UCRL-MI-127096

PCMDI Report No. 58

AN OBJECTIVE METHOD FOR GENERATING LAND/SEA MASKS FOR USE IN GCM SIMULATIONS

by

Karl E. Taylor and Charles Doutriaux

August 2000

PROGRAM FOR CLIMATE MODEL DIAGNOSIS AND INTERCOMPARISON
UNIVERSITY OF CALIFORNIA, LAWRENCE LIVERMORE NATIONAL LABORATORY,
LIVERMORE, CA 94550

UCRL-MI-127096

PCMDI Report No. 58

An Objective Method for Generating Land/Sea Masks for Use in GCM Simulations

Karl E. Taylor and Charles Doutriaux

Program for Climate Model Diagnosis and Intercomparison (PCMDI)

Lawrence Livermore National Laboratory

Livermore, CA 94550, USA

July 2000

Abstract

Most general circulation models (GCM's) require prescription of certain surface conditions, in particular a land/sea mask indicating whether grid cells are underlain by soil and vegetation or by water. Here we describe an algorithm that produces highly realistic land/sea masks at any desired resolution. Based on a high-resolution (1/6-degree longitude/latitude) data set that provides the fraction land in each grid cell, we determine, as a first approximation, whether the land fraction in each grid cell of the model exceeds 50%. Those grid cells meeting this criterion are tentatively designated land. We then adjust this first approximation to obtain a more accurate representation of islands and lakes that are about the same size of a single model grid cell but that straddle two or more model grid cells. Under the first approximation, these small islands and lakes tend to be omitted, but under an iterative scheme, they are included, though slightly displaced in position. This scheme also tends to better represent locally the relative fraction of land and ocean along coastlines.

1. Introduction

Heat, water, and momentum fluxes between the atmosphere and underlying surface are strongly affected by the nature of the surface, in particular whether or not it is water covered. In climate models these fluxes and other surface processes are usually calculated on a longitude-latitude grid. For models that assume a grid cell is either completely land covered or completely ocean covered (i.e., no fractional coverage), a land/sea "mask" is required. Here we describe an objective method for creating such masks.

2. Method for creating land/sea masks

The observed data set used here as the basis for creating realistic model land/sea masks was obtained from the U.S. Navy on a 1/6 degree longitude-latitude grid (see http://www.scd.ucar.edu/dss/datasets/ds754.0.html). Figure 1 shows the land fraction in each grid cell of the North American region, according to this data set. The objective is to create on the much coarser grids typical of most climate models land/sea masks that approximate the true land/sea distribution. As will be shown in the next section, the following procedure leads to highly realistic land/sea masks.

As a first step, the Navy data set is mapped to the model's grid (the "target" grid), using a grid-cell averaging algorithm (i.e., an "area-weighted" averaging scheme) that locally and globally preserves area-mean land fraction. Target grid cells with land fraction exceeding 50% are tentatively considered to be land cells, and all other cells are tentatively considered to be water. As we shall see, this "first approximation" land/sea mask turns out to be the same as the final land/sea mask except near the boundaries between water and land where a few grid cells will be switched to avoid missing some small islands and lakes and to better represent coast lines.

