Digital Simulation of Daily Rainfall in the tropics

Dissertation
Zur Erlangung des Doktorgrades
der Naturwissenschaften im Fachbereich
Geowissenschaften
der Universität Hamburg

vorgelegt von
Edgar Djoumessi-Tatsangue
aus Dschang, Kamerun

Hamburg
2003

Als Dissertation angenommen vom Fachbereich Geowissenschaften

der Universität Hamburg

Auf Grund der Gutachten von Prof. Dr. R. Schwarz
und Prof. Dr. H. von Storch

Hamburg, den 19 November 2003

Prof. Dr. H. Schleicher
Dekan
des Fachbereichs Geowissenschaften

Contents

List of Tables

List of Figures

List of symbols

Symbols	Meaning
$a, \mu_1, s_1, \mu_1, s_1$	parameters of double normal fonction
$a_i, \mu_{i1}, s_{i1}, \mu_{i2}, s_{i2}$	spatial parameters of double normal fonction of $Y_{i1}(m)$ $i = 0, 1$
$b(m)$	scale parameter of Weibull distribution
e_i	error function of $z(m)$ estimating $Y_{i1}(m)$
e_x	error function estimating Weibull parameters
E_j	estimated frequency belonging to class j
$f(r)$	probability density function
$g(m)$	shape parameter of Weibull distribution
h	horizontal direction
L	length of a spell
m	month argument
n_{ij}	transition numbers $i, j = 0, 1$
O_j	observed frequency belonging to class j
P	arbitrary point in the area of study
p_{i1}, p_{i0}	transition probability $i = 0, 1$
R	random number in the interval [0,1]
r	rain amount on day t
$v(r)$	overlay normal distribution adds to Weibull distribution
$w(r)$	Weibull distribution probability density function of rain amount r
$Y_{i1}(m)$	monthly logits of transition probability with $i = 0, 1$
$z(m)$	double normal function for month m
χ^2	chi square statistic
$\delta(h_k)$	relief gradient in direction h_k
$\varepsilon(g(m), b(m))$	error function of parameters g and b
ϑ	relief exposition
$\kappa(P)$	rain exposition at a point P
λ	longitude
μ	mean value
$\nu(t+1)$	daily rainfall occurrence
ω	altitude

φ	latitude
σ^2	variance of monthly rain amount
θ	yearly rain amount

Abstract

The generation of rainfall data needs a range of models depending on time and on spatial scales involved. Cox and Isham (1994) presented three types of rainfall models, namely, empirical statistical models, models of dynamic meteorology and intermediate stochastic models, a classification based on the amount of physical realism incorporated into the model structure. This study is based on the empirical statistical model type, where empirical stochastic models are fitted to the daily rainfall data range from 1951 to 1993 of 28 stations. The models for generating a long sequence of daily rainfall series are required increasingly, not only for the hydrological purposes but also to provide inputs for models for crops growth, landfills, tailing dams, land disposal of liquid waste and environmentally sensitive projects. It also provides the means of extending the simulation of rainfall to unobserved locations.

The present study develops models for simulating daily rainfall series i.e. occurrence processes and rainfall amounts on wet days using the spatially distributed predictors such as latitude, longitude, altitude, exposition effect etc.

Two-state first-order Markov chain is used to model wet-dry and wet-wet occurrence processes; transition probability between dry or wet and wet days are calculated. The Weibull distribution function is used to model the daily rainfall amounts on a wet day. The seasonal variation in rainfall is an important factor and several approaches have dealt with the seasonality; in this study, we assume that parameters vary either as a step function for each month and use a double normal function to describe the seasonal variation of parameters. The idea behind the double normal function is based on the assumption of two different rainy seasons during a year, each one with a peak similar to that of a normal function. If there is only one rainy season, we expected that this could also be represented by near zero values for one of the two normal distribution. Parameters of double normal function are estimated by a minimization of errors. Stepwise regression analysis is then used to approximate parameters of the double normal function from spatially distributed predictors. As the outcome of the regression analysis, 5 % significance level is chosen to decide the inclusion or the exclusion of a predictor. Approximated parameters of the double normal function are used to write the equations of the models allowing the generation of daily rainfall series (occurrence processes and daily rainfall amounts) by the following procedures:

- chose an arbitrary location P in Cameroon where you want to simulate daily rainfall series
- determine the necessary information on $P's$ position and relief surrounding for input in equations for estimating wet/dry and wet/wet occurrence processes and the user will get a time series of daily precipitation occurrence at P.
- determine the necessary information on $P's$ position and relief surrounding for input in equations for estimating the Weibull parameters g, b and simulate daily rain amounts by generating random numbers R from pseudo-random number generator. Each random number then inserted in equation gives daily rain amount r time series at P using the Monte Carlo method.

Regrettably, we were not able to use the observation data that are independent from the ones we used for calibration, i.e. estimating all the parameters. The main results are as follows: The models reproduced well the seasonal variation of dry-wet and wet-wet transition probabilities at all the stations tested in spite of the gap existing between the observed and the estimated probabilities. It is also noted that there is a discrepancy between the month where the maximum (peak) and the minimum of estimated probabilities occurred against the observed ones. The generator seems to be able to generate rain series (number of wet days) for inland stations with statistics resembling the observations whereas the coastal area (Douala and Ngambe) shows the contrary. The simulated series present the same configuration as the observed series with the maximum of wet days occurring in the rainy season. In Douala, Ngambe and Ngaoundere, the maximum number of wet days occurs in August for the observed series whereas it occurs in September for the estimated series. The test of the frequency distribution of generated and observed rain amount series shows that the generator produces good estimate of daily rainfall amounts in the northern area (Kaele) and in the southern plateau (Sangmelima). Weak rainfall amounts are generated at the stations situated on highlands (Koundja) and in the coastal area (Douala and Ngambe). The test of the mean rainfall indicates that the significance value is above .05 in Kaele, Koundja, Ngambe, Sangmelima for all the 12 months suggesting that there is no difference between the mean of observed and simulated rainfall series. In Douala, the significance is less than .05 in August; in Ngaoundere the significance is also less than .05 in December. Both cases suggest a difference between the mean of observed and simulated data in these months. The situation is the same when testing the monthly variance; the value under .05 indicates that there is a difference between the variance of observed and simulated series in August, October, and November in Douala.

1 GENERAL INTRODUCTION

1.1 Problems and objective

Stochastic weather generators are tools to create weather inputs for other environmental and agricultural models. The output of a stochastic weather generator is designed to have the same statistical properties as the data used to fit it, and consists of a set of sequential random values of the weather variable of interest. Some of the environmental variables that have been used by those stochastic weather generators include daily precipitation, daily temperature, and solar irradiance.

There are several reasons for the development of stochastic rainfall generators and the use of synthetic rainfall data instead of observed data in human activities.

The first is the provision of rainfall time-series long enough to be used in an assessment of environmental and agricultural risk. Observed daily rainfall data are major inputs into agriculture simulation models, but the length of time series is often insufficient to allow good estimates of the probability of extreme events, such as for example the quantity of rain that may effect crop yield. Moreover, observed time series represent a single realization of the climate, whereas a rainfall generator can simulate many 'realizations', thus providing a wider range of feasible situations where constructed models need to be tested.

The second purpose is to provide the means of extending the simulation of rainfall time series to unobserved locations. Environmental risk assessments are now often made on high resolution grids or at multiple sites across a region where observed weather records are not available (Harrison et al. 1995). Several techniques such as kriging (Phillips et al. 1992), thin plate smoothing splines (Hutchinson 1995) or precipitation elevation regression on independent slope models (PRISM) (Daly et al. 1994) have been developed to interpolate the monthly means of climate variables, with the emphasis on the estimation of rainfall in mountainous regions. However, many impact models require daily rainfall data and so a different approach is required. Rather than estimating the climate variable (rainfall) directly (see, for example, Running et al. (1987), the parameters of a stochastic rainfall generator for each of the observed sites can be estimated, with the resulting parameters being used by the rainfall generator to produce synthetic daily rainfall data for the unobserved locations. Spatially estimated sets of parameters are available for different weather generators and for some countries in the northern hemisphere, for example, WGEN in the USA (Richardson and Wright 1984; Hanson et al. 1994) and LARS-WG in the UK (Semenov and Brook 1999).

A third reason for needing stochastic rainfall generators has recently arisen from the studies of climate change impacts. A weather generator can serve as a computationally inexpensive tool to produce site-specific climate change scenarios at the time step. The changes in both climatic means and climate variability predicted by GCM experiments can be applied to the parameters of the rainfall generator for the current climate at the site in question. Daily scenario data can then be obtained by running rainfall generators using the revised set of parameters (Wilks 1992; Semenov and Barrow 1997). To test the reliability of crop simulation models, sensitivity analyses to climatic variable are required, both to changes in mean and climatic variability. The relative importance of changes in rainfall variability on simulated crop yields for example has been demonstrated recently (Semenov and Brook 1999). Such sensitivity analyses have been done using stochastic weather generators. Parameters of the weather generator responsible for mean and variability of weather variables were pertubated and then used to produce time series of synthetic weather.

Stochastic weather generators are becoming an integral part of decision support systems. They are used to assess the effects of management decisions for a variety of scenarios (Tsuji et al. 1998). Usually the decision in question is tested for many possible weather situations, generated by a stochastic weather generator. A temporal distribution of the model performance characteristics, such as, for example, yield, is constructed for every decision. These distributions are compared with one another to select the optimal management. Stochastic weather generators can also be used in real time simulations.

In many tropical developing countries, rainfall is a major influence on human activity, particularly agriculture. A number of rainfall parameters, such as amounts over a range of time periods, frequencies of occurrence of particular amounts, the nature of wet and dry spells, need consideration in order to fully understand this influence. The general problem can be summarized as follows: Which future amount of rainfall is to be expected at a given location within a certain time span and by which probability? How reliable is the estimation?

Thus, in this work the precipitation series are modeled by a separate process, so that rainfall is a two-step process:

1. determination of whether precipitation occurs, and

2. if precipitation does occur, determination of how much fell.

This work is focused on the building of simulation models which can be used to digitally generate rainfall series on a daily time scale; the main objective is to develop models for simulating daily rainfall occurrence processes and the nonzero precipitation amounts on wet days at a point P in Cameroon using spatially distributed predictors. The approach here is to estimate parameters of models describing daily rainfall series from the characteristics of terrain such as latitude, longitude, altitude, relief exposition etc; best terrain characteristics are chosen for using stepwise regression analysis.

1.2 Background of the study

Most effort in the construction of weather generators has been devoted to daily precipitation processes. Not only precipitation is the most critical meteorological variable for many applications, but the presence or the absence of precipitation also typically affects the statistics of many nonprecipitation variables to be simulated. Precipitation data are difficult to be modelled because they exhibit distinctive and difficult characteristics. Most rainfall generators contain separate treatments of precipitation occurrence and intensity. The precipitation occurrence process manifests itself in two weather states, wet or dry. A key aspect of stochastic rainfall models is their representation of the tendency of wet and dry days to exhibit persistence, or positive serial autocorrelation, so that wet and dry days tend to clump together in time more strongly than could be expected by chance. The precipitation intensity pertains to the modelling of nonzero precipitation amounts. These are typically strongly skewed to the right, with many small values and few but quite important large precipitation amounts. Although these concepts appear straightforward it has taken more than a century to formalize many of the processes within stochastic precipitation models.

1.2.1 Precipitation occurrence

Apparently the earliest published work on probabilistic modelling of precipitation occurrence was that of Quetelet, who reported in 1852 (Katz 1985) that runs of consecutive rainy and dry days at Brussels for 1833-1850 exhibited the persistence. Another expression of this tendency for wet and dry weather to persist was provided by Newnham (1916), who used daily rainfall records at Kew, Aberdeen, Valencia and Greenwich, UK, to demonstrate that the probability of a "rainy day" is greater if the preceding day was wet rather than dry. These two approaches considering run lengths and day-to-day statistical dependence to describe the temporal dependence of precipitation occurrences were pursued further by Besson (1924), Gold (1929) and Cochran (1938). Williams (1952) used geometric series to model dry and wet spell (i.e., consecutive runs of dry or wet days) lengths at Rothamsted Experimental Station, Harpenden, UK. Longley (1953) subsequently improved geometric series fit to observed wet and dry spell lengths in five Canadian cities by differentiating between the months in which spell fell.

