

Daily gridded meteorological variables in Brazil (1980-2013)

SUPPLEMENTAL INFORMATION

Alexandre C. Xavier¹, Carey W. King^{*,2}, and Bridget R. Scanlon³

1. Federal University of Espirito Santo, Alegre, Brazil
2. Energy Institute, The University of Texas at Austin, Texas
3. Bureau of Economic Geology, The Jackson School of Geosciences, The University of Texas at Austin, Texas

* careyking@mail.utexas.edu

Contents

Acronyms.....	1
Tables	2
Figures	5
README for using netcdf (.nc) files of gridded meteorological data.....	13
Overview	13
List of files.....	14
Sample code in MATLAB to read and plot data	17

Acronyms

Interpolation methods:

IDW = inverse distance weighting

OPK = ordinary point kriging

THINPLATE = thin plate spline

ADW = angular distance weighting

NATURAL = natural interpolation

AVERAGE = arithmetic average

Statistical metrics used for skill scores:

RMSE = root mean square error

R = coefficient of correlation

CRE = compound relative error

MAE = mean absolute error

CSIL = critical success index for low extreme values

CSIH = critical success index for high extreme values

Tables

Table S1. Statistics and their respective skill score for each model and variable.

Var	Method	AV rank	R	#	bias	#	RMSE	#	CRE	#	MAE	#	CSIL	#	CSIH	
<i>T_{max}</i>	IDW	1.571	0.91	1	-0.026	4	1.904	1	0.174	1	1.332	1	0.629	2	0.469	1
	OPK	2.286	0.908	3	-0.002	2	1.921	3	0.177	3	1.337	2	0.631	1	0.469	2
	THINPLATE	3.429	0.902	4	-0.019	3	1.983	4	0.189	4	1.388	3	0.62	3	0.46	3
	ADW	3.571	0.908	2	-0.033	6	1.913	2	0.176	2	1.395	4	0.615	4	0.449	5
	NATURAL	4.286	0.895	5	0.001	1	2.059	5	0.203	5	1.421	5	0.608	5	0.452	4
	AVERAGE	5.857	0.871	6	-0.031	5	2.328	6	0.26	6	1.612	6	0.576	6	0.417	6
<i>T_{min}</i>	IDW	1.571	0.906	1	-0.063	4	1.873	1	0.182	1	1.333	1	0.577	2	0.25	1
	OPK	1.857	0.906	2	-0.033	2	1.874	2	0.182	2	1.334	2	0.584	1	0.248	2
	THINPLATE	3.286	0.901	4	-0.019	1	1.922	4	0.191	4	1.372	3	0.573	3	0.238	4
	ADW	4	0.905	3	-0.108	5	1.877	3	0.182	3	1.39	4	0.561	4	0.221	6
	NATURAL	4.429	0.893	5	-0.037	3	2.002	5	0.208	5	1.405	5	0.559	5	0.245	3
	AVERAGE	5.857	0.87	6	-0.113	6	2.251	6	0.262	6	1.583	6	0.523	6	0.222	5
<i>R_s</i>	ADW	1	0.869	1	0.001	1	2.824	1	0.245	1	2.047	1	0.511	1	0.429	1
	OPK	2.571	0.863	3	0.003	2	2.902	3	0.259	3	2.076	2	0.5	2	0.419	3
	IDW	2.714	0.863	2	0.011	5	2.893	2	0.257	2	2.076	3	0.498	3	0.422	2
	THINPLATE	4.286	0.854	4	-0.02	6	2.995	4	0.276	4	2.131	4	0.493	4	0.412	4
	NATURAL	4.857	0.845	5	0.008	4	3.103	5	0.296	5	2.195	5	0.478	5	0.404	5
	AVERAGE	5.571	0.814	6	-0.003	3	3.474	6	0.371	6	2.449	6	0.452	6	0.378	6
<i>R_H</i>	ADW	1.429	0.851	1	-0.088	4	7.403	1	0.277	1	5.528	1	0.468	1	0.263	1
	IDW	2.571	0.841	2	-0.134	6	7.666	2	0.297	2	5.69	2	0.448	2	0.252	2
	OPK	3.286	0.84	3	-0.093	5	7.727	3	0.302	3	5.724	3	0.448	3	0.25	3
	THINPLATE	3.714	0.831	4	-0.032	2	7.935	4	0.318	4	5.867	4	0.443	4	0.24	4
	NATURAL	4.714	0.823	5	-0.036	3	8.162	5	0.337	5	6.033	5	0.432	5	0.228	5
	AVERAGE	5.286	0.786	6	0.005	1	9.168	6	0.425	6	6.775	6	0.4	6	0.201	6
<i>u₂</i>	ADW	1.286	0.572	1	-0.007	2	1.132	1	0.693	1	0.834	1	0.078	1	0.196	2
	IDW	2.143	0.55	2	-0.019	4	1.178	2	0.751	2	0.867	2	0.067	2	0.196	1
	OPK	3	0.535	3	-0.017	3	1.211	3	0.793	3	0.889	3	0.061	3	0.189	3
	THINPLATE	3.571	0.518	4	0.001	1	1.243	4	0.836	4	0.911	4	0.059	4	0.186	4
	NATURAL	5.143	0.498	5	-0.026	6	1.28	5	0.887	5	0.933	5	0.053	5	0.177	5
	AVERAGE	5.857	0.457	6	-0.023	5	1.404	6	1.067	6	1.025	6	0.047	6	0.158	6

Table S2. Cross validation results for interpolation methods per variable and per basin.

