Memo: how to compute statistical moments of a cell centred field using conservative regridding

Alex Pletzer, Jorge Bornemann and Wolfgang Hayek

April 5, 2017

Abstract

Given a source field on a grid, we show how to compute the statistical moments (mean, variance, skewness and kurtosis) of the source field projected on a target grid. This process requires conservative regridding only. Any software package computing the conservative interpolation weights can be used.

1 Motivation

A common use case in post- and pre-processing tasks is to take a source field on a very high resolution grid and project it onto a much coarser resolution. For a cell-centred field, this is generally accomplished via conservative regridding. In conservative regridding, field values on the target grid cells are assigned by summing area weighted source field contributions coming from all the source cells that overlap with the target cell. The so obtained target field values represent the mean value of the source field over target cells.

In addition to the mean source field on the target cell, it is often desirable to compute statistics, e.g. the variance, skew, kurtosis, etc., on each target cell. We show how this can be accomplished using conservative regridding only. Specifically, the interpolation weights need not be known – we only assume that the source field can be regridded. This distinction is important because there exists a number of conservative regridding tools but few, if any, expose the interpolation weights to the user. The proposed approach will work with any of these packages.

2 Conservative regridding computes averages

Let X be the source field defined on N cells and M the number of target cells. Conservative interpolation is the projection of X onto the target grid, which can be expressed as the matrix-vector multiplication

$$Y = W \cdot X \tag{1}$$

where W is the $M \times N$ sized matrix of weights and Y is the vector of size M representing the interpolated values . Each element of W represents the the overlap between target (A_t) and source (A_s) cell areas normalized to the target cell area, that is $(A_t \cap A_s)/A_t$.

3 Computing statistical moments

Below we show how to compute the mean, the variance, the skewness and the kurtosis from (1) and the source field alone. Each moment can be computed recursively using lower order moments and by "averaging" the field to some power. In doing so, the interpolation weights need only be computed once. Therefore, the proposed method should lead to a very fast implementation of source field statistics on target grid cells.

3.1 Mean

As mentioned above, regridding is an averaging operation,

$$< X >\equiv Y = W \cdot X \tag{2}$$

3.2 Variance

$$Var(X) \equiv \sigma^2 \equiv < (X - < X >)^2 > = < X^2 - 2X < X > + < X >^2 > = < X^2 > - < X >^2 = W \cdot (X^2) - Y^2 \quad (3)$$

where $Y^2 = (W \cdot X).\star(W \cdot X)$ and $.\star$ is the element-wise multiplication. The variance requires the regridding of X^2 and Y.

3.3 Skewness

$$Skew(X) \equiv < (X - < X >)^3 > ./\sigma^3 = (< X^3 > -3\sigma^2 \cdot Y - Y^3)./\sigma^3 \quad (4)$$

where $./$ is the element-wise division and σ^3 is shorthand notation for $(\sigma^2)^{3/2}$. The skewness requires the regridding $< X^3 >= W \cdot X^3$, the variance and Y.

3.4 Excess kurtosis

$$Kurt(X) \equiv < (X - < X >)^4 ./(\sigma^2)^2 - 3 = (< X^4 > -4Skew(X)Y \cdot \sigma^3 - 6Y^2 \cdot \sigma^2 - Y^4)./(\sigma^2)^2 - 3 \quad (5)$$

The computation of the excess kurtosis requires regridding $< X^4 >= W \cdot X^4$, the skewness, the variance and Y.