Gabriel and Neumann (1962) are generally credited with presenting the first statistical model of daily rainfall occurrence. In their work using rainfall data for Tel Aviv, Israel, the authors recognized that the frequency distributions for wet and dry-days length of the types identified by Williams (1952) and Longley (1953) may arise from a simple Markov chain model. In particular, Gabriel and Neumann (1962) proposed the use of first-order Markov chain for precipitation occurrence, assuming that the probability of rainfall on any day depends only on whether the previous day was dry or wet. This model can be fully defined by the two conditional probabilities $p_{01}(t)$ being the probability for having precipitation on the day t and no precipitation on day t_{-1} and $p_{11}(t)$ being the probability for having precipitation on the day t and precipitation on day t_{-1} which are called transition probabilities. Since there are only two possible states on a given day, the two complementary transition probabilities are $p_{00} = 1 - p_{01}$ dry day following a dry day and $p_{10} = 1 - p_{11}$ dry day following by a wet day. It was noted by Gabriel and Neumann (1962) that this simple model for rainfall occurrence was able to describe closely the persistent nature of daily precipitation occurrence patterns, and that certain other properties of the occurrence series could be derived from the transition probabilities.

p_{01} and p_{11} illustrate the use of the first-order Markov model to characterize important aspects of the precipitation occurrence climate. This model also provides a convenient and efficient means of generating sequences of random numbers that simulate the corresponding real rainfall data. For each simulated day, a random number R is drawn from the interval [0,1] in such a way that any real number in that interval is equally likely to be picked. In practice, these are usually produced by widely available computer algorithms called uniform pseudo-random number generators (Press et al. 1986; Bratley et al. 1987).

Once a random number R has been generated, whether the next day in the sequence is wet or dry, is determined using p_{01} and p_{11}. If the previous day $(t-1)$ was dry, then day t is simulated to be wet if $R \leq p_{01}$, and otherwise it is also dry. If the previous day was wet, then the current day is simulated to be wet if $R \leq p_{11}$, and is dry otherwise. As first-order Markov models fit to daily rainfall data, the simulations yield sequences of wet and dry days that are more persistent than independent draws.

For temperate climates, it has been found that the simple first-order Markov model generates synthetic rainfall series with too few dry spells (e.g.,Buishand (1977); Buishand (1978); Racsko et al. (1991); Guttorp (1995); Jones and Thornton (1997); Wilks (1999) addressed this deficiency by considering Markov chains of higher order. These techniques increase the length of the Markov model's "memory" of antecedent wet and dry days. For example, second-order Markov chains use the wet/dry state on both the preceding day, and two days prior, such that eight transition probabilities p_{ijk} must be defined. Here each of the indices i,j,k may be either one (wet) or zero (dry). Hence p_{101} would be probability of a wet day given that the previous day was dry, and the day before that was wet. Third and higher-order Markov chain can be similarly defined, although the number of transition probabilities required increases exponentially as the order increase, being 2^k for k^{th} -order chain. When only the dry days are not adequately modeled by the first order Markov model it is possible to improve the statistics of the simulated dry days using hybrid order Markov models, in which the Markov "memory" extends further back in time for the dry days only (Stern and Coe 1984; Wilks 1999).

When deciding between models having different degrees of complexity, one must judge how elaborate a model is justified by the data. Gabriel and Neumann (1962) compared the first-order Markov model for precipitation occurrence in Tel Aviv with the next simplest model, namely independent Bernoulli (i.e., binomial) occurrences, using Chi-square tests. While this is a reasonable approach when only two alternatives are being considered, ambiguities in such statistical tests arise when multiple comparisons are made, for example when choosing among Markov models of zeroth (binomial distribution), first and second orders. The usual approach in circumstances like this is to employ an objective order-selection criterion such as Akaike's information criterion AIC (Akaike 1974) or Bayesian information criteria BIC (Schwarz 1978). Both AIC and BIC are likelihood-based criteria, in that they choose the model having the largest maximized likelihood, after application of a penalty that increases with the number of free parameters allowed by each of the models considered. The likelihood function is notationally analogous to the probability distribution function or the probability density function; yet, the data are considered as fixed, while values of the parameters associated with the global maximum of this function are the fitted maximum likelihood estimates. Gates and Tong (1976) concluded that second-order Markov dependence was justified according to the AIC for Tel Aviv precipitation data considered by Gabriel and Neumann (1962). However, Katz (1981) concluded that first-order dependence for these data was adequate on the basis of the BIC, which he also showed to be asymptotically consistent (i.e., the BIC is correct on average for sufficiently large data samples).

An alternative to Markov chain models for simulating precipitation is the use of spell length models. Rather than simulating rainfall occurrences day by day, spell length models operate by fitting probability distribution to observed relative frequencies of wet and dry spell lengths. This kind of model is sometimes called an 'alternating renewal process' (Buishand 1977; Buishand 1978; Roldan and Woolhiser 1982), in that random numbers are generated alternately for the wet and dry spell length distributions. That is, a new spell length L is generated only when a run of consecutive wet or dry days has come to an end, at which point a new spell of the opposite site is simulated. Of course, if geometric distributions are used to model the length of wet and dry spells, the resulting synthetic series will exhibit the same characteristics as the equivalent first-order Markov process. Higher-order Markov chain have spell length distributions associated with them that are generalizations of the geometric distribution. Precipitation occurrence sequences with different statistical characteristics can be obtained using different distributions for the frequencies of spell lengths. Such distributions include the truncated negative binomial distribution (Buishand 1977; Buishand 1978; Roldan and Woolhiser 1982), the negative binomial distribution (Wilks 1999) and the mixed geometric distribution (Racsko et al. 1991). For climates where the above distributions yield very long dry spells with insufficient frequency, these more elaborate choices for modelling precipitation occurrence have been found to yield more realistic results (Racsko et al. 1991; Wilks 1999; Buishand 1977; Buishand 1978). However, this method is susceptible to poor parameter estimates in arid regions (Roldan and Woolhiser 1982). As part of Cameroon being located in the arid area, the previous limitation suggests the rejection of the use of spell-length models in this study.

A number of methods such as Markov chain models, Alternating Renewal Process (ARP), spell length models are among some approaches that have been used to simulate daily rainfall occurrence in temperate regions. However, Alternating Renewal Process presents some limitations (Roldan and Woolhiser 1982): it is susceptible to poor parameter estimates in wet regions, therefore many years of observations are necessary; dry regions are mostly found in tropical areas and tropical areas are characterized by the scarcity of longer observations. Because of these limitations, Alternating Renewal Process and spell length models are discarded in this study.

The two-state first-order Markov chain model as defined by Gabriel and Neumann (1962) is used in this study to model daily rainfall occurrence; it provides a convenient and efficient means of generating sequences of random numbers that simulate the corresponding real rainfall data. The application of Markov chain to study daily dry and wet sequences in tropics is a way of examining the persistent nature of rainfall patterns of this area. A two-state first-order Markov chain model is chosen over other stochastic processes because it is conceptually simple; it has been previously used in the context of analysing daily rainfall occurrence in most part of the world. Another reason is that the results of analysis are relevant to management (transition probabilities) and are easy to derive from successional data.

Comparatively little work on this topic has been carried out in tropical areas but in view of its significance, inevitably attempts will be made to generalize from one area to another. Although it is possible to make some broad statements about daily characteristics, the extremely varied mechanisms of tropical rainfall production, the characteristics of rainfall(large spatial and temporal variability) in the tropics make such generalization dangerous.

For the world as a whole, the terms "summer" and "winter" are commonly used to denote respectively the season of high sun (maximum temperature) and low sun (minimum temperature) in the tropical area. In middle and high latitudes a traditional four seasons year of spring, summer, autumn and winter is chiefly based on temperature contrasts and the length of daylight. However, in the tropics, smaller monthly temperature constract combined with a more equal length of daylight has made rainfall the most common criterion for seasonal subdivision. On this basis, a distinction is usually made between a wet season and a dry season, sometimes separated by two transitional seasons. Clearly, previous temperate area analyses are not necessarily a guide to the tropical situation because of the differentiation into two seasons. In view of the significance of modelling daily rainfall series and the potential utility of Markov processes to characterise a variety of wet/dry situations, modelling rainfall occurrence under the tropical conditions is important.

Technical details in this part of the work consist in

1. computing transition probabilities p_{11} and p_{01}
2. using a function to describe the seasonal variation of parameters
3. estimate parameters from the spatially distributed predictors which are latitude, longitude, altitude, relief position, relief exposition, and mean of yearly rainfall
4. generating the synthetic occurrence series of wet and dry days series,
5. generating synthetic transition probabilities (p_{11},p_{01}, p_{00}, p_{10})
6. compare the generated probabilities series against observed ones.

The scheme of the process followed in this study to analyse and to model the rainfall data series is presented in figure 1. The first decision taken at this level is to deal with the occurrence (wet/dry) process and to describe the distribution of rainfall amounts independently on the wet days. Many studies have examined rainfall occurrence by a two-state first-order Markov model. But choosing a proper order for the Markov models is problematic, and has important implications from the stand point of model effectiveness and parsimony. Following tabulation of frequencies of occurrence within each class and calculation of probabilities, initial probabilities are plotted as monthly averages for each of 28 stations. Data are then grouped on this basis for testing dependence/independence. Contingency tables are used to test a range of hypotheses about the order of Markov chain. Although later analyses may involve a range of thresholds, the classification methods presented by Jackson (1981) and Stern et al. (1982) is used. Briefly, for each site each day is classified as dry (wet) day depending whether the site received less (more) than 1mm of rain. The probability of a day being dry will depend on the state (dry or wet) of the preceding day.

Transition probabilities p_{11} and p_{01} are computed from the historical data. They are then transformed using a logit function

$$Y_{i1}(m) = ln(\frac{1}{1 - p_{i1}(m)^{0.125}}) \qquad i = 0, 1 \tag{1}$$

.

We then obtain the logit $Y_{i1}(m)$ of daily rainfall occurrence. The seasonal variation $Y_{i1}(m)$ is decribed using a double normal function with five parameters.

Spatial distribution of the probabilities is studied by applying a stepwise regression analysis of pameters of the double normal function on spatially distributed predictors; by doing so, we estimate parameters of the double normal function spatially so that we can use them to write the equations of models needed for generating wet-dry occurrence and wet-wet occurrence in a point P in Cameroon. The last step consists in estimating probabilities by inverting $Y_{i1}(m)$ to Y_{i1}^{-1}.

1.2.2 Daily rain amounts

The next necessary element of a rainfall generator is a model for nonzero precipitation amounts on wet days. The most prominent statistical feature of daily precipitation amounts is that their distribution is strongly skewed. That is, very small daily precipitation amounts are quite common, while the large daily precipitation amounts that are most important to hydrological, agricultural and engineering impacts are comparatively rare. Todorovic and Woolhiser (1975) were the first to produce a daily stochastic precipitation generator by combining the first-order Markov model for daily precipitation occurrence with a statistical model for daily nonzero precipitation amounts. Their choice for modelling the daily rainfall amounts r was the exponential distribution whose probability density function with mean μ is

$$f(r) = \frac{1}{\mu}\ exp \quad [\frac{-r}{\mu}] \tag{2}$$

The exponential distribution is probably the simplest reasonable model for daily precipitation amounts, as it requires specification of only one parameter, μ , yet reproduces qualitatively the strong positive skewness exhibited by daily precipitation data. The average nonzero precipitation amounts is μ , and the corresponding variance $\sigma^2 = \mu^2$. Exponential distributions have also been used by Richardson (1981); Wilby (1994) among others.

A number of more elaborate models have also been proposed for the distribution of daily precipitation amounts given the occurrence of a wet day. The two parameter gamma distribution has been the most popular choice (Thom 1958; Katz 1977; Buishand 1977; Buishand 1978; Stern and Coe 1984; Wilks 1989; Wilks 1992), and has the probability density function

$$f(r) = \frac{(r/\beta)^{\alpha-1}\ exp\ [-r/\beta]}{\beta\,\Gamma(\alpha)} \tag{3}$$

This distribution involves two parameters: the shape parameter α and the scale parameter β. Factor $\Gamma(\alpha)$ is the gamma function evaluated at α (see, e.g. Abramowitz and Stegun (1984); Wilks (1995)). This function has mean $\mu = \alpha\beta$ and variance $\sigma^2 = \alpha\beta^2$. For $\alpha \leq 1$ gamma distributions are qualitatively similar to exponential distribution in concentrating most of the probability near zero and producing large precipitation amounts only rarely. For $\alpha = 1$ gamma distribution reduces to exponential distribution, but in general the additional parameter allows more flexible accommodation of rainfall amount frequencies, and thus improves the realism of stochastic precipitation models.