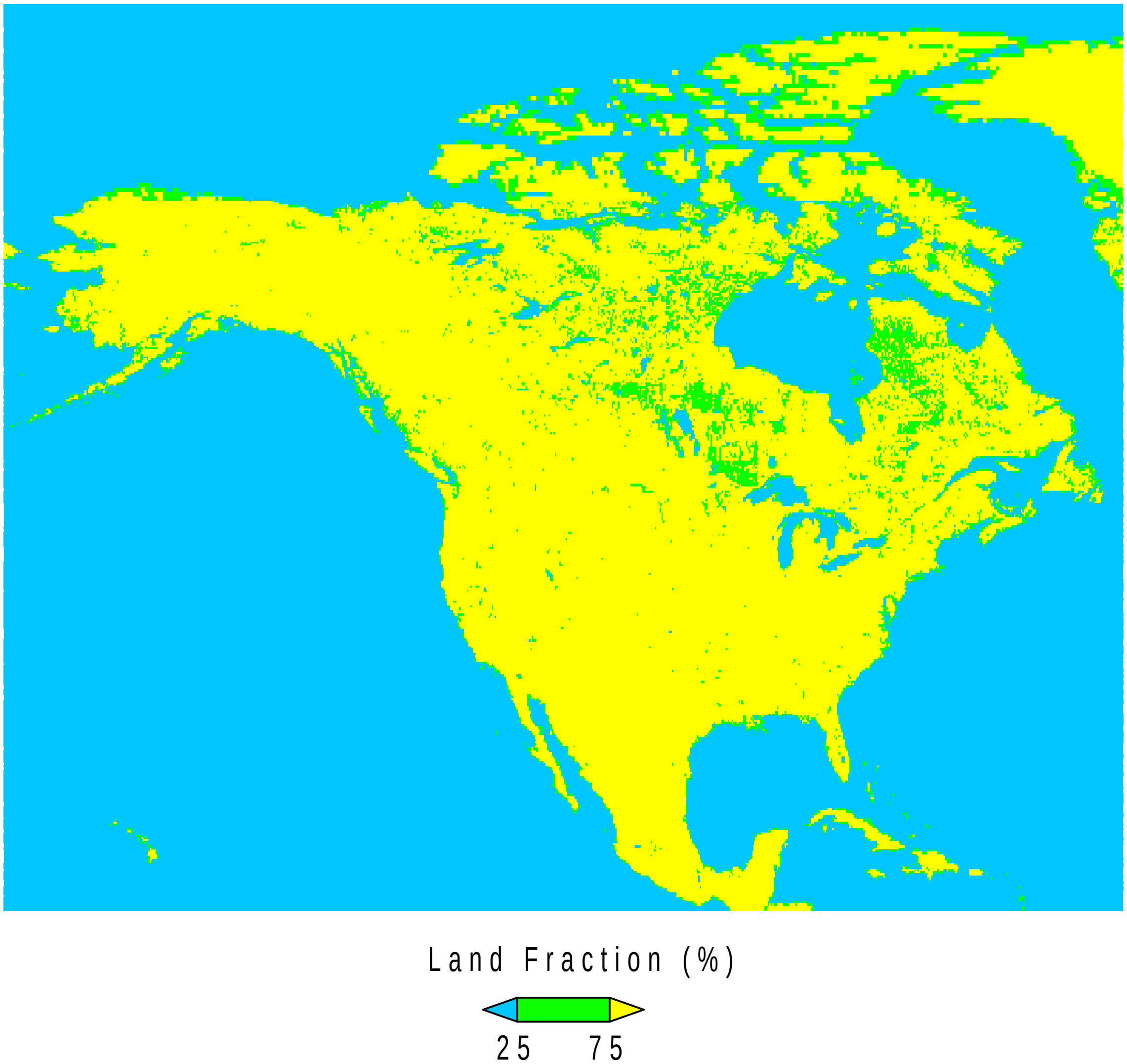

Fig. 1: Percent land in each grid cell of the U.S. Navy's data set (1/6 degree longitude/latitude grid) for the North American region.

Figure 2 shows the results of the first step for a model with a longitude-latitude resolution of approximately 3 x 3 degrees (more precisely, a T42 Gaussian grid with 128 longitudes and 64 latitudes). Although the "first approximation" minimizes (in a root-mean-square sense) the difference between the target mask land fraction (0% or 100%) and the Navy land fraction data averaged over each target grid cell, it may omit some islands and lakes that are about the same size as a grid cell.

