

Regional Frequency Analysis of Extreme Climate Events.

Theoretical part of REFRAN-CV

Course outline

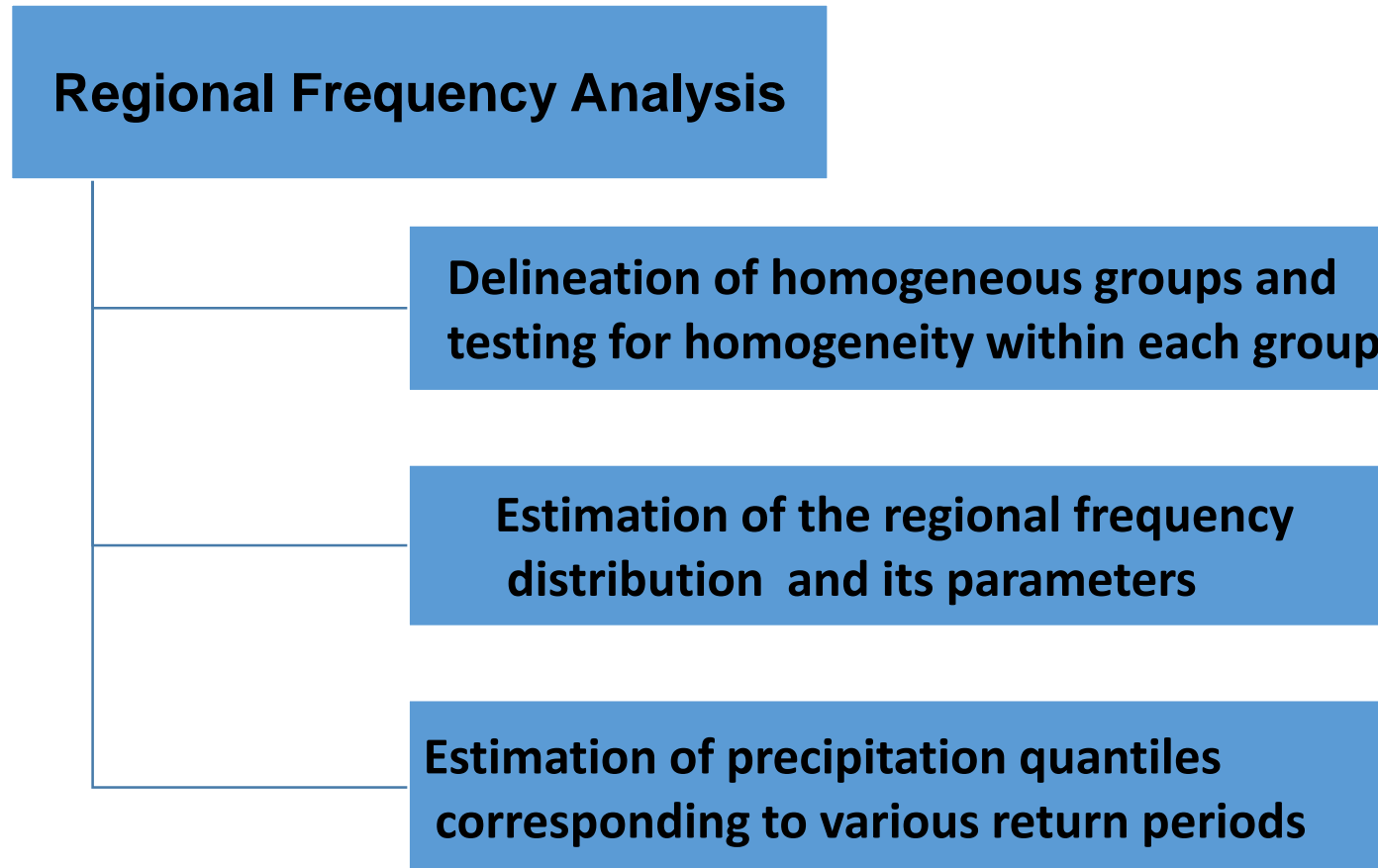
- Introduction
- L-moment statistics
- Identification of Homogeneous Regions
- L-moment ratio diagrams
- Example of the Regional Frequency Analysis with the NOAA Dataset across Africa
- Bibliography

Introduction

- The estimation of extreme climate events (precipitation/temperature) can be approached with the **regional frequency analysis** [Hosking, 1990].
- Regional frequency analysis **includes information from nearby stations exhibiting similar statistical behavior as at the site under consideration** in order to obtain more reliable estimates.

Introduction

- Regional frequency analysis is applied when **no local data are available** at a site of interest or the **data are insufficient** for a reliable estimation of the required return period.



L-moment statistics

Hosking (1990) has defined L-moments to summarize theoretical distribution and observed samples. **L-moments are the pillars of the regional frequency analysis.** Let $X_{(i|n)}$ be i^{th} largest obs. in sample of size n .

- L-Moment:
$$\lambda_r = r^{-1} \sum_{k=0}^{r-1} (-1)^k \binom{r-1}{k} E(X_{r-k:r})$$

A linear combination of order statistics

- Specifically, for the first 4 L-moments:

$$\lambda_1 = E(X_{1:n})$$

$$\lambda_2 = \frac{1}{2} E(X_{2:n} - X_{1:n})$$

$$\lambda_3 = \frac{1}{3} E(X_{3:n} - 2X_{2:n} + X_{1:n}) = \frac{1}{3} \{ E(X_{3:n} - X_{2:n}) - E(X_{2:n} + X_{1:n}) \}$$

$$\lambda_4 = \frac{1}{4} E(X_{4:n} - 3X_{3:n} + 3X_{2:n} - X_{1:n}) = \frac{1}{4} \{ \{ E[(X_{4:n} - X_{3:n}) + (X_{2:n} - X_{1:n})] \} - 2E(X_{3:n} - X_{2:n}) \}$$

L-moment statistics

Finally, the **L-moment ratios** are calculated as:

L-moment mean (L-mean):

$$\text{L-mean} = \tau_1 = \lambda_1$$

L-moment Coefficient of variation (L-CV):

$$\text{L-CV} = \tau_2 = \lambda_2/\lambda_1$$

L-moment coef. of skew (L-Skewness)

$$\text{L-Skewness} = \tau_3 = \lambda_3/\lambda_2$$

L-moment coef. of kurtosis (L-Kurtosis)

$$\text{L-Kurtosis} = \tau_4 = \lambda_4/\lambda_2$$

L-moment statistics

To wrap up...

- The **L-mean** is identical to the conventional statistical mean
- The **L-cv** measures a variable's dispersion, i.e. the expected difference between two random samples
- The **L-skewness** quantifies the asymmetry of the samples distribution
- The **L-kurtosis** measures whether the samples are peaked or flat relative to a normal distribution





L-moment statistics

Advantages of L-moment approach

- Less susceptible to the presence of outliers (Because L-moments avoid squaring and cubing the data)
- Perform better with small sample sizes

Identification of Homogeneous Regions

- A **homogeneous region** is considered as an **area** within which rescaled variables in different sites **have approximately the same probability distributions**.
- All sites can be described by one common probability distribution after the site data are rescaled by their at-site mean.

Identification of Homogeneous Regions

- Homogeneous regions (grouping of sites/gages) can be determined based on the similarity of the physical and/or meteorological characteristics of the sites. This can be done by performing **cluster analysis**.
- Hosking and Wallis (1997) proposed a **statistical test for testing the heterogeneity** of the proposed homogeneous regions.
- **L-moment** statistics can then be used to estimate the variability and skewness of the regional data and to test for heterogeneity as a basis for accepting or rejecting the proposed region formulation.

Identification of Homogeneous Regions

How to do it?

1. Calculate the weighted standard deviation of the at-site sample L -CVs,

$$V = \left\{ \frac{\sum_{i=1}^N n_i (\tau^{(i)} - \tau^R)^2}{\sum_{i=1}^N n_i} \right\}^{1/2}$$

2. Fit a four-parameter **kappa distribution** to the regional average L -moment ratios

$$\tau^R, \tau_3^R, \text{ and } \tau_4^R.$$

Identification of Homogeneous Regions

How to do it?