Another natural generalization of the exponential distribution is the mixed exponential distribution which is simply a probability mixture of two one-parameter exponential distributions. Its probability density function is

$$f(r) = \frac{\alpha}{\mu_1}\ exp\quad [\frac{-r}{\mu_1}] + \frac{1-\alpha}{\mu_2}\ exp\quad [\frac{-r}{\mu_2}] \tag{4}$$

Mathematically, this probability density indicates a superposition of two ordinary exponential distributions whose respective means are μ_1 and μ_2. From the standpoint of simulation, a natural generalization of exponential distribution is used to generate the precipitation amounts with probability α, and the second is used with probability $1-\alpha$. The mixed exponential distribution has mean $\mu = \alpha\mu_1 + (1-\alpha)\mu_2$ and variance $\sigma^2 = \alpha\mu_1^2 + (1-\alpha)\mu_2^2 + \alpha(1-\alpha)(\mu_1 - \mu_2)^2$. First suggested as a model for daily precipitation amounts by Woolhiser and Pegram (1979), the mixed exponential distribution has been rarely used comparatively. However, it has been reported to provide substantially better overall fits to daily precipitation data than the gamma distribution (Roldan and Woolhiser 1982; Foufoula-Georgiou and Lettenmaier 1987; Wilks 1998; Wilks 1999), and in particular Wilks (1999) reports better representation of the frequencies of the very largest precipitation amounts.

Given a distribution to represent the nonzero precipitation amounts, simulations are accomplished through computer algorithms (Bratley et al. 1987) that generate random numbers according to the fitted distribution. For each day the precipitation occurrence model simulates wet conditions. A new random variate for nonzero precipitation amounts is generated from the fitted distribution.

Most stochastic rainfall generators make the assumption that precipitation amounts on wet days are independent, and follow the same distribution. Allowing different probability distributions for precipitation amounts depending on that day's position in a wet spell (e.g., the mean rainfall on a wet day following a wet day might be greater than on a wet day following a dry day) has been considered by Katz (1977); Buishand (1977); Buishand (1978); Chin and Miller (1980). Similarly, the autocorrelation between successive nonzero precipitation amounts in daily series is sometimes statistically significantly different from zero, but is typically quite small and usually of little practical importance (Katz 1977; Buishand 1977; Buishand 1978; Foufoula-Georgiou and Lettenmaier 1987). In contrast, the accounting for serial correlation of nonzero precipitation amounts is essential if the precipitation model has an hourly(or smaller) rather than a daily time step (Katz 1995).

Another approach to the accounting for correlation in the nonzero precipitation amounts is the use of multistate (i.e., greater than the 2 states) Markov models. These Markov models simulate both precipitation occurrence and amounts, by defining different ranges of precipitation amounts as constituting distinct states. Transition probabilities among all possible pairs of states are estimated from data, and used in simulation. For example, for southeast England in spring, Gregory et al. (1993) found that there is a 16 % probability that a rainfall total of greater than 6.62 mm will be equalled or exceeded on the following day, but only a 5.7 % chance that such a day will be followed by a dry day. Gregory et al. (1993) also found that first-order Markov models yield smaller discrepancies in seasonal precipitation total if specified ranges of daily precipitation totals are used to condition the rainfall amounts on the following day. Haan et al. (1976) followed a similar approach. The validity of this multistate Markov approach clearly rests on the choice of the number of states and their ranges (i.e., the upper and lower rainfall thresholds) and on the distributions used for wet-day amounts in any given state. These models involve comparatively large numbers of parameters, and thus required quite long data records in order to be estimated well.

Many probability density functions have been used to simulate nonzero precipitation amounts, among them: the exponential distribution, the two-parameter gamma distribution, the mixed exponential distribution, the use of multistate models. These models involve comparatively large numbers of parameters, and thus require quite long data records in order to be estimated well (Wilks and Wilby 1999). Such limitations bring us to use Weibull distribution to simulate nonzero precipitation amounts in this study.

Technical details to be used in this part of the study consist in:

1. setting on some statistical parameters and patterns to nonzero rainfall
2. choosing a suitable distribution to fit the daily rainfall amounts
3. using a function to describe the seasonal variation of parameters of the chosen distribution function
4. estimating parameters of function describing the seasonal variation from the spatially distributed predictors
5. developing models for simulating daily rainfall amounts
6. generating the synthetic rain amounts on a wet day using Monte Carlo method
7. testing the frequency distribution of generated rainfall series against observed series
8. testing the mean and the monthly variance of the generated rainfall amounts.

The scheme of the process, followed in this study to analyse and to model daily rainfall amount is presented on the right side of figure 1. The whole data is divided into two parts; the first part from 1951 to 1993 is used to calibrate the model whereas to test the models, the series from 1994 to 2000 are used.

Basic statistical parameters (frequencies, monthly and yearly means etc.) and plots (histograms, curve) are computed followed by the description of patterns in the observed daily, monthly, seasonal, daily and yearly precipitation series. Histograms of monthly rainfall enable us to define the precipitation regimes whereas plots exhibit seasonal variation of the series.

The seasonal variation of precipitation series throughout the year is an important factor in the construction of models: we assume that parameters vary as step functions for each month and we use double normal function to provide daily variation of parameters. It is chosen because it can fit both bimodal or unimodal seasonal patterns easily. It has the advantage that the fitted probabilities at the beginning and end of the year are the same.

The next step deals with the spatial distribution of the daily rain amounts. This is done by applying stepwise regression analysis to parameters of the double normal function and spatially distributed predictors. By doing so, we estimate parameters of the double normal function spatially; the estimated parameters are then used to write equations of the models necessary for simulating daily rainfall amounts on wet day. The Monte Carlo method is used to generate a rainfall series.

Another step is the validation of generated precipitation series against observed precipitation series. The seasonal number of generated and observed wet days on the one hand, the frequency distribution of generated and observed daily rainfall amounts on the other hand are compared by using the Chi-square test. The monthly variance and the mean of the observed and generated rainfall amount are tested using F-test and t-test respectively.

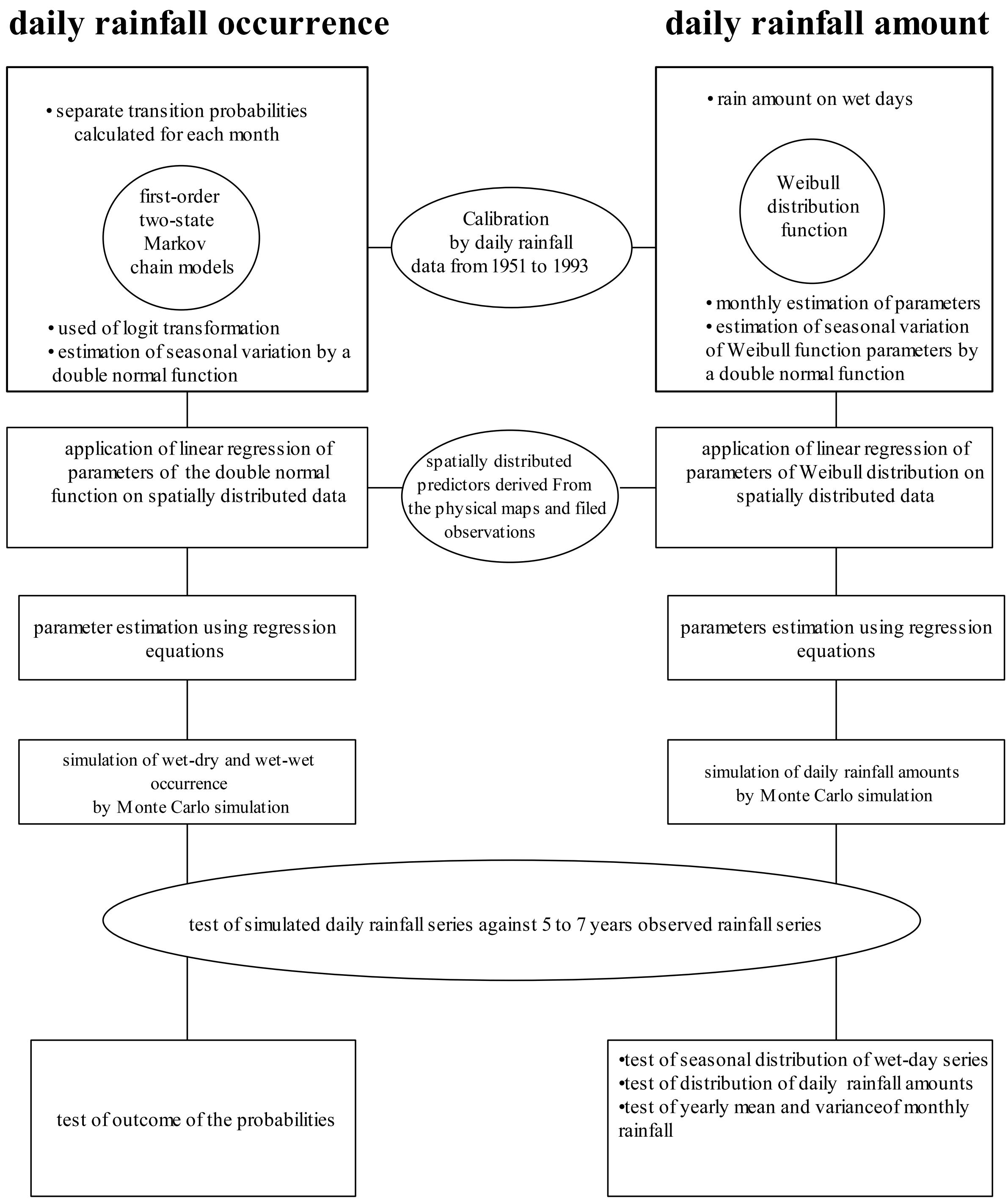


Figure 1: Scheme of the process followed in this study

2 THE AREA OF STUDY AND DATA PRESENTATION

2.1 Generalities on relief and climate of Cameroon

Cameroon is located in western Africa, bordering the Bight of Biafra between longitude 8° and 16° East and between latitudes 2° and 13° North of the Equator (figure 2). A look at the physical map of Cameroon (figure 3) shows that it is a country of varied landscape full of plains, plateau and highlands. The following major physical units can be distinguished: coastal plain in the southwest, dissected plateau in the centre, the Adamawa plateau, mountains in the west, plain in the north. These physical units play an important role on the climate, particularly in the repartition of precipitation in Cameroon represented in the figure 6.

By far the highest precipitation totals are those recorded where highlands interrupt the onshore passage of maritime air, in particular the Cameroon Highlands. Debundscha, near the base of mount Cameroon records some of the world's highest annual precipitation totals. It receives the full force of the moist southwesterlies as they ascend the mount Cameroon, recording an annual mean as high as more than 10,000 mm. In complete contrast, Ekona, in the rainshadow on the far side, has a mean of less than 1,500 mm. High totals are also recorded along the coastal plain (Douala, Kribi). Such high totals on the low-lying coast are related to the effects of frictional convergence as streams meet and cross onshore (Buckle 1996). From the coast to inland, rainfall decreases.

Precipitation totals distribution in Cameroon result also from the latitudinal extension of the country and its location at the Gulf of Guinea. In West Africa the east-west alignment of the Guinea coast, with the warm equatorial Atlantic to the south, extends the zone of humid tropical climates much further to Cameroon. Rainfall decreases from the coast to the interior, both in amount and in duration. At the coast, annual totals can exceed 4,000 mm (Douala), while in the northern part they drop to less 800 mm (Maroua). Rainfall, however, remains highly seasonal, the result of the alternating influence of two contrasting airstreams: humid maritime air from the Atlantic, and dry continental air from the Sahara.

As the tropical maritime air of the Atlantic draw to the equators it acquires more and more moisture until eventually it is transformed into the warm, moist, highly unstable equatorial air of the Inter-Tropical Convergence Zone (ITCZ). Moving onshore from the Atlantic, the lower layer of this air is warmed, enhancing thereby its instability still further. It is an equatorial air that brings the greatest proportion of rain to Cameroon. Within the body of this unstable airstream, convectional storms develop readily, while wherever it is forced upwards, such as on Mount-Cameroon and the adjacent offshore islands, rainfall is extremely heavy. Precipitation here originated from the thermal convection and squall lines, the most violent of which appears to result from the distortion of the Inter-Tropical Front under the surge of one of the converging fluxes, before thrusting into the mass of damp air as back draft of the surface circulation (Suchel 1988).

The great deserts of Sahara and Arabia north of the equator are the source regions of tropical continental air. The subsidence associated with the high pressure cells of these regions creates air that is hot, dry, and very stable. The dryness and subsidence limit cloud development and precipitation. In fact, despite the steep lapse rates and extreme instability of the lowest layers produced by the intense surface heating in the hottest months, the extremely low humidities rarely encourage more than dust storms to develop. Between December and March it is this dry, dusty air that reaches the country as the Harmattan haze. Tropical continental air brings dry season wherever in the country.