Variable (method)	Basin	R	Bias	RMSE	CRE	MAE	CSIL	CSIH
<i>Tmax</i> (IDW)	Amazon river	0.757	0.016	1.567	0.439	1.163	0.325	0.348
	Tocantins river	0.855	-0.198	1.53	0.281	1.114	0.392	0.524
	North Atlantic region	0.827	0.074	1.759	0.32	1.178	0.306	0.435
	Sao Francisco river	0.843	-0.186	1.905	0.295	1.374	0.362	0.43
	Central Atlantic region	0.793	-0.092	2.361	0.391	1.693	0.344	0.305
	Parana river	0.92	0.103	1.81	0.155	1.267	0.556	0.546
	Uruguay river	0.941	0.106	2.059	0.114	1.418	0.575	0.502
	South Atlantic region	0.925	-0.349	2.159	0.149	1.509	0.598	0.434
<i>Tmin</i> (IDW)	Amazon river	0.755	0.001	1.488	0.445	1.073	0.561	0.177
	Tocantins river	0.789	-0.282	1.745	0.423	1.22	0.356	0.252
	North Atlantic region	0.646	0.015	1.732	0.611	1.239	0.29	0.185
	Sao Francisco river	0.853	-0.076	1.874	0.278	1.355	0.404	0.354
	Central Atlantic region	0.817	-0.167	2.162	0.349	1.604	0.399	0.268
	Parana river	0.904	0.156	1.883	0.184	1.318	0.499	0.415
	Uruguay river	0.943	0.123	1.755	0.112	1.249	0.625	0.444
	South Atlantic region	0.928	-0.291	1.903	0.141	1.375	0.581	0.397
<i>Rs</i> (ADW)	Amazon river	0.71	0.132	3.409	0.508	2.599	0.237	0.253
	Tocantins river	0.788	0.018	2.79	0.384	2.039	0.331	0.391
	North Atlantic region	0.83	0.028	2.695	0.311	1.973	0.386	0.302
	Sao Francisco river	0.875	-0.246	2.54	0.236	1.835	0.411	0.363
	Central Atlantic region	0.859	0.109	3.049	0.263	2.192	0.4	0.339
	Parana river	0.888	-0.021	2.632	0.212	1.886	0.534	0.437
	Uruguay river	0.945	-0.045	2.475	0.106	1.77	0.467	0.44
	South Atlantic region	0.94	0.164	2.563	0.117	1.838	0.42	0.482

Table S2. Cross validation results for interpolation methods per variable and per basin (Continued).

Variable (method)	Basin	R	Bias	RMSE	CRE	MAE	CSIL	CSIH
<i>RH</i> (ADW)	Amazon river	0.746	0.01	6.859	0.452	5.041	0.497	0.125
	Tocantins river	0.896	-0.032	6.956	0.197	5.207	0.456	0.232
	North Atlantic region	0.84	-0.262	7.931	0.294	5.95	0.44	0.241
	Sao Francisco river	0.833	0.527	8.211	0.308	6.259	0.378	0.324
	Central Atlantic region	0.756	-0.683	7.254	0.441	5.526	0.356	0.247
	Parana river	0.873	0.043	6.888	0.239	5.151	0.504	0.338
	Uruguay river	0.844	0.557	7.028	0.29	5.233	0.318	0.314
	South Atlantic region	0.802	-0.858	6.722	0.368	5.078	0.347	0.293
<i>u2</i> (ADW)	Amazon river	0.253	0.03	0.933	1.044	0.697	0	0.106
	Tocantins river	0.326	0.003	0.902	0.963	0.701	0	0.142
	North Atlantic region	0.652	0.021	1.108	0.583	0.824	0.092	0.142
	Sao Francisco river	0.596	-0.011	1.066	0.66	0.788	0.04	0.208
	Central Atlantic region	0.452	-0.012	1.181	0.832	0.861	0.071	0.127
	Parana river	0.432	-0.022	1.059	0.853	0.82	0	0.181
	Uruguay river	0.376	0.003	1.461	0.927	1.049	0.071	0.124
	South Atlantic region	0.399	0.17	1.512	0.921	1.148	0.064	0.111

Figures

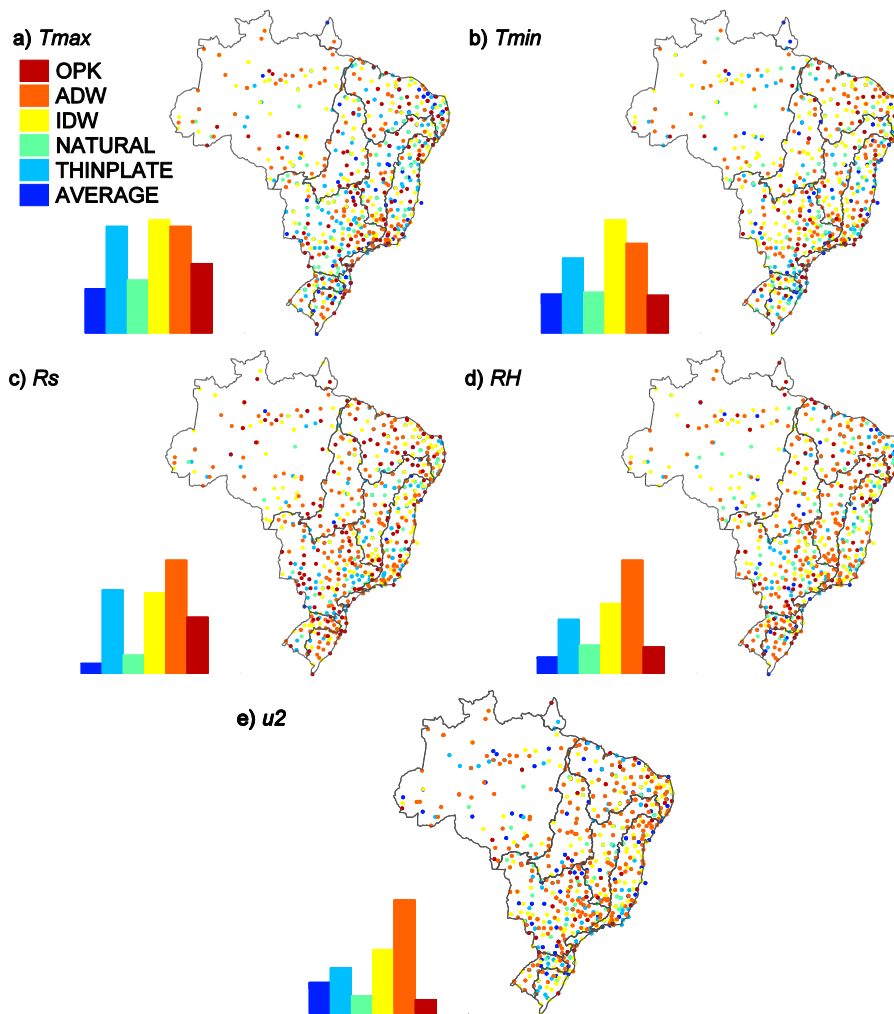


Figure S1. The most precise interpolation method, as measured by highest average skill score, is indicated for Brazil overall (relative frequency histogram in lower left of each plot) and for each station (colored dots on the map) for T_{max} (a), T_{min} (b), R_s (c), RH (d) and u_2 (e). Each interpolation method from top to bottom in the legend is represented from right to left in the histograms.

