NCAR_Topo (v1.0): NCAR global model topography generation software for unstructured grids

P. H. Lauritzen¹, J. T. Bacmeister¹, P. F. Callaghan¹, and M. A. Taylor²
¹National Center for Atmospheric Research, 1850 Table Mesa Drive, Boulder, Colorado, USA
²Sandia National Laboratories, Albuquerque, New Mexico, USA

Correspondence to: P. H. Lauritzen (pel@ucar.edu)

Received: 12 May 2015 – Published in Geosci. Model Dev. Discuss.: 22 June 2015
Revised: 30 September 2015 – Accepted: 1 December 2015 – Published: 14 December 2015

Abstract. It is the purpose of this paper to document the NCAR global model topography generation software for unstructured grids (NCAR_Topo (v1.0)). Given a model grid, the software computes the fraction of the grid box covered by land, the grid-box mean elevation (deviation from a geoid that defines nominal sea level surface), and associated sub-grid-scale variances commonly used for gravity wave and turbulent mountain stress parameterizations. The software supports regular latitude–longitude grids as well as unstructured grids, e.g., icosahedral, Voronoi, cubed-sphere and variable-resolution grids.

1 Introduction

Accurate representation of the impact of topography on atmospheric flow is crucial for Earth system modeling. For example, the hydrological cycle is closely linked to topography and, on the planetary scale, waves associated with the mid-latitude jets are very susceptible to the effective drag caused by mountains (e.g., Lott, 1999). Despite the fact that surface elevation is known globally with a high level of precision, the representation of its impact on atmospheric flow in numerical models remains a challenge.

When performing a spectral analysis of high-resolution elevation data (e.g., black line in Fig. 1), it is clear that Earth’s topography decreases quite slowly with increasing wave number (see also Balmino, 1993; Uhrner, 2001; Gagnon et al., 2006). Consequently, at any practical model resolution there will always be a non-negligible spectral component of topography present near the grid scale and there will always be a non-negligible spectral component of topography below the grid scale (sub-grid-scale component). The resolved scale component is the mean elevation in each grid box \( h \) (deviation from a geoid that defines nominal sea level surface), or, equivalently (if gravity is assumed constant), the surface geopotential \( \Phi_s \). For models using terrain-following coordinates it is common practice not to force the highest wave numbers directly in the model in order to alleviate obvious spurious noise (e.g., Navarra et al., 1994; Lander and Hoskins, 1997). Hence, for that class of models, \( \Phi_s \) is usually smoothed. Figure 1 shows the power spectrum for surface elevation for different levels of smoothing of topography in the NCAR-DOE CESM (Community Earth System Model) CAM (Community Atmosphere Model; Neale et al., 2010) SE (spectral element dynamical core; Thomas and Loft, 2005; Dennis et al., 2005) and CAM-FV (finite-volume dynamical core; Lin, 2004). Figure 2 shows the associated elevations for a cross section through the Andes mountain range. The amount of smoothing necessary is intrinsically linked to the numerical methods and discretization choices in the dynamical core. Further discussion on \( \Phi_s \) smoothing is given in Sect. 2.2.3.

The component of topography that cannot be represented by \( \Phi_s \) is referred to as sub-grid-scale topography. Sub-grid-scale processes may be sub-grid flow blocking, flow splitting, sheltering effects, generation of turbulence by roughness and gravity wave breaking (e.g., Bougeault et al., 1990). These processes can have an important impact on the resolved scale flow. Global models usually have a parameterization for gravity wave drag (GWD) (e.g., Eckermann and Chun, 2003) and turbulent effects referred to as turbulent mountain stress (TMS). An increasing number of models also have a low-level blocking parameterization (e.g., Lott and Miller, 1997;
According to linear theory, gravity waves can propagate in the vertical only when their intrinsic frequency is lower than the Brunt–Väisälä frequency $N$ (e.g., Durran, 2003). For orographically forced gravity waves the intrinsic frequency is set by the obstacle horizontal scale and the wind speed. When obstacle scales are too small to generate propagating waves, we expect drag to be produced by unstratified turbulent flow, a process which is typically parameterized in models’ TMS schemes. For larger obstacles we expect both drag and vertically propagating waves to result, processes which are dealt with by GWD schemes. Unfortunately the scale separating TMS and GWD processes is flow dependent. For typical midlatitude values of low-level wind ($10 \text{ m s}^{-1}$) and $N (10^{-2} \text{ s}^{-1})$, waves with wavelengths less than around 6000 m will not propagate in the vertical. A separation scale of 5000 m has been used by ECMWF (1997) and Beljaars et al. (2004). Here we will generate two sub-grid-scale variables derived from the topography data: the variance of topography below the 6000 m scale (referred to as $\text{Var}^{\text{TMS}}$) and the variance of topography with a scale longer than 6000 m and less than the grid scale (referred to as $\text{Var}^{\text{GWD}}$).