During the northern hemisphere low-sun period (November-March), dry stable Sahara air, carried equatorwards by tropical easterlies, dominates Cameroon. During this period, the Harmattan winds become stronger than humid Atlantic air so that the Inter-Tropical Front is pushed more to the south and is situated around 5°N (figure 4); the entire North Cameroon is invaded by the Harmattan. The effects of the Harmattan are very drastic in the north but become less intense as one moves southwards. Rainfall is rare and some small streams dry up completely while many river sizes reduce.

From April to October, however, when much of the country is overlaid by a wedge of humid Atlantic air, Cameroon experiences its rainy season. The Inter-Tropical Front is above Lake Chad and the humid Atlantic air is found practically above all Cameroon (figure 5). Maximum totals are recorded in July and August, but the nature and the quantity of rainfall vary significantly with the depth of moist air.

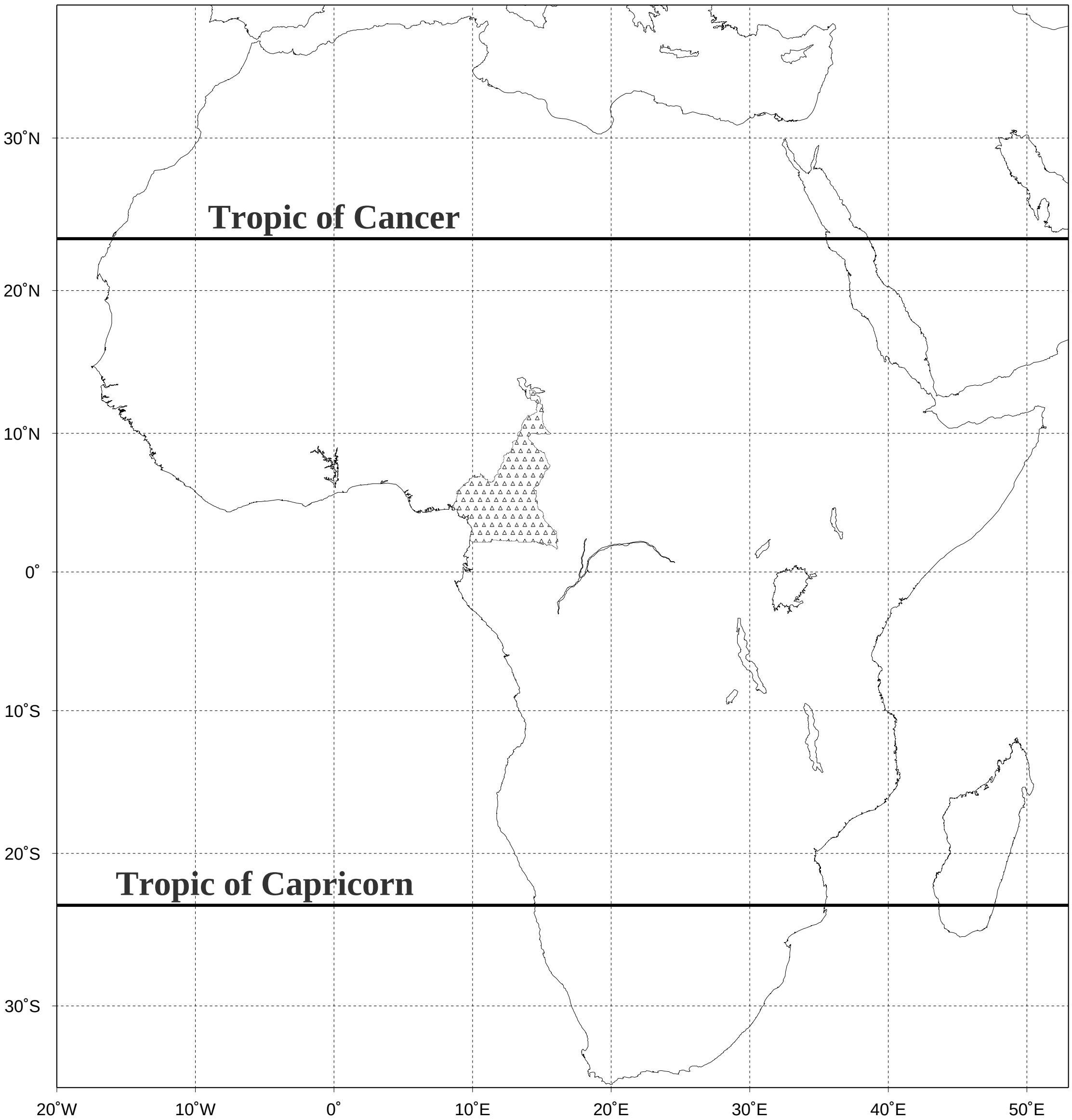
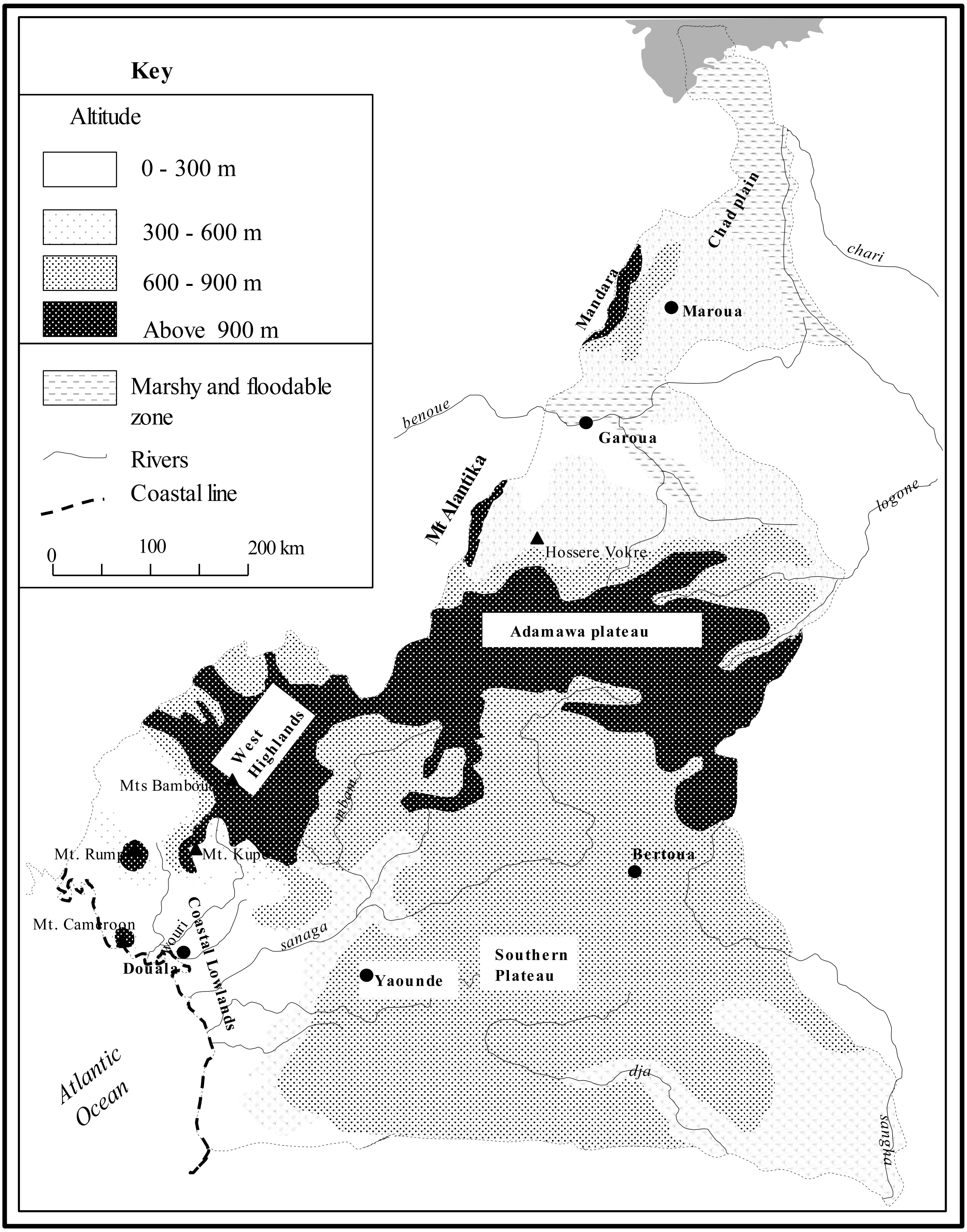


Figure 2: Geographical location of Cameroon



Adapted from Suchel J.B.