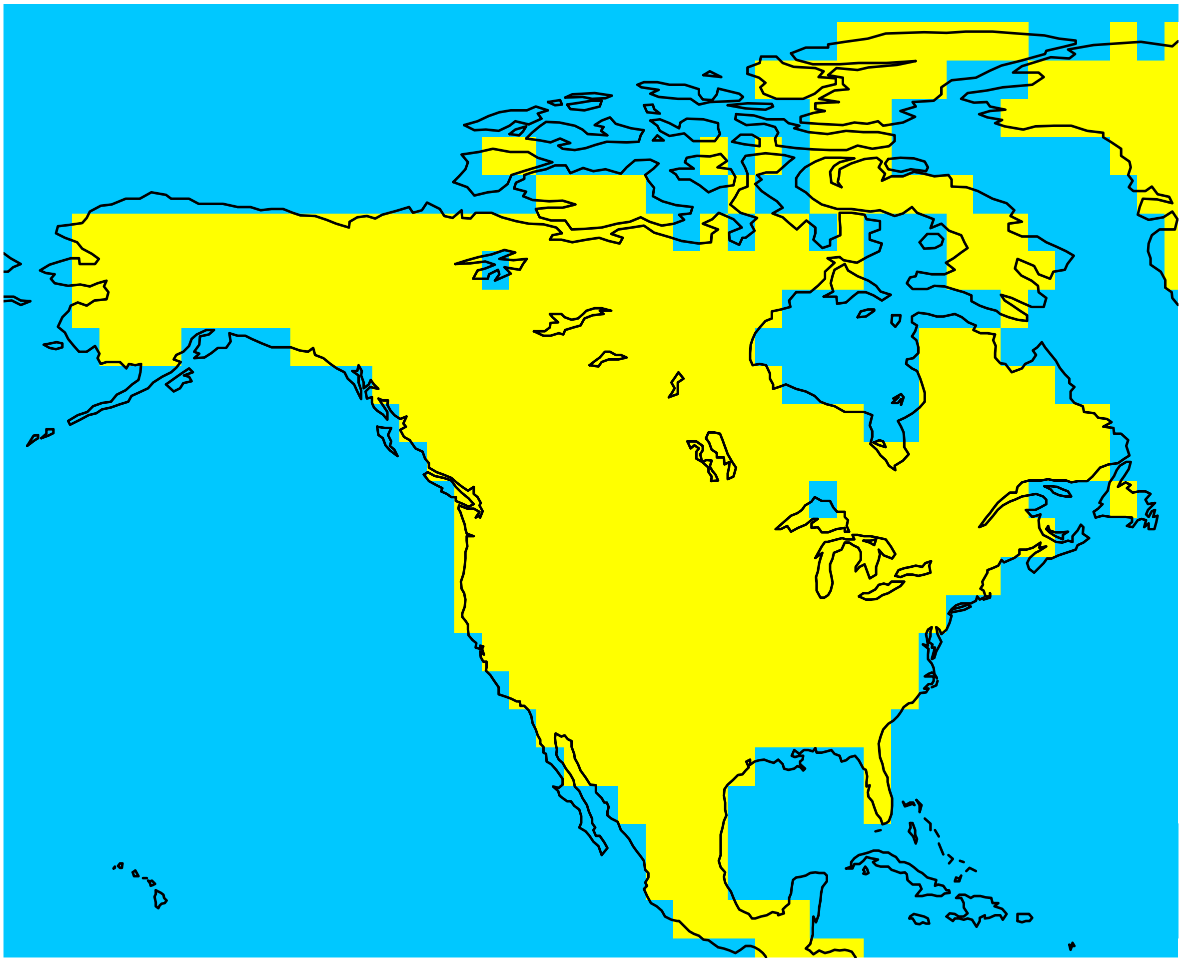

Fig. 2: "First approximation" land/sea mask for a T42 Gaussian grid.

Consider, for example, the hypothetical case shown in figure 3 in which a pair of islands (colored yellow) with total area 1.56 times the area of a single model grid cell are located near the intersection of the boundaries of certain model grid cells.

The model resolution shown in figure 3 is about 1.33 degrees longitude by 0.67 degrees latitude. The yellow regions are land covered and the blue regions are water covered. For simplicity, all grid cells in this example are assumed to be the same size (although in the more general application of our procedure the grid cell areas can vary spatially). In addition, all grid cells on the high-resolution grid are assumed to either be 100% land or 100% ocean (although the Navy data set actually contains fractional land amounts).

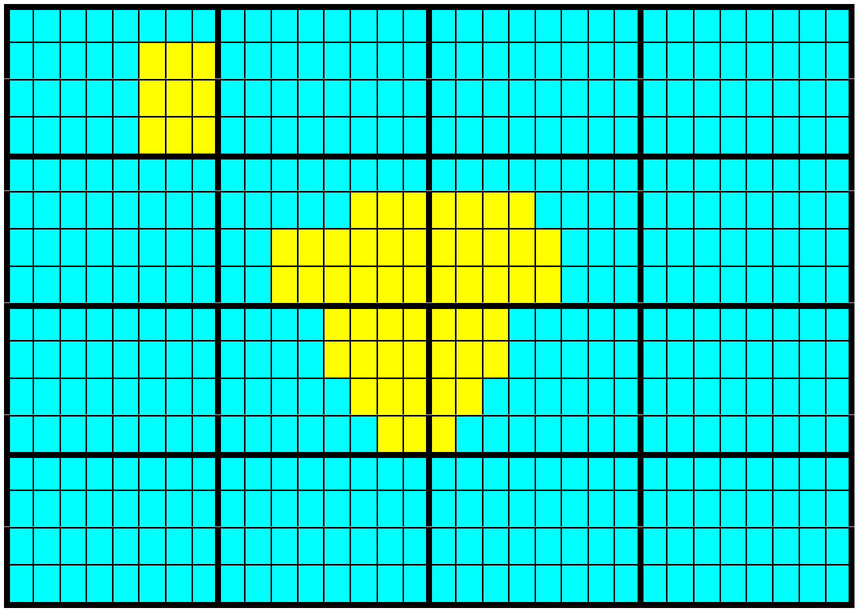

Fig. 3: Land cells (yellow) on the Navy's 1/6-degree grid with a coarser target grid (heavy lines) superposed.

In this example the result of the first step of our procedure would be to assign ocean to all of the target grid cells in the domain shown (because each individual target cell is greater than 50% water), which therefore yields a target land sea mask with the islands completely missing. (Note that another grid of the same resolution, but shifted in longitude and/or latitude by half a grid cell relative to the grid shown in the example, would result in land being assigned to one or two of the grid cells, depending on the shift.)

In this example the islands together cover an area larger than a grid cell (closer, in fact, to the area of two grid cells), so ideally our method would assign land to two of the four grid cells shown. Then the island, which occupies 9.8% of the domain shown, would appear also on the target grid and would occupy 12.5% of the area on that grid (rather than not appearing at all, which is the result of the first step of our procedure).