It is the purpose of this paper to document a software package (NCAR_Topo (v1.0)) that, given a “raw” high-resolution global elevation data set, maps elevation data to any unstructured global grid and separates the scales needed for TMS and GWD parameterizations. This separation of scales is done through an intermediate mapping of the raw elevation data to a 3000 m cubed-sphere grid ($\Omega \rightarrow A$) before mapping fields to the target model grid ($A \rightarrow \Omega$) as schematically shown in Fig. 3. This two-step process is described in Sect. 2.2 after a mathematical definition of the scale separation (Sect. 2.1). In Sect. 2.2.3 we briefly discuss $\Phi_\xi$ smoothing (if applicable). Output from the topography software is shown and discussed in Sect. 3.

2 Method

2.1 Continuous: separation of scales

The separation of scales is, in continuous space, conveniently introduced using spherical harmonics. Assume that elevation (above sea level) is a smooth continuous function, in
which case it can be represented by a convergent expansion of spherical harmonic functions of the form

\[ h(\lambda, \theta) = \sum_{m=-\infty}^{\infty} \sum_{n=-|m|}^{\infty} \psi_{m,n} Y_{m,n}(\lambda, \theta), \]  

(1)

(e.g., Durran, 2010) where \( \lambda \) and \( \theta \) are longitude and latitude, respectively, and \( \psi_{m,n} \) are the spherical harmonic coefficients. Each spherical harmonic function is given in terms of the associated Legendre polynomial \( P_{m,n}(\theta) \):

\[ Y_{m,n} = P_{m,n}(\theta) e^{im\lambda}, \]  

(2)

where \( m \) is the zonal wave number and \( m - |n| \) is the number of zeros between the poles and can therefore be interpreted as meridional wave number.

For the separation of scales the spherical harmonic expansion is truncated at wave number \( M \),

\[ h^{(M)}(\lambda, \theta) = \sum_{m=-M}^{M} \sum_{n=-|m|}^{M} \psi_{m,n}^{(M)} Y_{m,n}(\lambda, \theta), \]  

(3)

where a triangular truncation, which provides a uniform spatial resolution over the entire sphere, is used.

Let \( h^{(tg)}(\lambda, \theta) \) be a continuous representation of the elevation containing the spatial scales of the target grid. We do not write \( h^{(tg)}(\lambda, \theta) \) explicitly in terms of spherical harmonics as the target grid may be variable resolution and therefore contain different spatial scales in different parts of the domain.

For each target grid cell \( \Omega_k, k = 1, \ldots, N_t \), where \( N_t \) is the number of target grid cells, define the variances

\[ \text{Var}^{\text{(tms)}}_{\Omega_k} = \int \int_{\Omega_k} [h^{(M)}(\lambda, \theta) - h(\lambda, \theta)]^2 \cos(\theta) \, d\lambda \, d\theta, \]  

(4)

\[ \text{Var}^{\text{(gwd)}}_{\Omega_k} = \int \int_{\Omega_k} [h^{(tg)}(\lambda, \theta) - h^{(M)}(\lambda, \theta)]^2 \cos(\theta) \, d\lambda \, d\theta. \]  

(5)

Thus \( \text{Var}^{\text{(tms)}}_{\Omega_k} \) is the variance of elevation on scales below wave number \( M \) and \( \text{Var}^{\text{(gwd)}}_{\Omega_k} \) is the variance of elevation on scales larger than wave number \( M \) and below the target grid scale.

