MAHASRI Pathfinder Gridded Precipitation Data of the Asian Region

1. IDENTIFICATION INFORMATION

Name	MAHASRI Pathfinder Gridded Precipitation Data of the Asian Region
Edition	data version 1.1 – document 1.0en
Abbreviation	MAHAPGP1.1_doc1.0en
DOI	doi:10.20783/DIAS.38 [https://doi.org/10.20783/DIAS.38]
Metadata Identifier	MAHAPGP20230727051153-DIAS20221121113753-en

2. CONTACT

2.1 CONTACT on DATASET

Name	Kooiti Masuda
Organization	Japan Agency for Marine-Earth Science and Technology
Address	3173-25 Showa-machi, Kanazawa-ku, Yokohama, Kanagawa, 236-0001, Japan
TEL	+81-45-778-5538
FAX	+81-45-778-5706
E-mail	masuda@jamstec.go.jp

2.2 CONTACT on PROJECT

2.2.1 Data Integration and Analysis System

Name	DIAS Office
Organization	Japan Agency for Marine-Earth Science and Technology
Address	3173-25, Showa-Cho, Kanazawa-ku, Yokohama-shi, Kanagawa, 236-0001, Japan
E-mail	dias-office@diasjp.net

3. DOCUMENT AUTHOR

Name	Kooiti Masuda
Organization	Japan Agency for Marine-Earth Science and Technology
E-mail	masuda@jamstec.go.jp

4. DATASET CREATOR

Name	Kooiti Masuda
Organization	Japan Agency for Marine-Earth Science and Technology

E-mail	masuda@jamstec.go.jp

5. DATE OF THIS DOCUMENT

2023-07-27

6. DATE OF DATASET

creation : 2008-03-31

revision : 2008-11-28

7. DATASET OVERVIEW

7.1 Abstract

This is a gridded data set of precipitation made from station precipitation included in the GAME Phase 2 Collected Data, by first summing up for each month at each station and then interpolating spatially. We have not taken account of topographical effects, nor systematic errors of raingauges. In our product, data are given on grid boxes over land only. Furthermore, those grid boxes which are far from any observing stations to be void of data. (The threshold distance is predetermined by subjective decision of the analyst.) The production of this data set is also a part of MAHASRI (Monsoon Asian Hydro-Atmosphere Scientific Research and Prediction Initiative) which has evolved from the legacy of GAME.

7.2 Topic Category(ISO19139)

climatologyMeteorologyAtmosphere

7.3 Temporal Extent

Begin Date	1997-01-01
End Date	2002-12-31
Temporal Characteristics	Monthly

7.4 Geographic Bounding Box

North bound latitude	50
West bound longitude	60
Eastbound longitude	150
South bound latitude	-5

7.5 Grid

Dimension Name	Dimension Size (slice number of the dimension)	Resolution Unit

column	90	1 (deg)
row	55	1 (deg)

7.6 Geographic Description

Asia

7.7 Keywords

7.7.1 Keywords on Dataset

Keyword Type	Keyword	Keyword thesaurus Name
theme	Atmosphere > Precipitation > Precipitation Amount	GCMD_science
theme	HYDROLOGY > Precipitation, ATMOSPHERIC PROCESSES > Precipitation	AGU
theme	Water, Climate	GEOSS
place	Asia	others
place	Asia > Eastern Asia > China, Asia > Eastern Asia > Japan, Asia > Eastern Asia > Republic of Korea, Asia > South Eastern Asia > Philippines, Asia > South Eastern Asia > Viet Nam, Asia > South Eastern Asia > Lao People's Democratic Republic, Asia > South Eastern Asia > Cambodia, Asia > South Eastern Asia > Thailand, Asia > South Eastern Asia > Malaysia, Asia > South Eastern Asia > Singapore, Asia > South Eastern Asia > Myanmar, Asia > Southern Asia > Bangladesh, Asia > Southern Asia > Nepal	Country

7.7.2 Keywords on Project

7.7.2.1 Data Integration and Analysis System

Keyword Type	Keyword	Keyword thesaurus Name
theme	DIAS > Data Integration and Analysis System	No_Dictionary