To capture small islands and lakes that are missed in the first approximation, we have devised a series of operations, which after iteration leads to what we think is a more realistic land/sea mask. First, for each target grid cell we consider a roughly rectangular region (centered on the grid cell) that is four times the size of the grid cell itself. Considering, for example, the grid cell second from the left and second from the top of the domain of figure 3, we focus on the region shown in figure 4. This region encompasses part of each of the adjacent cells: one-half of the cells above, below, to the right and to the left, and one-fourth of the four other cells that join it at its corners.

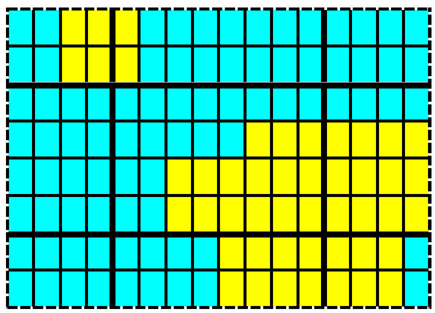

Fig. 4: Sample region four times larger than the central grid cell over which the average land fraction is calculated to help guide correction of the initial target land/sea mask.

The fraction of the sub-domain shown in figure 4 that is occupied by land is 47/128 (= 0.367). Fractions can be computed for all the other grid cells in an analogous way, and these fractions can then be displayed as the upper number of each pair shown in figure 5. In this figure the boxes represent the same grid cells shown in figure 3, and they are arrange identically. Also shown in figure 5 (lower number of each pair) is the land fraction in each target grid cell, as determined from the Navy's land fraction data set.

To determine whether the "first approximation" can be improved, we consider first the fractions calculated at half the model resolution. At this resolution we examine the difference between the observed land fractions shown in figure 5 (upper numbers) and the comparable land fractions taken from the "first approximation" land/sea mask, which in our example are all 0%. Thus, the upper numbers in figure 5 not only represent the observed land fraction but also the difference between the observed and "first approximation" land fraction. In our example, all numbers are non-negative, so in this case we search for candidate cells that might be switched

.070 .188	.125 .094	.055 .000	.000 .000
.078 .000	.367 .469	.305 .438	.016 .000
.031 .000	.328 .406	.313 .281	.016 .000
.000 .000	.063 .000	.063 .000	.000 .000

Fig. 5: Percent land on the model grid (lower numbers) and on a coarser grid comprising overlapping rectangular areas four times the size of individual grid cells (upper numbers).

from ocean to land in order to improve our "first approximation." We look for grid cells in figure 5 where the difference at half the model resolution (upper number) exceeds 0.2 and where the land fraction at full model resolution (lower number) exceeds 0.3. Once a cell is found satisfying these fundamental criteria, we examine all eight neighboring cells that also satisfy these criteria. If the central cell contains more land (according to the lower number) than any of its neighbors that satisfy the fundamental criteria, then we switch this cell to land. In practice, of course, the procedure is generalized to allow also for switching land cells to ocean.

The critical fractions (0.2 and 0.3) were chosen somewhat subjectively after exploring the land sea masks produced by a variety of other critical values. The fraction 0.2 means that a cell will not be a candidate for change unless there is enough unaccounted for land in the region four times the area of the cell to fill at least 80% of the grid cell. The fraction 0.3 means that a cell will not be a candidate for change unless at least 30% of the cell itself is actually land covered. For cells that might be changed from land to ocean, the corresponding critical fractions are -0.2 and 0.7 (i.e., the difference between the observed land fraction and the land fraction taken from the "first approximation" land sea mask must be less than -0.2 and the land fraction in the central grid cell must be less than 0.7).

In our example, only three of the cells satisfy the fundamental criteria, and the cell second from the left and second from the top contains the highest land fraction (0.469). Thus, in this case this cell is switched from ocean to land, yielding the next approximation shown in figure 6.

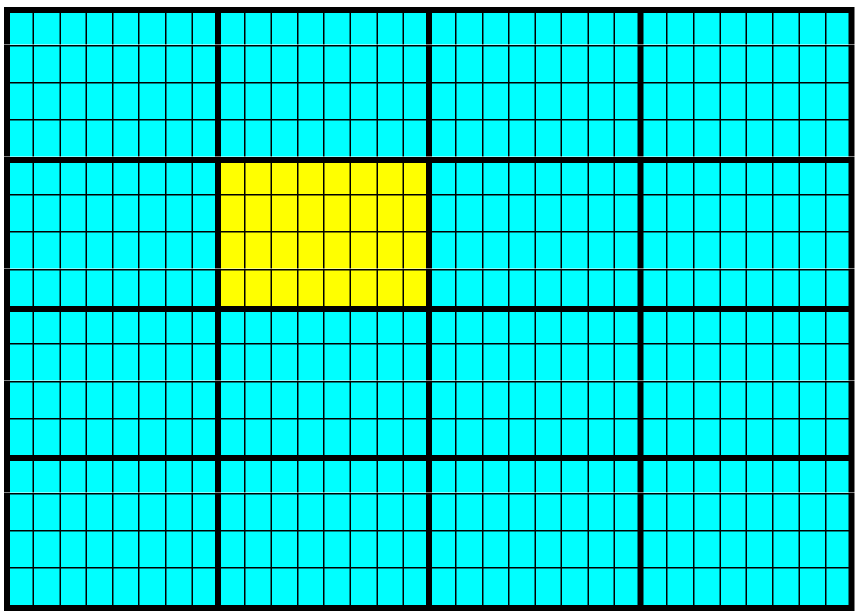

Fig. 6: The target land/sea mask after the first correction.