### 2.2 Discrete: separation of scales

The separation of scales is done through the use of a quasi-isotropic gnomonic cubed-sphere grid in a two-step regridding procedure: binning from source grid \( \Lambda \) to intermediate grid \( A \) (separation of scales) and then rigorously remapping variables to the target grid \( \Omega \).
Any quasi-uniform spherical grid could, in theory, be used for the separation of scales or, given the discontinuous/fractal nature of topography, a more rigorous scale separation method such as wavelets or other techniques could have been used. It is noted that the separation of scales through an intermediate grid does not correspond exactly to a spectral transform truncation (used in the previous section). The intent is to approximately eliminate scales below what can be resolved on the intermediate grid. For reasons that will become clear, we have chosen to use a gnomonic cubed-sphere grid (see Fig. 3) resulting from an equiangular gnomonic (central) projection,

\[
x = r \tan \alpha \quad \text{and} \quad y = r \tan \beta; \quad \alpha, \beta \in \left[\frac{-\pi}{4}, \frac{\pi}{4}\right],
\]

(Ronchi et al., 1996) where \(\alpha\) and \(\beta\) are central angles in each coordinate direction, \(r = R/\sqrt{3}\) and \(R\) is the radius of the Earth. A point on the sphere is identified with the three-element vector \((x, y, \nu)\), where \(\nu\) is the panel index. Hence the physical domain \(S\) (sphere) is represented by the gnomonic (central) projection of the cubed-sphere faces, \(\Omega^{(v)} = [-1, 1]^2, \nu = 1, 2, \ldots, 6\), and the non-overlapping panel domains \(\Omega^{(\nu)}\) span the entire sphere: \(S = \bigcup_{\nu=1}^{6} \Omega^{(\nu)}\). The cube edges, however, are discontinuous. Note that any straight line on the gnomonic projection \((x, y, \nu)\) corresponds to a great-circle arc on the sphere. In the discretized scheme we let the number of cells along a coordinate axis be \(N_c\) so that the total number of cells in the global domain is \(6 \times N_c^2\). The grid lines are separated by the same angle \(\Delta \alpha = \Delta \beta = \frac{\pi}{N_c}\) in each coordinate direction.

For notational simplicity the cubed-sphere cells are identified with one index \(i\), and the relationship between \(i\) and \((icube, jcube, \nu)\) is given by

\[
i = icube + (jcube - 1) N_c + (\nu - 1) N_c^2,
\]

where \((icube, jcube)\in[1,\ldots,N_c]^2\) and \(\nu \in [1, 2, \ldots, 6]\). In terms of central angles \((\alpha, \beta)\) the cubed-sphere grid cell \(i\) is defined as

\[
A_i = [(icube - 1)\Delta \alpha - \frac{\pi}{4}, icube\Delta \alpha - \frac{\pi}{4}]
\times \left[ (jcube - 1)\Delta \beta - \frac{\pi}{4}, jcube\Delta \beta - \frac{\pi}{4} \right],
\]

and \(\Delta A_i\) denotes the associated spherical area. A formula for the spherical area \(\Delta A_i\) of a grid cell on the gnomonic cubed-sphere grid can be found in Appendix C of Lauritzen and Nair (2008) (note that Eq. C3 is missing arccos on the right-hand side). A quasi-uniform approximately 3000 m resolution is obtained by using \(N_c = 3000\), which results in a scale separation of roughly 6000 m (based on a simple “2\(\Delta x\) assumption”). For more details on the construction of the gnomonic grid see, for example, Lauritzen et al. (2010).

### 2.2.1 Step 1: raw elevation data to intermediate cubed-sphere grid \((A \rightarrow A)\)

The raw elevation data are provided on a geoid such as the World Geodetic System 1984 (WGS84) ellipsoid, whose center is located at the Earth’s center. Latitude and longitudes use the WGS84 geodetic datum. Similarly, elevation above sea level is defined as the deviation from the WGS84 geoid or WGM96 (Earth Geopotential Model 1996) geoid for newer data sets. Global weather/climate models typically assume that Earth is a sphere, so the elevations at a certain latitude–longitude pair on the geoid are taken to be the same on the sphere. The consequences of this approximation in the context of limited-area models are discussed in Monaghan et al. (2013) and in the context of shallow-water experiments in Bénard (2015). Derivation of the equations of motion in more accurate coordinates is starting to emerge in the literature (e.g., White et al., 2008; Gates, 2004; Staniforth and White, 2015).

The “raw” elevation data are usually from a digital elevation model (DEM) such as the GTOPO30, a 30 arcsecond global data set from the United States Geological Survey (USGS; Gesch and Larson, 1998) defined on an approximately 1 km regular latitude–longitude grid. Several newer, more accurate, and locally higher resolution elevation data sets are available such as GMTED2010 (Global Multi-Resolution Terrain Elevation Data 2010; Danielson and Gesch, 2011), GLOBE (GLOBE Task Team and others, 1999) and the NASA Shuttle Radar Topographic Mission (SRTM) DEM data (Farr et al., 2007). The SRTM data, however, are only near-global (up to 60° north and south). Here we use GTOPO30 and GMTED2010. The GTOPO30 data come in 33 tiles (separate files) in 16 bit binary format. Fortran code is provided to convert the data into one NetCDF file. Even though the elevation is stored as integers, the size of the NetCDF file is 7 GB. The GTOPO file contains height \(h\) and land fraction \(\text{LANDFRAC}\) (for the mathematical operations below we use \(f\) to refer to land fraction). The GMTED2010 data do not contain \(\text{LANDFRAC}\), so the land fraction is obtained from Moderate Resolution Imaging Spectroradiometer (MODIS) 1 km land–water mask data downloaded from ftp://ladscl1.nascom.nasa.gov/pub/outgoing/c6_dem/.

