

## How do I interpret and use the PFT and FPCGRID variables?

gus@ideo\_columbia.edu · Feb 3, 2010



gus@ideo\_columbia.edu  
Member

Feb 3, 2010

Dear CLM experts

I and my colleagues are very confused about the meaning of variables PFT and FPCGRID in our CAM3+CLM3+DGVM run (files camrun.clm2.hv.YYYY-01-01-00000.nc). We cannot relate the PFT and FPCGRID data with the list of tree, grass, and crop types listed in our pft-physiology file, or translate them into information about the geographic distribution and percent coverage of those various plant types.

We are using cam-3.1.p2, the latest official release of CAM3, and the default pft-physiology file that lists 16 tree, grass, and crop types.

I read the DGVM, CLM3, and CAM3 User Guides, and all CLM postings on this bulletin board, but they don't clarify this matter. I need your help.

Variables PFT and FPCGRID have the same four dimensions: lon, lat, pft, and time, where the pft dimension size is 10 (range 0-9).

Variable PFT is an integer number with a range 0-14, showing "plant functional type".  
Variable FPCGRID is a real number with a range 0-100%, showing "fraction of vegetated area".

Questions:

- 1) The pft dimension has size 10, but there are 16 plant types listed on my pft-physiology file, and there are 15 possible PFT values (0-14) on variable PFT. How do I make sense of these three disparate numbers associated to plant functional types: 10, 16, 15? How do I relate them to each other? How to I relate the data on variables PFT and FPCGRID to the plant types listed on the pft-physiology file?
- 2) Is there a table/mapping associating the 0-14 range (15 values) of variable PFT to a subset of the 16 plant types listed in our pft-physiology file? Could somebody please post this table? Who's number 0? Who's number 1? ... Who's number 14?
- 3) What do the pft levels of variable FPCGRID point to? The meaning doesn't seem to be the same as the pft dimension in the PFT variable, but what is it?

Do I need to find the PFT variable "value" at each grid point, and then use that value to find out which plant type covers the percent area shown "at the same grid point" by FPCGRID?

4) How can use the data on variables PFT and FPCGRID to produce a simple figure showing the geographic distribution and percent cover of, say, broadleaf\_evergreen\_tropical\_tree?

I want and need to make plots similar to figures 3,4, 8 and 9 in the DGVM Technical Description and User's Guide, do diagnose if our spin up run reached steady state.

5) Is this all just hopeless? Did I make a mistake on my model setup? Do I have to use a pft-physiology file with exactly 10 plant types, no more, no less?

Thank you,  
Gus Correa

#####  
my pft-physiology file, the default one that comes with cam-3.1.p2  
#####

```

needleleaf_evergreen_temperate_tree 0.055 0.67 0.04 1. 51. 6. 0.06 0.07 0.35 0.16 0.39 0.05 0.10 0.001 0.001 0.01 7.0 2.0
needleleaf_evergreen_boreal_tree 0.055 0.67 0.04 1. 43. 6. 0.06 0.07 0.35 0.16 0.39 0.05 0.10 0.001 0.001 0.01 7.0 2.0
needleleaf_deciduous_boreal_tree 0.055 0.67 0.04 1. 43. 6. 0.06 0.07 0.35 0.16 0.39 0.05 0.10 0.001 0.001 0.01 7.0 2.0
broadleaf_evergreen_tropical_tree 0.075 0.67 0.04 1. 75. 9. 0.06 0.10 0.45 0.16 0.39 0.05 0.25 0.001 0.01 7.0 1.0
broadleaf_evergreen_temperate_tree 0.075 0.67 0.04 1. 69. 9. 0.06 0.10 0.45 0.16 0.39 0.05 0.25 0.001 0.001 0.10 7.0 1.0
broadleaf_deciduous_tropical_tree 0.055 0.67 0.04 1. 40. 9. 0.06 0.10 0.45 0.16 0.39 0.05 0.25 0.001 0.001 0.01 6.0 2.0
broadleaf_deciduous_temperate_tree 0.055 0.67 0.04 1. 51. 9. 0.06 0.10 0.45 0.16 0.39 0.05 0.25 0.001 0.001 0.25 6.0 2.0
broadleaf_deciduous_boreal_tree 0.055 0.67 0.04 1. 51. 9. 0.06 0.10 0.45 0.16 0.39 0.05 0.25 0.001 0.001 0.25 6.0 2.0
broadleaf_evergreen_shrub 0.120 0.68 0.04 1. 17. 9. 0.06 0.07 0.35 0.16 0.39 0.05 0.10 0.001 0.001 0.01 7.0 1.5
broadleaf_deciduous_temperate_shrub 0.120 0.68 0.04 1. 17. 9. 0.06 0.10 0.45 0.16 0.39 0.05 0.25 0.001 0.001 0.25 7.0 1.5
broadleaf_deciduous_boreal_shrub 0.120 0.68 0.04 1. 33. 9. 0.06 0.10 0.45 0.16 0.39 0.05 0.25 0.001 0.001 0.25 7.0 1.5
c3_arctic_grass 0.120 0.68 0.04 1. 43. 9. 0.06 0.11 0.58 0.36 0.58 0.07 0.25 0.220 0.380 -0.30 11.0 2.0
c3_non-arctic_grass 0.120 0.68 0.04 1. 43. 9. 0.06 0.11 0.58 0.36 0.58 0.07 0.25 0.220 0.380 -0.30 11.0 2.0
c4_grass 0.120 0.68 0.04 0. 24. 5. 0.04 0.11 0.58 0.36 0.58 0.07 0.25 0.220 0.380 -0.30 11.0 2.0
corn 0.120 0.68 0.04 1. 50. 9. 0.06 0.11 0.58 0.36 0.58 0.07 0.25 0.220 0.380 -0.30 6.0 3.0
wheat 0.120 0.68 0.04 1. 50. 9. 0.06 0.11 0.58 0.36 0.58 0.07 0.25 0.220 0.380 -0.30 6.0 3.0
    
```



