EMULATE[1] Extremes Indices Software

-- User Information --

Alexander Walther[*], November 2004

[*] Earth sciences centre, Regional Climate Group, Gothenburg University.
contact: alex@gvc.gu.se

The EMULATE extremes indices can be computed using a self-programmed Fortran routine "*EMULATE-SEAS-EXTREMES.f90*". The routine was developed within EMULATE Work Package 4 by Alexander Walther (University of Wuerzburg, 2003/2004). This documentation will give an overview of the software, the included indices and the output matrices.

(1) Station files

There must be one separate file for each station containing up to four climate variables. If one or more variables are not available from any station the particular column has to be filled with missing values. The required standard format for the station files is as follows. The first line is the header containing details on the location:

No.	Parameter	Format:	MV-code
1	Wmo code	I7	-999999
2	Latitude	I5	-9999
3	Longitude	I6	-99999
4	Elevation	I5	-9999
5	Stationname	A20	
6	Country	A13	
7	Startyear of any variable	I4	
8	Endyear of any variable	I4	
9	GCM-Box	I7	-999999
10	Local ID	I9	-99999999

Table 1. Headerline format in station files

Starting in line two there is one data line for each available day:

No.	Parameter	Format:	MV-code
1	Year	I4	
2	Month	I2	
3	Day	I2	
4	Tmin	I5	-9999
5	Tmax	I5	-9999
6	Tmean	I5	-9999
7	Precipitation	I5	-9999

Table 2. Dataline format in station files

The terms in a data line are delimited by *SPACE*. In this context each station file can be read using the following formats:

1. header: *FORMAT(i7,i5,i6,i5,a20,a13,2i4,i7,i9)*
2. data lines: *FORMAT(i4,2(x,i2),4(x,i5))*

[1] acronym for EU-funded project: **E**uropean and North Atlantic daily to **MUL**tidecadal clim**ATE** variability (http://www.cru.uea.ac.uk/projects/emulate).

(2) Fortran routine

The routine calculates all the indices for a certain number of locations. The locations to be processed have to be listed in a separate *location file*. Two different files are predefined in the code by default.

1. *"EMULATE_LOC"* contains the pathnames to the station files in the regular EMULATE station database (e.g. *'+austria/graz-uni.dat'*)
2. *"ECAONLY_LOC"* contains the pathnames to locations in the ECA-ONLY directory (complete ECA database minus locations which are already included in the EMULATE database)

There must be one filename per line. The file to be used can be selected after running the routine. After selecting the *location file*, the locations will be processed in the same order they are listed in the file selected.

(3) Configuring the code

The compiled Fortran code offers no command line options. So any user specific configuration has to be done by changing the source code directly. There are two sections in the source code where a number of parameters can be set. The main intention for developing this software was the application within EMULATE. In this context the integration of a sophisticated user-interface was not the in focus during developing the code. Probably, the available version is not the final version of this software. Any kind of feedback could lead to improvements and the correction of (hopefully no) bugs. First of all the software has been tested with the default parameter settings which are needed for computing the EMULATE extremes indices.
Since the routine covers a relatively wide field of established climate extremes indices, an advanced configuration probably be avoided. The following table shows a list with the default parameters which can be adapted.

parameter	identifier	default value
climate variables	1: TMIN, 2: TMAX, 3: TMEAN, 4: PREC	
Processed variables	VARS	4
sufficient no. and length of filenames	FNAME('x')*'y'	no.=1000, length=200
sufficient range of years (start year)	YRfrom	1700
sufficient range of years (end year)	YRto	2100
MV-code in station files	MV	-9999
No. of Temperature Percentiles	PERCno(1:3)	6
No. of Precipitation Percentiles	PERCno(4)	3
Max no. of percentiles for any variable	maxPERC	6
Precipitation threshold for raindays	RAINTHR	1.0 mm
Temperature threshold for frostdays	FTHR	0 °C
Tmean threshold for Growing Season Length	GSLTHR	5.0 °C
Start of reference period	REFstart	1961/1/1
End of reference period	REFend	1990/12/31
Internal Percentile numbers	PRC(<VAR> , <no>)	1: 90th, 2: 95th, 3: 98th, 4: 2nd, 5: 5th, 6: 10th

Table 3. Default parameters

(4) Calculating Percentiles

A number of the calculated indices are percentiles itself, whereas others are percentile-based. All percentiles will be computed using the following formula:

$$PERC = \frac{(100-q)*x_k + q*x_{k+1}}{100} \; ; \; with \;\; k = \mathrm{int}\left(\frac{n*q}{100}\right)$$

Therefore the particular series has to be sorted in ascending order. q is the percentile (*e.g. 90*) and n the sample size. x_k and x_{k+1} are two neighbouring positions within the sorted series. This method is an empirical approach and can be found in BAHRENBERG[2)] (1990, p.41).