We now begin the second iteration of our correction procedure by calculating once again the difference between the observed land fraction data set and the most recent approximation (shown above). Figure 7 shows the difference (projected on the Navy's grid) between the Navy's data set and our land/sea mask. A '+' indicates a land fraction difference of 1.0 (i.e., 100% land on the Navy's grid, 0% on the model grid) and a '-' indicate a difference of -1.0 (i.e., 100% water on the Navy's grid, 0% on the model grid).

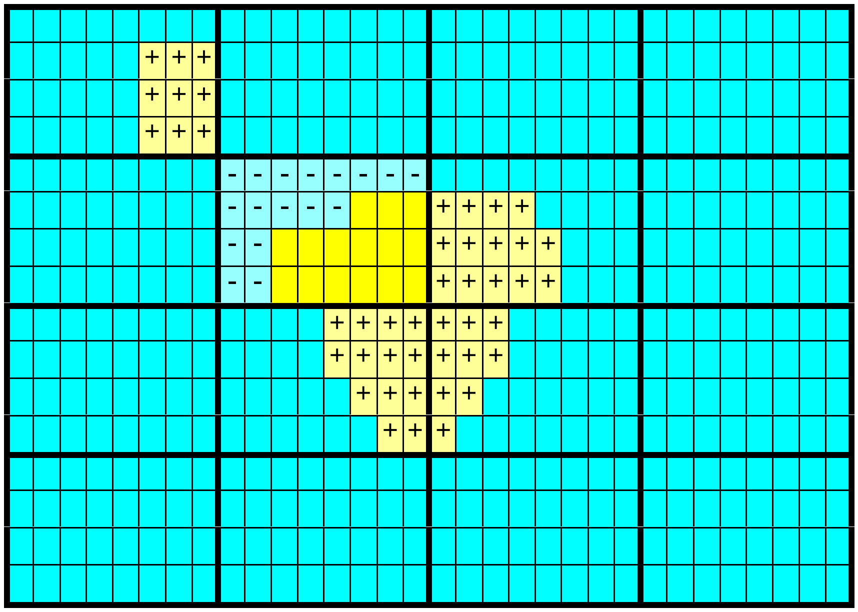

Fig. 7: The difference between the observed land fraction data set and the target land/sea mask after the first correction. Plus signs indicate where land is observed but water appears on the target grid. Minus signs indicate where water is observed but land appears on the target grid. Everywhere else the surface type is the same on both masks.

As before, we compute at half the model resolution the difference between the target land/sea mask and the observed land fraction for each target grid cell. Examining, for example, the same region as in figure 4, we consider the differences shown in figure 8.

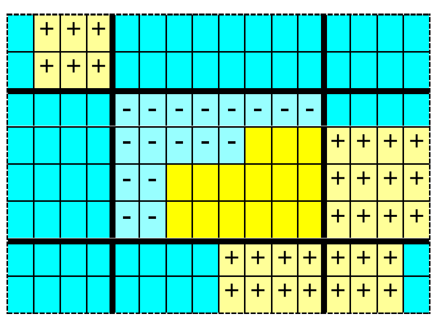

Fig. 8: The same region shown in figure 4, but after the first iteration of our correction procedure.

The error in land fraction for the region shown in figure 8 is obtained by calculating the difference between the number of pluses and minuses and dividing by the number of Navy grid cells in the region: (32-17)/128 = 0.117. A similar calculation can be performed for all the other grid cells, and the respective errors can then be assigned to the appropriate cells, as is shown in figure 9.

.008 .188	.000 .094	-.008 .000	.000 .000
-.047 .000	.117 .469	.180 .438	.016 .000
-.031 .000	.203 .406	.250 .281	.016 .000
.000 .000	.063 .000	.063 .000	.000 .000

Fig. 9: Same as figure 5, but after the first correction.

Now, as before, we search for grid cells where the upper number exceeds 0.2 and the lower number exceeds 0.3. (In the more general case we would also look for cells where the upper number is less than -0.2 and the lower number is less than 0.7.) After identifying a cell satisfying the fundamental criteria, we determine whether this cell contains a higher fraction of the appropriate surface type than the surrounding cells that satisfy the fundamental criteria. If so, we change this cell appropriately. In our example, we see that only one grid cell satisfies the fundamental criteria: the cell second from the bottom and second from the left. Switching this cell from ocean to land, we obtain the next approximation to the target land/sea mask shown in figure 10.

