momentum diffusion. This GLOB16 simulation suggests that more efforts shall be dedicated to sensitivity experiments for detecting the optimal configuration of horizontal and vertical dynamics. Unfortunately, re-running 10-year experiments is not faceable at the moment. To improve this aspect of GLOB16, we are currently performing short (2year) test experiments.

Line 617 – Southern Ocean MLD maxima appear also not too good in Southern Ocean?
Already answered (please see above the comment to line 430)

Line 633 – As it appears that you change both vertical and spatial resolution it is hard to conclusively attribute GLOB16 versus GLOB4 differences to spatial resolution. Could you state all configuration differences between GLOB4 and GLOB12 configurations as bullet points. In an ideal world one would mimimise the differences, e.g. use same number of vertical levels and vary only viscosity, isopyncal diffusion and perhaps slip between the two runs?
Key differences in parameter settings between the two configurations are now listed in the end of section 2.
Vertical resolution may be adequate for the horizontal scales one is physically concerned with. Lindzen and Fox-Rabinovitz (1989) showed the substantial benefits in refining both horizontal and vertical resolution give some support to scaling arguments deduced from quasi-geostrophic theory implying that horizontal and vertical resolution ought to be chosen consistently. They developed simple physical criteria, based on Rossby radius and gravity wave dispersion, for the vertical resolution consistent with horizontal resolution.
Increasing the vertical resolution in the ocean model can provide a better representation of the upper ocean physics, improve the properties of dense water formation as well as the dynamics of overflow between ocean basins, and more accurately represent the bottom boundary layer physics. The DRAKKAR group showed, for example, improvements in ORCA12 representation of overflows when the vertical levels are changed from 75 to 300.
Both horizontal and vertical resolutions are needed to more accurately depict the 3D structure of the ocean. A multi-step approach to isolate the effects of any specific modification from GLOB4 to GLOB16 was, unfortunately, not viable. Therefore, GLOB16 includes both vertical and horizontal refined grids, and there are no twin runs in which the two configurations do share the same vertical discretization. We attempted to achieve a consistency between horizontal and vertical resolution in our GLOB16 configuration. Applying what we considered to be the most adequate choice (also considering the real limitation due to computational cost), we moved from 75 vertical levels in GLOB4 to 98 in GLOB16, with level spacing from 1 m near the surface to 160 in the deep ocean.

Lindzen, R. S., and M. S. Fox-Rabinovitz, 1989: Consistent vertical and horizontal resolution. Mon. Wea. Rev., 117, 2575– 2583.

We have made the following additional changes to the paper:
In the list of References, we updated Stepanov et al. 2016, Storto et al. 2016.

We corrected the Mozambique Channel transport in Table 2 and accordingly added the following reference
Ridderinkhof, H., P. M. van der Werf, J. E. Ullgren, H. M. van Aken, P. J. van Leeuwen, and W. P. M. de Ruijter: Seasonal and interannual variability in the Mozambique Channel from moored current observations, J. Geophys. Res., 115, C06010, 2010.