The extremes indices nos. 5. – 25. (in section 6) are percentiles calculated for the particular season (e.g. for JJA, sample-size 92 days).

For all percentile-based indices the particular reference percentile will be computed from a predefined base period (by default: 1961-1990). There are 365 reference percentiles (1 per day) computed using a 5-day window centred on the particular day of interest throughout the whole base period. So each reference percentile will be computed from a sample-size of *5days x 30years = 150 values*. All 29 February in leap years within the base period will be included in the reference for 28 February. In those cases the sample-size increases to 157 (7 leap years in base period).

(5) Avoiding inhomogeneities at start/end of base periods

The use of a base period for computing reference percentiles could cause inhomogeneities in index time series that count the number of days when a climatic variable exceeds a percentile-based threshold. According to ZHANG[3)] et al. (2004), such threshold-exceedance counts for a particular year in the base period may be biased if data from that same year is included in the sample from which the percentile is estimated. This may lead to inhomogeneities at the beginning and at the end of the base period. By the application of a special bootstrap technique presented in their paper, those breaks can be avoided.

The whole base-period (here the 30 years 1961-1990 is used) is divided into one "out-of-base" year (for which the exceedance is to be estimated) and the "almost-complete-base-period" (the remaining 29 years, which constitutes the sample from which percentiles are estimated). To obtain a complete 30-year base period, one year of the remaining 29 years will be duplicated. This (30-year)-period is then used to estimate percentiles. This procedure is repeated for each of the remaining 29 years – each time leading to a slightly different samples and hence slightly different percentiles. The final reference percentile for the "out-of-base" year is computed by averaging the 29 estimates.

[2)] BAHRENBERG, G., GIESE, E., NIPPER, J. (1990): Statistische Methoden in der Geographie. Teubner, Stuttgart, Germany, pp.233.

[3)] ZHANG et al.(2004) : Avoiding inhomogeneity in percentile-based indices of temperature extremes. J. Climate, submitted.

This procedure has to be repeated for every single day during the whole base-period. By doing so we obtain one individual reference percentile for each single day during the base-period. Threshold-exceedance counts for years outside the base period are made without this resampling technique, i.e. by using all data in the 30-year base period.

The bootstrap method introduced by ZHANG et al. (2004) is applied to all extremes indices dealing with threshold exceedances above reference percentiles computed from the predefined base-period (i.e. indices nos. 26.-55. and 61.-62.).

(6) Extremes Indices

1. Mean daily minimum temperature (*MEANTN*).

- mean of Tmin
- $MEANTN_j = \sum_{i=1}^{N} (TN_{ij}) \Big/ N$
- Let TN_{ij} be the daily minimum temperature for day i of period j.

2. Mean daily maximum temperature (*MEANTX*).

- mean of Tmax
- $MEANTX_j = \sum_{i=1}^{N} (TX_{ij}) \Big/ N$
- Let TX_{ij} be the daily maximum temperature for day i of period j.

3. Mean daily mean temperature (*MEANTG*).

- mean of Tmean
- $MEANTG_j = \sum_{i=1}^{N} (TG_{ij}) \Big/ N$
- Let TG_{ij} be the daily mean temperature for day i of period j.

4. Precipitation Total (*PRECTOT*).

- $PRECTOT_j = \sum_{i=1}^{N} (PREC_{ij})$
- Let $PREC_{ij}$ be the daily precipitation for day i of period j.

5. – 10. Tmin percentiles (*TNxP*).

5. 2nd percentile of daily minimum temperatures
6. 5th percentile of daily minimum temperatures
7. 10th percentile of daily minimum temperatures
8. 90th percentile of daily minimum temperatures
9. 95th percentile of daily minimum temperatures
10. 98th percentile of daily minimum temperatures

11. – 16. Tmax percentiles (*TXxP*).