Figure 3: Physical map of Cameroon

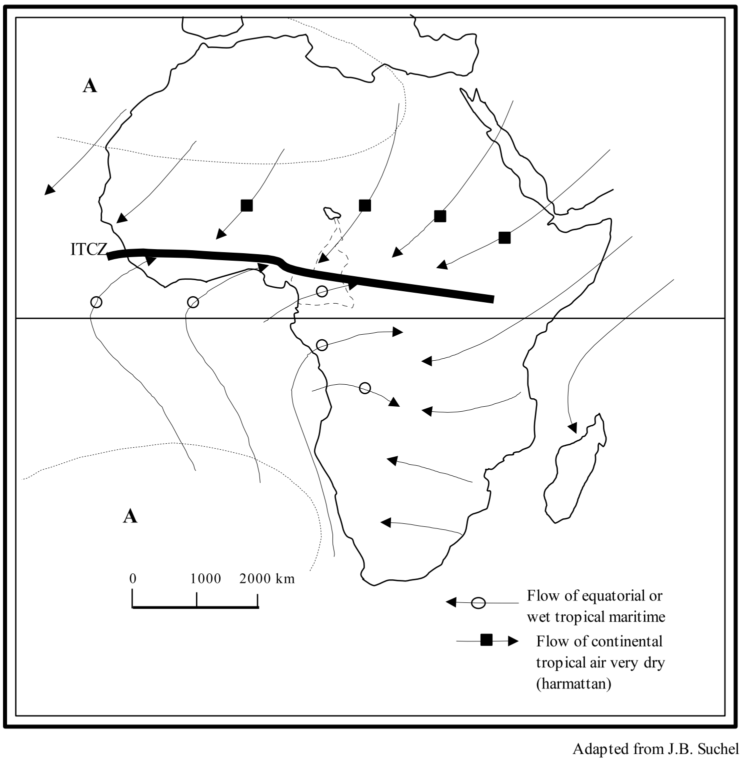


Figure 4: Centres of action and flow: average situation of January

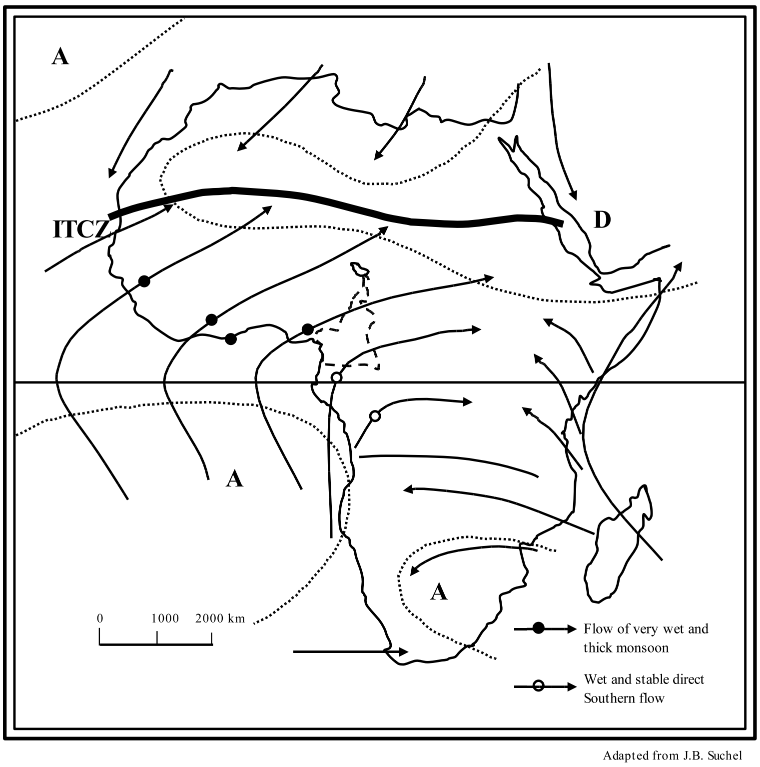


Figure 5: Centres of action and flow: average situation of July

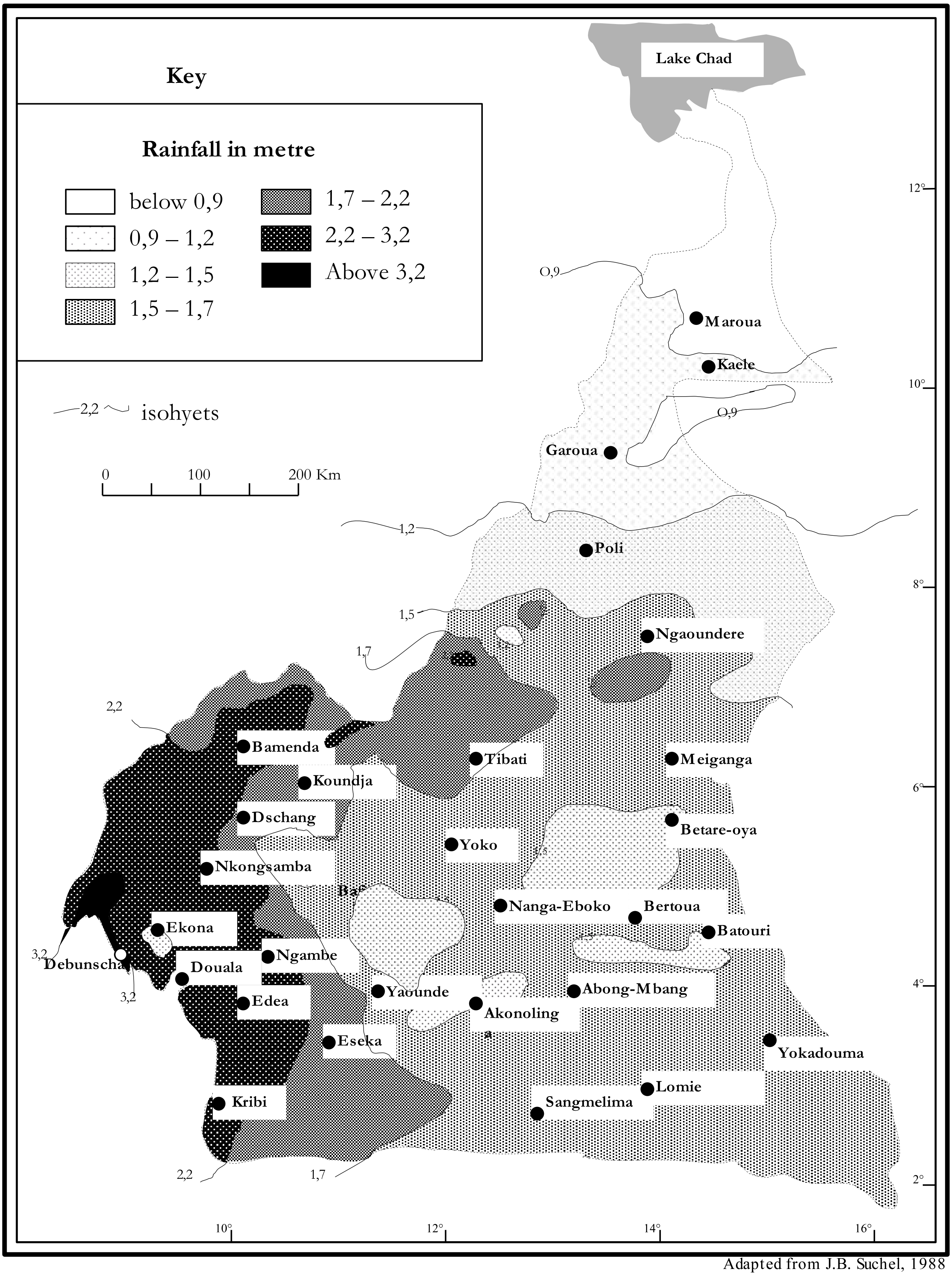


Figure 6: Spatial distribution of the mean yearly rainfall

2.2 Data

2.2.1 Daily rainfall data

Data from 28 meteorological stations, indicated by stars in Figure 7, are used in this study. All these stations are part of observer networks operated by the national department of meteorology in Cameroon which reports daily rainfall. For most of these stations, data from the years 1951 through 2000 are used. The data are taken as is relying on the routine checks done by the meteorological service; the length of the series varies from one station to another, with a lot of lacuna within the series as reported in table 1. The frequency of daily observation is the same throughout the network (twice per day). When rainfall occurs during the day (6H00:18H00), observations are reported at 18.00H. When it occurs in the night (18:00H-6:00H), observations are reported at 6H00. The measurements are carried out with Hellmann rain gauge type at all the 28 stations.

Like all measurements, the measurement of the rain is prone to error. The average level of urbanization in Cameroon in the years eighties was around 38 % especially in 1985; during the year 1990, it progressed to 42 % and reached 52% in year 2000 (UNEP 2002). The urban population who was 9.8 % in 1950, increased at 44.9 % in 1995 (UNEP 2002). The consequences of the rapid urbanization and urban population growth include intensifying pressure on natural habitats and resources to satisfy the growing demand of space and housing. Municipalities are unable to provide space and housing quickly enough to meet this demand; therefore, unplanned settlements emerged even nearby the meteorological stations. The construction of the houses nearby the stations is a source of systematic errors present in rainfall series in the sense that this forms a screen around the station and thus there is always less rain measured than has fallen.

The observation time is also a significant factor introducing "noise" into both the occurrence process and in the distribution of rainfall depth; in rain gauges there is always, a delay between the rain falling, collecting in the gauge, and recording. In Cameroon's stations, the observers read the gauge twice a day as stated before. Between the first (morning) and second observation (evening), some rain may have evaporated because of the exposition of gauges; the consequence is the series present a reduced number of wet days or greater mean daily amounts; the latter can be explained by the fact that an observer has missed making an observation on one day and the total accumulation is recorded on the following day.

Another problem is the split phenomenon; when one looks on the synoptic table of rainfall, there is continuous rain over several days. All these days have high rainfall amounts of nearly the same order of magnitude. This problem becomes serious when testing a distribution to describe the daily rainfall amount; the long rain periods overlapping several days to cause the misfit of the Weibull distribution (see chapter 5). Hence, the distribution predicts much fewer big rain amounts the higher the rainfall is.

In Cameroon the 1987's economic crisis, lead personal of meteorological service to a certain "laissez aller" because the government cut down salaries; the qualities of observations have been deteriorated since then. The effect of these methodological factors and errors is primarily reflected in the occurrence process and in the smaller amounts of rainfall. In spite of these problems, daily rainfall data for 28 stations are used in this study.

2.2.2 Classification of daily rainfall data

Before starting the analysis of daily rainfall data, the first work is to use the existing literature to define a wet or a dry day according to a certain threshold.

In tropics, because of high evaporative demand, a fairly high threshold might seem appropriate. However, choice of such threshold is still a problem and it may be unwise to ignore light rain that may be important for plant survival. Marked spatial variation is a characteristic of tropical rainfall. Although amounts at one point may be low, a very short distance away they may be considerable or zero. Particularly in the applied sense, concern is with rainfall over an area rather a point. This perhaps suggests that choice of appropriate threshold may not be so critical. Although later analyses may involve a range of thresholds, in this study we use the method of classification presented by Jackson (1981) and Stern et al. (1982). Briefly, for each site each day is classified as dry (wet) day depending whether the site received less (more) than 1mm of rain. In general, the probability of a day being dry will depend on the state (dry or wet) of the preceding day. If the dependence extends to the two previous days, the process is second order, etc. Only two states have been specified, but one could divide days into dry and various states of wetness, with appropriate rainfall amount to separate them. For N years data having classified each day as dry or wet, they are then relabelled as the case may be as: dry following dry, wet following dry, dry following wet and wet following wet. For every date of the civil calendar year, the number of occurrences i.e. transition numbers of each combination is tabulated for each month in all the years. We note them $n_{00}(m)$, $n_{01}(m)$, $n_{10}(m)$, $n_{11}(m)$ respectively.

2.2.3 Spatial data

The geographer main concern is to study the spatial distribution of phenomenon on the earth surface; spatial distribution of the different seasonal course of rainfall occurrence and rainfall amount are our main preoccupation in this study. This is why we try to describe the seasonal variation by time invariant parameters. It's in order to describe its spatial variation and for the latter we need the spatial predictors. The study of areal distribution of rainfall is important especially when physiographic and other factors vary over the same region causing different rainfall patterns. As the geographical factors are time invariant, that is their patterns cannot experience any temporal change and are easy to observe, we use them to predict rainfall in Cameroon. These data are derived from the topographic maps and/or direct observations in the field. This category of data are time invariant because their patterns cannot experience any temporal change. Spatially distributed data in this study are known as predictors. They include latitude, longitude, altitude, local topography (relief orientation, rain exposition). Mean yearly rainfall amounts are also used as spatial data.

It is well known that when latitude increases, rainfall decreases; that is the higher the latitude, the less the rain. The longitude has the same effects as the latitude, with lower rain when the longitude increases.

As previously stated, the topography of Cameroon is dominated by the chain of highlands. The highlands strongly influence precipitation patterns, generating high levels of orographic rainfall when dominant winds are uplifted as they encounter the relief. Incursion of the moisture from the coastal area, the orientation of the hills are the geographical factors influencing the spatial distribution of rainfall. The orientation and the altitude of a station above sea level are the additional factors responsible for exceptional rainfall. Thus precipitation is highest along the windward whereas the leeward lies in the shadow with low precipitation. This influence of the topography on rainfall leads us to construct a measure called "rain exposition" $r(P)$ at the measurement point P in altitude $\omega(P)$. By whole numbers from 1 to 8 it measures the degree of rain exposition to $k = 1, 2, ..., 8$

equally distributed wind directions around P. In each of eight directions k we take the surface point H_k with altitude $\omega(H_k)$ where the gradient

$$\delta(H_k) = \frac{\omega(P) - \omega(H_k)}{d(P, H_k)} = max, \qquad (5)$$

related to the distance $d(P, H_k)$ between P and H_k is at maximum.

We define a function E_k of directional rain exposition.

$$E_k = \begin{cases} 0 & \text{if} \quad \delta(H_k) \geq 0 \\ 1 & \text{if} \quad \delta(H_k) < 0 \end{cases} \qquad (6)$$

Then the rain exposition $\kappa(P)$ is defined as the sum of all the eight directions E_k

$$\kappa(P) = \sum_{k=0}^{8} E_k \qquad (7)$$

The selection of good predictors is guided by stepwise linear regression displaying the strength of a relationship and the direction of this relationship between parameters of a distribution and the spatially distributed predictors. The dependent variable here are parameters of functions used to describe daily rainfall series.

Table 1: Geographical location of meteorological stations and the presentation of the time series

	Latitude N	Longitude E	Altitude (m)	Length of lacuna (Year)	Length of time Series (Years)
Abong-Mbang	3°58'	13°12'	693	1	42
Akonolinga	3°46'	12°14'	671	6	37
Bafia	4°44'	11°15'	500	5	38
Bamenda	6°13'	10°07'	1239	5	38
Batouri	4°28'	14°22'	653	complete	43
Bertoua	4°33'	13°43'	650	complete	43
Betare	5°36'	14°04'	815	13	30
Douala	4°00'	9°44'	5	complete	43
Dschang	5°20'	10°03'	1407	12	31
Edea	3°46'	10°04'	31	3	40
Eseka	3°37'	10°44'	228	6	37
Garoua	9°20'	13°23	242	3	40
Kaele	10°05	14°24	389	complete	43
Koundja	5°39'	10°45'	1208	1	42
Kribi	2°57'	9°54'	10	1	42
Lomie	3°09'	13°37'	624	4	39
Maroua	10°27'	14°15'	421	4	39
Meiganga	6°32'	14°22'	1027	1	42
Nanga-Eboko	4°39'	12°24'	622	4	39
Ngambe	4°16'	10°36'	610	1	42
Ngaoundere	7°21'	13°13'	1114	1	42
Nkongsamba	4°57'	9°56'	816	2	41
Poli	8°28'	13°15'	436	1	42
Sangmelima	2°56'	11°32'	712	complete	43
Tibati	6°29'	12°36'	873	3	40
Yaounde	3°52'	11°32'	759	complete	43
Yokadouma	3°31'	15°06'	534	1	42
Yoko	5°33'	12°22'	1027	1	42