slevis  
Member

Feb 4, 2010

Dear Gus,

I will try to clarify. Let's start with information from your posting:

"Variables PFT and FPCGRID have the same four dimensions: lon, lat, pft, and time, where the pft dimension size is 10 (range 0-9).  
Variable PFT is an integer number with a range 0-14, showing "plant functional type".  
Variable FPCGRID is a real number with a range 0-100%, showing "fraction of vegetated area"."

The first cause for confusion is the pft dimension present in both variables. This dimension represents sub-grid heterogeneity in the soil land unit of the grid cell. It represents 10 slots that plant functional types are allowed to occupy if they happen to exist in a grid cell. Of all clm plant functional types, you would not expect more than 10 to coexist in a grid cell due to climate considerations. So 10 was a choice that worked. Contrast this to 4 slots in the default clm3 (dgvn not active) where plant functional types and their %cover originates in clm's surface data set. Bottom line here is that this dimension could have been referred to as "soil land unit sub-grid heterogeneity" instead of "pft".

Given this new information, it seems relatively intuitive that each of these slots contains a PFT (range 0-14) with corresponding % or fractional cover FPCGRID. PFT values start with zero for bare ground and then follow the list in the pft-physiology file in the same order. No need to change this file. The clm uses the same file with or without dgvn active.

So, for example, to plot FPCGRID for PFT 4 (tropical broadleaf evergreen tree), you will need to isolate that PFT from one of the 10 slots in each grid cell (not necessarily slot #4) and return the FPCGRID associated with that slot. Once you come up with the algorithm to do that, you can write a script for your plotting software that will automate the process for you.

This is how things worked in clm3 and it was not a random decision nor a decision intended to confuse the user. I hope that my response makes sense and that you can make progress in your work now.

Sam Levis  
CLM Science Liaison



gus@ideo\_columbia.edu  
Member

Feb 5, 2010

Dear Sam Levis

Thank you for your helpful clarifications!

They confirm what I figured out meanwhile, and reassure me that what I did on Wednesday to build the plant cover maps is correct.

Also, thank you for confirming that the order of the PFTs listed on the pft-physiology file is honored by the program. I assumed this was true in what I did, and it is a relief to know that that was right.

Basically, for each of the 0-14 valid PFT types, I used the PFT variable to build a mask for that particular plant type on each pft index. Then I applied the mask to the FPCGRID variable e to get the plant cover maps. The algorithm goes along the line of my comments after my original question #3, and matches what you explained.

The results look correct now, with the various PFTs progressively occupying the areas where they are expected to be: tropical stuff in the Tropics (not in Greenland), Arctic grasses in Siberia, Canada and Alaska, etc.

It all makes sense now.

---

I have some extra questions:

These four PFTs:

- 3 (needleleaf\_deciduous\_boreal\_tree),
- 9 (broadleaf\_evergreen\_shrub)
- 10 (broadleaf\_deciduous\_temperate\_shrub)
- 11 (broadleaf\_deciduous\_boreal\_shrub)

don't seem to have been "seeded" during the "arbitrary initialization" of our run. Therefore, their cover remains equal to zero everywhere from the beginning to forever.

Other plant types, even those with a very restricted initial geographic distribution (e.g. PFT 1, needleleaf\_evergreen\_temperate\_tree), thrive and expand their area and cover.

Questions:

1) Is there any particular reason for the choice to "sacrifice" these particular four types in the "arbitrary initialization"?

(The program takes care internally of the "arbitrary initialization", there is no user intervention there. Are those four plant types less important or less representative than the others?)

2) Are the initial areal and cover distribution of PFTs fixed (hardwired) inside the program, or are they based on a random number generator, or other probabilistic technique that might change the "elected" PFTs?

3) As a consequence, should I expect a different set of PFTs to thrive if I run another spinup, or will I get the same group of PFTs that I already have?

4) Is there any way to tweak/change the "arbitrary initialization" and how it "seeds" the various PFTs?

Thank you again.  
Gus Correa



slevis  
Member

Feb 12, 2010

1) When I coupled the LPJ-dgvn into the CLM, the dgvn did not "know" how to simulate these pfts. So I kept them out using the table in subroutine DGVMEcosystemDynini.

2 and 3) The outcome will be the same every time.

4) Definitely. First pls read the clm and dgvn technical descriptions (not just the user's) available on the clm distribution pages. These should answer most of your questions. It's also worth browsing through studies that used this model before. If, as you suggest, you intend to make changes to the model, I also recommend becoming very familiar with the code, so that you know how it works before you change it.

Good luck with your work.

Sam

You must log in or register to reply here.