A variable number of neighbours for OASIS-SCRIP nearest-neighbours interpolations

E. Maisonnave

TR/CMGC/16/29508

Table of Contents

Light modifications in the legacy SCRIP/OASIS code added an option to the popular nearest neighbours (NN) interpolations: the number of NN is adapted to the local ratio of areas between target and source grid points. This enhancement reduces the already quite satisfactory errors when the spatial variation of the resolution ratio is strong, near the coast line and also, but more marginally, on offshore areas.

Discussion

A large community, encompassing climate modelling laboratories all over the world, regularly uses the OASIS coupler [1] to link geophysical models into larger ensembles like Earth System models [2,3,4,5]. Coupling fields, exchanged at model interface, are discretized over different grids. An important function of couplers consists in interpolating this information from source grid to target grid. OASIS relies on the SCRIP interpolation package [6] to perform such operation. Basically, users have to determine which interpolation better fits their needs between (i) nearest-neighbours (Gaussian or distance inverse weighted), (ii) bilinear or bicubic and (iii) locally conservative interpolations. The new fully parallel OASIS3-MCT library slightly modified the choice constraints. As experienced by groups in charge of CMIP6 coupled system implementation [7], the bicubic and 2nd order conservative interpolations require additional arguments (basically field derivatives) that are not necessarily easy to provide in a parallel context. First order conservative interpolation also shows limitations in case of mismatches between coast line represented in source and target grids and its use could lead to strong errors in these regions. In most of the cases, the bilinear interpolation is not a satisfactory rescue choice, due to the constant minimal number of source grid point involved (only 4). That's why nearest-neighbours (NN) interpolations stay the unavoidable choice for users who don't want to spend time in off-line calculation of their own weight and address (W&A) file (usable through the OASIS “MAPPING” option). However, in some cases, the NN OASIS interpolations are not fully satisfactory.

Among the grids used in the large OASIS user community, most of them exhibit dispersion in their (unmasked) grid point areas (see table 1).

Grid (Model)	Acronym	Normalized standard deviation
T127-Gaussian (ARPEGE)	t127	19.2%
T135-Gaussian- 2.5 Stretched (ARPEGE)	s135	107.5 %
T359-Gaussian (ARPEGE)	t359	6.7%
GL05 (NICAM)	gl05	~ 0%
ORCA2 (NEMO)	or02	56.0%
ORCA1 (NEMO)	or01	53.5%
ORCA025 (NEMO)	or25	59.1%

Table 1: Grid point area variation (normalized standard deviation)
for OASIS coupled atmosphere and ocean grids

Variation of grid point areas from icosahedral (GL) and Gaussian grids is, by definition, close to zero or at least, reduced. It is not the case for ORCA class grids, the global grids of NEMO [8] model. Like other so called lat-lon grids (regular in latitude-longitude degrees), grid point areas differ at pole and in tropical regions. The maximum of variance is reached with the ARPEGE stretched grid [9]: its geometry reduces resolution near the pole of interest and increases it at its antipode.

It is quite usual that two components of an OASIS based coupled system rely on at least one resolution varying discretization. In this case, the ratio between areas of source and target grid point is also geographically varying. But, the OASIS interpolation parameter that defines the number of neighbours in all OASIS NN interpolations must be constant. From this we deduce that, in case of large variation in resolution ratio, a NN interpolation must be computed with too few source grid points (sub-sampling) in some regions and/or with too much source grid points in others. The first error leads to over-estimated gradients, the second has a smoothing effect. In this report, we describe a solution implemented in the OASIS code (including the SCRIP library) that removes the fixed NN number effect. We show its efficiency in selected examples.

Unlike more recent couplers [10], OASIS W&A calculations are sequential. Ample modifications would have been necessary to implement our new NN varying number algorithm with the same kind of parallelism as described in [11]. For this reason, our interpolation is also sequential. Nevertheless, considering that (i) a large part of OASIS based coupled systems have moderate resolution and parallelism, or enough time to perform the W&A operation – static interpolation - (ii) the automatic calculation of interpolation W&A is a function appreciated by the OASIS users that avoid to install, compile, parametrize and test an external solution which W&A results must, most of the time, be made conformable with SCRIP/OASIS format and (iii) due to its large diffusion among climate modelling community, the OASIS standard is supposed to stay standard in the years to come, we intent to deliver a solution that can be easily included in the next OASIS release and be easily tested (and reviewed) by users of the broad and stable OASIS community.