3. Simulate a large number N_{sim} of realizations of a region with N sites, each having this kappa distribution as its frequency distribution.
4. For each simulated region, calculate V .
5. From the simulations determine mean and standard deviation of the N_{sim} values of V . Call these μ_V and σ_V

Identification of Homogeneous Regions

H : is the discrepancy between L-Moments of observed samples and L-Moments of simulated samples assessed in a series of N_{sim} Monte Carlo simulation :

$$H = \frac{V - \mu_v}{\sigma_v}$$

$H \geq 2$: Region is **heterogeneous**.

$1 \leq H < 2$: Region is possibly **heterogeneous**.

$H < 1$: Region is acceptably **homogeneous**.

L-moment ratio diagrams

- Once homogeneous regions are defined, a single probability distribution is applied to all sites within a homogeneous region. Thus, it is necessary to choose a best-fit distribution from a set of candidate distributions.
- The goodness-of-fit can be judged by how well the L -skewness and L -kurtosis of the fitted distribution match the regional average L -skewness and L -kurtosis of the observed data.

L-moment ratio diagrams

TABLE 18.1.2 Values of L Moments and Relationships for the Inverse of the cdf for Several Distributions

Distribution and inverse cdf	L moments
Uniform: $x = \alpha + (\beta - \alpha)F$	$\lambda_1 = \frac{\beta + \alpha}{2}$ $\lambda_2 = \frac{\beta - \alpha}{6}$ $\tau_3 = \tau_4 = 0$
Exponential:* $x = \xi - \frac{\ln [1 - F]}{\beta}$	$\lambda_1 = \xi + \frac{1}{\beta}$ $\lambda_2 = \frac{1}{2\beta}$ $\tau_3 = \frac{1}{3}$ $\tau_4 = \frac{1}{6}$
Normal† $x = \mu + \sigma\Phi^{-1}[F]$	$\lambda_1 = \mu$ $\lambda_2 = \frac{\sigma}{\sqrt{\pi}}$ $\tau_3 = 0$ $\tau_4 = 0.1226$
Gumbel: $x = \xi - \alpha \ln [-\ln F]$	$\lambda_1 = \xi + 0.5772 \alpha$ $\lambda_2 = \alpha \ln 2$ $\tau_3 = 0.1699$ $\tau_4 = 0.1504$
GEV: $x = \xi + \frac{\alpha}{\kappa} (1 - [-\ln F]^\kappa)$	$\lambda_1 = \xi + \frac{\alpha}{\kappa} (1 - \Gamma[1 + \kappa])$ $\lambda_2 = \frac{\alpha}{\kappa} (1 - 2^{-\kappa}) \Gamma(1 + \kappa)$ $\tau_3 = \left\{ \frac{2(1 - 3^{-\kappa})}{(1 - 2^{-\kappa})} - 3 \right\}$ $\tau_4 = \frac{1 - 5(4^{-\kappa}) + 10(3^{-\kappa}) - 6(2^{-\kappa})}{1 - 2^{-\kappa}}$
Generalized Pareto: $x = \xi + \frac{\alpha}{\kappa} (1 - [1 - F]^\kappa)$	$\lambda_1 = \xi + \frac{\alpha}{1 + \kappa}$ $\lambda_2 = \frac{\alpha}{(1 + \kappa)(2 + \kappa)}$ $\tau_3 = \frac{1 - \kappa}{3 + \kappa}$ $\tau_4 = \frac{(1 - \kappa)(2 - \kappa)}{(3 + \kappa)(4 + \kappa)}$
Lognormal	See Eqs. (18.2.12), (18.2.13)
Gamma	See Eqs. (18.2.30), (18.2.31)

L-moments & Distribution Parameter Relations

From “Frequency Analysis of Extreme Events,” Chapter 8 in Handbook of Hydrology, By Stedinger, Vogel, and Foufoula-Georgiou, McGraw-Hill Book Company, New York, 1993

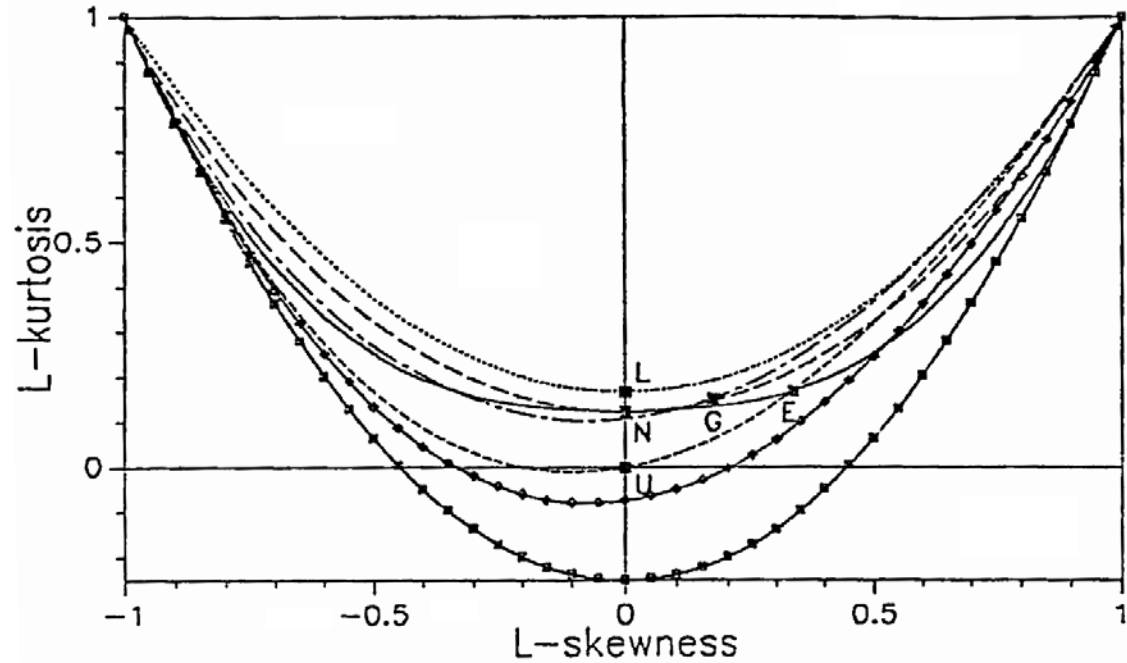
* Alternative parameterization consistent with that for Pareto and GEV distributions is:
 $x = \xi - \alpha \ln[1 - F]$ yielding $\lambda_1 = \xi + \alpha$; $\lambda_2 = \alpha/2$.

† Φ^{-1} denotes the inverse of the standard normal distribution (see Sec. 18.2.1).

Note: F denotes cdf $F_X(x)$.

Source: Adapted from Ref. 72, with corrections.

