Soil temperature simulation results in Alaska (1980 - 2014) – Data archive for “Evaluation and enhancement of permafrost modeling with the NASA Catchment Land Surface Model”

Jing Tao (JingTao@umd.edu)[1], Rolf H. Reichle[2], Randal D. Koster[2], Barton A. Forman[3], Yuan Xue[3]

1 - Earth System Science Interdisciplinary Center, University of Maryland, College Park, Maryland

2 - Global Modeling and Assimilation Office, NASA Goddard Space Flight Center, Greenbelt, Maryland

3 - Department of Civil and Environmental Engineering, University of Maryland, College Park, Maryland

Contents

1. Abstract

The datasets archived here include simulation results discussed in the paper, "Evaluation and enhancement of permafrost modeling with the NASA Catchment Land Surface Model", to be published in Journal of Advances in Modeling Earth Systems. Specifically, subsurface soil temperatures for 1980-2014 across Alaska were produced by a baseline simulation with the NASA Catchment Land Surface Model (CLSM). Five sets of point simulations were also conducted at permafrost sites in Alaska, including 1) T1BC - the top layer temperature is prescribed to observations, 2) T1BC_OrgC – repeat of the T1BC simulation but using the updated model version that incorporates soil thermal impacts of organic carbon content, 3) T2BC - the temperatures of both the 1st and 2nd layer are prescribed to observations, 4) T2BC_OrgC – repeat of the T2BC simulation but using the updated model version, and 5) M2_OrgC – simulations with the updated model version driven by MERRA-2 forcing. Details about the model configuration and the changes defining the updated model version can be found in the paper. The major findings in this paper include: a) profile-average RMSE of simulated soil temperature versus in situ observations is reduced by using corrected local forcing and land cover; b) subsurface heat transport is mostly realistic, and when not, it is improved via treatment of soil organic carbon-related thermal properties; and c) mean bias and RMSE of climatological ALT between simulations and observations are significantly reduced with updated model version.

2. Data description

The datasets provide 3-hourly simulation results with the NASA CLSM for the period 1980-2014. The hour and minute contained in the binary filename indicate the center of a 3-hourly span. For instance, the file

`SMAP_Nature_v05_PT.ens_avg.ldas_tile_xhourly_out.20040303_0730z.bin`

contains the 3-hourly average results from 0600z to 0900z.

Results from six simulations are archived here, including Baseline, T1BC, T1BC_OrgC, T2BC, T2BC_OrgC and M2_OrgC. Soil temperatures from each simulation are provided at six model layers. The soil depths of the six soil layers are 0~0.1m, 0.1~0.3m, 0.3~0.7m, 0.7~1.4m, 1.4~3m, and 3~13m from top to bottom, respectively. Besides soil temperature and ice fraction in each soil layer (necessary for computing the depth of thaw-to-frozen condition), many other variables are also stored in the binary files. Each binary file contains 46 variables as listed in Table 1.

Table 1 – Variables stored in binary file.

Field	**Variables**	**Units**
1	Soil moisture in surface layer	m3 m-3
2	Soil moisture in rootzone	m3 m-3
3	Soil moisture in profile	m3 m-3
4	Soil wetness (degree of saturation) in surface layer	dimensionless
5	Soil wetness in rootzone	dimensionless
6	Soil wetness in profile	dimensionless
7	Surface temperature	K
8	**Soil temperature in layer 1**	**K**
9	**Soil temperature in layer 2**	**K**
10	**Soil temperature in layer 3**	**K**
11	**Soil temperature in layer 4**	**K**
12	**Soil temperature in layer 5**	**K**
13	**Soil temperature in layer 6**	**K**
14	Snow mass	kg m-2
15	Snow depth	m