Principle & implementation

We propose to calculate, *for each unmasked target grid point*, a variable NN number among the source grid point. This number is a function of the ratio between the area of the target grid point (At) and the area of the the first neighbour in the source grid (As) :

$$NN = max\left(\left|At/As + 4\times\sqrt{At/As}\right|; 4\right)$$

In addition to the simple ratio *At/As*, the number of source grid points that cover the area of the target grid point, we add source grid points that can possibly be intersected by the *four* sides of the target grid point, assuming that grid cells only have 4 sides. We bound the minimum NN number to 4. This assumption potentially smooths gradients in areas where source grid point areas are bigger than target grid points.

OASIS code modifications are rather limited:

- the content of `areas.nc` OASIS auxiliary file must be read by the PSMILE library

(`mod_oasis_map.F90`)

- and transmitted to the SCRIP library (`scrip.F`)
- where they are necessary to calculate the NN number (`remap_distwgt.F90` but also `remap_gauswgt.F90`, the same algorithm being suitable for the Gaussian NN interpolation)

The prior calculation of the NN number requires a modification of the W&A algorithm. In a first step, the nearest neighbour position is searched and its area used in the NN number calculation. Then, the first source/target grid point distance is stored. The search of the (NN-1) other neighbours and computations of their distances to the target grid point can start as previously. The final weights normalization remains unchanged.

Results

An arithmetic coupled field (sinus of slowly varying latitude and longitude) is interpolated with two simple inverse distance NN interpolations, with constant and variable NN number, from/to a couple of ocean/atmosphere grids including land/sea masks. These grids are currently operated in the community. The constant NN is deduced from the mean *At/As* ratio and cannot be lower than 4. Two toy models are used to perform the W&A and error calculations without involving real climate models. The corresponding error (averaged over all unmasked target grid points) is shown on table 2.

Source/Target Grid	Constant NN: Mean error (o/oo)		Variable NN: Mean error (o/oo)
or01 / t127	4	0.50	0.46
or02 / gl05	4	0.86	0.77
or25 / t359	4	0.33	0.19
or01 / s135	4	0.36	0.30
s135 / or01	4	0.88	0.90

Table 2: Mean errors of interpolations with constant or variable NN #

The variable NN interpolation gives always better results, except for the last line, where resolution ratio is practically always lower than 1. Nevertheless, this performance must be moderated, considering that the mean error over the global domain is relatively small. It is smaller when resolution increases (or25/t359).

Mean error is a biased indicator of interpolation error in offshore regions. On Fig 1 and 2, we show the spatial map of errors in case of constant (bottom left) or variable (bottom right) NN interpolations for two couples of source/target grids On Fig 1, the interpolated field (top) is added.