The centers of the regular latitude–longitude grid cells for the “raw” topographic data are denoted \((\lambda_{i\text{lon}}, \theta_{j\text{lat}})\), \(i\text{lon} = 1, \ldots, n\text{lon}\), \(j\text{lat} = 1, \ldots, n\text{lat}\). For the data sets used here \(n\text{lon} = 43\,200\) and \(n\text{lat} = 21\,600\). As for the cubed sphere, we use one index \(j\) to reference the grid cells,

\[
j = i\text{lon} + (j\text{lat} - 1) \times n\text{lon}, \quad j \in [1, \ldots, N_r],
\]

where \(N_r = n\text{lon} \times n\text{lat}\). The spherical area of grid cell \(\Lambda_j\) is denoted \(\Delta \Lambda_j\) and the average elevation in cell \(j\) is given by \(h_{j\text{raw}}^{\text{avg}}\).

These data are binned to the cubed-sphere intermediate grid by identifying in which gnomonic cubed-sphere grid cell
\((\lambda_{i\text{lon}}, \theta_{i\text{lat}})\) is located. Due to the “Cartesian-like” structure of the cubed-sphere grid, the search algorithm is straightforward (following Nair et al., 2005):

- Transform \((\lambda_{i\text{lon}}, \theta_{i\text{lat}})\) coordinate to Cartesian coordinates \((x, y, z)\):
  \[
  x = \cos(\lambda_{i\text{lon}}) \cos(\theta_{i\text{lat}}), \\
  y = \sin(\lambda_{i\text{lon}}) \cos(\theta_{i\text{lat}}), \\
  z = \sin(\theta_{i\text{lat}}).
  \]
  \hspace{1cm} (10)  \hspace{1cm} (11)  \hspace{1cm} (12)

- Locate which cubed-sphere panel \((\lambda_{i\text{lon}}, \theta_{i\text{lat}})\) is located on through a “coordinate maximality” algorithm, i.e., let
  \[
  \text{pm} = \max(|x|, |y|, |z|),
  \]
  \hspace{2cm} (13)

  then if

  - \(\text{pm} = |x| \) and \(x > 0\), then \((\lambda_{i\text{lon}}, \theta_{i\text{lat}})\) is on \(v = 1\);
  - \(\text{pm} = |y| \) and \(y > 0\), then \((\lambda_{i\text{lon}}, \theta_{i\text{lat}})\) is on \(v = 2\);
  - \(\text{pm} = |x| \) and \(x < 0\), then \((\lambda_{i\text{lon}}, \theta_{i\text{lat}})\) is on \(v = 3\);
  - \(\text{pm} = |y| \) and \(y < 0\), then \((\lambda_{i\text{lon}}, \theta_{i\text{lat}})\) is on \(v = 4\);
  - \(\text{pm} = |z| \) and \(z \leq 0\), then \((\lambda_{i\text{lon}}, \theta_{i\text{lat}})\) is located on the bottom panel, \(v = 5\);
  - \(\text{pm} = |z| \) and \(z > 0\), then \((\lambda_{i\text{lon}}, \theta_{i\text{lat}})\) is located on the top panel, \(v = 6\).

- Given the panel number the associated central angles \((\alpha, \beta)\) are given by
  
  - \(v = 1: (\alpha, \beta) = \left[\arctan\left(\frac{x}{z}\right), \arctan\left(\frac{y}{z}\right)\right]\);
  - \(v = 2: (\alpha, \beta) = \left[\arctan\left(\frac{-x}{z}\right), \arctan\left(\frac{y}{z}\right)\right]\);
  - \(v = 3: (\alpha, \beta) = \left[\arctan\left(\frac{x}{z}\right), \arctan\left(\frac{-y}{z}\right)\right]\);
  - \(v = 4: (\alpha, \beta) = \left[\arctan\left(\frac{x}{z}\right), \arctan\left(\frac{-y}{z}\right)\right]\);
  - \(v = 5: (\alpha, \beta) = \left[\arctan\left(\frac{-x}{z}\right), \arctan\left(\frac{-y}{z}\right)\right]\);
  - \(v = 6: (\alpha, \beta) = \left[\arctan\left(\frac{x}{z}\right), \arctan\left(\frac{-y}{z}\right)\right]\).