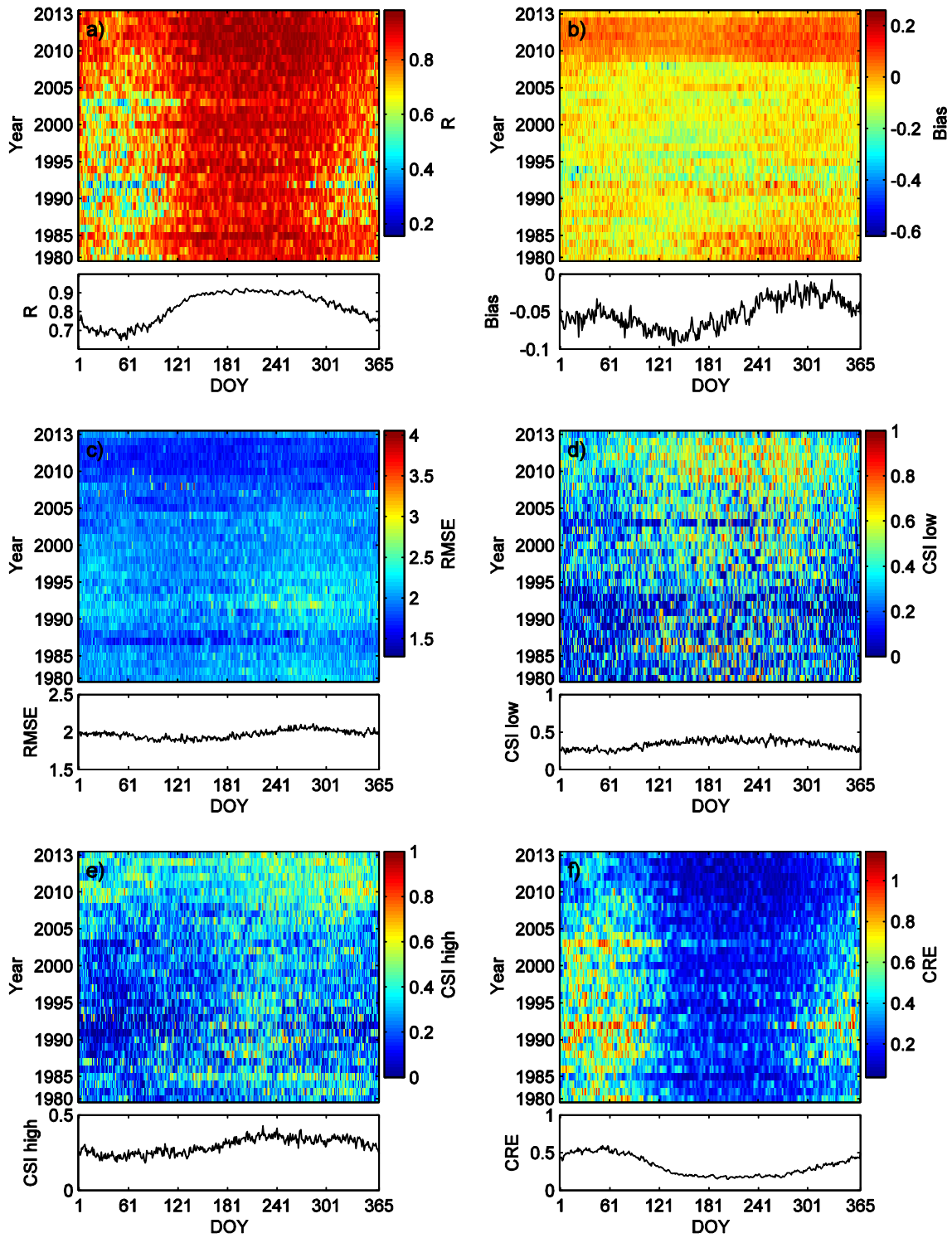


Figure S2. Daily skill scores of the relationship between observed and estimated T_{max} when interpolating using IDW.

