Central Data Archive for STARDEX (D14 and D17)

2004-10-27, Version 0.4.1

Rationale

The central data archive provides easy access to observations, reanalysis, model and downscaled data for STARDEX partners and end users. It includes seasonal time series of the STARDEX indices and daily data for precipitation, minimum and maximum temperature, and circulation indices. The data will be stored in NetCDF files conforming to the Climate and Forecast (CF) metadata conventions and made available through a OPeNAP/DODS server.
This document describes the organization of the data archive and the conventions used to store the data in the NetCDF files.

Links:
NetCDF homepage: http://my.unidata.ucar.edu/content/software/netcdf/
CF conventions: http://www.cgd.ucar.edu/cms/eaton/cf-metadata/

1 Directory Hierarchy

`<data-type>/<region>/`

	values
`<data-type>`	data or indices
`<region>`	STARDEX region (alps, europe, germany, greece, iberia, italy, uk)

All directory names should consist of lowercase letters only.

2 File Naming Convention

A generic filename has the following form:
`<variable>.<region>.<type>.<institution>.<source>[.<exp>].nc`

component	description
`<variable>`	variable name
`<region>`	region and location of the stations
`<type>`	type of data: station (st), gridded (gp), region (reg), time series (ts)
`<institution>`	Specifies who produced the data.
`<source>`	The method of production of the data (e.g. downscaling method).
`<exp>`	Description of the experiment (e.g. driving GCM, scenario).

The file names should consist of lowercase letters only. Consult appendix A for further details and some example file names.

3 File Contents

In this section we define the metadata, the coordinate systems, and the variables for the different types of data to be stored. The types of data include: station observations, upscaled observations, downscaled data, model data on the one hand and preciptation, minimum and maximum temperature and the STARDEX indices on the other hand.

3.1 Metadata

The metadata should contain all relevant information about the dataset.

name	contents	mandatory
Conventions	CF-1.0	x
title	A short description of what is in the dataset.	x
source	The method of production of the data (e.g. downscaling method).	x
experiment	Description of the experiment: Driving GCM, which scenario.	
institution	Specifies where the data was produced.	x
comment	Additional information about the data or methods used to produce it.	
history	Provides an audit trail for modifications to the original data.	

The most important metadata related to downscaling are `source` and `experiment`. `source` should contain a concise description of the downscaling method, whereas `experiment` specifies the data source used for the predictors. Consult appendix B for further details.

3.2 Coordinate Systemes

The coordinate system should provide an accurate description of the location of the data in space and time. Three different coordinate systems are indroduced: one for station data, one for gridded data on a regular lat/lon grid, and one for gridded data on a rotated grid. The table below lists the variables required to describe this coordinate systems.

type of data	coordinate variables	auxillary coordinate variables	grid mapping variable
station data	station, time	lon, lat, hgt, station_id, station_name	-
lat/lon grid	lon, lat, hgt, time	-	-
rotated grid	rlon, rlat, hgt, time	lon2d, lat2d	rotated_pole

For ensembles an additional dimensions can be added to the above coordinate systems, e.g.

type of data	coordinate variables	auxillary coordinate variables	grid mapping variable
station data	station, time, member	lon, lat, hgt, station_id, station_name	-

3.3 Variables

Here we define the data variable names and their standard attributes. See appendix C for a complete variable list, including all the STARDEX indices.

variable	long_name	units
pre	precipitation	mm d-1
tmin	minimum temperature	degC
tmax	maximum temperature	degC

The long_name should be, despite its name, a short name for the variable which can be used for labeling plots.