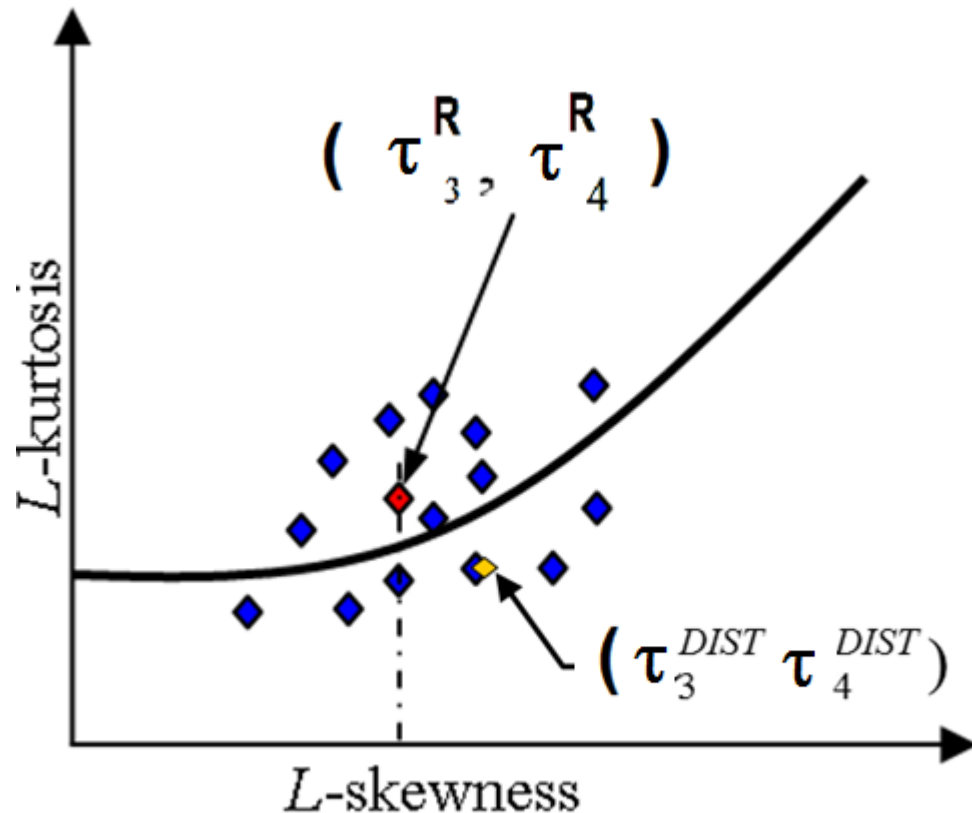
L-moment ratio diagrams



E	exponential	—	generalized logistic	—●—	lower bound for Wakeby
G	Gumbel	- - -	generalized extreme-value	—■—	lower bound for all distributions
L	logistic	· · ·	generalized Pareto		
N	Normal	- · -	lognormal		
U	uniform	—	gamma		

L-moment ratio diagrams

- The GEV distribution fitted by the method of L-moments has L -skewness equal to the regional average L -skewness.



- We thus judge the quality of fit by the difference between the L -kurtosis of the fitted GEV distribution and the regional average L -kurtosis.

Goodness-of-fit test

- Given a set of candidate three-parameter distributions (Pearson type III, GEV, lognormal, generalized Pareto, etc.). We need to fit each distribution to the regional average L -moment ratios .

$$Z^{\text{DIST}} = (\tau_4^{\text{DIST}} - \tau_4 + \beta_4) / \sigma_4$$

“Dist” refers to the candidate distribution,

τ_4^{DIST} is the average L-Kurtosis value computed from simulation for a fitted distribution.

τ_4 is the average L-Kurtosis value computed from the data of a given region,

β_4 is the bias of the regional average sample L-Kurtosis,

σ_v is standard deviation.

A given distribution is declared a good fit if $|Z^{\text{Dist}}| \leq 1.64$



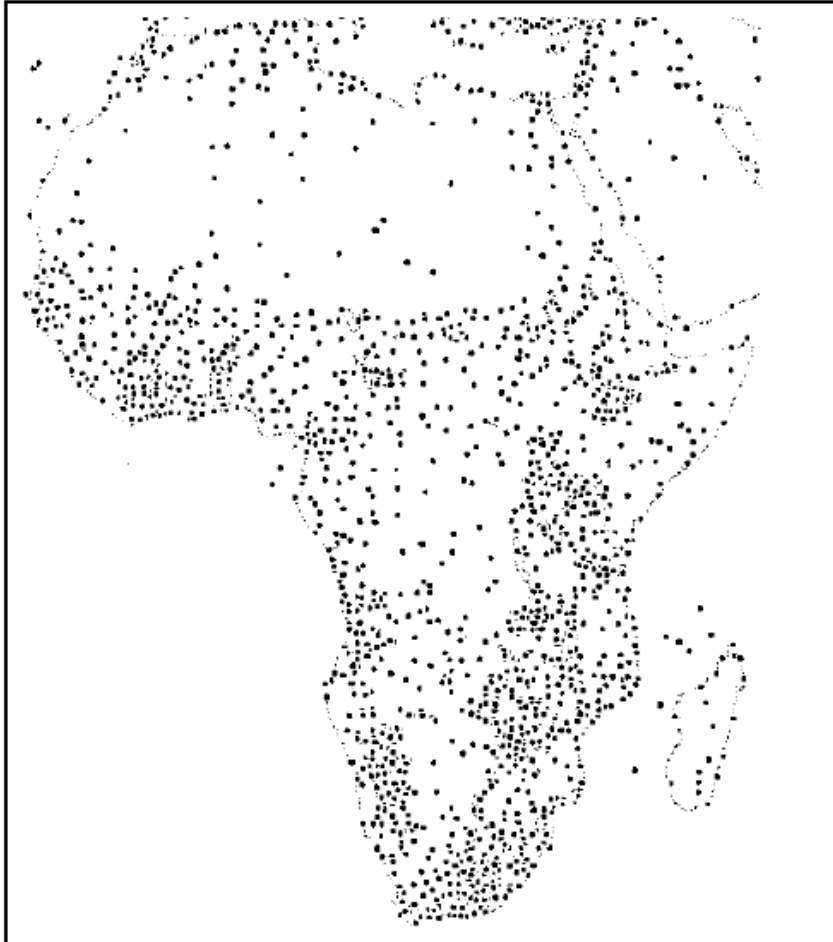
Example of the regional frequency analysis with the NOAA Dataset

- 1.L-moments calculation at each pixel
- 2.Identification parameters of the precipitation probability distribution function (Annual precipitation)
- 3.Evaluation of the most suitable distribution function for each pixel (Annual precipitation)
- 4.Identification of best distribution: L-skewness and L-kurtosis were used to perform a goodness-of-fit assessment
- 5.Generation of precipitation maps associated with different return periods



Data coverage

NOAA's PREC/L



- NOAA's Precipitation over land:
- Monthly records from 1951-1991 (Gauge stations with 10-yr-or-longer recording periods)
- Spatial resolution: 0.5 degrees
- Reconstruction obtained using thousands stations from the Global Historical Climatology Network (GHCN) and the Climate Anomaly Monitoring System (CAMS)

(Adapted from Chen et al., 2002)

<https://gis.ncdc.noaa.gov/maps/ncei/summaries/monthly>

NOAA Climate Data Portal



NOAA Climate Data Portal

The screenshot displays the NOAA Climate Data Portal interface. On the left, a sidebar titled 'Monthly Summaries Map' contains a 'Layers' section with 'Monthly Summary' checked. Below this, a 'Date' selector is set to 'February 2016'. Under 'Observations', 'Temperature Average' is selected. Under 'Selected Observation Units', '°F' is chosen. An 'Options' section includes a checked box for 'Display stations with zero amounts'. An 'UPDATE MAP' button is located at the bottom of the sidebar. The main map area shows a map of Europe and Africa with numerous climate stations marked by colored dots. A large blue arrow points from the map towards the right-hand content area. The right-hand content area features the NOAA logo and navigation links. Below the logo is a 'Request Submitted' confirmation message: 'Your request was successfully submitted. An email with a link to the requested data should be sent shortly.' This is followed by an 'ORDER INFORMATION' table, a 'PERIOD OF REQUEST' table, and a 'REQUESTED DATA' table listing station names and IDs. On the far right, there are sections for 'Order questions' and 'Help'.

Monthly Summaries Map

Layers Results

Monthly Summary

Date

February 2016

Observations

- Temperature Average
- Temperature Minimum
- Temperature Maximum
- Precipitation
- Snowfall
- Snow Depth

Selected Observation Units

- °F
- °C

Options

- Display stations with zero amounts

UPDATE MAP

NOAA NATIONAL CENTERS FOR ENVIRONMENTAL INFORMATION

Home Climate Information Data Access Customer Support Contact About

Home > Climate Data Online > Order Complete

Request Submitted

Step 1: Choose Options Step 2: Review Order Step 3: Order Complete

Your request was successfully submitted. An email with a link to the requested data should be sent shortly.

