

# User's Guide for China Meteorological Forcing Dataset

(Data version: 01.05.0016)

Kun Yang, Prof., Dr. ([yangk@itpcas.ac.cn](mailto:yangk@itpcas.ac.cn))

Data Assimilation and Modeling Center for Tibetan Multi-spheres (DAM)  
Institute of Tibetan Plateau Research (ITP), Chinese Academy of Sciences (CAS)

## 1. Introduction

The China Meteorological Forcing Dataset was produced by merging a variety of data sources. This dataset currently covers the period 1979-2010. Its spatial resolution is 0.1 deg. and its temporal resolution is 3-hr. The data set can be used for hydrological modeling, land surface modeling, land data assimilation, and other terrestrial modeling.

This data sources used to produce the forcing data include:

- (1) CMA (China Meteorological Administration) station data (wind, air temperature, relative humidity, sunshine duration, precipitation, and surface pressure) for the period of 1979-2010. The observed data together with the radiation data estimated from the observed sunshine duration are used as true values of the meteorological parameters.
- (2) TRMM satellite precipitation analysis data (3B42) for the period of 1998-2008 and GLDAS precipitation for the period of 1979-2010. The GLDAS precipitation data during 1998-2008 were only used to replace TRMM 3B42 data that are not available beyond 40°N.
- (3) GEWEX-SRB downward shortwave radiation for the period of 1983.07-2007.12, and GLDAS downward shortwave radiation data for the period of 1979.01-1983.06 and 2008.01-2010.12.
- (4) Princeton forcing data (wind, air temperature, relative humidity, surface pressure) for the period of 1979-2008.
- (5) GLDAS data (wind, air temperature, relative humidity, surface pressure) for the period of 2009-2010.

The data fusion technique is to be introduced in a manuscript in preparation.

## 2. Variables

Variables	Variable name	Unit	Physical meaning
Temperature	temp	K	Instantaneous near surface (2 m) air temperature.
Pressure	pres	Pa	Instantaneous near surface (2 m) air pressure.
Specific humidity	shum	kg kg <sup>-1</sup>	Instantaneous near surface (2 m) air specific humidity.
Wind speed	wind	m s <sup>-1</sup>	Instantaneous near surface (10 m) wind speed.
Downward shortwave radiation	srad	W m <sup>-2</sup>	3-hourly mean (from -1.5 hr to +1.5 hr) surface downward shortwave radiation.
Downward longwave radiation	lrاد	W m <sup>-2</sup>	3-hourly mean (from -1.5 hr to +1.5 hr) surface downward longwave radiation.
Precipitation rate	prec	mm hr <sup>-1</sup>	3-hourly mean (from -3.0 hr to 0.0 hr) precipitation rate.

## 3. Data archive and access

This dataset can be obtained at

<http://westdc.westgis.ac.cn/data/7a35329c-c53f-4267-aa07-e0037d913a21> .

The data archive is arranged as follows:

- (1) Each data file is named as *VVVV\_ITPCAS-CMFD\_Vvvvv\_CCCC\_YYYYMM.nc* (e.g. *temp\_ITPCAS-CMFD\_V0105\_B-01\_197901.nc*), where *VVVV* is the abbreviation of variable name, *vvvv* is the data version, *CCCC* is the category of dataset, *YYYY* is the four-digital year, *MM* is the two-digital month, that means one file contains data of only one variable for only one month, additionally, *.nc* is the file name suffix which indicates that the data were stored as NetCDF format;

- (2) Each file is compressed as .gz format in order to save disk space; users must uncompress these data files before using them. Recommended software to do this is gunzip (on UNIX/Linux) or WinRAR (on Windows);
- (3) Another data-compressing method is used in NetCDF files via packing 32-bits float values to 16-bits short integer. Two parameters, scale\_factor and add\_offset, are crucial to this technic. Most of high-level software (e.g. GrADS, Ferret) can automatically unpack data using these two parameters so users don't need to care them, but when users want to write programs themselves to read data from NetCDF files, the following formula should be used to unpack 16-bits short integer data to 32-bits float:

$$\text{unpacked\_data\_value} = \text{packed\_data\_value} * \text{scale\_factor} + \text{add\_offset}$$

The following table is scale\_factor and add\_offset of each variables.