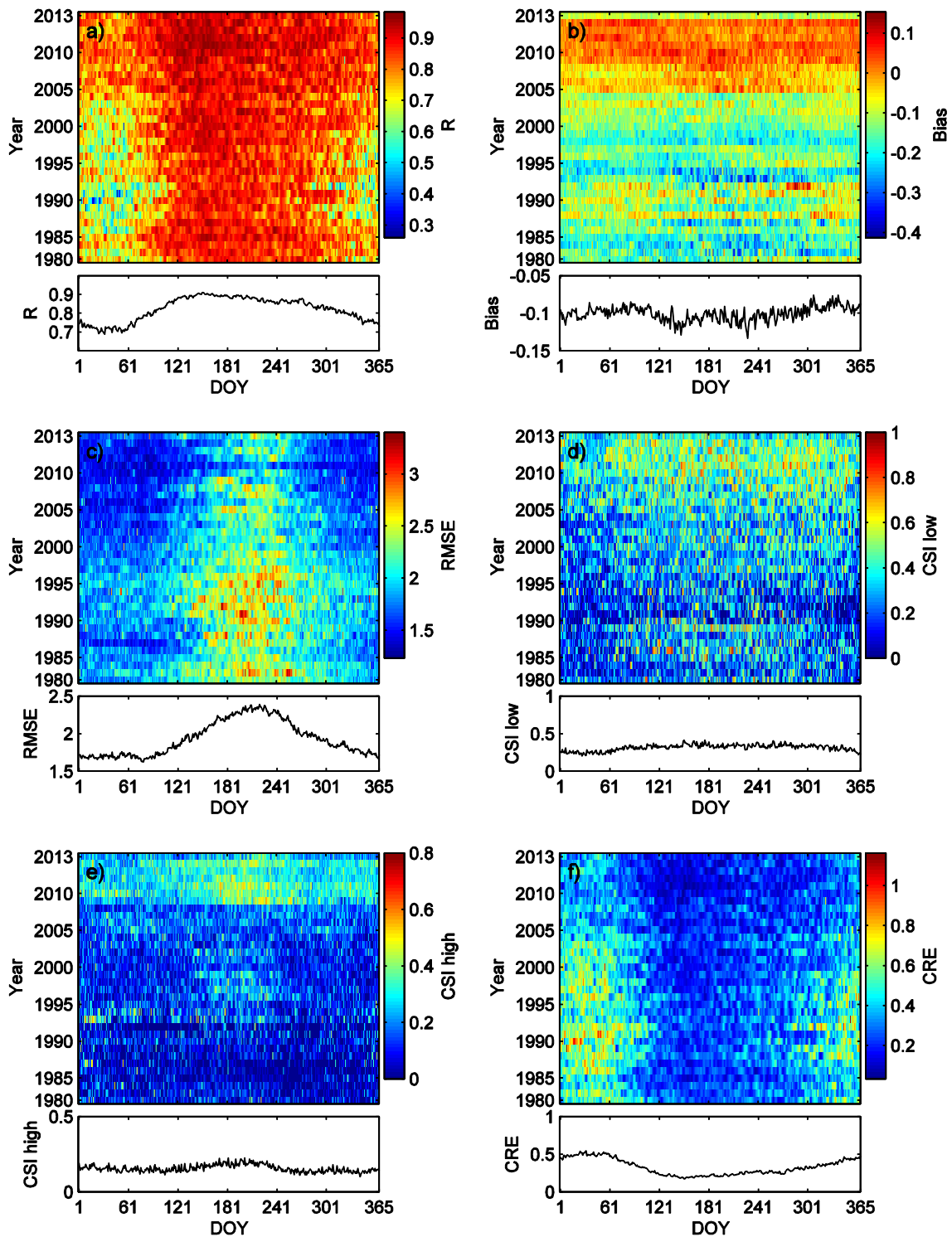


Figure S3. Daily skill scores of the relationship between observed and estimated T_{min} when interpolating using IDW.

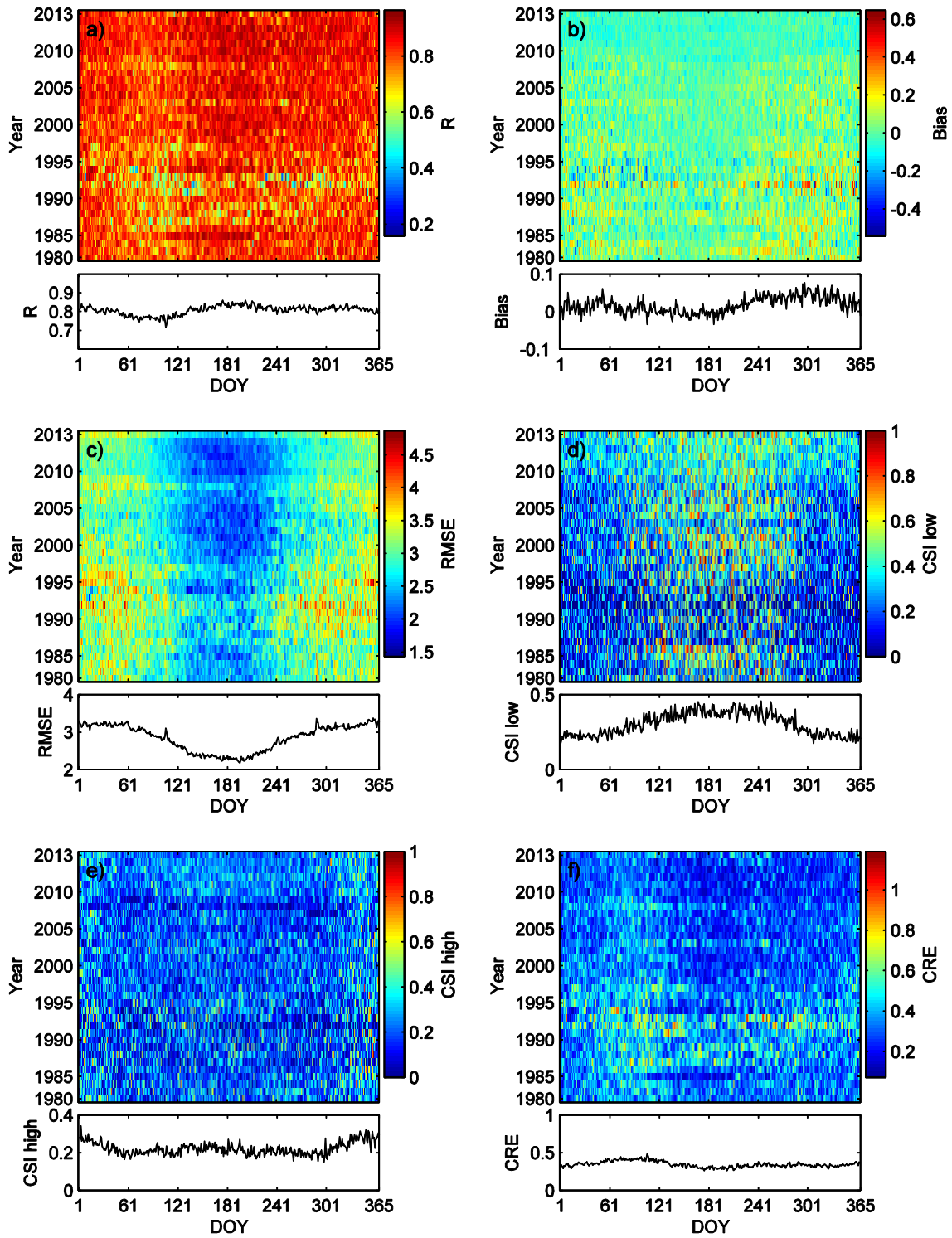


Figure S4. Daily skill scores of the relationship between observed and estimated R_s when interpolating using ADW.