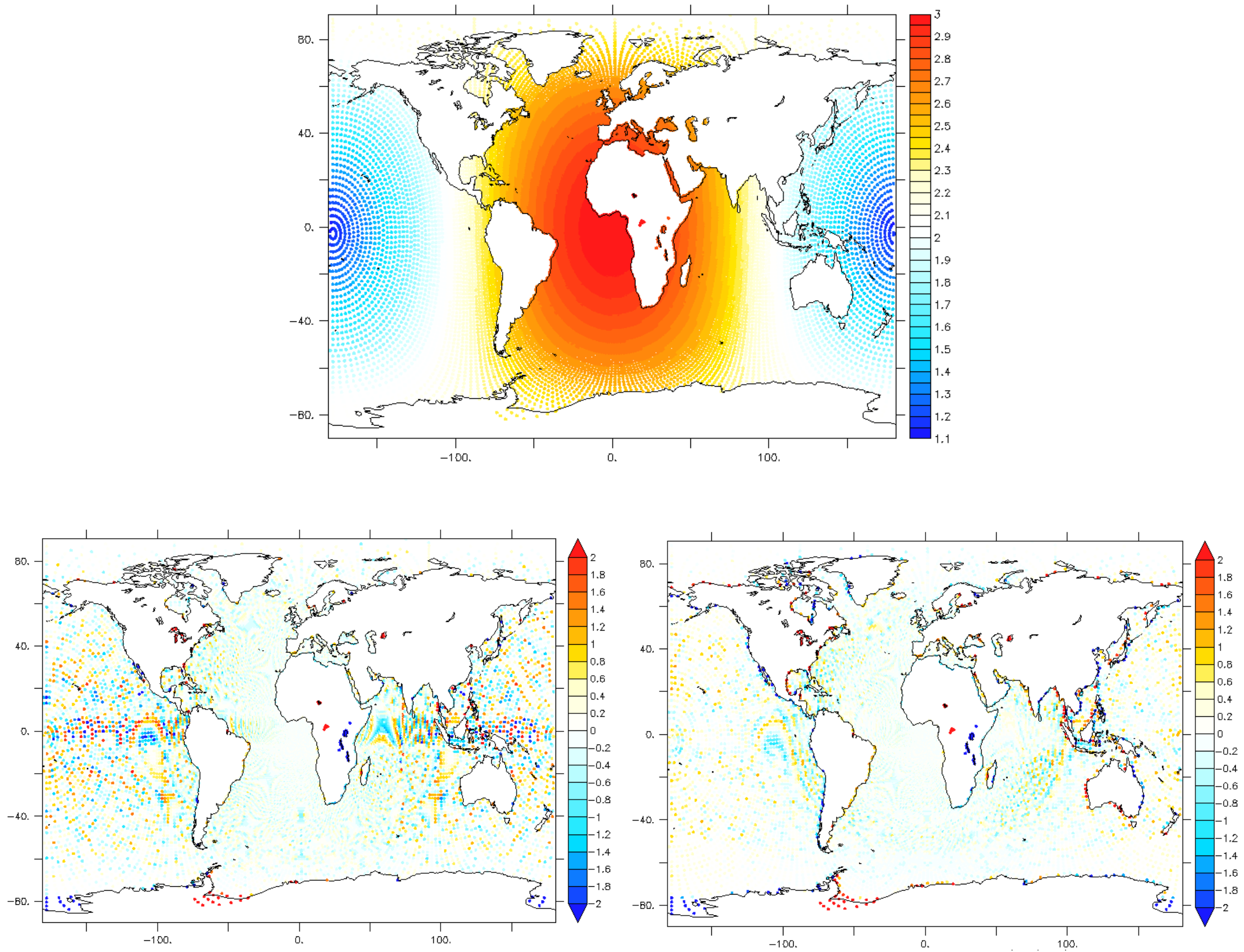

Figure 1: Values (top) and errors (in ‰) of constant (bottom left) and variable (bottom right) NN interpolations of a large scale arithmetic field from NEMO ORCA 1 degree grid to an ARPEGE stretched grid (T135 centred on Guinea gulf, stretched factor 2.5)

The major errors (greater than 2 or lower than -2 per thousand) are concentrated near the coastline (or in lake or inner seas) and their origin is the land/sea mask mismatch in the source and target grids. On Figure 2-left, the difference between variable and constant interpolation (negative when variable interpolation is better) shows that the offshore errors are corrected by the variable NN number and compensate a coastline downgrading in the mean error reported on table 2. Also visible comparing Fig 1-left and right, the mean correction on offshore areas should not exceed a few tenths of per thousands.

The coastline downgrading of interpolation performances is due to the difference of NN number. In the constant case, a single unmasked nearest neighbour is searched when all nearest neighbours are masked. In the variable case, a variable number of (not lower than 4) unmasked neighbours is searched anyway. The resulting smoothing effect significantly increases the error in these areas. To address this issue, we could implement in the variable NN interpolation the same kind of algorithm that attributes a single unmasked neighbour to a target point with all NN masked. But we must be careful in our conclusion: the error estimation in this area is only possible for arithmetic coupling field and it is not obvious that this basically wrong operation (how to interpolate values in areas where coupling field *is not calculated at all* by the source model ?) would be more wrong with our new method.

Figure 1 also emphasizes the capacity of the new interpolation to reduce errors when gradient are moderates. On strong gradient areas (Peru coast, offshore Indian ocean and more generally speaking on the white areas of the interpolated field represented in Fig 1-top), it reveals its limitations: using a larger number of source grid points implies a smoothing of the interpolated field. It should be possible to attenuate this drawback with the appropriate coefficient of a Gaussian weighted interpolation.