- The indices of the cubed-sphere cell in which the center of the latitude–longitude grid cell is located are given by
  
  \[
  \text{icube} = \text{CEILING}\left(\frac{\alpha + \frac{x}{z}}{\Delta\alpha}\right), \\
  \text{jcube} = \text{CEILING}\left(\frac{\beta + \frac{y}{z}}{\Delta\beta}\right),
  \]
  \hspace{1cm} (14)  \hspace{1cm} (15)

  where \text{CEILING}(\cdot) function returns the smallest integer not less than the argument.

The set of indices for which center points of regular latitude–longitude grid cells are located in gnomonic cubed-sphere cell \(A_i\) is denoted \(S_i\). Note that, since the raw data sets are higher resolution, the cubed-sphere \(S_i\) is guaranteed to be non-empty. Through this binning process the approximate average elevation in cubed-sphere cell \(i\) becomes

\[
\bar{h}^{(\text{cube})}_{A_i} = \frac{1}{\Delta A_i} \sum_{j \in S_i} h^{(\text{raw})}_{F_{ij}} \Delta A_j.
\]

When \(\bar{h}^{(\text{cube})}_{A_i}\) is known we can compute the sub-grid variance on the intermediate cubed-sphere grid \(A\):

\[
\text{Var}^{(\text{ms})}_{A_i} = \frac{1}{\Delta A_i} \sum_{j \in S_i} \left(\bar{h}^{(\text{cube})}_{A_i} - h^{(\text{raw})}_{F_{ij}}\right)^2 \Delta A_j.
\]

The land fraction is also binned to the intermediate cubed-sphere grid,

\[
\bar{f}^{(\text{cube})}_{A_i} = \frac{1}{\Delta A_i} \sum_{j \in S_i} f^{(\text{raw})}_{F_{ij}} \Delta A_j.
\]

To remain consistent with previous GTOP030-based topography generation software for CAM, all land fractions south of 79°S are set to 1 for topography data based on GTOP030, which effectively extends the land for the Ross Ice Shelf. There are eight types of masks in the MODIS data: shallow ocean, land, shoreline, shallow inland water, ephemeral water, deep inland water, moderate ocean, and deep ocean. Here we set all masks to 0 except land, shoreline and ephemeral water, which are set to 1. For topography data based on MODIS, no further alternations are made to the land fraction.

The binning process is straightforward since the cubed-sphere grid is essentially an equidistant Cartesian grid on each panel in terms of the central angle coordinates. This step could be replaced by rigorous remapping in terms of overlap areas between the regular latitude–longitude grid and the
2.2.2 Step 2: cubed-sphere grid to target grid \((A \rightarrow \Omega)\)

The cell-averaged values of elevation and sub-grid-scale variances \(\text{Var}_\Omega^{(\text{tms})}\) and \(\text{Var}_\Omega^{(\text{gwd})}\) on the target grid are computed by rigorously remapping the variables from the cubed-sphere grid to the target grid. As for step 1 of the algorithm, several methods could be used such as SCRIP and Kritsikis et al. (2015). The latter has the advantage of more accurately representing both small- and great-circle arc sides of the control volumes, whereas SCRIP approximates cells’ sides in latitude–longitude space. Here the remapping is performed using CSLAM (Conservative Semi-Lagrangian Multi-tracer transport scheme) technology (Lauritzen et al., 2010), which approximates the cell sides with great-circle arcs. That said, the algorithm can still operate on grids where the control volumes do not consist of great-circle arcs, such as regular latitude–longitude grids where the control volume sides parallel with latitudes are small-circle arcs, but there will be an additional truncation error in the geometric approximation to the control volumes. It is possible to use large parts of the CSLAM technology since the source grid is a gnomonic cubed-sphere grid; hence, instead of remapping between the gnomonic cubed-sphere grid and a deformed Lagrangian grid, as done in CSLAM transport, the remapping is from the gnomonic cubed-sphere grid to any target grid constructed from great-circle arcs (the target grid “plays the role” of the Lagrangian grid). However, a couple of modifications were made to the CSLAM search algorithm. First of all, the target grid cells can have an arbitrary number of vertices, whereas the CSLAM transport search algorithm assumes that the target grid consists of quadrilaterals and the number of overlap areas is determined by the deformation of the transporting velocity field. In the case of the remapping needed in this application, the target grid consists of polygons with any number of vertices and the search is not constrained by the physical relation between regular and deformed upstream quadrilaterals. Secondly, the CSLAM search algorithm for transport assumes that the target grid cells are convex, which is not necessarily the case for target grids. The CSLAM search algorithm has been modified to support non-convex cells that are, for example, encountered in variable-resolution CAM-SE (see Fig. 4); essentially that means that any target grid cell may cross a gnomonic isoline (source grid line) more than twice.