11. 2nd percentile of daily maximum temperatures
12. 5th percentile of daily maximum temperatures
13. 10th percentile of daily maximum temperatures
14. 90th percentile of daily maximum temperatures
15. 95th percentile of daily maximum temperatures
16. 98th percentile of daily maximum temperatures

17. – 22. Tmean percentiles (*TGxP*).

17. 2nd percentile of daily mean temperatures
18. 5th percentile of daily mean temperatures
19. 10th percentile of daily mean temperatures
20. 90th percentile of daily mean temperatures
21. 95th percentile of daily mean temperatures
22. 98th percentile of daily mean temperatures

23. – 25. Precipitation percentiles (*PREC'x'P*).

23. 90^{th} percentile of daily precipitation amount (*PREC90P*)
24. 95^{th} percentile of daily precipitation amount (*PREC95P*)
25. 98^{th} percentile of daily precipitation amount (*PREC98P*)

26. – 31. No. of events exceeding/falling below Tmin percentiles (*TNxN*)

- using reference percentile from base period (by default: 1961-1990)

26. no. of events < 2^{nd} percentile of Tmin (*TN2N*)
27. no. of events < 5^{th} percentile of Tmin (*TN5N*)
28. no. of events < 10^{th} percentile of Tmin (*TN10N*)
29. no. of events > 90^{th} percentile of Tmin (*TN90N*)
30. no. of events > 95^{th} percentile of Tmin (*TN95N*)
31. no. of events > 98^{th} percentile of Tmin (*TN98N*)

32. – 37. No. of events exceeding/falling below Tmax percentiles (*TXxN*)

- using reference percentile from base period (by default: 1961-1990)

32. no. of events < 2^{nd} percentile of Tmax (*TX2N*)
33. no. of events < 5^{th} percentile of Tmax (*TX5N*)
34. no. of events < 10^{th} percentile of Tmax (*TX10N*)
35. no. of events > 90^{th} percentile of Tmax (*TX90N*)
36. no. of events > 95^{th} percentile of Tmax (*TX95N*)
37. no. of events > 98^{th} percentile of Tmax (*TX98N*)

38. – 43. No. of events exceeding/falling below Tmean percentiles (*TGxN*)

- using reference percentile from base period (by default: 1961-1990)

38. no. of events < 2^{nd} percentile of Tmean (*TG2N*)
39. no. of events < 5^{th} percentile of Tmean (*TG5N*)
40. no. of events < 10^{th} percentile of Tmean (*TG10N*)
41. no. of events > 90^{th} percentile of Tmean (*TG90N*)
42. no. of events > 95^{th} percentile of Tmean (*TG95N*)
43. no. of events > 98^{th} percentile of Tmean (*TG98N*)

44. – 52. Further precipitation extremes indices (*RxN, RxT, RxAM*)

- using reference percentile from base period (by default: 1961-1990)

44. no. of events > 90^{th} percentile of precipitation (*R90N*)
45. % of total rainfall from events > 90^{th} percentile (*R90T*)
46. precipitation total from events > 90^{th} percentile (*R90AM*)
47. no. of events > 95^{th} percentile of precipitation (*R95N)*
48. % of total rainfall from events > 95^{th} percentile (*R95T*)
49. precipitation total from events > 95^{th} percentile (*R95AM*)
50. no. of events > 98^{th} percentile of precipitation (*R98N*)
51. % of total rainfall from events > 98^{th} percentile (*R98T*)
52. precipitation total from events > 98^{th} percentile (*R98AM*)

53. – 55. Mean precip. from events > x^{th} long-term percentile (*SDII'x'P*)

- modified Simple Daily Intensity Index
- $SDIIxP_j = \sum_{i=1}^{E} (PREC_{ij}) \Big/ E$

- Let $PREC_{ij}$ be the daily precipitation amount on wet days (>= 1mm by default) of period *j* exceeding the long-term xth percentile. *E* is the total number of percentile exceedings during period *j*.
- Long-term percentiles are calculated for a predefined base period (by default: 1961-1990)

53. mean precipitation amount from events > 90th long-term percentile (*SDII90P*)

54. mean precipitation amount from events > 95th long-term percentile (*SDII95P*)

55. mean precipitation amount from events > 98th long-term percentile (*SDII98P*)

56. Simple Daily Intensity Index (*SDII*)

- $SDII_j = \sum_{i=1}^{N}(PREC_{ij}) \Big/ N$
- Let $PREC_{ij}$ be the daily precipitation amount on wet days (>= 1mm by default) of period *j*. *N* is the total number of wet days.