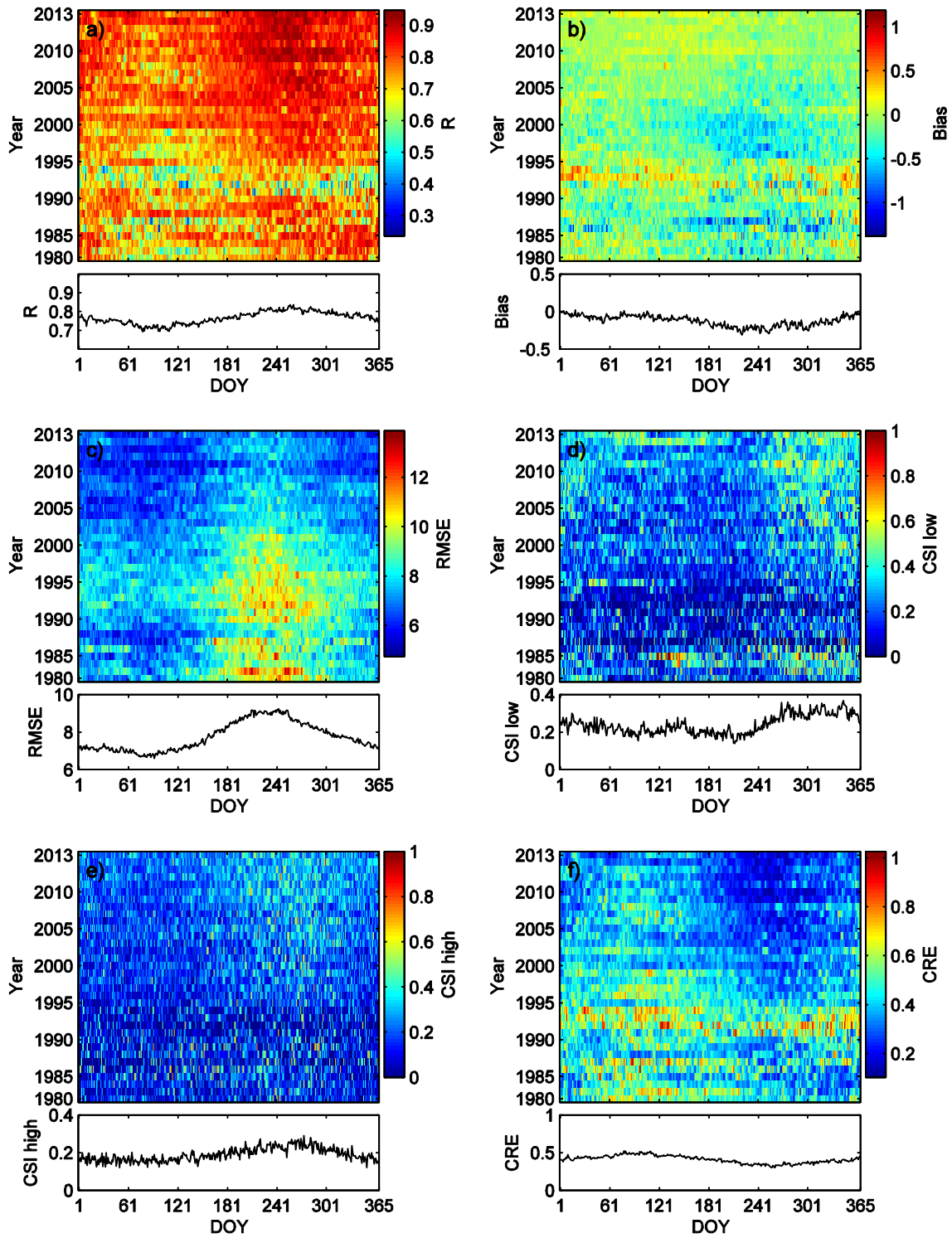


Figure S5. Daily skill scores of the relationship between observed and estimated RH when interpolating using ADW.