Print Receipt

ORDER INFORMATION	
Order Number	705939
Order Format	Custom Monthly Summaries of GHCN-Daily CSV
Email Address	guido.ceccherini@gmail.com
Date Submitted	2016-3-17 11:34:39 EST
Check Order Status	CHECK ORDER STATUS

PERIOD OF REQUEST	
Start Date	1990 -01-01
End Date	2016 -02-01

REQUESTED DATA	
Stations	CRAIOVA, RO (GHCND:ROE00108893) VARNA, BU (GHCND:BUM00015552) NIS, RB (GHCND:RIE00100814) SOFIA, BU (GHCND:BUM00015614) SKOPJE ZAJCEV RID, MK (GHCND:MKM00013588) BELGRADE OBSERVATORY, RB (GHCND:RIE00100818) BUCURESTI BANEASA, RO (GHCND:ROE00108889) VIDIN, BU (GHCND:BUM00015502)

Order questions

How will my data be delivered?
Your data request will have a confirmation delivered via email with links to access the files via FTP.

When will my data be delivered?
Most orders only take a few minutes to process but larger orders take longer and high volumes of traffic may cause delays.

What if my order doesn't complete?
1. Check your spam folder and ensure that no-reply@noaa.gov is on your approved list
2. Check order status online
3. Contact customer support

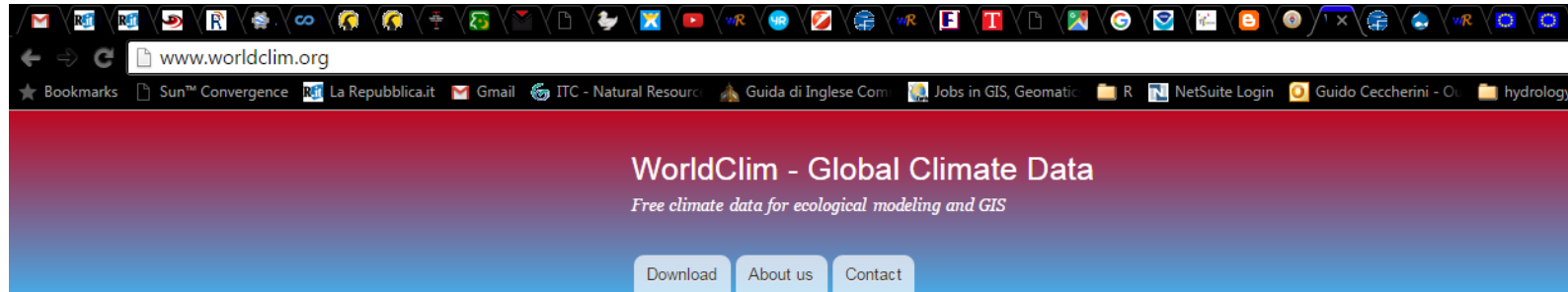
Help

Have questions about the data? Need some assistance? Use the links below to quickly find the answers you need.

[Online help](#)
[Check request status](#)
[Request assistance](#)

Temperature Avg: February 2016

WorldClim portal



Here it is possible to download the Mean Annual Precipitation (MAP) map required by REFRAN-CV

WorldClim

WorldClim is a set of global climate layers (climate grids) with a spatial resolution of about 1 square kilometer. The data can be used for mapping and spatial modeling in a GIS or with other computer programs. If you are not familiar with such programs, you can try [DIVA-GIS](#) or the *R* raster package.

The current version is Version 1.4 (release 3). Please [write us](#) if you find any problems.

---> [Download data](#)

Information about the [methods](#) used to generate the climate layers, and the [units and formats](#) of the data. You can find more info in the **preferred citation**:

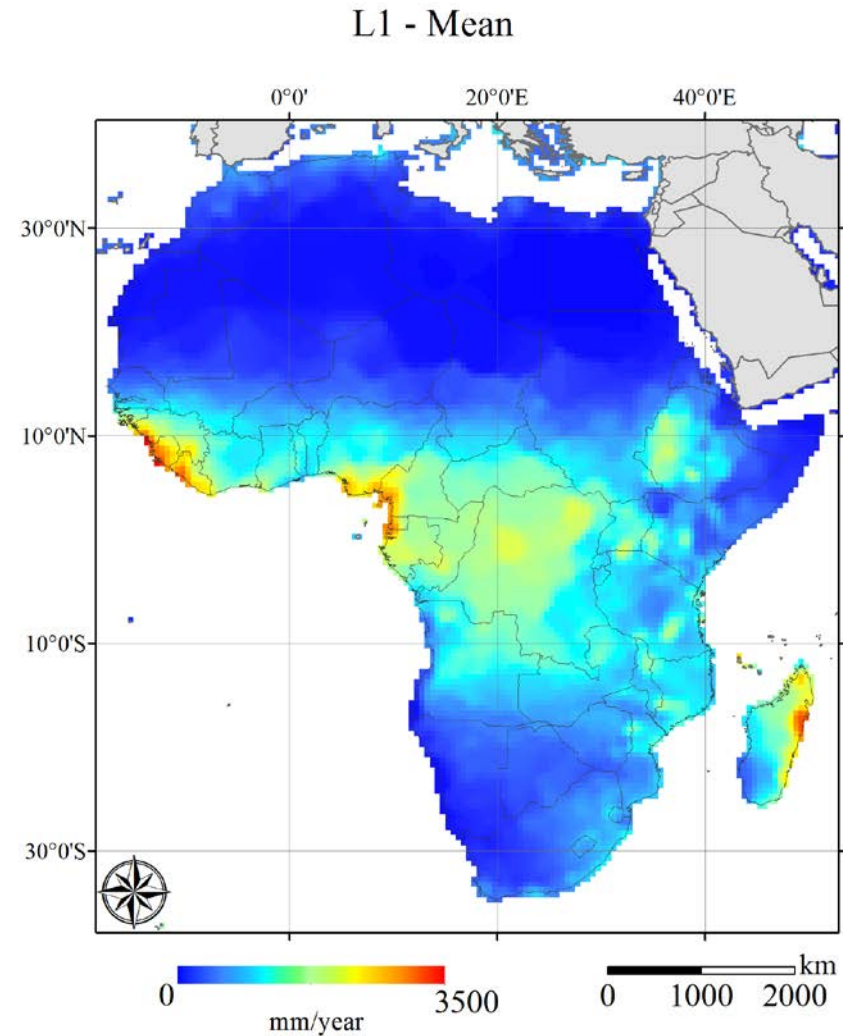
Hijmans, R.J., S.E. Cameron, J.L. Parra, P.G. Jones and A. Jarvis, 2005. Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology* 25: 1965-1978.

Frequently asked [question](#) and some 'known issues'.

This dataset is freely available for academic and other non-commercial use. Redistribution, or commercial use is not allowed without prior permission. You are free to create maps and use the data in other ways for publication in academic journals, books, reports, etc.

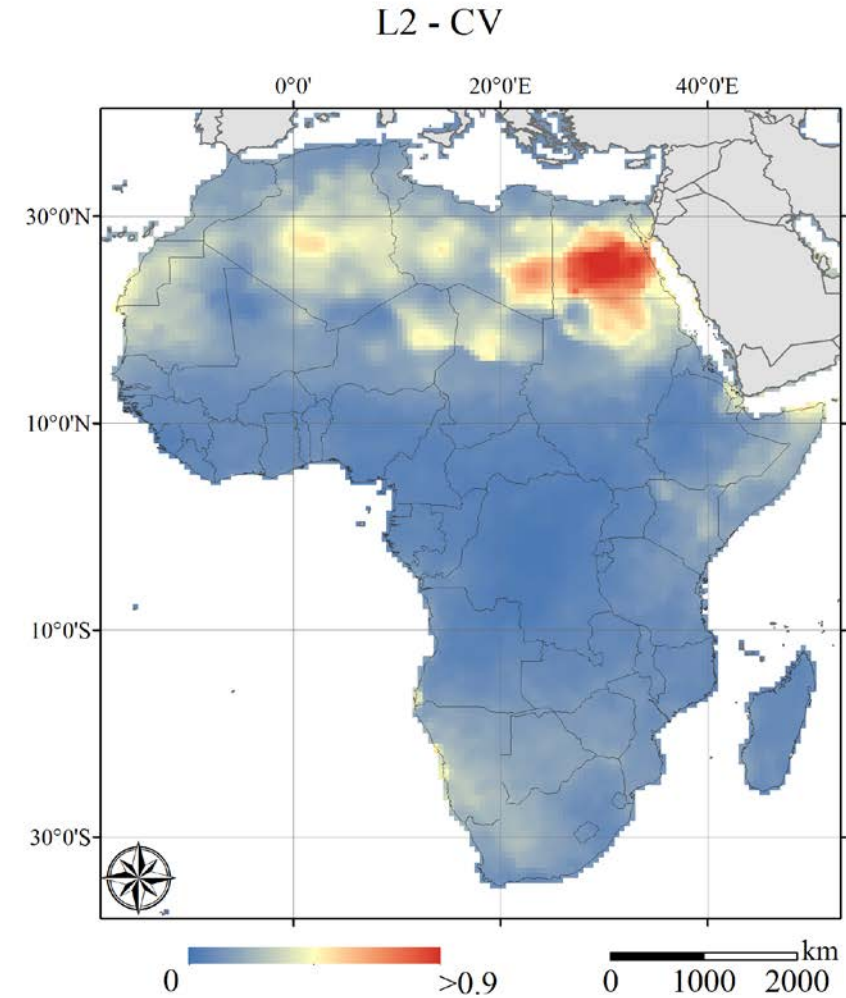
1st L-Mean

The mean total annual precipitation distribution of the first L-moment follows the well-known distribution of global total annual precipitation.