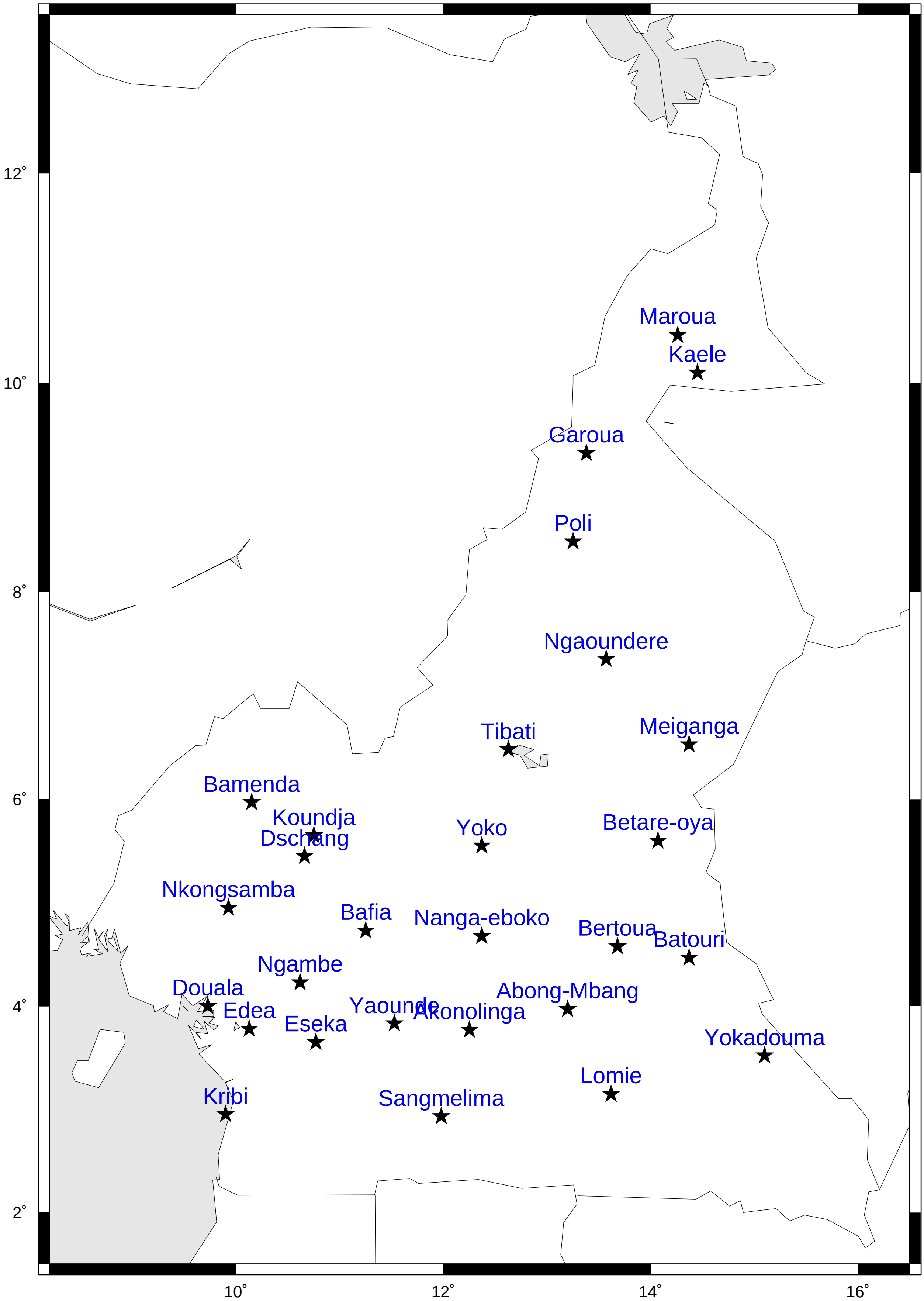


Figure 7: Spatial distribution of meteorological stations in Cameroon

3 DISTRIBUTION OF MONTHLY MEAN RAINFALL

3.1 Spatial distribution of mean rainfall in some significant months

The spatial distribution of rainfall in January, April; July and October presents an idea on the rainfall regime in Cameroon. These months are significative because January represents the heart of the dry season, April the transitional month between the dry season and the start of the rain; July represents the rainy season and corresponds also to the "little dry season" in the southern plateau; October represents the second rainy season in the southern plateau and is also a transitional month between the end of the rainy season and the begining of the dry season. The analysis here is based on the maps drawn by Suchel (1988).

December and January are the months with the least amounts of rainfall in Cameroon. The distribution of rainfall during the month of January is presented in figure 8. Generally, rainfall amounts reduce from the south towards the north in the inland regions when the ITCZ is around the 5^{th} parallel. When the south of the country is fairly dry in January, in the same month most of Adamawa and the rest of the north are dry.

From figure 9, in April, while the major part of the country receives much rainfall, the northern part keeps partly dry (less rain). The Douala meteorological station measures an average rainfall of 236 mm, which is not so much more than that of Eseka (242 mm) or that of Bamenda (200 mm). The reason is that the thermal depressions are not yet significantly or fully established. Most of the rains are as a result of frontal activities. During this time, there is a reduction of rainfall amounts not only from the south towards the north but from west towards east (Bamenda 200 mm, Betare-oya 131 mm) which is an indication that the ocean winds have enough power of perturbation over the influence of altitude. Rainfall distribution in May is relatively stable compared with that of April. These two months represent a characteristic of the seasonal cycle in most of the major climatic regions of Cameroon: in the centre-south, this is a season of little rain, in the littoral and Western Highlands, there is moderate rain which precedes the monsoons; in the north, it is the beginning of the rainy season; characterized by the fall of isolated heavy drops of rains which are very violent and sometimes very destructive.

August has a very different distribution of rainfall. The contrast is remarkable especially in the western part of the country. The coastal region is the Cameroonian "ridge " (High pressure ridge) which experiences strong monsoon during this time. The moisture-trade winds are very heavy around the littoral regions, which reach high level of humidity(Douala 760 mm). This heavy rains resulting from the monsoon winds continue till October. It is in the month of July (figure 10) and August that the western edge of south Cameroon acts as a boundary between southwest monsoon and the Northeast trade winds. This zone is relatively stable and cool provoking a "small dry season" in the centre-south. While the August rainfall is still as high as 520 mm in Ngambe, rainfall in Akonolinga does not measure more than 108mm. In North Cameroon, it is a full rainy season. Maximum rainfall occurs in the month of August in this region.

In October (figure 11), the distribution of rainfall presents original characteristics and seems to resemble that of the month of April. Traits of the ITCZ indicating the end of the rainy season present a transitional situation similar to that of April. However, the month of October has much rain wherever compared with April. The differences in the distribution of rainfall during these months are more distinct. In the humid air, the storms are concentrated in the coastal areas where raindrops intervene in the last monsoons. This is the period of high rainfall and of a peak in the rainy season in the entire Centre South plateau from the Adamawa to the Eastern boundaries. Yaounde, Lomie, Nanga-Eboko etc. collect more than 300 mm of rainfall.

3.2 Rainfall regimes

This section is not an exhaustive description of rainfall regimes in Cameroon. This has been well developed by Suchel (1988). It is however important to identify some important points for better understanding of rainfall patterns in Cameroon. The focal interest is to have a general idea on monthly distributions of rainfall in Cameroon.

The study of rainfall regimes (figure 12) shows two maximums and two minimums of rainfall records during a year in Yaounde, Sangmelima, Lomie, Bertoua, Abong-Mbang, Akonolinga, Bafia, Eseka, Betare Oya, Nanga-Eboko, Yoko, Kribi and Batouri. The main maximum occurs during the principal rainy season with the peak occurring in October. The secondary maximum, which is longer than the first but less intensive, occurs in May and sometimes in April and exceptionally in June; it occurs during the small rainy season. Between these two maximums, there is an intermediary season linking the two.

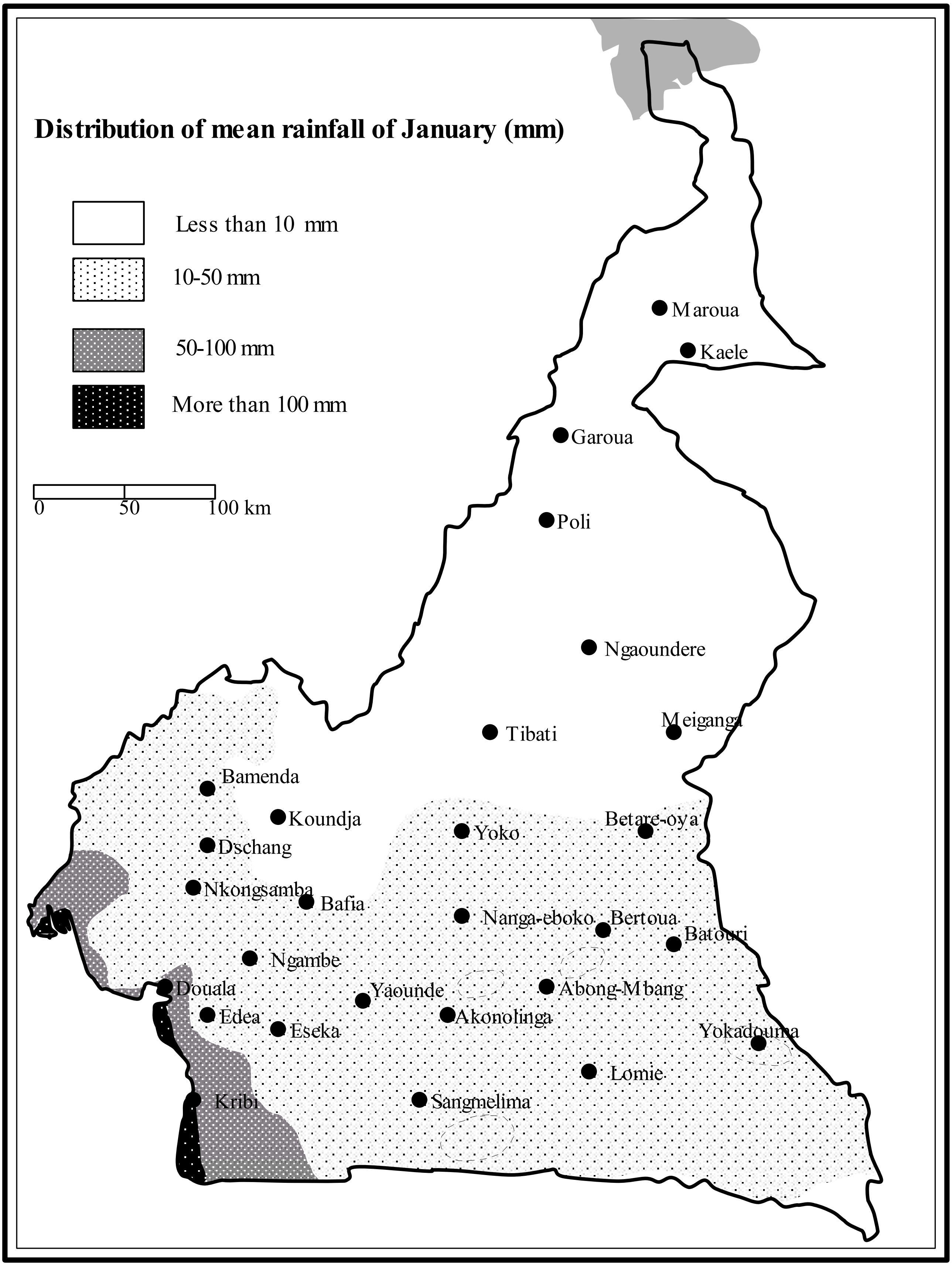
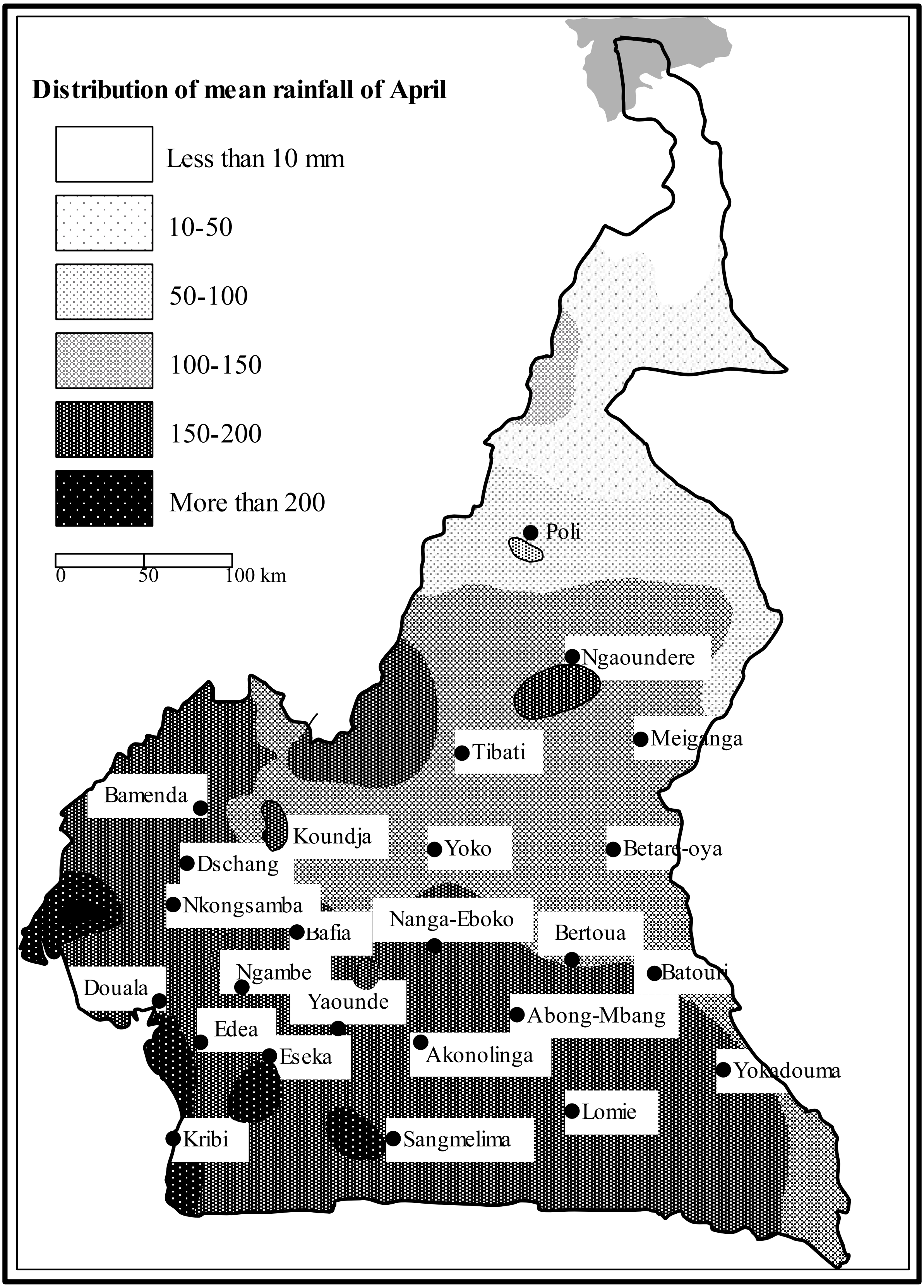