Let the target grid consist of \(N_t\) grid cells \(\Omega_k, k = 1, \ldots, N_t\) with associated spherical area \(\Delta\Omega_k\). The search algorithm for CSLAM is used to identify overlap areas between the target grid cell \(\Omega_k\) and the cubed-sphere grid cells \(A_\ell, \ell = 1, \ldots, N_c\). Denote the overlap area between \(\Omega_k\) and \(A_\ell\)

\[
\Omega_{k\ell} = \Omega_k \cap A_\ell, \quad \ell = 1, \ldots, N_c
\]

(see Fig. 5) and let \(L_k\) denote the set of indices for which \(\Omega_k \cap A_\ell\) is non-empty.

Note that the computation of overlap areas \(\Omega_{k\ell}\) is facilitated by the “Cartesian” layout of the equiangular cubed-sphere grid in computational space. The target grid vertex locations within the cubed-sphere source grid are computed by first locating which panel the vertices are on (using the “coordinate maximality” algorithm Eq. 13), and their locations thereafter trivially result when the vertex coordinates are converted into the particular panels equiangular coordinates (Eqs. 14 and 15). The overlap areas are computed by finding intersection between the \((x, y)\) gnomonic isolines (“Cartesian” layout) and the target grid cell sides. Since topographic variables are computed just once for each horizontal grid resolution, computational performance is less critical. On a standard laptop even high-resolution (e.g., 25 km) grid overlaps are computed on the order of minutes irrespective of the target grid topology (Voronoi, cubed sphere, etc.).

The average elevation and variance used for TMS in target grid cell \(k\) are respectively given by

\[
\bar{h}_{\Omega_k}^{(\text{tgr, raw})} = \frac{1}{\Delta\Omega_k} \sum_{\ell \in L_k} h_{A_\ell}^{(\text{cuble})} \Delta\Omega_{k\ell},
\]

(20)

\[
\text{Var}_{\Omega_k}^{(\text{tms})} = \frac{1}{\Delta\Omega_k} \sum_{\ell \in L_k} \text{Var}_{\Omega_k}^{(\text{tms})} \Delta\Omega_{k\ell}.
\]

(21)

The variance of the cubed-sphere data \(h_{\Omega_k}^{(\text{cuble})}\) with respect to the target grid cell average values \(\bar{h}_{\Omega_k}^{(\text{tgr, raw})}\) is given by

\[
\text{Var}_{\Omega_k}^{(\text{gwd, raw})} = \frac{1}{\Delta\Omega_k} \sum_{\ell \in L_k} \left( h_{\Omega_k}^{(\text{tgr, raw})} - \bar{h}_{A_\ell}^{(\text{cuble})} \right)^2 \Delta\Omega_{k\ell}.
\]

(22)

Figure 5. Schematic of the notation used to define the overlap between target grid cell \(\Omega_k\) (red polygon) and high-resolution cubed-sphere grid cell \(A_\ell\) (Cartesian-like grid in figure). The overlap area, \(\Omega_{k\ell} = \Omega_k \cap A_\ell\), is shaded with a cross-hatch pattern in the figure.
The appended superscript “raw” in $h^{(tgt, raw)}$ and $\text{Var}^{(gwd, raw)}$ refers to the fact that the elevation and GWD variance is based on unsmoothed elevation. As mentioned in the Introduction, the elevation (for models based on terrain-following coordinates) is usually smoothed – i.e., the highest wave numbers are removed. This smoothed elevation is denoted $h^{(tgt, smooth)}$.

After smoothing the target grid elevation, the sub-grid-scale variance for GWD could be recomputed as the smoothing operation will add energy to the smallest wavelengths:

$$\text{Var}^{(gwd, smooth)} = \frac{1}{\Delta \Omega_k} \sum_{\ell \in \Omega_k} \left( h^{(tgt, smooth)} - h^{(cube)} \right)^2 \Delta \Omega_k \ell. \tag{23}$$

See discussion on the difference between using $\text{Var}^{(gwd, raw)}$ or $h^{(tgt, smooth)}$ in Sect. 3. The smoothing of elevation is discussed in some detail in the next section.