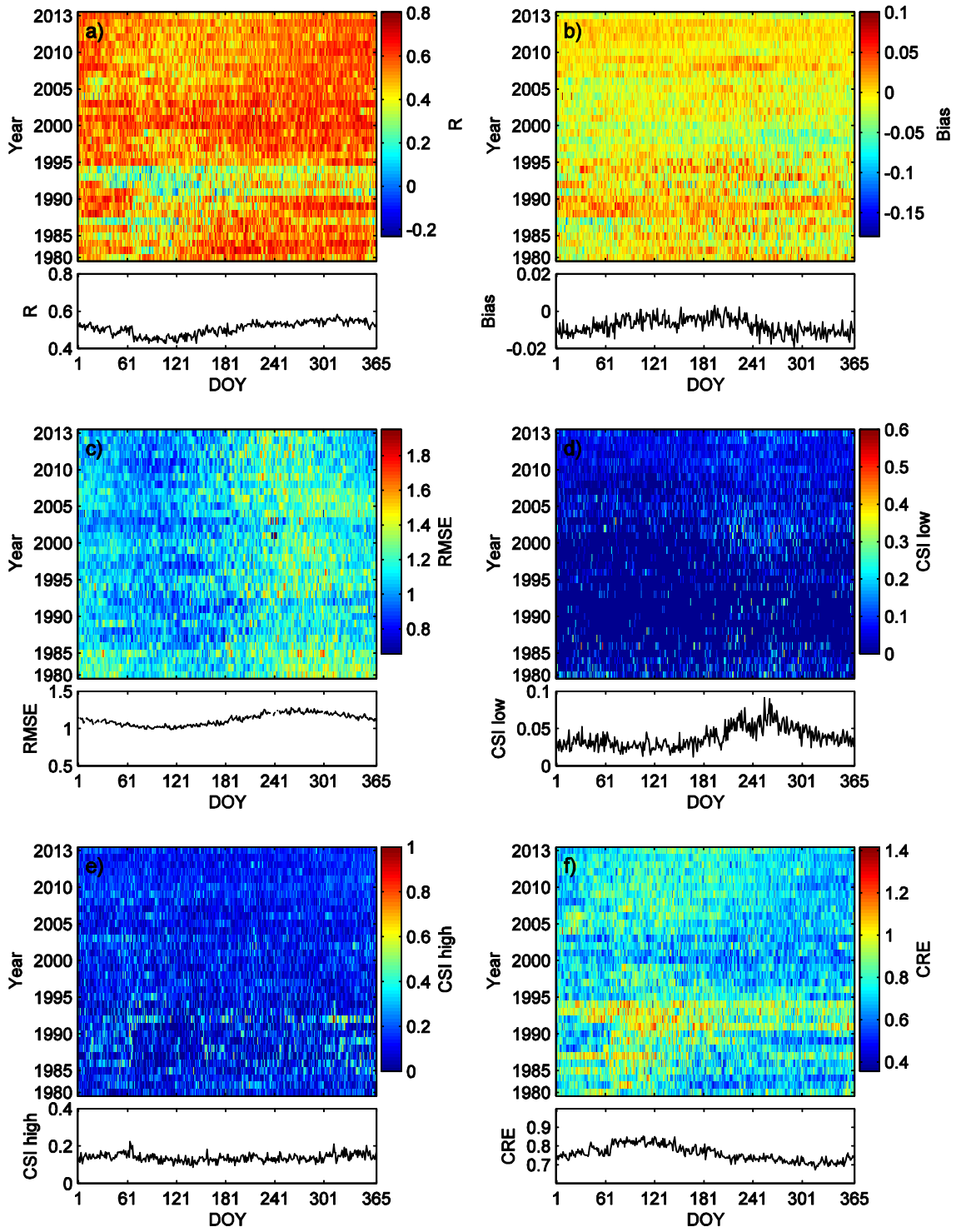


Figure S6. Daily skill scores of the relationship between observed and estimated u_2 when interpolating using ADW.

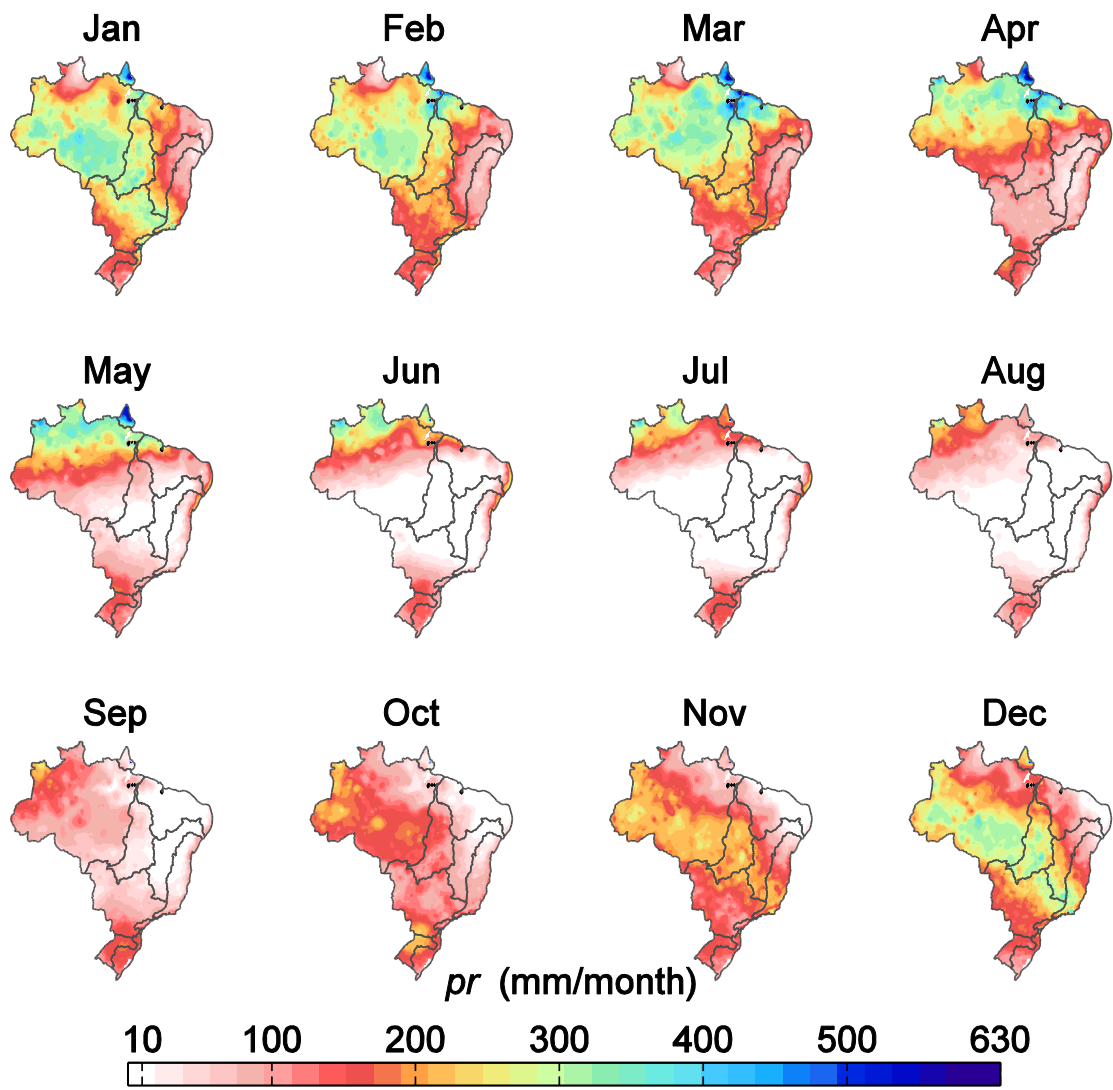


Figure S7. Maps generated from gridded data of monthly average of pr for the period 1980-2013.