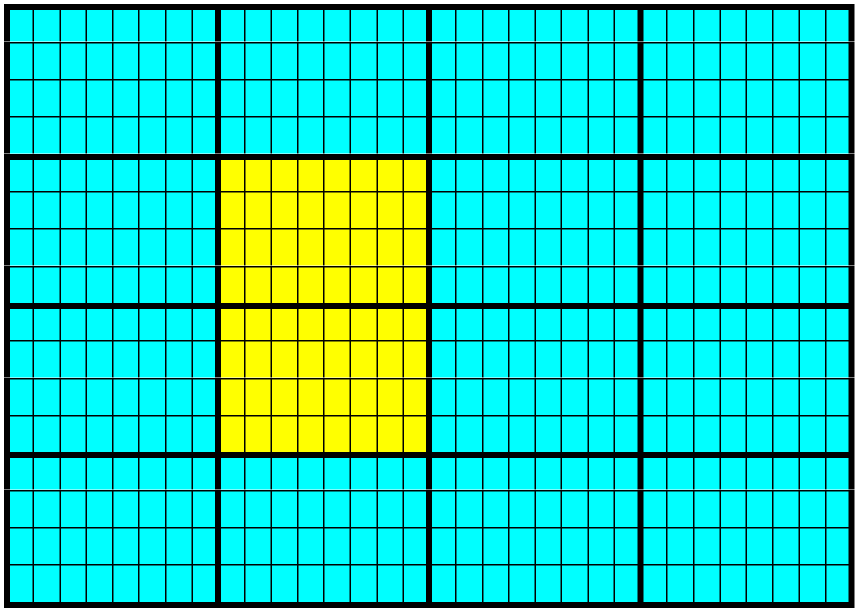

Fig. 10: The target land/sea mask after the second (and final) correction.

This iterative procedure is repeated until none of the grid cells satisfies the fundamental criteria (i.e., for no cell is the upper number greater than 0.2 and the lower number greater than 0.3, and for no cell is the upper number less than -0.2 and the lower number less than 0.7). In our example, this criterion is met after the second correction as can be seen in figure 11.

In our example, the iterative procedure has resulted in two grid cells being switched from ocean to land. In general the procedure reduces the difference (in an root-mean-square sense) recorded by the upper numbers shown in figures 5, 9, and 11 (i.e., the difference between the target land/sea data averaged over overlapping rectangular regions of four times the area of an individual cell and the corresponding averages of the high resolution observational data over the same regions). So although at the actual resolution of the target grid, the root-mean-square (RMS) difference between the original and target land/sea data may be slightly increased, the RMS difference decreases when evaluated at half the resolution. This is because islands and lakes that are missed, but are as large as a target grid cell, are accounted for, although they may be offset (by about a half grid cell, or so) from their true locations. Their inclusion in the final

.008 .188	.000 .094	-.008 .000	.000 .000
-.141 .000	-.008 .469	.117 .438	.016 .000
-.156 .000	-.047 .406	.125 .281	.016 .000
-.063 .000	-.063 .000	.000 .000	.000 .000

Fig. 11: Same as figures 5 and 9, but after the second (and last) correction.

land/sea mask reduces the RMS difference at the scale of half the model resolution, but because they may be slightly offset, the RMS difference increases at the scale of the full resolution of the model.

For a T42 Gaussian grid, figure 12 shows the difference in the North American region between the final land/sea mask (upon completion of our procedure) and the "first approximation" land/sea mask resulting from the first step (and shown in figure 2). It is evident that only a few grid cells have been altered by our procedure and these appear to better approximate the true land/sea distribution.