### 2.2.3 Smoothing of elevation $h^{(tgt)}$ (if applicable)

As discussed in the Introduction, mapping the elevation to the target grid without further filtering to remove the highest wave numbers usually leads to excessive spurious noise in the simulations when using terrain-following coordinates. Non-terrain-following coordinates such as the $\eta$ coordinates (Mesinger, 1984; Mesinger et al., 1988) and $z$ coordinates using shaved cells or cut cells (Adcroft et al., 1997; Walko and Avisar, 2008) do not require filtering of topography. For models that do require topography smoothing, there seems to be no standardized procedure, for example a test case suite, to objectively select the level of smoothing and the filtering method. The amount of smoothing necessary to remove spurious noise in, for example, vertical velocity depends on the amount of inherent or explicit numerical diffusion in the dynamical core (e.g., Lauritzen et al., 2011). It may be considered important that the topographic smoothing is done with discrete operators consistent with the dynamical core.

While it is necessary to smooth topography to remove spurious grid-scale noise, it potentially introduces two problems. Filtering will typically raise ocean points near step topography to non-zero elevation. Perhaps the most striking example is the Andes mountain range, which extends one or two grid cells into the Pacific after the filtering operation (Fig. 2). Ocean and land points are treated separately in weather/climate models, so raised sea points may potentially be problematic. Secondly, the filtering will generally reduce the height of local topographic maxima, and given the importance of barrier heights in atmospheric dynamics, this could be a problem for the global angular momentum budget and could fundamentally change the flow (unless a parameterization for blocking is used). To capture more of the barrier effect (blocking), two approaches have been put forward in the literature to deepen valleys and increase peak
Figure 7. (a) Idealized topography $1000 + 1000 \cdot \cos(2\pi \cdot x/300) + 100 \cdot \sin(2\pi/20)$ for $-150 \text{ km} < x < 150 \text{ km}$ (thin black line), topography averaged in 50 km grid boxes (thick black line), and topography smoothed with boxcar smoother using a 500 km window (thick blue line). (b) Deviation of topography from 50 km grid-box mean (thin black line), and deviations from 500 km boxcar-smoothed topography (thin blue line).

The naming conventions for the topographic variables in the software and NetCDF files are as follows:

\begin{align*}
\text{PHIS} &= g \tilde{h}^{(\text{gt})}_\Omega, \\
\text{SGH} &= \sqrt{\text{Var}^{(\text{gwd})}_\Omega}, \\
\text{SGH30} &= \sqrt{\text{Var}^{(\text{tms})}_\Omega}, \\
\text{LANDFRAC} &= f^{(\text{gt})}_\Omega,
\end{align*}

where $g$ is the gravitational constant.

3 Results

Below the topography software has been applied to different dynamical core grids in CAM, for example CAM-SE, which uses the spectral element dynamical core from HOMME (High-Order Method Modeling Environment; Thomas and Loft, 2005; Dennis et al., 2005), as well as CAM-FV. CAM-SE is defined on a cubed-sphere grid, and we show results for a configuration of approximately $1^\circ$ resolution, which is the NE30NP4 setup with 30 x 30 elements on each cubed-sphere panel and 4 x 4 Gauss–Lobatto–Legendre (GLL) quadrature points in each element. CAM-FV topography data are shown for the $0.9^\circ \times 1.25^\circ$ resolution. The topography software has also been used to generate topography data for variable-resolution MPAS (Model for Prediction Across Scales; Skamarock et al., 2012) as well as variable-resolution CAM-SE (not shown).
3.1 Example: smoothing methods used in CAM-FV and CAM-SE

As mentioned in Sect. 2.2.3 dynamical cores may need smoothing of elevation data. In the CAM-FV the highest wave numbers are removed by mapping $\Phi_s$ (surface geopotential) to a regular latitude–longitude grid that is half the resolution of the desired model resolution and then mapping back to the model grid by one-dimensional remaps along latitudes and longitudes, respectively, using the PPM (piecewise parabolic method) with monotone filtering. In CAM-SE the surface geopotential is smoothed by multiple applications of the CAM-SE Laplace operator combined with a bound-preserving limiter. Figure 2 shows different levels of smoothing of surface height for CAM-SE and, for comparison, CAM-FV. It can clearly be seen that there are large differences between the height of the mountains with different smoothing operators and smoothing strength. Figure 6 shows the topography variables for CAM-SE based on GTOPO30 raw data and smoothing with 16 applications of the Laplace operator.