16	Land evapotranspiration flux	kg m-2 s-1
17	Overland runoff flux	kg m-2 s-1
18	Baseflow flux	kg m-2 s-1
19	Snow melt flux	kg m-2 s-1
20	Soil water infiltration flux	kg m-2 s-1
21	Land fraction of saturated area	dimensionless
22	Land fraction of unsaturated but non-wilting area	dimensionless
23	Land fraction over which soil moisture is below wilting point	dimensionless
24	Land fraction of snow covered area	dimensionless
25	Heat flux sensible	W m-2
26	Heat flux latent	W m-2
27	Heat flux ground	W m-2
28	Net downward shortwave flux	W m-2
29	Net downward longwave flux	W m-2
30	Radiation shortwave downward flux	W m-2
31	Radiation longwave absorbed flux	W m-2
32	Total precipitation (surface flux)	kg m-2 s-1
33	Snowfall (surface flux)	kg m-2 s-1
34	Surface pressure	Pa
35	Height of lowest model layer of atmospheric model	m
36	Temperature at the lowest atmospheric model layer	K
37	Specific humidity at the lowest atmospheric model layer	kg kg-1
38	Wind speed at the lowest atmospheric model layer	m s-1
39	Vegetation greenness fraction	dimensionless
40	Leaf area index	m2 m-2
41	**Ice fraction in soil layer 1**	**dimensionless**
42	**Ice fraction in soil layer 2**	**dimensionless**
43	**Ice fraction in soil layer 3**	**dimensionless**
44	**Ice fraction in soil layer 4**	**dimensionless**
45	**Ice fraction in soil layer 5**	**dimensionless**
46	**Ice fraction in soil layer 6**	**dimensionless**

All simulations were conducted at the EASE-Grid 9km resolution. Both the baseline and M2_OrgC simulations used MERRA2 forcing and thus contain results spanning 1980 to 2014. But for the other four series of experiments, i.e. T1BC, T1BC_OrgC, T2BC and T2BC_OrgC, the observed soil temperature in the 1st or both the 1st and 2nd layers were used as upper boundary conditions, and thus the results are specifically restricted to the observing period at each site.

The simulations with the updated model version that incorporates soil thermal impacts of organic carbon content, i.e. T1BC_OrgC, T2BC_OrgC and M2_OrgC, contain two sets of carbon data, namely IGBP and NCSCD which are explicitly included in folder names for each site. M2_OrgC contains results at 86 permafrost sites that are maintained by the Permafrost Laboratory at University of Alaska Fairbanks, while T1BC, T1BC_OrgC, T2BC and T2BC_OrgC only contains results at 7 sites. Table 2 provides site locations and names for all the 86 sites.

Table 2 – Locations and names of all the 86 permafrost sites[1].

No.	Site Code	Site Name	Latitude	Longitude	Elevation
1	AG1	Atigun Pass 1	68.1339°	-149.4630°	1,345m
2	AG2	Atigun Pass 2	68.1298°	-149.4786°	1,473m
3	AKR	Akulik River	64.9175°	-160.7281°	62m
4	BL1	Birch Lake	64.3235°	-146.6883°	262m
5	BR1	Barrow 1 (N. Meadow Lake No.1 / NML-1)	71.3106°	-156.6543°	5m
6	BR2	Barrow 2 (N. Meadow Lake No.2 / NML-2)	71.3090°	-156.6615°	5m
7	BR3	Barrow Road	71.3217°	-156.6714°	5m
8	BZ1	Bonanza Creek 1	64.7069°	-148.2913°	125m
9	BZ2	Bonanza Creek 2	64.7150°	-148.2885°	132m
10	CF1	Coldfoot	67.2373°	-150.1614°	335m
11	CFS	Coldfoot South	67.2012°	-150.2702°	354m
12	COF	Council 2 (Forest)	64.9069°	-163.6778°	70m
13	COS	Council 1 (Shrub)	64.9348°	-163.7380°	84m
14	COT	Council 4 (Tundra)	64.8412°	-163.6963°	34m
15	COW	Council 3 (Woodland)	64.8999°	-163.6644°	41m
16	CPR	Caribou-Poker creek (C4new, Poker1)	65.1800°	-147.4400°	295m
17	CPT	College Peat	64.8678°	-147.7849°	137m
18	CS1	Chandalar Shelf	68.0691°	-149.5804°	976m
19	DH1 / DH2	Deadhorse	70.1613°	-148.4653°	17m
20	DN1	Donnelly	63.7500°	-145.8333°	758m
21	EA1	Eagle	64.7784°	-141.1802°	267m
22	FB1	Franklin Bluffs (borehole)	69.6739°	-148.7219°	88m
23	FBD	Franklin Bluffs (dry)	69.6741°	-148.7208°	122m
24	FBW	Franklin Bluffs (wet)	69.6746°	-148.7196°	122m
25	FX1	Fox	64.9506°	-147.6177°	240m
26	GI1	GI #1	64.8753°	-147.8465°	170m
27	GI2	GI #2	64.8711°	-147.8468°	170m