7.8 Online Resource

: http://www.jamstec.go.jp/e/medid/dias/kadai/mon/mon_ar.html

file download : https://data.diasjp.net/dl/storages/filelist/dataset:38

7.9 Data Environmental Information

[Processing environment]: programs written by the data set creator in Fortran 77 and Awk, Fortran 77 program "Spheremap" version 99.8a from the University of Delaware, and Generic Mapping Tools (GMT) version 4.3.1 (Wessel and Smith), on Intel PC running Linux OS. [Data format, file names, size]: The essentially equivalent data are stored in 4 different formats. (1) ASCII format with explicit longitude and latitude values, (2) ASCII raster-type format, (3) Simple binary raster-type format, and (4) NetCDF. The files are gathered by "tar" and compressed by "gzip". The name of the archive files

are as follows (where YYYY is replaced by 4-digit expression of year). (1) xyvgstnp_YYYY.tar.gz, (2) ascgstnp_YYYY.tar.gz, (3) bingstnp_YYYY.tar.gz, (4) grdgstnp_YYYY.tar.gz . The amount of data (in one of the formats) are approximately 0.4 Mbytes when compressed 4 Mbytes when uncompressed.

7.10 Distribution Information

name	version	specification
ASCII format with explicit coordinates	no informaation	ASCII. Each line contains longitude, latitude and value, space-separated. The "xyz" format of GMT software.
ASCII raster format	no information	ASCII. Each line contains 90 values (west to east) and there are 55 lines (north to south). Line separaters are single LFs.
simple binary raster format	BSQ	4-byte floating point number (little-endian). Each line 90 values from west to east and there are 55 lines from north to south). No separaters between lines.
NetCDF	version 3.6	Produced by program "xyz2grd" of GMT 4.3.1.

8. DATA PROCESSING

8.1 Data Processing (1)

8.1.1 General Explanation of the data producer's knowledge about the lineage of a dataset

This is a gridded data set of precipitation made from station precipitation included in the GAME Phase 2 Collected Data, by first summing up for each month at each station and then interpolating spatially.

As the algorithm for interpolation, "Spheremap" of Willmott et al. (1985) is used. This is originally the two-dimensional spatial interpolation algorithm by Shepard (1968) which is a kind of weighted average method, but modified to work with the spherical coordinates. The algorithm is also used by Global

Precipitation Climatology Centre (GPCC, located in the German Weather Bureau) in production of their gridded data set of precipitation.

We have not taken account of topographical effects, nor systematic errors of raingauges.

In our product, data values are given in grid boxes over land only. Furthermore, those grid boxes which are far from any observing stations to be void of data. (The threshold distance is predetermined by subjective decision of the analyst.)

The current Version 1.1 has basically the same specification as the Version 1 which was released previously, but it has incorporated a few modifications in the details of the production algorithm (mentioned in the section of "Data Remarks").

8.1.2 Data Source

Data Source Citation Name	Description of derived parameters and processing techniques used

GAME Phase 2 Collected Data [Takahashi K. and Agata Y. eds., 2006: GAME (GEWEX Asian Monsoon Experiment) Phase 2 Collected Data. GAME CD-ROM No. 12.]	GAME Phase 2 Collected Data (part of precipitation). GAME (GEWEX Asian Monsoon Experiment) was an international collaborative research project conducted from Japanese Fiscal Years 1996 to 2004 as a part of GEWEX (Global Energy and Water Cycle Experiment) of WCRP (World Climate Research Program). As an item of the Phase 2 (JFY 2002 - 2004) of GAME, meteorological data for the 6-year period from 1997 to 2002 were collected through collaboration of operational meteorological agencies of Asian countries. This collection of data includes daily values of precipitation observed by raingauges at approximately 1000 stations. The following documentation refers to this part. [Data format:] ASCII. [Organization of data:] Each file contains a record of a station in a year. Such files are gathered in archive files with respect to agencies that provided the data. [Coordinate system:] latitude and longitude of stations. [Spatial extent:] Countries in East, Southeast and South Asia (Bangladesh, Cambodia, China (mainland, Taiwan), Japan, Korea (south), Laos, Malaysia, Myanmar, Nepal, Philippines, Singapore, Thailand, Vietnam). [Number of stations:] approximately 1000 stations. [Time period:] 1 January 1997 to 31 December 2002 (6 years). [Time resolution:] 1 day. [Size of data:] approximately 17 MB as archive files. [Availability:] The data are available on-line as a CD-ROM image from http://game.suiri.tsukuba.ac.jp/cdroms/CD-ROM.html . More information about this data set is in the section "Routine (Station) Observations" of http://hydro.iis.u-tokyo.ac.jp/GAME-T/GAIN-T/.

