We thank Dr Danilov for his careful consideration of this paper and his helpful suggestions on how to improve it. He has some concerns, which we will aim to address in a revised version. The general criticism raised is that the work is insufficiently deep and lacks practical recommendations with respect to numeric approaches. The aim of this paper is to clearly articulate the problem, assess possible solutions and how practical they are on particular time scales. In the revision, we aim to improve the depth of this analysis by being more quantitative in our assessment, particularly by drawing more on the observational base (e.g. fig 4), how the current generation of global models (e.g. CMIP5 and where we have data new CMIP6 models) performs against this in the coastal regions and, whether the developments suggested (in fig 5) improve on this, using a direct comparison with observations. It is not the objective of this paper to offer
new algorithms to provide solutions to this problem. Describing specific, novel, numerical solutions to the issues presented is substantially beyond the scope of this work and could not be addressed in a single paper. We can, however, be more detailed in our description and assessment of solution options in the literature, and this will be a key aspect of the revision. He suggests several aspects that deserve greater attention, namely reducing spurious mixing, factors effecting stability and scale-aware mixing parameterisation. We agree these are all important and will ensure they are appropriately covered in the revision.

Dr Danilov very helpfully points out that we neglect the important, related, issues of time stepping, scalability and throughput, and this will be addressed in the revision. Essentially we focus on resource (cpu.hrs), rather than time to complete a simulation (throughput). The former is what is metered by our computer centres, but the latter limits how much science we can do with the resource. As grids are refined, throughput will reduce irrespective of available resource unless scalability can be approved (fewer grid cells efficiently allocated to each MPI process) or time stepping made more efficient. We will consider if these can be included in our cost model, and will add to the discussion.

Dr Danilov is correct that our considerations of unstructured mesh models is somewhat out of date, and this will be addressed in the revision. Particularly we will look in the literature for better estimates for the addition costs of unstructured v’s structured models to improve on the factor of 5 we use in our cost model. We will also consider in more detail what the factors affecting this ratio are (complexity of code, indirect memory addressing etc), and how they can be alleviated. It is not our intention to disparage unstructured mesh approaches here, but rather make a balanced, realistic assessment.

The more specific comments, which we generally agree with (with some minor exceptions), will all be addressed in the revision. Particularly sections 1.2 and 1.3 will be combined and reduced, with aim of making our motivation clearer. We agree that
figure 10 and the discussion around it can be deleted.

Interactive comment on Geosci. Model Dev. Discuss., doi:10.5194/gmd-2016-145, 2016.