Variables	Variable name	Scale_factor	Add_offset
Temperature	temp	0.01	273.15
Pressure	pres	2.00	63500.00
Specific humidity	shum	0.000001	0.025
Wind speed	wind	0.002	60.00
Downward shortwave radiation	srad	0.25	685.00
Downward longwave radiation	lrad	0.25	685.00
Precipitation rate	prec	0.0025	50.00

- (4) A simple FORTRAN program named nc\_read.f90 is provided along with the dataset as a sample to show how to read the NetCDF data files, but users might have to modify it to satisfy special needs. This program has been tested on Linux and Windows platform with netcdf-3.x libraries, and anyone who wants to compile this program must have NetCDF library installed first. For more information about NetCDF, please see <http://www.unidata.ucar.edu/software/netcdf>.
- (5) An mask file named mask.nc is also provided along with this dataset, which has same grid distribution as the data product. In this file, grid points locate at China's land give values of 1, while the other grid points give values of 0. Another file named talt.nc is identical to mask.nc, but for terrain altitude.

#### **4. Caution of usage (Known problems)**

The downward shortwave radiation is developed using more than one dataset as background field. As stated in section 1, we used Princeton LSM forcing data (downward shortwave radiation) for the period of 1979.01-1983.06 and 2008.01-2010.12, and GEWEX-SRB downward shortwave radiation during 1983.07-2007.12. This would have probably generated inconsistency in the data time series. Additionally, this problem also exists in precipitation data product and downward longwave radiation (because downward longwave radiation is calculated using the downward shortwave radiation and air temperature and humidity products).

#### **5. Data policy**

All data users are requested to follow these data release guidelines:

- (1) A data user is not allowed to transfer the data to a third party without the approval of the data provider.
- (2) Re-processing or reforming the data needs to inform the data provider, and the ITP contribution must be acknowledged and referenced in the re-processed data and relevant publication
- (3) The origin of the forcing data being used for publication of scientific results must be acknowledged and referenced in the publication.
- (4) All data users are strongly encouraged to establish direct contact with data providers for complete interpretation and analysis of data for publication purposes
- (5) Commercial use and exploitation of this data set is prohibited.

#### **6. Data Acknowledgement**

- (1) A formal citation of the data, whenever the ITP forcing data are being used for publication of scientific results, is as follows:

Jie He, Kun Yang, Jun Qin, Zuoqi Chen: in preparation.

Regarding the radiation estimation, you may refer to:

Kun Yang, Jie He, Wenjun Tang, Jun Qin, CCK Cheng, 2010: On downward shortwave and longwave radiations over high altitude regions: Observation and modeling in the Tibetan Plateau. *Agric. Forest. Meteorol.*, 150, 38-46.

- (2) To recognize the valuable role of the data provider, the data's origin must be acknowledged (“*The forcing dataset used in this study was developed by Data Assimilation and Modeling Center for Tibetan Multi-spheres, Institute of Tibetan Plateau Research, Chinese Academy of Sciences*”).
- (3) To facilitate the maintenance and improvements of the dataset, please send a copy of scientific publications to the data provider via email ([yangk@itpcas.ac.cn](mailto:yangk@itpcas.ac.cn)).

## **7. Disclaimer**

Although we have made great efforts to control input data and improve the forcing dataset, the accuracy and reliability of these data is subject to the availability of in situ data and is not guaranteed in any way. The provider disclaims liability of any kind whatsoever, including, without limitation, liability for quality, performance, and fitness for a particular purpose arising out of the use.