[1] http://permafrost.gi.alaska.edu/sites_map

28	GK1	Gulkana	62.1669°	-145.4707°	479m
29	GL1	Galbraith Lake	68.4774°	-149.5024°	823m
30	HI1	Howe Island	70.3151°	-147.9931°	0m
31	HL1	Healy	63.8784°	-149.2535°	673m
32	HP1	Gakona 1	62.3929°	-145.1453°	550m
33	HP2	Gakona 2	62.3913°	-145.1469°	548m
34	HP3	Gakona 3	62.3947°	-145.1569°	556m
35	HP4	Gakona 4	62.4009°	-145.1575°	565m
36	HV1	Happy Valley	69.1466°	-148.8483°	309m
37	HVO	Happy Valley (borehole)	69.1565°	-148.8374°	319m
38	IM1	Imnaviat 1	68.6397°	-149.3523°	871m
39	IM2	Imnaviat 2	68.6336°	-149.3459°	871m
40	IV1	Ivotuk-1 (Acidic tundra)	68.4873°	-155.7428°	571m
41	IV2	Ivotuk-2 (Non acidic tundra)	68.4797°	-155.7359°	574m
42	IV3	Ivotuk-3 (Shrub)	68.4789°	-155.7381°	575m
43	IV4	Ivotuk-4 (Moss)	68.4803°	-155.7437°	560m
44	KA1	Kaktovik (ANWR / Barter)	70.1165°	-143.6285°	-
45	KC1	Kugurak Cabin	66.5624°	-159.0046°	7m
46	KCF	Kugurak Cabin Forest	66.5617°	-159.0002°	15m
47	KCT	Kugurak Cabin Tundra	66.5621°	-159.0034°	15m
48	KZ1	Kotzebue	66.8518°	-162.6043°	20m
49	KZR	Kuzitin River (SePe-5)	65.2295°	-164.8285°	-
50	LG1	Livengood	65.5068°	-148.6015°	179m
51	NO1	Nome	64.5080°	-165.2961°	20m
52	OM1	Old Man	66.4502°	-150.6188°	390m
53	PHS	Pilgrim Hot Springs	65.0854°	-164.8952°	-
54	QZC	Quartz Creek	65.5475°	-161.4032°	181m
55	S1-BF	Selawik Upland Spruce-Aspen Forest	66.7636°	-160.0921°	89m
56	S1-WS	Selawik Upland White Spruce	66.8457°	-160.0170°	228m
57	S2-PB	Selawik Upland Burn 2010	66.5382°	-158.3628°	68m
58	S3-AWS	Selawik Upland Alder Willow Shrub	66.6113°	-158.6836°	148m
59	S3-BEW	Selawik Upland Birch-Ericace	66.6071°	-158.6795°	129m
60	S3-LSF	Selawik Lowland Sedge Fen	66.5846°	-158.7682°	19m
61	S3-TM	Selawik Tussock Meadow	66.6125°	-158.6554°	107m
62	S4-AWS	Selawik Lowland Alder Willow	66.6535°	-160.1482°	16m
63	S4-LS	Selawik Lowland Birch Ericaceous	66.6551°	-160.1362°	15m
64	S4-TM	Selawik Lowland Tussock Meadow	66.6593°	-160.1219°	10m
65	S8-PB	Selawik Old Post Burn	66.8912°	-158.7009°	101m
66	SC1	Spruce Creek 1	65.0592°	-147.5665°	414m
67	SC2	Spruce Creek 2	65.0603°	-147.5707°	359m
68	SG1	SagMNT (Moist Non-Acidic Tundra)	69.4330°	-148.6738°	277m
69	SG2	SagMAT (Moist Acidic Tundra)	69.4283°	-148.7001°	278m
70	SL1	Smith Lake 1	64.8694°	-147.8608°	160m