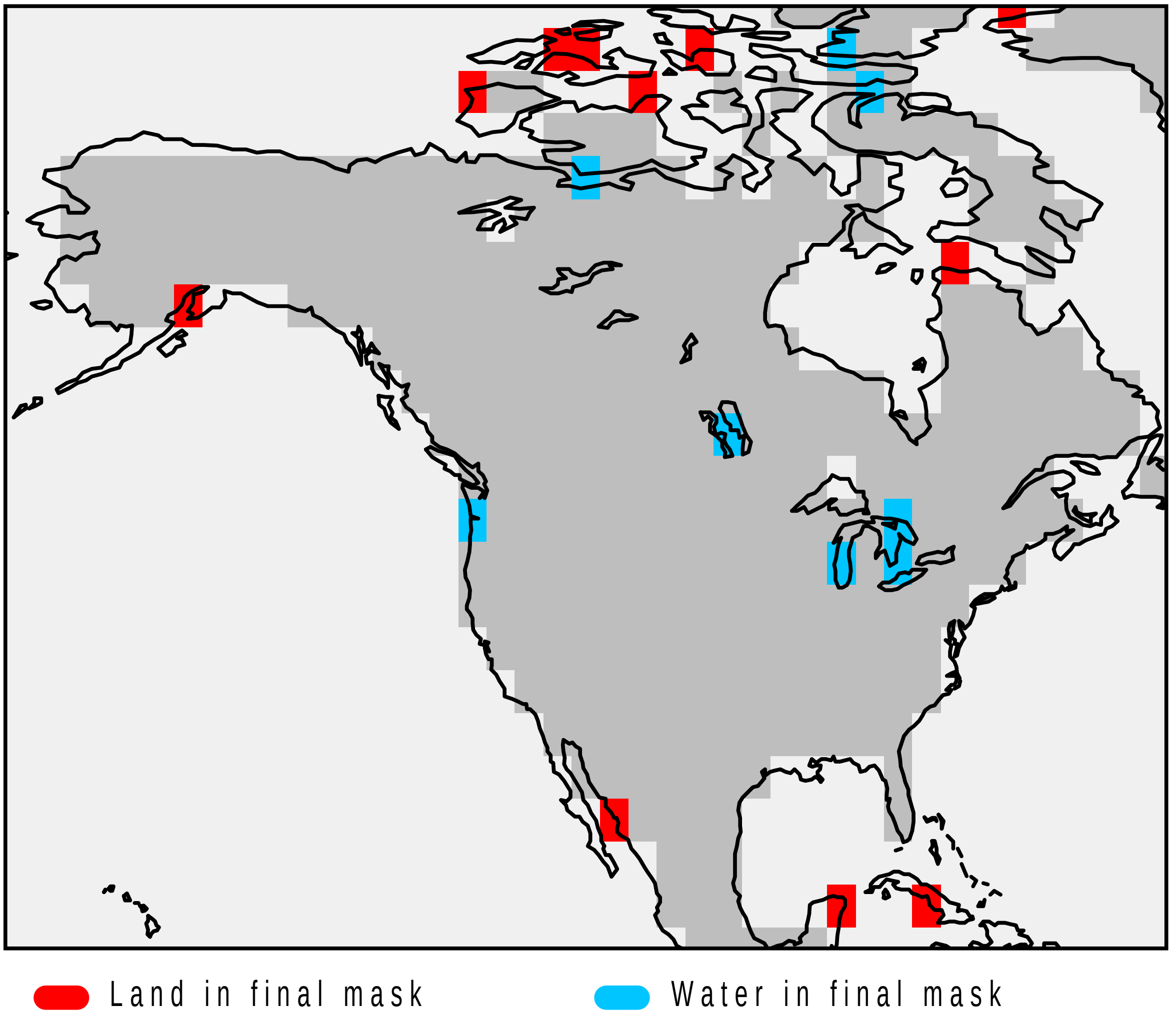

Fig. 12: Difference in the North American region between the final land/sea mask and the "first approximation" land/sea mask

3. Discussion and concluding remarks

There are various figures of merit that can be used to gauge how well a land/sea mask represents the true land/sea distribution. Here we consider three: 1) the global mean land fraction, 2) the root-mean-square (RMS) difference between the model's land fraction and the observations mapped to the model's grid (i.e., the RMS of the lower numbers appearing in figures 5, 9, and 11), and 3) the RMS difference between the model's land fraction and the observations on a grid of half the model's resolution (i.e., the RMS of the upper numbers appearing in figures 5, 9, and 11). In computing these figures of merit, all global means are computed with proper area-weighting of each grid cell.

Table 1 contains the three figures of merit for three different model resolutions after the "first approximation" and after the final iteration in our procedure. For comparison, the global mean land fraction obtained directly from the Navy's 1/6 x 1/6 degree data set is 28.02.

Table 1: Land fraction statistics for three different model resolutions (units: percent).

Resolution*	Iteration	Global Mean	RMS Difference (at the full model resolution)	RMS Difference (at half the model resolution)
R15 (40x48)	first approximation	28.18	13.39	6.38
	final	28.25	13.89	5.46
T42 (64x128)	first approximation	28.04	10.92	5.36
	final	28.17	11.29	4.47
T62 (94x192)	first approximation	28.40	9.71	4.84
	final	28.34	9.97	4.11

* Spherical harmonic truncation and number of latitude/longitude grid cells.

Our procedure captures a larger number of sub-grid scale islands than it does the sub-grid scale lakes, so the land fraction is overestimated on the model grids. (Some would argue, however, that many small lakes and rivers that are accounted for at a resolution of 1/6 degree should in fact be omitted entirely at the typical model resolution, so perhaps this overestimation can be justified.) In any case, the global mean land fraction is quite close to the observed at all resolutions. As noted earlier, the "first approximation" land/sea mask minimizes the RMS difference at the model's resolution. "Correcting" this mask through our iterative procedure increases this difference slightly, but at half the model resolution, the correction reduces the

RMS difference by a larger percentage. Thus, this method better represents the land/sea distribution when viewed at a resolution roughly half the model's resolution.