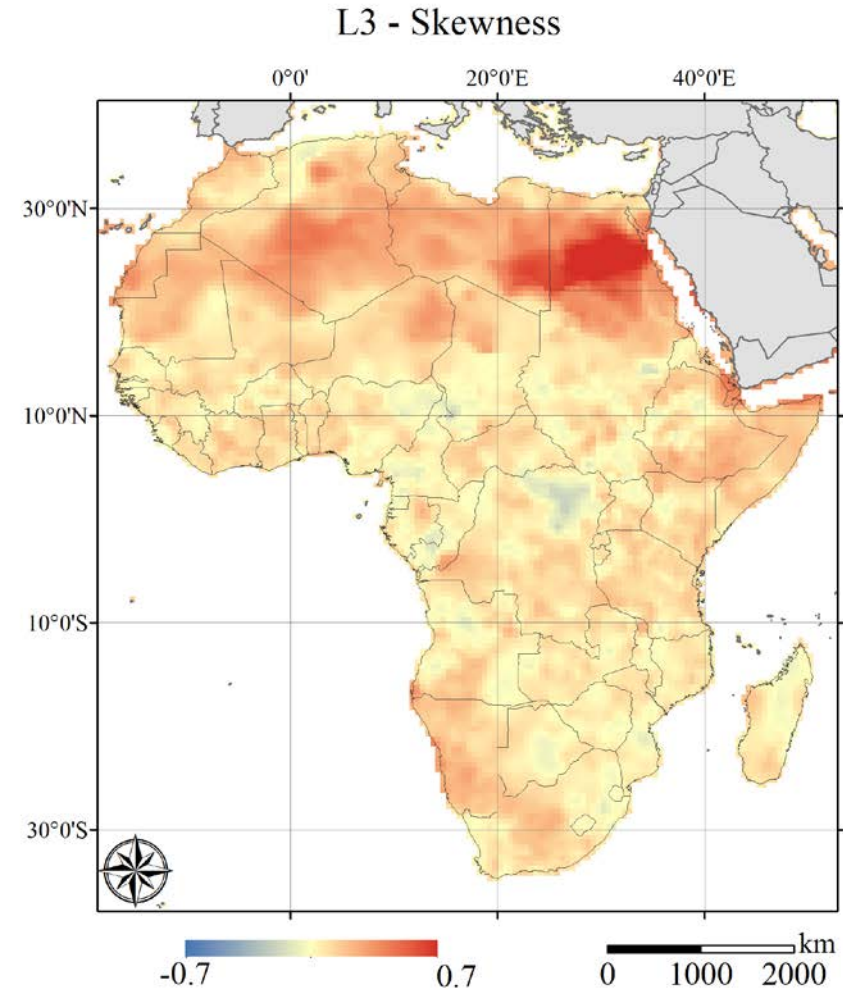
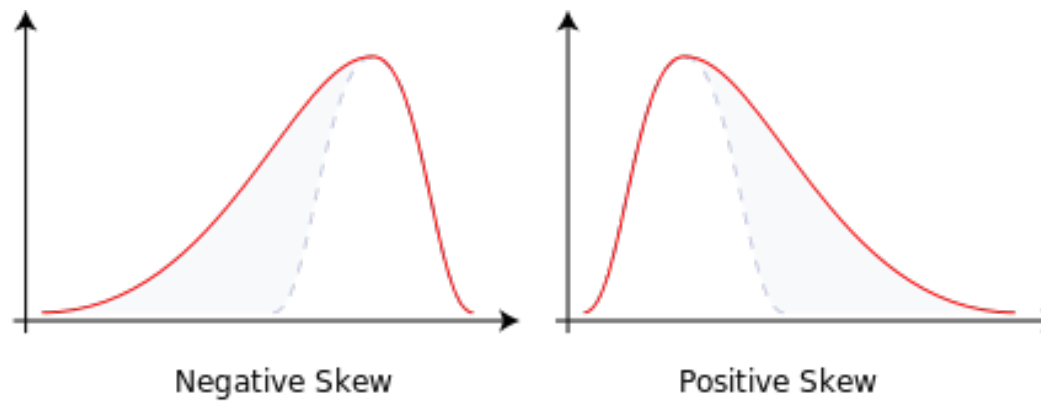
2nd Coefficient of L-variation

- The L-moment ratio (L-cv) measures a variable's dispersion, i.e. the expected difference between two random samples.



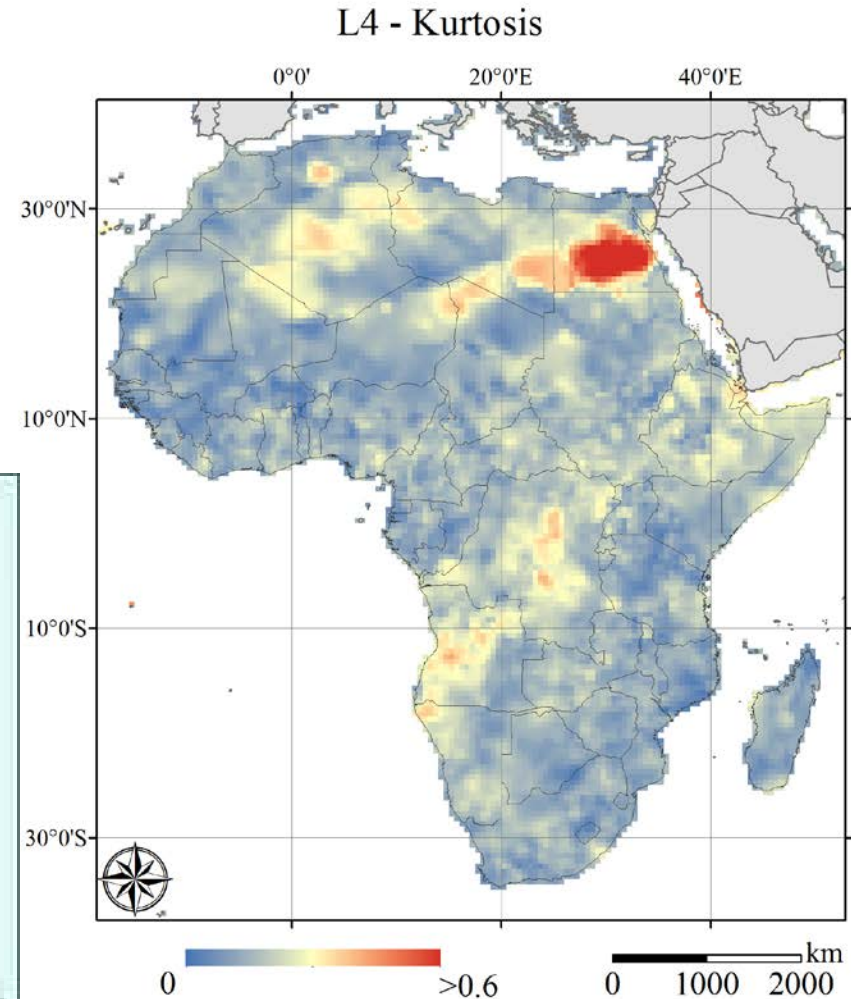
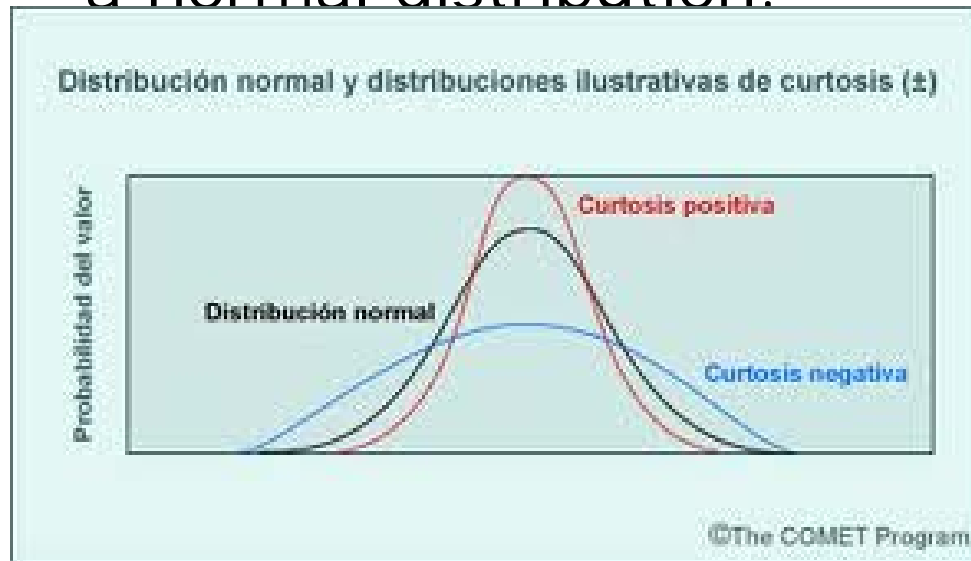
3rd L-Skewness