57. Greatest 5-day total rainfall (*R5d*)

- greatest precipitation total falling on 5 consecutive days during the period of interest

58. Greatest (1-day total) rainfall (*R1d*)

- greatest rainfall amount measured on a single day during the period of interest

59. No. of consecutive dry days (*CDD*)

- maximum number of consecutive days with precipitation amount falling below a predefined threshold (by default: <1mm)

60. Heat Wave Duration Index (*HWDI*)

- $Tx_{ij} > Tx_{inorm} + 5$
- Let Tx_{ij} be the daily maximum temperature at day *i* of period *j*. Let Tx_{inorm} be the calendar day mean calculated for a 5 day window centred on each calendar day during a specific base period (by default: 1961-1990).
- counted are the total number of at least 6 consecutive days exceeding $Tx_{inorm} + 5°C$

61. Warm Spell Duration Index (*WSDI90*)

- counted are the total number of at least 6 consecutive days with Tmax exceeding long-term 90th percentile (long-term: by default 1961-1990)

62. Cold Spell Duration Index (*CSDI10*)

- counted are the total number of at least 6 consecutive days with Tmin below long-term 10th percentile (long-term: by default 1961-1990)

63. No. of frost days during one year (*FD*)

- number of days with *TN* < *0°C* during a year

64. Growing season length (*GSL*)

- counted are the number of days between the first spell of the year of more than five days with *Tmean* > *threshold* and the first autumn-winter-spell of the year of more than five days with *Tmean* < *threshold* (threshold by default: 5°C)

(7) Missing value criterion

An overall missing value criterion will be applied. Up to one missing value per month is allowed. Otherwise a missing value will be written into the particular index column in the output matrix.

(8) Output matrices

Two output files will be created for every processed station file in the same directory where the particular station file is stored (see location file). Indices for the four natural seasons (DJF, MAM, JJA, SON) will be stored in *'<station>-3seas.ext'*. **NOTE: *The particular winter season (DJF) is assigned to the year in which the December falls*** (one example: "1900 12" means that the data for December 1900, January 1901 and February 1901 have been processed). The calculated indices for the 12 two month seasons (JF, FM, MA, AM, MJ, JJ, JA, AS, SO, ON, ND, DJ) will be stored in *'<station>-2seas.ext'*. The headerline contains the identifiers of the indices.

The columns in the output files are delimited by *SPACE*. The output matrices can be processed by using the following formats:

1. header: *FORMAT(a4,x,a2,64(x,a8))*
2. data: *FORMAT(i4,x,i2,25(x,f8.2),18(x,i8),3(x,i8,2(x,f8.2)),6(x,f8.2),6(x,i8))*

A complete list is shown in table 4.

PLEASE NOTE: Indices nos. 63 (FD) and 64 (GSL) have an annual resolution (indices nos. 1 – 62 seasonal). In this context these indices are listed as often as the particular year appears in the first column of the output matrix. E.g.: In the output matrix listing the indices for 3-month-seasons, index no. 63 (FD) appears 4 times for each year (yyyy:3, yyyy:6, yyyy:9, yyyy:12) etc.

(9) References

BAHRENBERG, G., GIESE, E., NIPPER, J. (1990): Statistische Methoden in der Geographie. Teubner, Stuttgart, Germany, pp.233.

ECA: European Climate Assessment (http://eca.knmi.nl)

EMULATE: Emulate EU project (http://www.cru.uea.ac.uk/projects/emulate)

STARDEX: Stardex EU project (http://www.cru.uea.ac.uk/projects/stardex)

ZHANG et al.(2004) : Avoiding inhomogeneity in percentile-based indices of temperature extremes. J. Climate, submitted.