A File names

A.1 Examples

filename	description
pre.al-fic.st.eth.obs.nc	FIC observations at the 10 common alpine stations
pre.al-fic.st.eth.dloci.ha3p-a2a.nc	downscaled FIC OBS from HADAM3P, Scenario A2a
pre.al-ch.st.eth.obs.nc	Swiss precipitation stations (high-resolution dataset)
pre.al-alra.gp.eth.obs.nc	Alpine Reanalysis Observations
pre.al-fic.gp.eth.alra-obs.nc	Observed gridpoint precipitation at the FIC stations
pre.al-nit.reg.adgb.obs.nc	Observed regional precipitation series for northern Italy
pre.gr.st.auth.uobs.nc	Upscaled observed greek station data
pre.gr-west.reg.auth.obs.nc	observed precipitation for the western Greek region
ct500.gr.ts.auth.ncep.nc	500 hPa circulation types for Greece from NCEP
ctthick.gr.ts.auth.ncep.nc	1000-500 hPa thickness circulation types for Greece from NCEP
pre.gr.st.auth.ct.ncep.nc	DS precip for Greek stations using NCEP circulation types
pind.al-fic.st.eth.obs.nc	precipitation indices for the 10 alpine FIC stations
pind.al-fic.st.eth.dloci.ha3p-a2a.nc	precipitation indices for downscaled station data
tind.al-fic.st.uibe.obs.nc	temperature indices for the 10 alpine FIC stations

A.2 Variable Name

abbreviation	description
pre	precipitation
tmin	minimum temperature
tmax	maximum temperature
pind	precipitation indices
tind	temperature indices
ct500	500 hPa circulation types
ctthick	1000-500 hPa thickness circulation types

A.3 Region/Location

abbreviation	description
al	Alps
eu	Europe
ge	Germany
gr	Greece
ib	Iberia
it	Italy
uk	United Kingdom
al-fic	the common FIC stations for the Alps
al-ch	the high-resolution Swiss station dataset
al-alra	the Alpine Reanalysis
it-er	Emilia-Romagna
gr-west	western greek region
gr-east	eastern greek region

A.4 Type of Data

abbreviation	description
st	station data
gp	gridpoint data
reg	regionally averaged data
ts	time series

A.5 Institution

abbreviation	description
adgb	University of Bologna, Italy
arpa	Servizio Meteorologico Regionale, ARPA-SMR Emilia-Romagna, Italy
auth	University of Thessaloniki, Greece
cnrs	Centre National de la Recherche Scientifique, France
dmi	Danish Meteorological Institute, Denmark
eth	Swiss Federal Institute of Technology, Switzerland
fic	Fundación para la Investigación del Clima, Spain
fts	Fachhochschule Stuttgart - Hochschule für Technik, Germany
kcl	King's College London, UK
unibe	University of Berne, Switzerland
ustutt	Institut für Wasserbau, Germany
uea	University of East Anglia, UK

A.6 Source

abbreviation	description	institution
obs	observations	
ncep	NCEP reanalysis	
ha3p-cta	HADAM3P, control, member a (see also appendix A.7)	
...		
anal2	two-step analogue method	FIC
ann	artificial neural network	AUTH
ann-gamma	ANN using hybrid Bernoulli/Gamma data misfit term	UEA
ann-gammamc	as ann-gamma but with Monte Carlo simulation	UEA
ann-irbf	individual radial basis fn ANN	KCL
ann-mlp	multi layer perceptron ANN	KCL
ann-rbf	radial basis fn ANN	KCL
ann-sse	ANN using sum-of-squares data misfit term	UEA
cca	canonical correlation analysis	ARPA, AUTH, UEA, UNIBE
cr	conditional resampling	KCL
cwg	conditional weather generator	DMI
dloci	dynamical local rescaling of GCM pre intensity	ETH
hyper4	random sampling within 4-dim hyperspace	ADGB
loc	local rescaling of GCM precipitation	ETH
loci	local rescaling of GCM precipitation intensity	ETH
mar	multivariate auto-regressive model	USTUTT
mlr	multiple linear regression	ARPA, USTUTT
mlr-ct	MLR using circulation types	AUTH
ppci	potential precipitation circulation index	CNRS

Note that the abbreviations for the downscaling methods do not include the institution, as the institution is a standard part of the file name.

A.7 Experiment

abbreviation	description
ha3p-cta	HADAM3P, Control, Member a
ha3p-a2b	HADAM3P, Scenario A2, Member b
ech4-a2a	ECHAM4, Scenario A2, Member a
ncep	NCEP Reanalysis
...	