- The L-skewness quantifies the asymmetry of the samples distribution



4th L-Kurtosis

- L-kurtosis measures whether the samples are peaked or flat relative to a normal distribution.

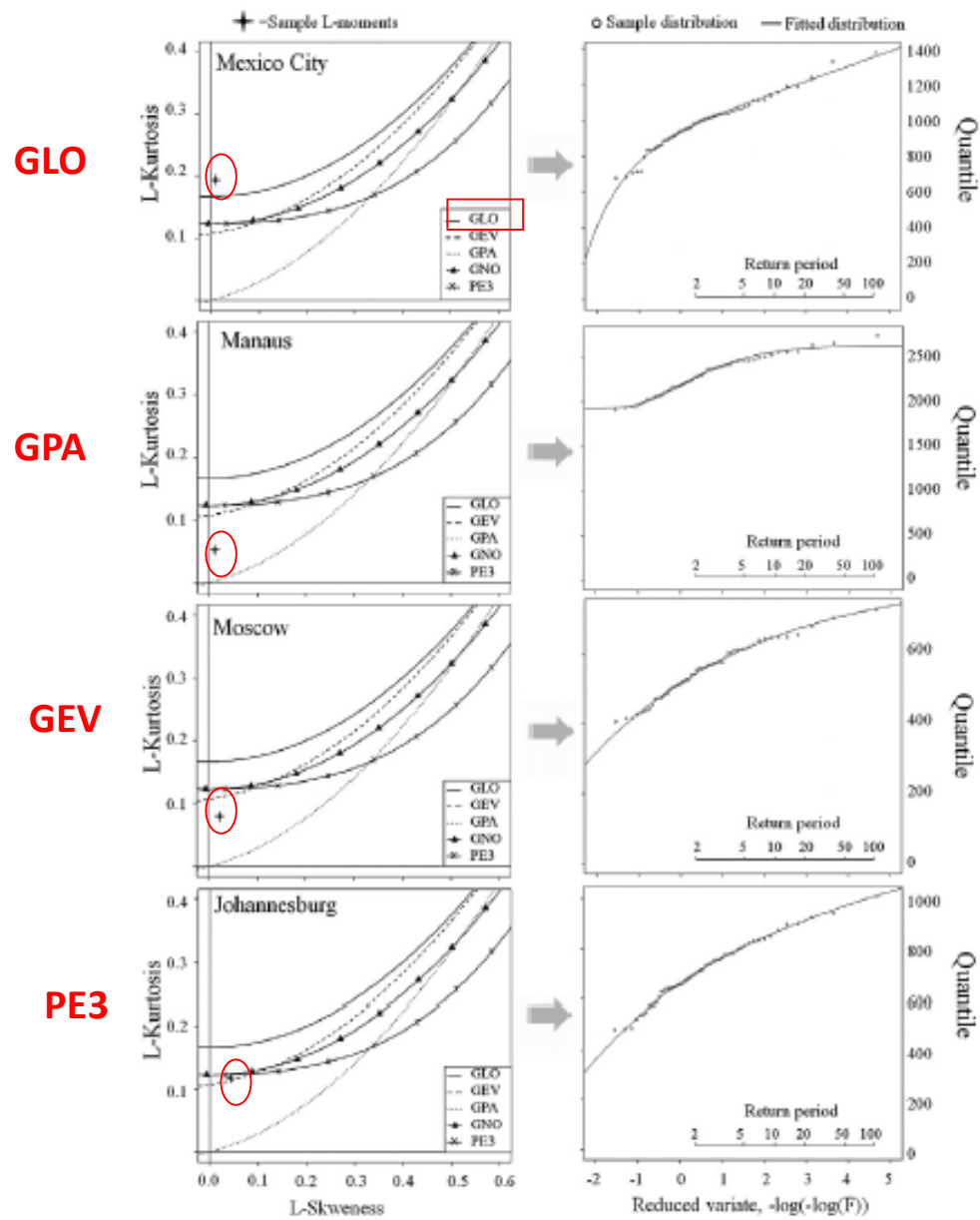


Probability distribution functions

Table 1 Name, acronym and equation of the distribution functions tested in the assessment

Distribution function	Acronym	Equation
Generalised logistic	GLO	$F(x) = \frac{1}{(1 + \exp(-y))}$, where: $y = \left(\frac{-1}{k}\right) \log\left(\frac{1 - k(x - \xi)}{\alpha}\right)$
Generalised extreme-value	GEV	$F(x) = \exp(-\exp(-y))$, where: $y = \left(\frac{-1}{k}\right) \log\left(\frac{1 - k(x - \xi)}{\alpha}\right)$
Generalised normal	GNO	$F(x) = \text{Phi}(y)$, where: $y = \left(\frac{-1}{k}\right) \log\left(\frac{1 - k(x - \xi)}{\alpha}\right)$
Three-parameter lognormal	PE3	$f(x) = \frac{1}{\alpha} \frac{e^{-\frac{(\ln x - \xi)^2}{2\sigma^2}}}{x}$ $y = \left(\frac{-1}{k}\right) \log\left(\frac{1 - k(x - \xi)}{\alpha}\right)$ if $k \neq 0$, $y = \frac{(x - \xi)}{\alpha}$ if $k = 0$
Generalised Pareto	GPA	$F(x) = 1 - \exp(-y)$, where: $y = \left(\frac{-1}{k}\right) \log\left(\frac{1 - k(x - \xi)}{\alpha}\right)$