Figure 8: Spatial distribution of rainfall in January



Adapted from J.B. Suchel

Figure 9: Spatial distribution of rainfall in April

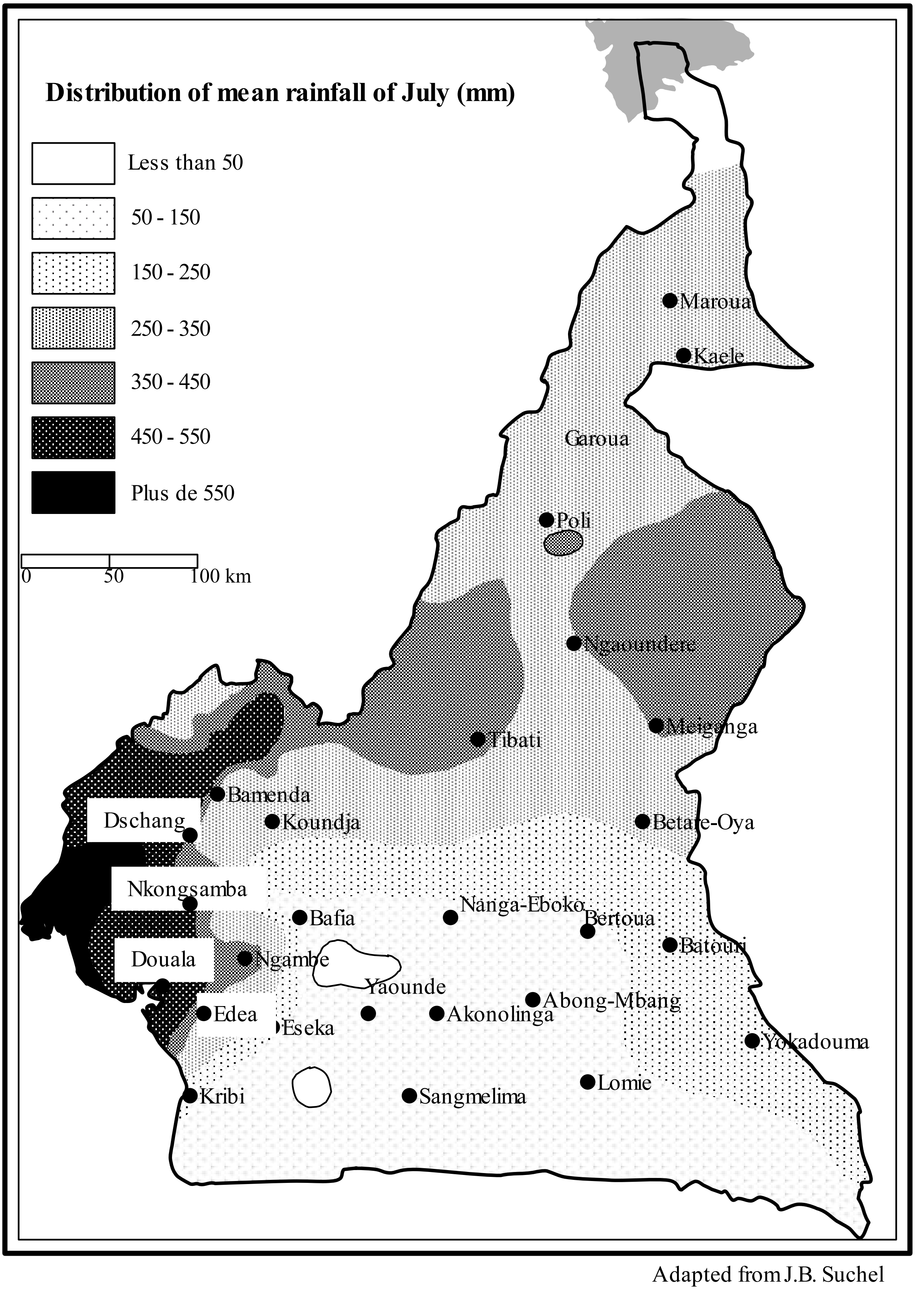
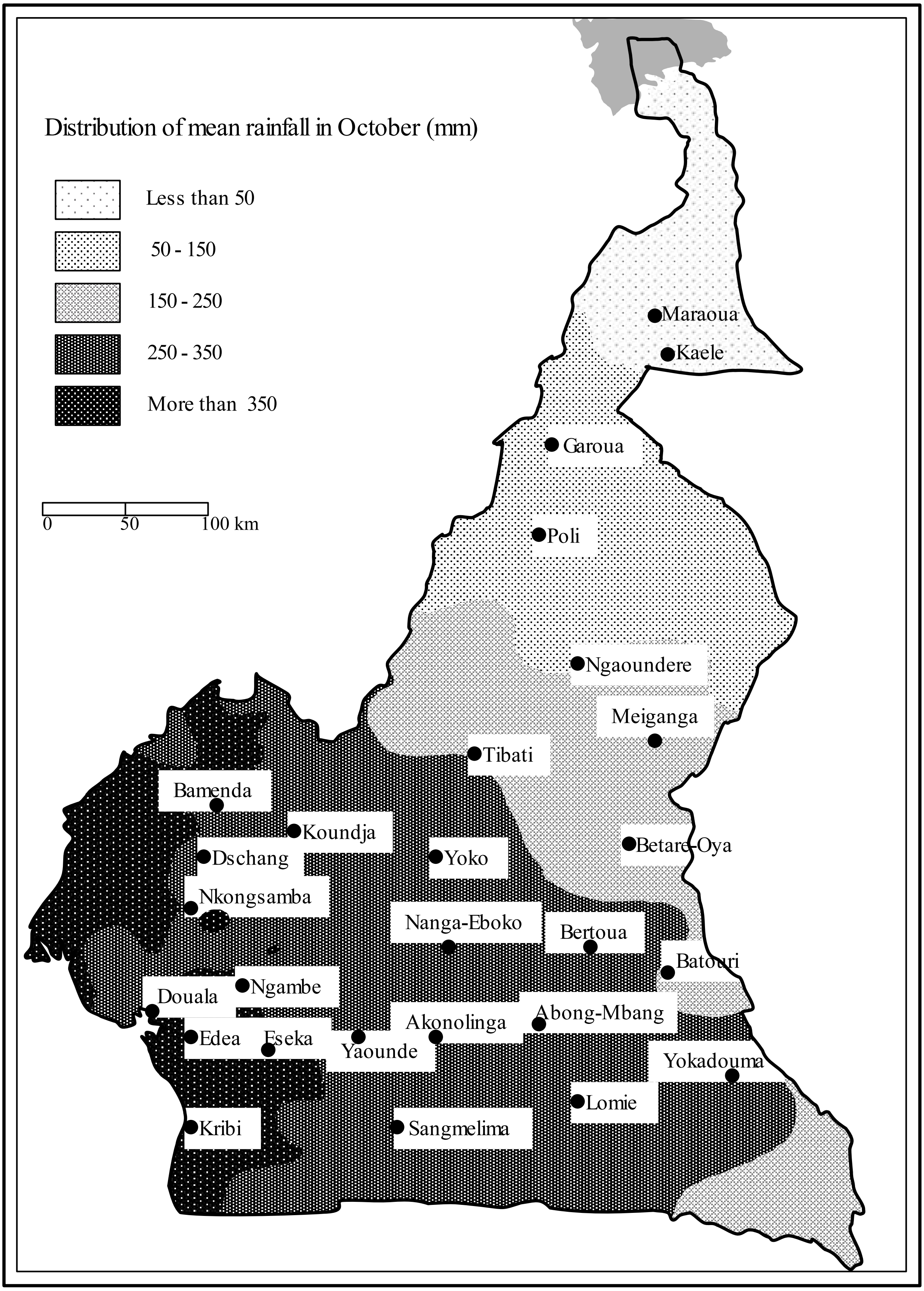


Figure 10: Spatial distribution of rainfall in July



Adapted from J.B. Suchel

Figure 11: Spatial distribution of rainfall in October

It is usually called a "small dry season" and extends from July to August.

Douala and Nkongsamba (figure 13) have a common regime with maximum rainfall in August. From November to January, there is a decrease in the rainfall depth. Koundja, Bamenda, Edea, Meiganga, Ngambe and Tibati, have also a unimodal distribution but the maximum occurs in the month of September. Less rainfall amount during the months of May and June gives rise to 'stairs' shape distribution. December to February is marked by a substantial fall in rainfall.

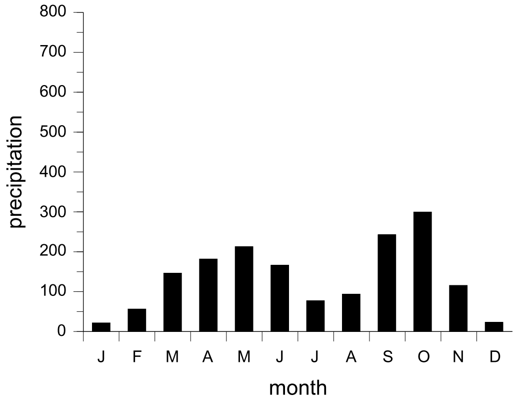
The five months (November to March) period of dry season at Ngaoundere is a characteristic of the Sudanese humid climate. The distribution of rainfall has a dome-shaped profile with a maximum occurring in August. The "frank" Sudanese climate type imposes at the Garoua and Poli a profile which is more vigorous (inselberg) with maximum rainfall occurring in August. The rainfall regime presents six wet months from May to August and six dry months from November to April.

At Kaele and Maroua, the rainy months are reduced to 5 against 7 dry months where 5 months record a zero amount of rainfall. It is from July to August that these two stations collect 37% of their rainfall; maximum rainfall occurs in August.

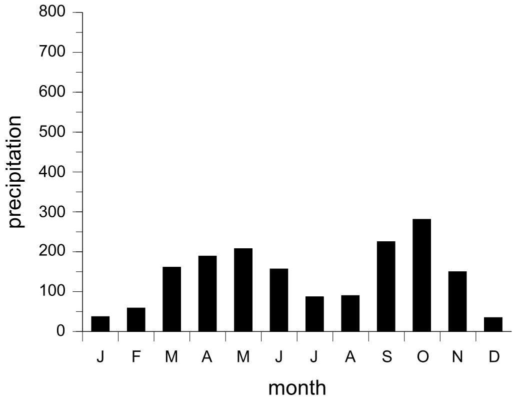
An analysis of spatial distribution of rainfall regimes (figure 14) shows two contrasting areas in the distribution of rainfall regimes in Cameroon. In a general view, from far north near lake Chad to the southern boundaries of the country there are zones of regimes conforming in its major aspects to a classical schema of gradation of Africa inter-tropical climate, from sahelian climate to equatorial climate, passing through Sudanese and subequatorial types. The second area is limited to the littoral area and western highlands, where oceanic monsoon effects and relief cause considerable instability. The regimes in these zones are not only complex in nature but there is a lot of variation on it.