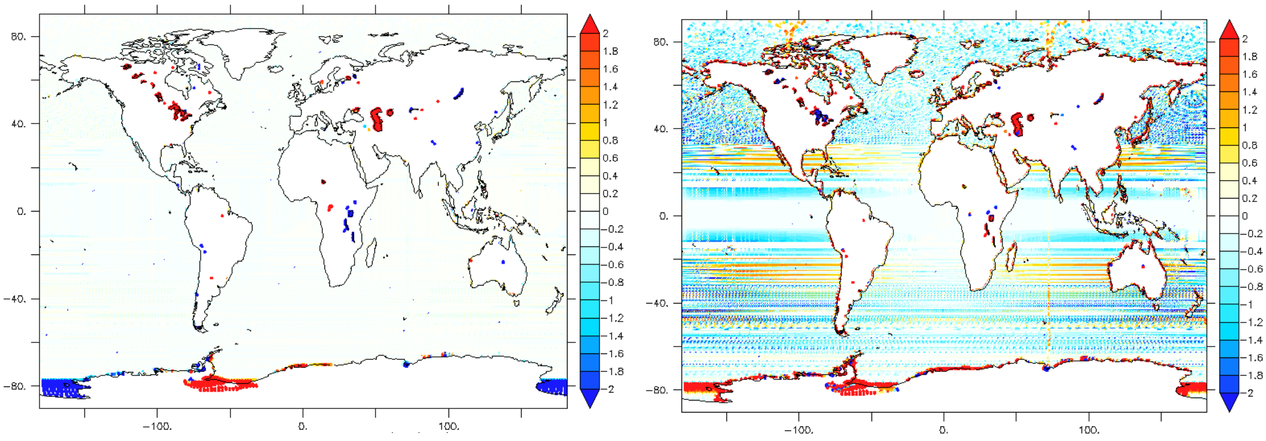

Figure 2: Error (in ‰) of constant NN interpolation (left) and error difference (in 1/10 of ‰) between variable and constant NN interpolations of a large scale arithmetic field from NEMO ORCA 1/4 degree grid to an ARPEGE Gaussian grid (T359)

Conclusion

Light modifications in the legacy SCRIP/OASIS code added an option to the popular nearest neighbours interpolations: the number of NN is adapted to the local ratio of areas between target and source grid points. This enhancement reduces the already quite satisfactory errors when the spatial variation of the resolution ratio is strong, near the coast line and also, but more marginally, on offshore areas.

References

[1] Valcke, S., Craig, T., Coquart, L., 2015: OASIS3-MCT User Guide, OASIS3-MCT 3.0, Technical Report, TR/CMGC/15/38, CERFACS/CNRS SUC URA No 1875, Toulouse, France
[2] Hazeleger, W., Wang, X., Severijns, C. et al, 2012: EC-Earth V2: description and validation of a new seamless Earth system prediction model., *Climate Dynamics*, 39, 2611-2629, doi: 10.1007/s00382-011-1228-5
[3] Voldoire, A., Sanchez-Gomez, E., Salas y Mélia, D. et al., 2012: The CNRM-CM5.1 global climate model: description and basic evaluation, *Climate Dynamics*, DOI:10.1007/s00382-011-1259-y
[4] Marti, O., Braconnot, P., Dufresne, J.-L. et al., 2010: Key features of the IPSL ocean atmosphere model and its sensitivity to atmospheric resolution, *Climate Dynamics*, VOL. 34, DOI:10.1007/s00382-009-0640-6

[5] Bi, D., Dix, M., Marsland, S. J. *et al.,* 2013: The ACCESS coupled model: description, control climate and evaluation. *Aust. Meteorol. Oceanogr. J*, 2013, vol. 63, no 1, p. 41-64.
[6] Jones, P. W., 1997: A user's guide for SCRIP: A spherical coordinate remapping and interpolation package. *Los Alamos National Laboratory*, *1*.
[7] Maisonnave, E., Coquart, L., Fladrich, U. et al., 2016: Model interface updates for compliance with new OASIS coupler, Technical Report, TR/CMGC/16/29506, CERFACS/CNRS CECI UMR No 5318, Toulouse, France
[8] Madec G. 2014: "NEMO ocean engine". Note du Pôle de modélisation, Institut Pierre-Simon Laplace (IPSL), France, No 27 ISSN No 1288-1619
[9] Courtier, P., Geleyn, J. F., 1988: A global numerical weather prediction model with variable resolution: Application to the shallow-water equations. *Quarterly Journal of the Royal Meteorological Society*, *114*(483), 1321-1346.
[10] Hanke, M., Redler, R., Holfeld, T., Yastremsky, M. YAC 1.2. 0: An extendable coupling software for Earth system modelling
[11] Maisonnave, E., 2015: Interpolation weights calculation distributed with OpenMP® Working Note, WN/CMGC/15/5, SUC au CERFACS, URA CERFACS/CNRS No1875, France