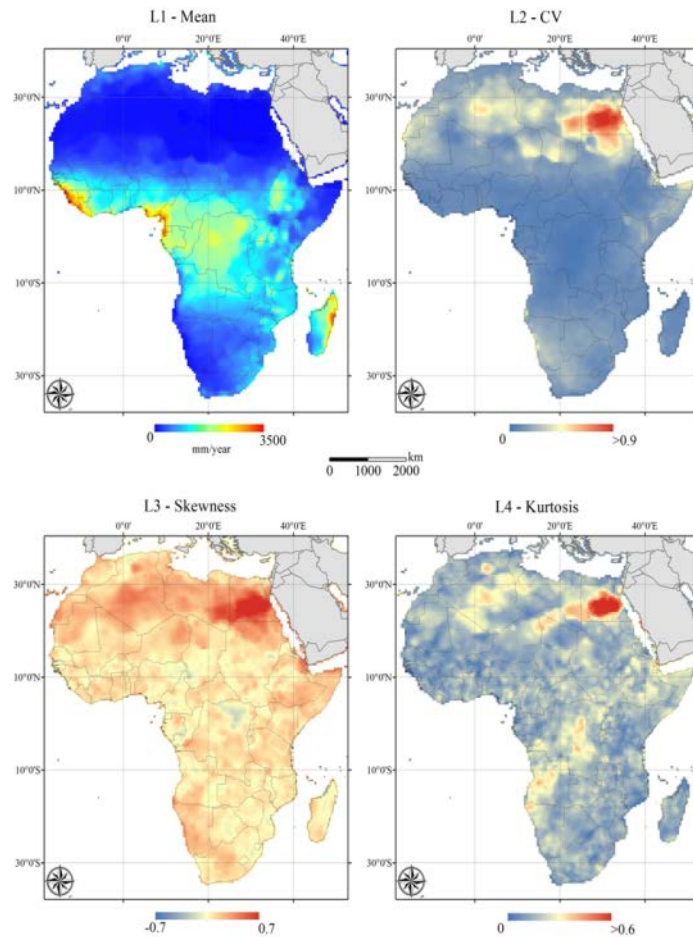
Note: In the equations, ξ represents the location parameter, α is the scale parameter and k is the shape parameter of the distribution functions. $\text{Phi}(y)$ is the distribution function of normal distribution, with x bounded by $\xi + \alpha/k$ from below if $k < 0$ and from above if $k > 0$



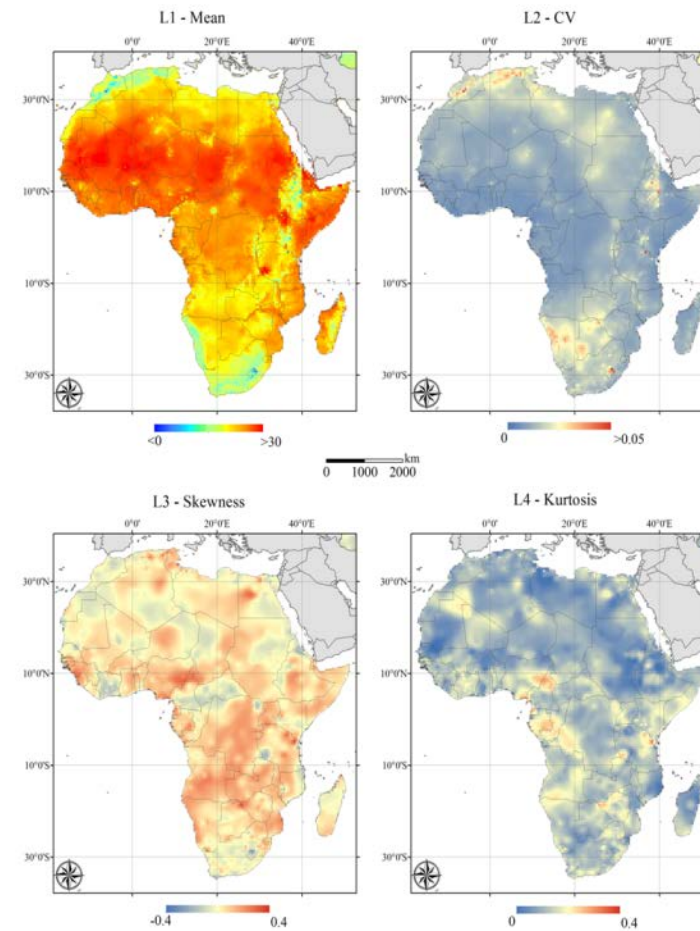
Results

Frequency analysis at annual time-scale

Annual precipitation L-Moments



Annual average temperature L-Moments



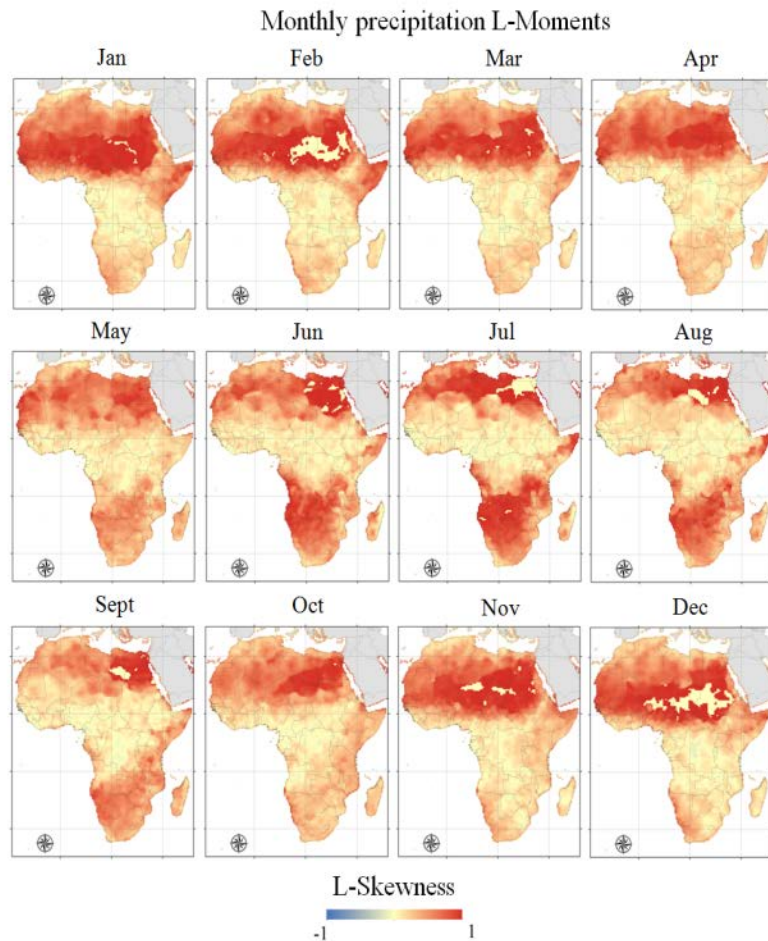
Available at
monthly or annual
scale

Source:

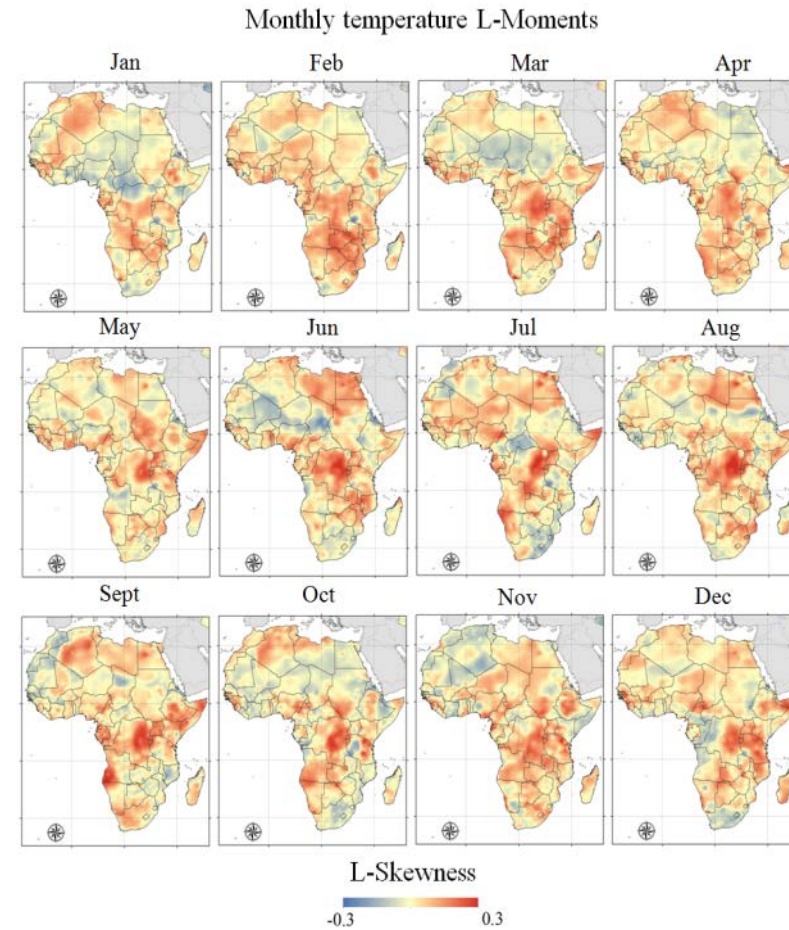
Maeda, E.E. et al "Characterization of global precipitation frequency through the L-moments approach". AREA- Royal Geographical Society