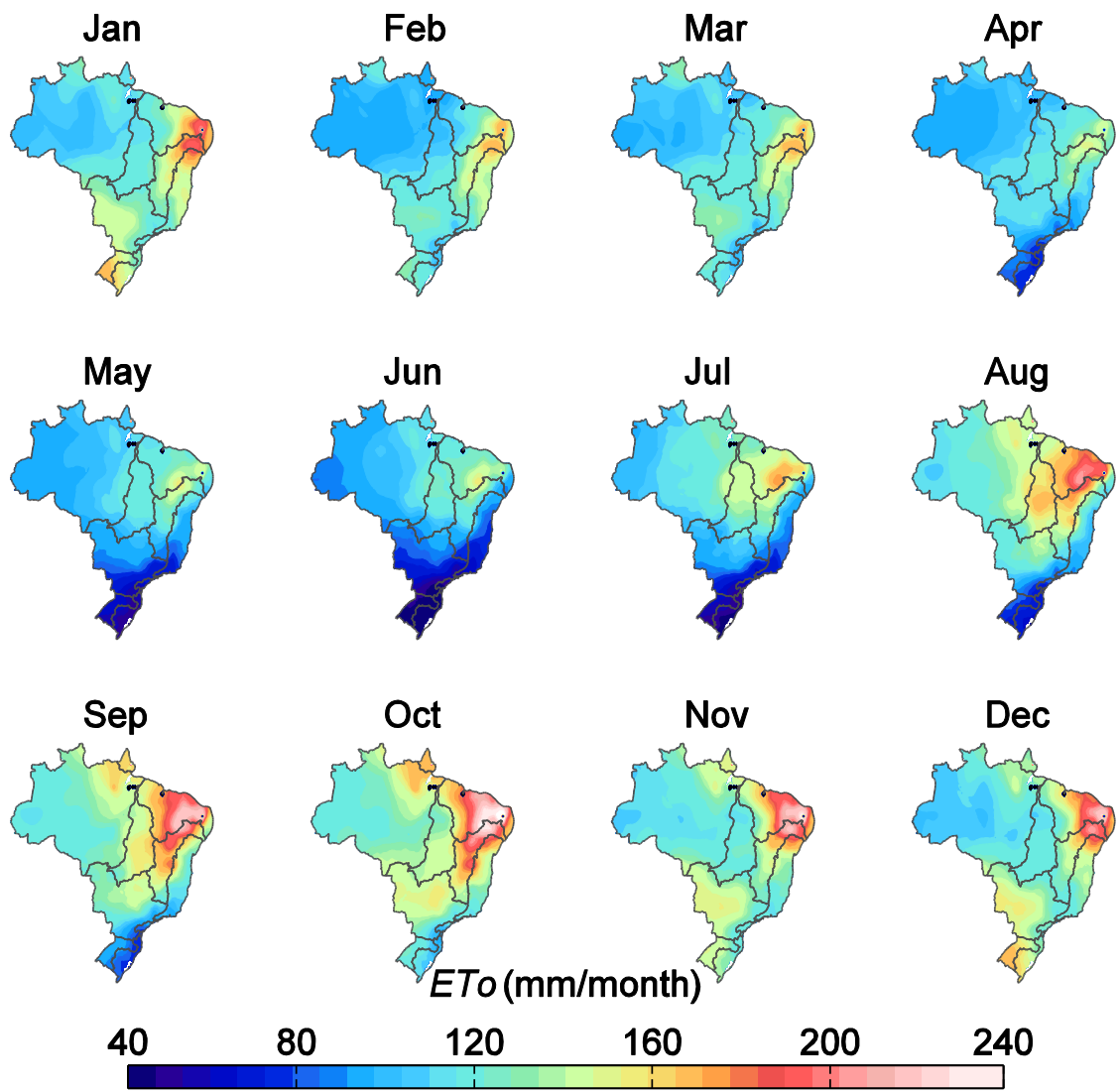


Figure S8. Maps generated from gridded data of monthly average of ET_0 for the period 1980-2013.

README for using netcdf (.nc) files of gridded meteorological data

Overview

This is a metafile description of how to download and use the gridded data sets generated for the paper.

Download location for gridded data set (URL):

<https://utexas.box.com/Xavier-et-al-IJOC-DATA>

The name of the files are formatted as:

VARNAME_TIMRESOLUTION_ABBREVIATION_COUNTRYNAME_VERSION_STARTEDDATE_END
DATE.nc

where:

VARNAME, the name of the variable (ETo, Tmax, Tmin, Rs, u2, RH and prec);

TIMRESOLUTION, daily or monthly;

ABBREVIATION: UT (University of Texas at Austin);

COUNTRYNAME: Brazil;

VERSION: v1;

STARTEDDATE: data file started on yyyyymmdd;

ENDDATE: data file finished on yyyyymmdd;

Example filename: ETo_daily_UT_Brazil_v1_20070101_20131231.nc

A second set of files end with the word "CONTROL". These files have only data about our controls. For each day or monthly grid (a matrix 162x168) we have two 'control' pieces of information that are in two other matrices:

- i) one with the count of stations into the cell and
- ii) the distance to the nearest station with data.

Example 'control' filename: ETo_daily_UT_Brazil_v1_20070101_20131231_Control.nc

In the monthly data files, the data of the interpolations and the control are in the same (single) monthly file.

List of files

Each variable was saved in:

- 1) four (4) files with daily data;
- 2) four (4) files with two daily controls: (1) distance of the nearest station with non-missing data and (2) count of station in the cell with available data;
- 3) and one (1) file with monthly data and its controls.

Thus, for each variable we have nine files. Because there are seven meteorological variables (ETo, Tmax, Tmin, Rs, u2, RH and prec), there are a total of 63 files. Below I show the name and the size of each file (MB). The TOTAL size of all files is 23.6 GB.

NAME	SIZE (MB)
'ETo_daily_UT_Brazil_v1_19800101_19891231.nc'	412.85
'ETo_daily_UT_Brazil_v1_19800101_19891231_Control.nc'	364.65
'ETo_daily_UT_Brazil_v1_19900101_19991231.nc'	412.85
'ETo_daily_UT_Brazil_v1_19900101_19991231_Control.nc'	364.47
'ETo_daily_UT_Brazil_v1_20000101_20061231.nc'	469.43
'ETo_daily_UT_Brazil_v1_20000101_20061231_Control.nc'	330.21
'ETo_daily_UT_Brazil_v1_20070101_20131231.nc'	469.43
'ETo_daily_UT_Brazil_v1_20070101_20131231_Control.nc'	330.21
'ETo_monthly_UT_Brazil_v1_198001_201312.nc'	313.78
'RH_daily_UT_Brazil_v1_19800101_19891231.nc'	412.85
'RH_daily_UT_Brazil_v1_19800101_19891231_Control.nc'	364.65
'RH_daily_UT_Brazil_v1_19900101_19991231.nc'	412.85
'RH_daily_UT_Brazil_v1_19900101_19991231_Control.nc'	364.47
'RH_daily_UT_Brazil_v1_20000101_20061231.nc'	469.43
'RH_daily_UT_Brazil_v1_20000101_20061231_Control.nc'	330.21
'RH_daily_UT_Brazil_v1_20070101_20131231.nc'	469.43
'RH_daily_UT_Brazil_v1_20070101_20131231_Control.nc'	330.21
'RH_monthly_UT_Brazil_v1_198001_201312.nc'	313.78
'Rs_daily_UT_Brazil_v1_19800101_19891231.nc'	412.85
'Rs_daily_UT_Brazil_v1_19800101_19891231_Control.nc'	364.65
'Rs_daily_UT_Brazil_v1_19900101_19991231.nc'	412.85
'Rs_daily_UT_Brazil_v1_19900101_19991231_Control.nc'	364.47
'Rs_daily_UT_Brazil_v1_20000101_20061231.nc'	469.43
'Rs_daily_UT_Brazil_v1_20000101_20061231_Control.nc'	330.21
'Rs_daily_UT_Brazil_v1_20070101_20131231.nc'	469.43