No.	Identifier	Parameter	Unit	Format
		Year		I4
		month	1st month of season	I2
1	MEANTN	Mean Tmin	°C	F8.2
2	MEANTX	Mean Tmax	°C	F8.2
3	MEANTG	Mean Tmean	°C	F8.2
4	PRECTOT	Precipitation total	[mm]	F8.2
5	TN2P	Tmin 2nd percentile	°C	F8.2
6	TN5P	Tmin 5th percentile	°C	F8.2
7	TN10P	Tmin 10th percentile	°C	F8.2
8	TN90P	Tmin 90th percentile	°C	F8.2
9	TN95P	Tmin 95th percentile	°C	F8.2
10	TN98P	Tmin 98th percentile	°C	F8.2
11	TX2P	Tmax 2nd percentile	°C	F8.2
12	TX5P	Tmax 5th percentile	°C	F8.2
13	TX10P	Tmax 10th percentile	°C	F8.2
14	TX90P	Tmax 90th percentile	°C	F8.2
15	TX95P	Tmax 95th percentile	°C	F8.2
16	TX98P	Tmax 98th percentile	°C	F8.2
17	TG2P	Tmean 2nd percentile	°C	F8.2
18	TG5P	Tmean 5th percentile	°C	F8.2
19	TG10P	Tmean 10th percentile	°C	F8.2
20	TG90P	Tmean 90th percentile	°C	F8.2
21	TG95P	Tmean 95th percentile	°C	F8.2
22	TG98P	Tmean 98th percentile	°C	F8.2
23	PREC90P	Precip. 90th percentile	[mm]	F8.2
24	PREC95P	Precip. 95th percentile	[mm]	F8.2
25	PREC98P	Precip. 98th percentile	[mm]	F8.2
26	TN2N	exc. Tmin 2nd percentile	[days]	I8
27	TN5N	exc. Tmin 5th percentile	[days]	I8
28	TN10N	exc. Tmin 10th percentile	[days]	I8
29	TN90N	exc. Tmin 90th percentile	[days]	I8
30	TN95N	exc. Tmin 95th percentile	[days]	I8
31	TN98N	exc. Tmin 98th percentile	[days]	I8
32	TX2N	exc. Tmax 2nd percentile	[days]	I8
33	TX5N	exc. Tmax 5th percentile	[days]	I8
34	TX10N	exc. Tmax 10th percentile	[days]	I8
35	TX90N	exc. Tmax 90th percentile	[days]	I8
36	TX95N	exc. Tmax 95th percentile	[days]	I8
37	TX98N	exc. Tmax 98th percentile	[days]	I8
38	TG2N	exc. Tmean 2nd percentile	[days]	I8
39	TG5N	exc. Tmean 5th percentile	[days]	I8
40	TG10N	exc. Tmean 10th percentile	[days]	I8
41	TG90N	exc. Tmean 90th percentile	[days]	I8
42	TG95N	exc. Tmean 95th percentile	[days]	I8
43	TG98N	exc. Tmean 98th percentile	[days]	I8
44	R90N	exc. Prec 90th percentile	[days]	I8
45	R90T	Fraction above precip. perc.	[%]	F8.2
46	R90AM	Prec. total above percentile	[mm]	F8.2
47	R95N	exc. Prec 95th percentile	[days]	I8
48	R95T	Fraction above precip. perc.	[%]	F8.2
49	R95AM	Prec. total above percentile	[mm]	F8.2
50	R98N	exc. Prec 98th percentile	[days]	I8
51	R98T	Fraction above precip. perc.	[%]	F8.2
52	R98AM	Prec. total above percentile	[mm]	F8.2
53	SDII90P	Daily intensity above 90th perc.	[mm]	F8.2
54	SDII95P	Daily intensity above 95th perc.	[mm]	F8.2
55	SDII98P	Daily intensity above 98th perc.	[mm]	F8.2
56	SDII	Simple daily intensity index	[mm]	F8.2
57	R5d	Greatest 5-day total rainfall	[mm]	F8.2
58	R1d	Greatest (1-day) total rainfall	[mm]	F8.2
59	CDD	No. of consecutive dry days	[days]	I8
60	HWDI	Heat wave duration index	[days]	I8
61	WSDI90	Warm spell duration index	[days]	I8
62	CSDI10	Cold spell duration index	[days]	I8
63	FD	No. of Frost days	[days]	I8
64	GSL	Growing season length	[days]	I8

Table 4. Emulate extremes indices output files.