The land/sea mask resulting from the procedure described here can potentially produce a few grid cells where some would argue the representation appears unrealistic. Consider, for example, the hypothetical group of islands shown in figure 13.

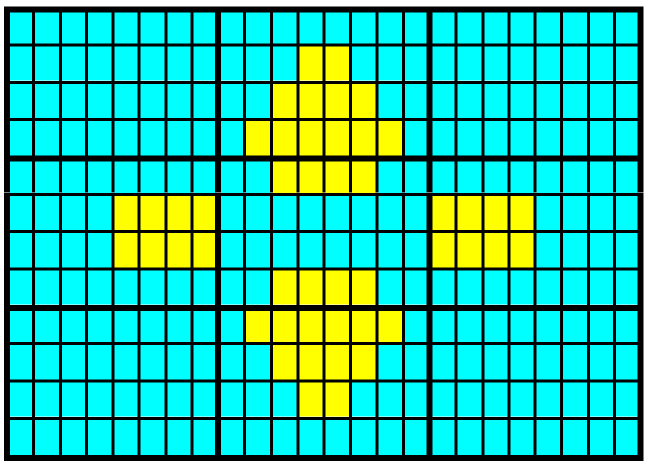

Fig. 13: A hypothetical distribution of islands.

The first approximation to the target land/sea mask would show ocean everywhere because the land fraction in each individual grid cell never exceeds 50%. The land fractions corresponding to the island distribution shown in figure 13 are given in figure 14.

Given the fractions in figure 14, it is clear that the "first approximation" would not be modified under our objective procedure because none of the grid cells satisfies the fundamental criteria that the upper number exceed 0.2 and the lower number exceed 0.3. Thus, no land associated with the four islands would appear on our final, objectively created land/sea mask. We note that in this case the land missed by our procedure would occupy an area 1.5 times the size of a model grid cell. Nevertheless, it is distributed over five grid cells in such a way that our procedure omits all land from the final land/sea mask.

.094 .000	.188 .375	.094 .000
.180 .250	.188 .375	.180 .250
.094 .000	.188 .375	.094 .250

Fig. 14: The land/sea fraction for the distribution of land shown in figure 13.

As a second example, consider a nine grid cell region populated by very small islands as shown in figure 15.

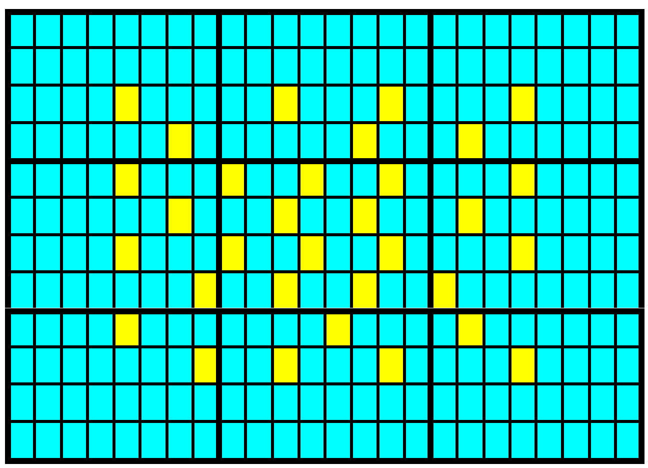

Fig. 15: Another hypothetical distribution of islands.

In this case the land fraction in each grid cell is less than 50%, which implies that the "first approximation" to the land/sea mask is everywhere water. However, once the correction procedure is applied, the central target grid cell becomes land (because it has a land fraction exceeding 0.3 and at a resolution half that of the model it has a land fraction exceeding 0.2). The final land/sea mask then contains a single land grid cell, which yields an average land fraction for the entire region of 11.1%. This agrees exactly with the original total land fraction. Some would argue, however, that the climate in this region dotted by tiny islands would hardly feel the effect of these islands, so the first approximation might be judged more realistic (i.e. the climate for an ocean dotted with tiny islands might be better simulated by a model with land-free oceans rather than a model with oceans punctuated by a few unrealistically large islands that supposedly represent the net influence of many small islands).

The above examples illustrate that objective procedures for creating land/sea masks on model grids, although capable of producing highly realistic masks, are inherently limited by the assumption that individual grid cells are either entirely land or entirely water. Moreover, there might be overriding considerations that would justify the use of a land/sea mask slightly different from the objectively generated mask. In coupled atmosphere/ocean models, for example, a very narrow strait could be missed by an objective procedure, but the strait might be critical for realistic simulation of the ocean circulation. In this case it might be preferable to override the objective procedure and change a grid cell that in reality is mostly land covered into an ocean grid cell. This again illustrates that the objective procedure described here for generating a land/sea mask for use in climate models may not in all cases be considered "best."

Acknowledgment. This work was performed under the auspices of the U.S. Department of Energy, Environmental Sciences Division, by the University of California under contract No. W-7405-Eng-48.