3.2 Topography smoothing and SGH

The implications of topography smoothing on SGH are problematic, as illustrated in Fig. 7. A strict decomposition of the topography into its grid-box average and deviations from that average is conceptually simple. In this ideal situation, the model dynamical core represents circulations forced by grid-box mean topography (thick black line, Fig. 7a), and
a parameterization scheme represents the drag and other effects induced by waves forced by "unresolved" topographic deviations (thin black line, Fig. 7b). The grid mean of these deviations is zero.

However, when additional smoothing is applied to the grid-box mean topography (thick blue line, Fig. 7a), the situation becomes less well defined. It is difficult to argue that the "unresolved" topographic forcing is fully captured by the deviations from the grid-box mean. On the other hand, the deviations from the smoothed topography (thin blue line, Fig. 7b) have properties that make it unclear how to incorporate them in a traditional gravity wave parameterization framework. The grid-box mean of these deviations is not zero. Non-zero deviations can exist far from actual topographic forcing. In the case of real-world topography, Fig. 8 shows the difference between the deviations (thin black line, Fig. 7b). The grid mean of these deviations is zero. Non-zero deviations can exist far from actual topographic forcing. In the case of real-world topography, Fig. 8 shows the difference between $\sqrt{\text{Var}(\text{gwd}, \text{smooth})}$ and $\sqrt{\text{Var}(\text{gwd}, \text{raw})}$. There are significant differences between SGH based on smoothed or unsmoothed topography. We do not advocate any particular strategy for parameterization of sub-grid orographic effects; our software is simply intended to make the generation of the forcing data sets for global models traceable and transparent.

### 3.3 GTOPO30 and GMTED2010 raw data

Many global models use GTOPO30 as the raw topography data; however, more accurate data sets are available, such as GMTED2010. NCAR_Topo (v1.0) is set up to use both GTOPO30 and GMTED2010 (with MODIS for land fraction) source data. The differences between the raw 1 km data sets are discussed in detail in Danielson and Gesch (2011). Here we focus on the differences between the data sets on the approximately 3 km cubed-sphere grid and standard climate model resolutions (approximately 100 km). Figure 9 shows the elevation differences as well as SGH30 differences on the intermediate (approximately 3 km) cubed-sphere grid. There are significant differences in elevation, especially over Greenland, Antarctica and South America, as well as associated changes to SGH30. Even at the 100 km scale and with topography smoothing, the differences are rather large (see Fig. 10). The differences in land fraction are mainly due to MODIS providing a more detailed land–water mask. There are eight types of masks in the MODIS data: shallow ocean, land, shoreline, shallow inland water, ephemeral water, deep inland water, moderate ocean, and deep ocean. Here we set all masks to 0 except land, shoreline and ephemeral water, which are set to 1.

### 4 Conclusions

The NCAR global model topography generation software for unstructured grids (NCAR_Topo (v1.0)) has been documented and example applications to CAM have been presented. The topography software computes sub-grid-scale variances using a quasi-isotropic separation of scales through the intermediate mapping of high-resolution elevation data to an equiangular cubed-sphere grid. The software supports structured or unstructured (e.g., variable resolution) global grids.

### Code availability

The source codes for the NCAR Global Model Topography Generation Software for Unstructured Grids are available through GitHub. The repository URL is https://github.com/UCAR/Topo/tags/. The repository also contains NCL scripts to plot the topography variables (Fig. 6) that the software generates.

### Acknowledgements

NCAR is sponsored by the National Science Foundation (NSF). M. Taylor was supported by the Department of Energy Office of Biological and Environmental Research, work package 12-015334 “Multiscale Methods for Accurate, Efficient, and Scale-Aware Models of the Earth System”. Thanks to Sang-hun Park (NCAR) for providing data for the GTOPO30 power spectrum. Thanks to Max Suarez (NASA) for encouraging upgrading to the GMTED2010 data. We thank the two reviewers for their useful comments. Partial support for this work was provided through the Scientific Discovery through Advanced Computing (SciDAC) program funded by U.S. Department of Energy, Office of Science, Advanced Scientific Computing Research.

Edited by: S. Valcke

### References


Geosci. Model Dev., 8, 3975–3986, 2015

www.geosci-model-dev.net/8/3975/2015/