B Metadata

B.1 Two Examples

name	example
Conventions	CF-1.0
title	STARDEX indices of precipitation at the 10 FIC stations in the Alps
source	FIC stations (OBS)
experiment	HADAM3P-A2a
institution	ETH
history	2004-06-16: generated by STARDEX indices software version 3.3.1

name	example
Conventions	CF-1.0
title	Precipitation on the Alpine grid
source	loci-eth (DS)
experiment	HADAM3P-CTa
institution	ETH

B.2 title

title
Precipitation at the 10 FIC stations in the Alps
STARDEX indices of precipitation at the 10 FIC stations in the Alps
…

B.3 source

source	description
FIC stations (OBS)	observations from the FIC stations
CHRM (MOD)	CHRM model data
NCEP (RA)	NCEP reanalysis data
dloci-eth (DS)	downscaled data, method dloci from ETH
cca-auth (DS)	downscaled data, using CCA developed by AUTH
ann-gamma-uea (DS)	downscaled data, using a neural network developed by UEA
…	

Note that the source metadata should include the type of data in parenthesis, i.e. observations (OBS), model data (MOD), reanalysis (RA), or downscaled (DS).
Note also that for downscaled data/indices the source metadata consists of two parts, the downscaling method and the institution that developed the method. See appendix A.6 for the abbreviations to be used for the downscaling method and appendix A.5 for the abbreviations to be used for the institution.

B.4 experiment

experiment	description
ERA15	ERA15 reanalysis
NCEP	NCEP reanalyis
HADAM3P-CTa	HADAM3P, control, member a
HADAM3P-A2b	HADAM3P, scenario A2, member b
ECHAM4-CTa	ECHAM4, control, member a
…	

B.5 institution

See Appendix A.5.

C List of NetCDF variables

C.1 Data Variables

variable	long_name	units
pre	precipitation	mm d-1
tmin	minimum temperature	degC
tmax	maximum temperature	degC
ct500	500 hPa circulation types	
...		

C.2 STARDEX Precipitation Indices

variable	long_name	units
pav	mean	mm d-1
pfre	frequency	1
pint	intensity	mm d-1
pqNN	NN% quantile	mm d-1
pfNN	fraction of total above NN% quantile	1
pn10mm	no events $>$ 10mm	1
pnlNN	no events $>$ clim NN% quantile	1
pfl90	fraction of total above clim 90% quantile	1
pxNd	maximum N-day total	mm
ppww	mean wet-day persistence	1
ppdd	mean dry-day persistence	1
ppcr	correlation of spell lengths	1
pxcdd	max no consecutive dry days	day
pxcwd	max no consecutive wet days	day
pwsav	mean wet-spell length	day
pwsmed	median wet-spell length	day
pwssdv	stddev wet-spell length	day
pdsav	mean dry-spell length	day
pdsmed	median dry-spell length	day
pdssdv	stddev dry-spell length	day

C.3 STARDEX Temperature Indices

variable	long_name	units
tnav	mean Tmin	degC
tnq10	Tmin 10% quantile	degC
tnq90	Tmin 90% quantile	degC
tnf10	fraction of days Tmin $<$ 10% quantile	1
tnf90	fraction of days Tmin $>$ 90% quantile	1
tnfd	number of frost days	1
tncwd	cold wave duration	day
tnfsl	frost season length	day
txav	mean Tmax	degC
txq10	Tmax 10% quantile	degC
txq90	Tmax 90% quantile	degC
txf10	fraction of days Tmax $<$ 10% quantile	1
txf90	fraction of days Tmax $>$ 90% quantile	1
txice	number of ice days	1
txhwd	heat wave duration	day
txhw90	90% quantile heat wave duration	day
trav	mean diurnal temperature range	degC
trq10	temperature range 10% quantile	degC
trq90	temperature range 90% quantile	degC
tiaetr	intra-annual extreme temperature range	degC
tgdd	growing degree days	degC day
tgsl	growing season length	day