NAME	SIZE (MB)
'Rs_daily_UT_Brazil_v1_20070101_20131231_Control.nc'	330.21
'Rs_monthly_UT_Brazil_v1_198001_201312.nc'	313.78
'Tmax_daily_UT_Brazil_v1_19800101_19891231.nc'	412.85
'Tmax_daily_UT_Brazil_v1_19800101_19891231_Control.nc'	364.65
'Tmax_daily_UT_Brazil_v1_19900101_19991231.nc'	412.85
'Tmax_daily_UT_Brazil_v1_19900101_19991231_Control.nc'	364.47
'Tmax_daily_UT_Brazil_v1_20000101_20061231.nc'	469.43
'Tmax_daily_UT_Brazil_v1_20000101_20061231_Control.nc'	330.21
'Tmax_daily_UT_Brazil_v1_20070101_20131231.nc'	469.43
'Tmax_daily_UT_Brazil_v1_20070101_20131231_Control.nc'	330.21
'Tmax_monthly_UT_Brazil_v1_198001_201312.nc'	313.78
'Tmin_daily_UT_Brazil_v1_19800101_19891231.nc'	412.85
'Tmin_daily_UT_Brazil_v1_19800101_19891231_Control.nc'	364.65
'Tmin_daily_UT_Brazil_v1_19900101_19991231.nc'	412.85
'Tmin_daily_UT_Brazil_v1_19900101_19991231_Control.nc'	364.47
'Tmin_daily_UT_Brazil_v1_20000101_20061231.nc'	469.43
'Tmin_daily_UT_Brazil_v1_20000101_20061231_Control.nc'	330.21
'Tmin_daily_UT_Brazil_v1_20070101_20131231.nc'	469.43
'Tmin_daily_UT_Brazil_v1_20070101_20131231_Control.nc'	330.21
'Tmin_monthly_UT_Brazil_v1_198001_201312.nc'	313.78
'prec_daily_UT_Brazil_v1_19800101_19891231.nc'	412.85
'prec_daily_UT_Brazil_v1_19800101_19891231_Control.nc'	364.65
'prec_daily_UT_Brazil_v1_19900101_19991231.nc'	412.85
'prec_daily_UT_Brazil_v1_19900101_19991231_Control.nc'	364.47
'prec_daily_UT_Brazil_v1_20000101_20061231.nc'	469.43
'prec_daily_UT_Brazil_v1_20000101_20061231_Control.nc'	330.21
'prec_daily_UT_Brazil_v1_20070101_20131231.nc'	469.43
'prec_daily_UT_Brazil_v1_20070101_20131231_Control.nc'	330.21
'prec_monthly_UT_Brazil_v1_198001_201312.nc'	313.78
'u2_daily_UT_Brazil_v1_19800101_19891231.nc'	412.85
'u2_daily_UT_Brazil_v1_19800101_19891231_Control.nc'	364.65
'u2_daily_UT_Brazil_v1_19900101_19991231.nc'	412.85
'u2_daily_UT_Brazil_v1_19900101_19991231_Control.nc'	364.47
'u2_daily_UT_Brazil_v1_20000101_20061231.nc'	469.43
'u2_daily_UT_Brazil_v1_20000101_20061231_Control.nc'	330.21

NAME	SIZE (MB)
'u2_daily_UT_Brazil_v1_20070101_20131231.nc'	469.43
'u2_daily_UT_Brazil_v1_20070101_20131231_Control.nc'	330.21
'u2_monthly_UT_Brazil_v1_198001_201312.nc'	313.78

Sample code in MATLAB to read and plot data

After you run this code, you can open the following two variables:

i) 'infor'

you will see the metadata about "ETo_daily_UT_Brazil_v1_20070101_20131231.nc"

ii) 'inforcontrol', you will see the metadata about

you will see the metadata about "ETo_daily_UT_Brazil_v1_20070101_20131231_Control.nc"

The code plots the first day of the .nc files (01/01/2007, or January 1, 2007). The first figure plots reference evapotranspiration, ETo; the second figure plots the count of weather stations in each grid cell; and the last figure plots the distance to the nearest station with data.

```
+++++++ CODE ++++++
+++++++
clear all, close all
Name = 'ETo_daily_UT_Brazil_v1_20070101_20131231';
NameFileData = [Name '.nc'];
infor = ncinfo(NameFileData); % request information on the file
data = double(ncread(NameFileData,'ETO')); % get data
dates = double(ncread(NameFileData,'time')); % get dates
lon = ncread(NameFileData,'lon'); % get lon
lat = ncread(NameFileData,'lat'); % get lat
lon = repmat(lon',length(lat),1); % vector2grid
lat = repmat(lat,1,length(lon)); %vector2grid
data2show = 1;
daymonthyear = datestr(dates(data2show)); %get the date string of the
"data2show"

%% plot data
limits = [-75 -30 -35 10]; %limits to plot
% plot variable
figure,surf(lon,lat,data(:,:,data2show))
view(0,90), shading interp, colorbar, axis(limits)
title(['ETO on ' daymonthyear])

%% read control
namefilecontrol = [Name '_Control.nc']; %name control
inforcontrol = ncinfo(namefilecontrol); %information about control

% plot count
control_count = double(ncread(namefilecontrol,'count'));
figure,surf(lon,lat,control_count(:,:,data2show))
view(0,90), shading interp, colorbar, axis(limits)
title(['# of stations into cell on ' daymonthyear])

% plot nearest station
control_distance = double(ncread(namefilecontrol,'dist_nearest'));
figure, surf(lon,lat,control_distance(:,:,data2show))
title(['Distance (km) of the nearest station to cell ' daymonthyear])
view(0,90), shading interp, colorbar, axis(limits)

+++++++ CODE ++++++
```