Results

Frequency analysis at monthly time-scale Precipitation

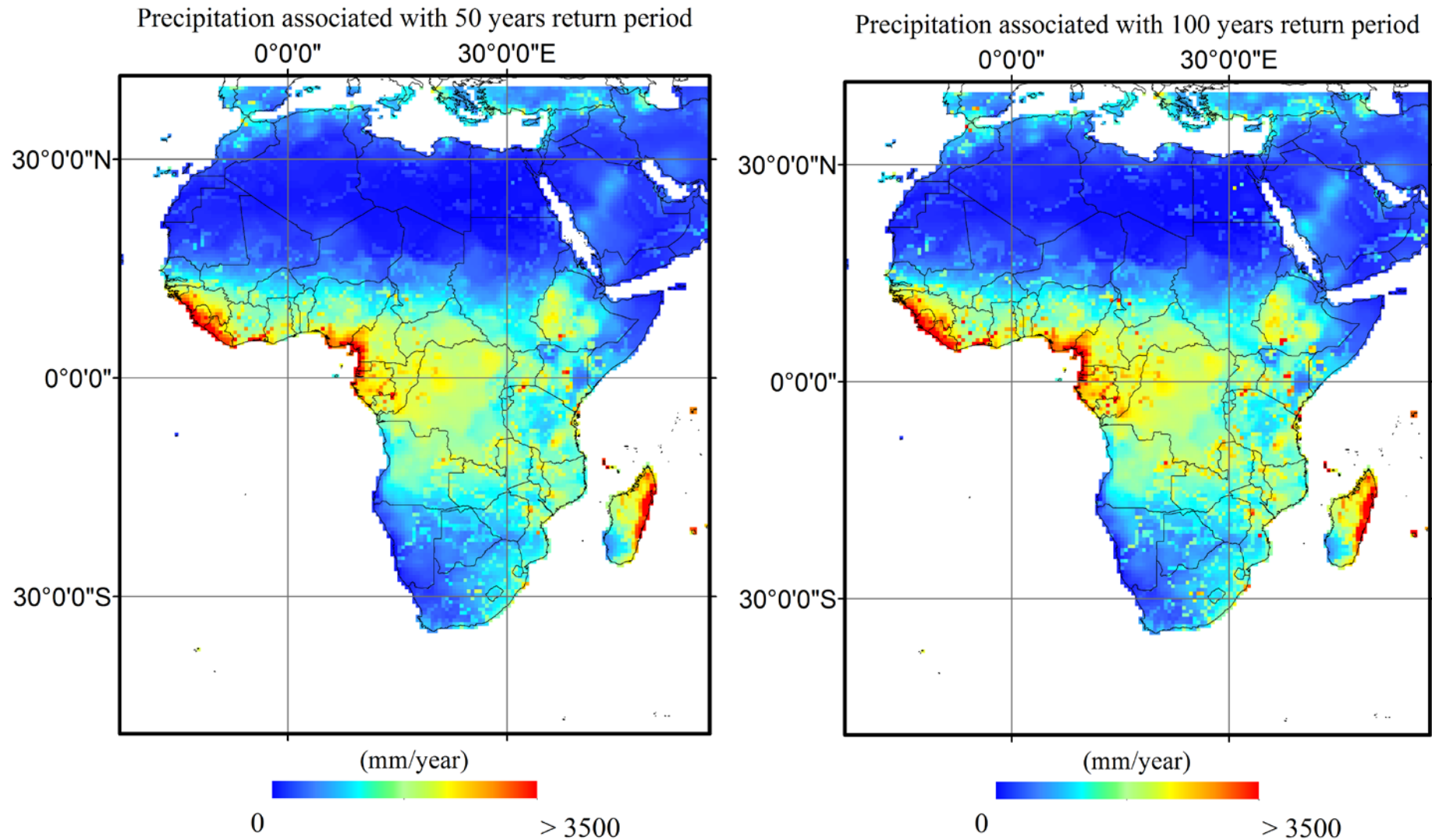


Temperature



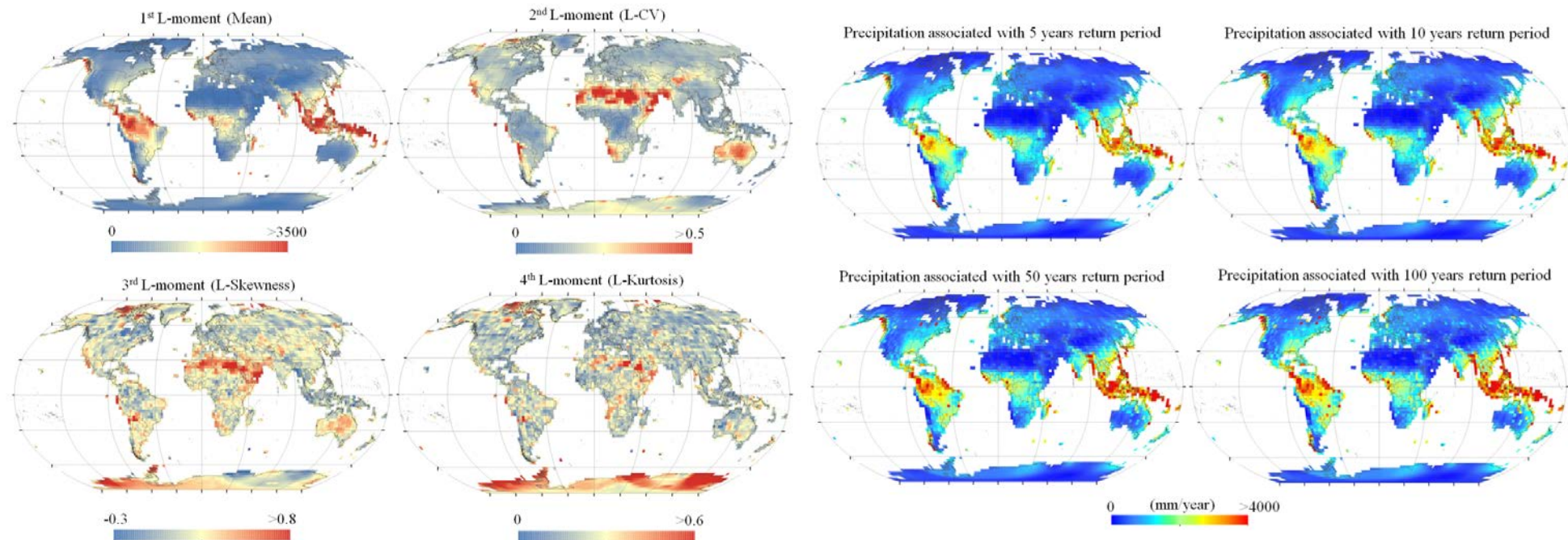
Source:
Maeda, E. E. et al "Characterization of global precipitation frequency through the L-moments approach". AREA- Royal Geographical Society

Annual precipitation return period



Bibliography

Maeda, E.E., Arévalo, J., Carmona-Moreno, C. (2012) "Characterization of global precipitation frequency through the L-moments approach". *Area-Royal Geographical Society*. doi: 10.1111/j.1475-4762.2012.01127.x JRC66941
<http://onlinelibrary.wiley.com/doi/10.1111/j.1475-4762.2012.01127.x/abstract>



Bibliography



Regional Frequency Analysis: An Approach Based on L- Moments

- Hosking, J. R. M., and Wallis, J. R. (1997). *Regional frequency analysis: an approach based on L-moments*. Cambridge University Press, Cambridge, U.K.