model documentation

Markus Dunkel

March 2020

1 How to use the model

The model can be fully initialized with the help of *init.py*. To run the model locally, *run.py* can be executed. To run the model on the server I use the *neccluster.nqs* file. Since $self.save\_output$ is set to False no output will be provided. However, the results can be viewed in the local copy of the Processing class.
The most important settings are:

- $self.max\_time\_step$, number of time steps
- $self.dt$, timestep length
- *domain_split*, number of domain splits for parallelization at first instance.
- *intern_split*, number of parallelization for the SWM integration for each subdomain.
- *sample_number*, number of stored timesteps. The very last time step is always included.
- $self.mode\_list$, included vertical modes, 0 mode is the first baroclinic mode and so on.

2 A few words about the parallelization

1. The model includes three parallelization strategies at the moment. At the first instance, there is the possibility of domain splitting according to Martins domain module realized with ray. Set number and organization with *domain_split* attribute.
2. On a second instance, stateful workes for the Shallow Water integration can also be initiated with ray. Set number with the attribute *intern_split*.
3. Computations beside the SWM integration can be separated into zonal momentum, meridional momentum and density equation. Here, with the help of ray, a parallelization is done automatically. However, the workers

are not stateful since no data need to be stored. At the moment this parallelization includes the vertical mixing module solely.

An exemplary setting:
$domain\_split = (1, 3)$,
$intern\_split = 3$

Three independent domains and three bunches of SWM models make at maximum 20 parallel processes at a time. The reason is that there is i.) always a stateful worker initiated for processing the output, ii.) everything besides the SWM integration is split into computations for density equation, zonal momentum and meridional momentum and iii.) the main instance is held independent. The exemplary setting makes at maximum 3(3+3)+2= 20 active worker at a time.
On top, the computational most expensive loops in the vertical mixing module are potentially parallelized with numba.
The optimal choice so far:
With the used resolution ($N\_y = 101, N\_x = 200, N\_z = 900$) the best choice is to set $domain\_split = (1, 1)$ and $intern\_split = 3$ The computational time for one step is then 0.6 s.