71	SL2	Smith Lake 2	64.8661°	-147.8568°	157m
72	SL3	Smith Lake 3	64.8675°	-147.8588°	157m
73	SL4	Smith Lake 4	64.8669°	-147.8584°	156m
74	SS-AWS	Selawik Thaw Slump Alder Willow	66.5014°	-157.6094°	150m
75	SS-TM	Selawik Thaw Slump Tussock Meadow	66.5012°	-157.6074°	482m
76	SS-WS	Selawik Thaw Slump White Spruce	66.4998°	-157.6042°	130m
77	SSE	Shishmaref SE (Shishmaref SE)	66.1012°	-165.5465°	-
78	SV1	Selawik Village	66.6056°	-160.0192°	55m
79	TKN104	Nome Creek	65.3460°	-146.6545°	-
80	TL1	Toolik Lake	68.6280°	-149.5954°	727m
81	TLO	Teshekpuk Lake Observatory	70.7229°	-153.8363°	-
82	UF1	UAF Farm	64.8529°	-147.8575°	150m
83	UUG	Upper Ungalik	65.0554°	-159.4734°	391m
84	WA1	Wainwright	70.6446°	-160.0223°	3m
85	WD1	West Dock (Old Instrumentation)	70.3745°	-148.5522°	3m
86	YB1	Yukon Bridge	65.8804°	-149.7099°	221m

3. Reading data

1) First, one should unzip all the compressed files using the command below:

```
tar -xvf  Filename.tar.gz
```

In each uncompressed folder, one should find similar data structure organized as:

```
./cat
    ../ens_avg/Y1980/M01/
./rc_out
./rs
```

The binary data are stored within the `cat/ens_avg/` folder. The folder `rs` includes restart fields for the model. The folder `rc_out` includes domain setup information and grid coordinates.

2) Four Matlab codes are provided for reading the binary data. One should use the code below to extract simulation results from a binary file.

```
[tile_coord ] = read_tilecoord(tilecoord_nm);
[ tile_grid_g, tile_grid_d ] = read_tilegrids(tilegrids_nm);

[tile_data] = read_tile_data(filename, tile_coord.N_tile, N_out_fields);
```

Here, one should specify absolute paths of the two coordinate files which are contained in folder `rc_out`, and the absolute path of the binary filename.

```
tilecoord_nm: *_tilecoord.bin
tilegrids_nm: *_tilegrids.bin
filename: *.ens_avg.ldas_tile_xhourly_out.YYYYMMDD_HHMMz.bin
```

For point simulation, i.e. T1BC, T1BC_OrgC, T2BC, T2BC_OrgC and M2_OrgC, all the variables are contained in `tile_data` which has a dimension of `Nfield`. Here `Nfield = 46` is the total number of variables as shown in Table 1.

For baseline (spatial) simulation, one should run extra two lines of command to map the `tile_data` to a three-dimensional global matrix.

```
[grid_data]  =  tile2grid( tile_data, tile_coord, tile_grid_g);

data =permute(grid_data,[2 1 3]);
```

The global matrix `data` has dimension `[Nrow, Ncol, Nfield]` where `Nrow = 1624` is the number of grid cells from north to south and `Ncol = 3856` is the number of grid cells from west to east. Note that for the baseline simulation, although the extracted data encompass a global grid, only grid cells over the Alaska area have valid data. For grid cells not simulated, `NaN` will be found for all the variables.

3) The coordinates of grid cells for each specific simulation can be extracted from the files within rc_out. For example,

```
[tile_coord ] = read_tilecoord(tilecoord_nm);
[ tile_grid_g, tile_grid_d ] = read_tilegrids(tilegrids_nm);


[ latc ]  =  tile2grid( tile_coord.com_lat', tile_coord, tile_grid_g);
[ lonc ]  =  tile2grid( tile_coord.com_lon', tile_coord, tile_grid_g);
```

Here `latc` and `lonc` is the latitude and longitude at the mass center of grid cells.

Please contact Jing Tao (jingtao@umd.edu) for any questions.