9. DATA REMARKS

* We checked the source data and partially modified before entering them to the production. (The GAME archive has not been corrected yet.) Modifications which have effects to the product are as follows.

    The data at a station in the Philippines (PHI98223, Laoag City) of years 1997   2001 were wrong. Replaced.

    Information of latitude and longitude is added to 25 stations in Vietnam and 2 stations in Thailand, which were missing in the GAME archive.

* The following convention is assumed when monthly precipitation is calculated.

    If some of the values of daily precipitation are missing, the monthly sum is calculated without them. It means that the missing values were assumed as if they were zeroes.

    Trace precipitation is assumed to be the same as zero.

* The selection of stations which are used for interpolation into the grid is as follows:

5

Only such stations that the count of valid observational values at 2/3 or more of the overall time steps (the number of days in the case of daily data, and eight times the number of days in the case of 3-hourly data) are considered valid.

Such stations that are off the region of the grid to be produced by more than 2 degrees of latitude or longitude are excluded.

* The following masking is done after spatial interpolation to the grid.

Only those grid boxes which correspond to land are considered valid. The distinction between land and sea follows that of the GPCC (Global Precipitation Climatology Centre) data set ("Monitoring Product" as of 2004). That is, we put the missing-value flag to grid boxes which are considered as oceanic boxes by GPCC.

The value in such grid boxes that are more than a certain distance farther from any stations with valid values are considered as missing. The threshold distance is defined to be equal to the difference of 1.6 degrees latitude along a meridian.

10. DATA POLICY

10.1 Data Policy by the Data Provider

The content of this dataset should not be used for commercial purposes. The source should be properly acknowledged in any work obtained with this dataset. The creators of this dataset are not responsible for any loss or damage caused by using this dataset. [Reference Requirement] This dataset should be referenced as the following statement. "Kooiti Masuda et al. (2008) : MAHASRI Pathfinder Gridded Precipitation Data of the Asian Region. Data Integration and Analysis System in Japan Agency for Marine-Earth Science and Technology, Yokohama, Japan. "

10.2 Data Policy by the Project

10.2.1 Data Integration and Analysis System

If data provider does not have data policy, DIAS Terms of Service (https://diasjp.net/en/terms/) and DIAS Privacy Policy (https://diasjp.net/en/privacy/) apply.

If there is a conflict between DIAS Terms of Service and data provider's policy, the data provider's policy shall prevail.

11. LICENSE

12. DATA SOURCE ACKNOWLEDGEMENT

12.1 Acknowledge the Data Provider

12.2 Acknowledge the Project

12.2.1 Data Integration and Analysis System

If you plan to use this dataset for a conference presentation, paper, journal article, or report etc., please include acknowledgments referred to following examples. If the data provider describes examples of acknowledgments, include them as well.

13. REFERENCES

Shepard, D., 1968. A two-dimensional interpolation function for irregularly-spaced data. Proceedings, 1968 ACM National Conference, 517 524. (Reference of spatial interpolation algorithm)

Willmott, C.J., Rowe, C.M. and Philpot, W.D. 1985. Small-scale climate maps: a sensitivity analysis of some common assumptions associated with grid-point interpolation and contouring. The American Cartographer, 12, 5 16. (Reference of spatial interpolation algorithm)