The motivations for building a digital simulation model for daily rainfall, the area of study, the description of the necessary data are presented in the previous sections, as well as the spatial and temporal distribution of the monthly mean rainfall. It is obvious that for simulating daily rainfall series, we need first of all to build specific models; thus developping models for simulating dry-wet occurrence and wet-wet occurrence is the main preoccupation in the next section of this study.

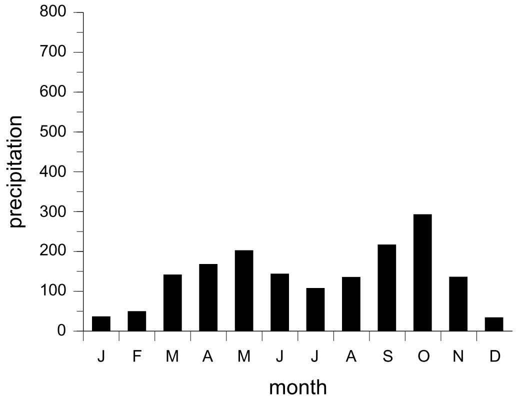
Yaounde



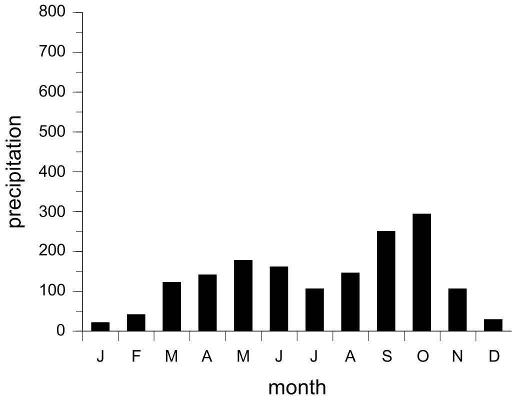
Sangmelima



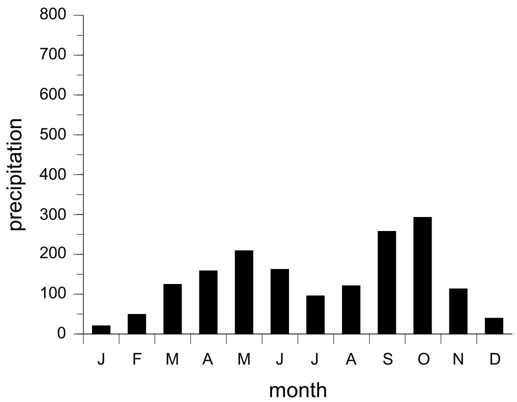
Lomie



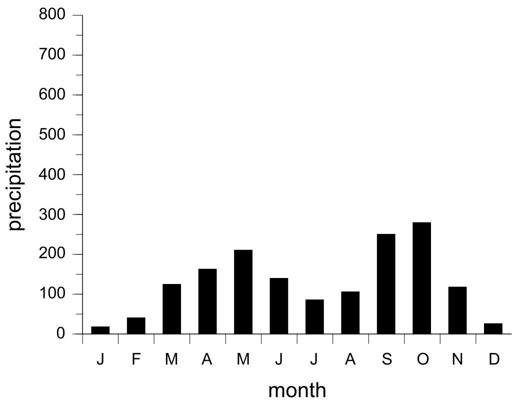
Bertoua



Abong-mbang



Akonolinga



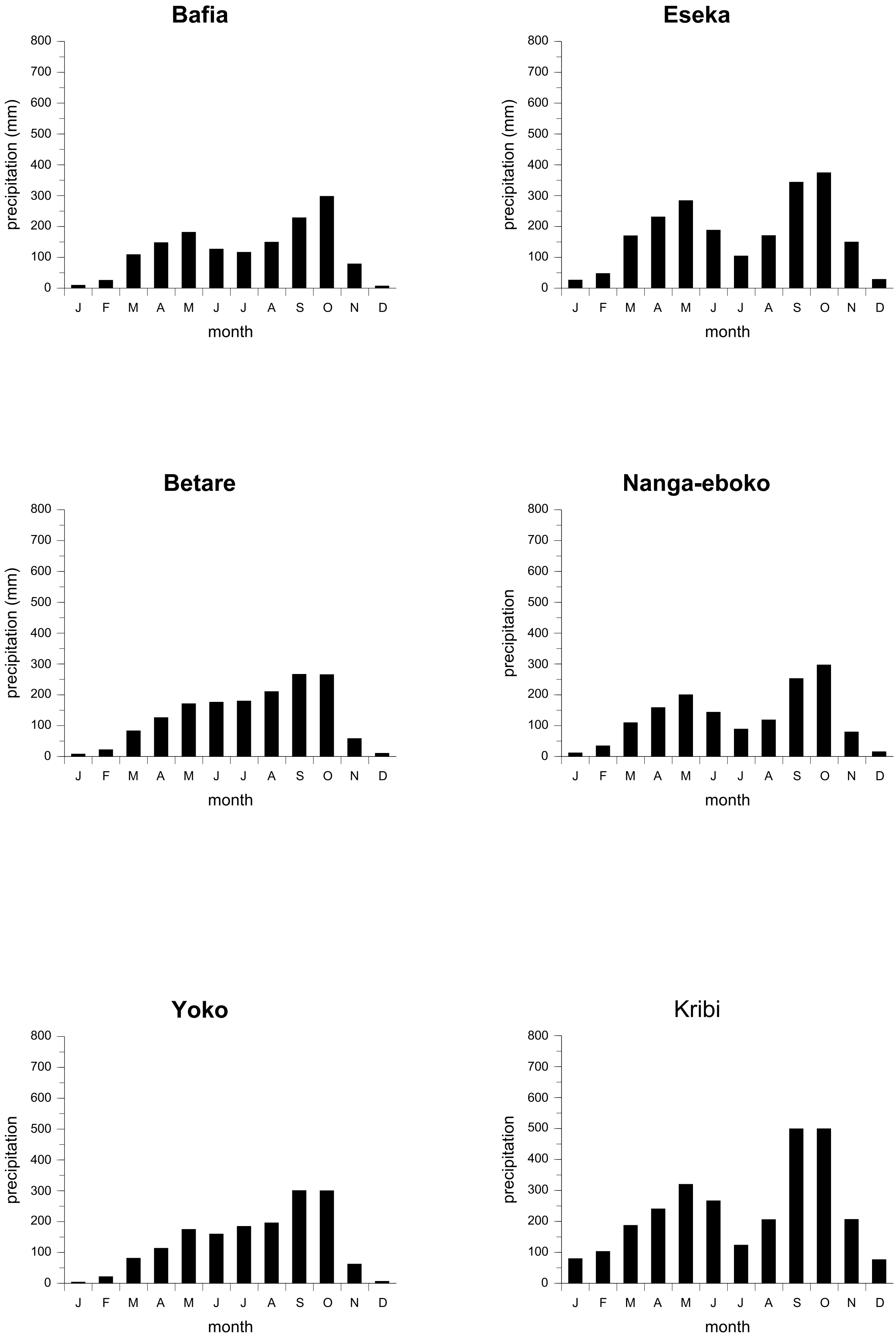
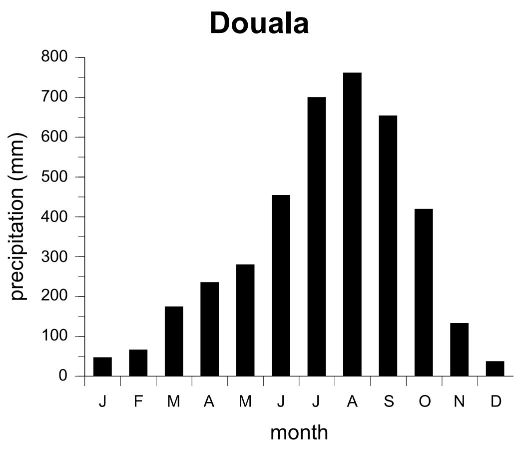
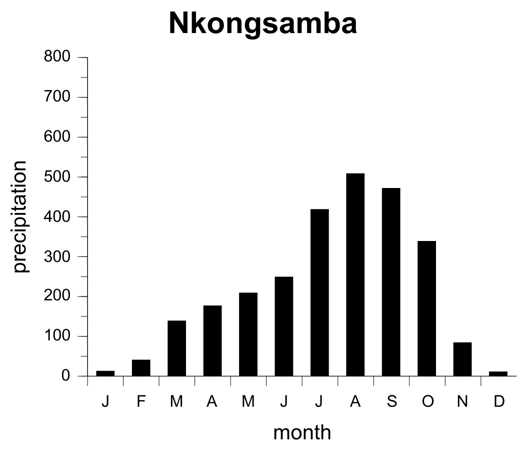
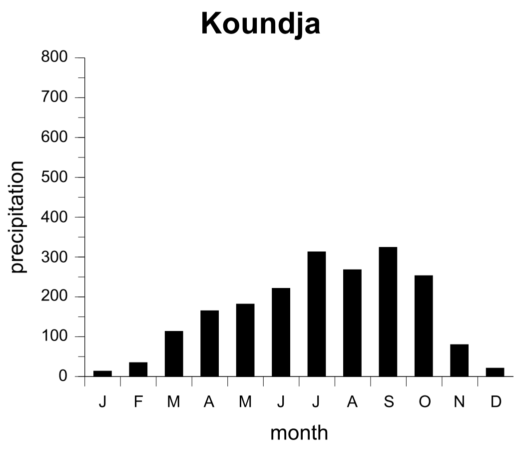
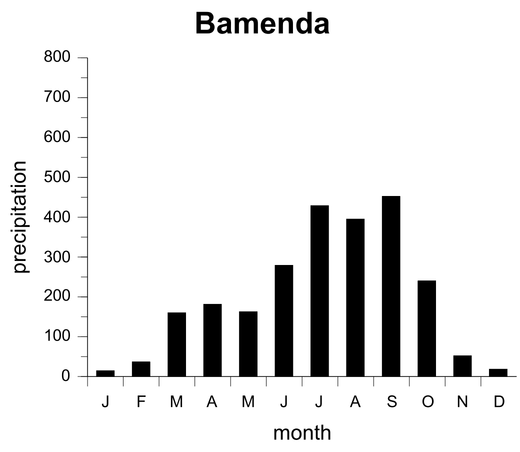


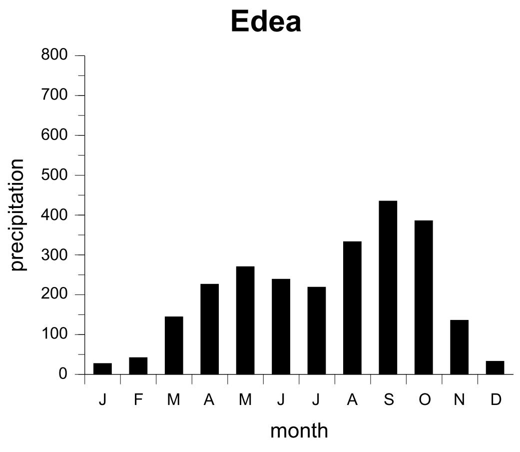
Figure 12: Rainfall regimes with four seasons

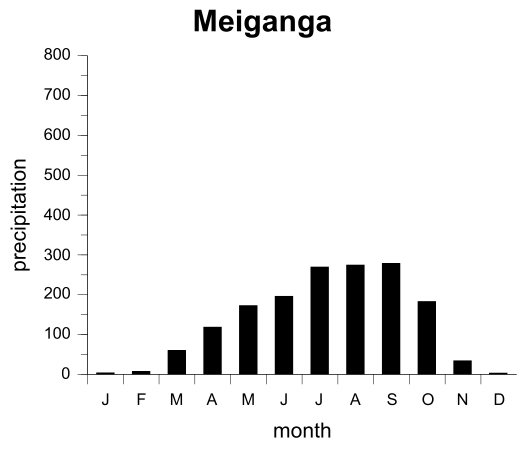

Douala
precipitation (mm)
0 100 200 300 400 500 600 700 800
J F M A M J J A S O N D
month


Nkongsamba
precipitation
0 100 200 300 400 500 600 700 800
J F M A M J J A S O N D
month


Koundja
precipitation
0 100 200 300 400 500 600 700 800
J F M A M J J A S O N D
month


Bamenda
precipitation
0 100 200 300 400 500 600 700 800
J F M A M J J A S O N D
month


Edea
precipitation
0 100 200 300 400 500 600 700 800
J F M A M J J A S O N D
month


Meiganga
precipitation
0 100 200 300 400 500 600 700 800
J F M A M J J A S O N D
month

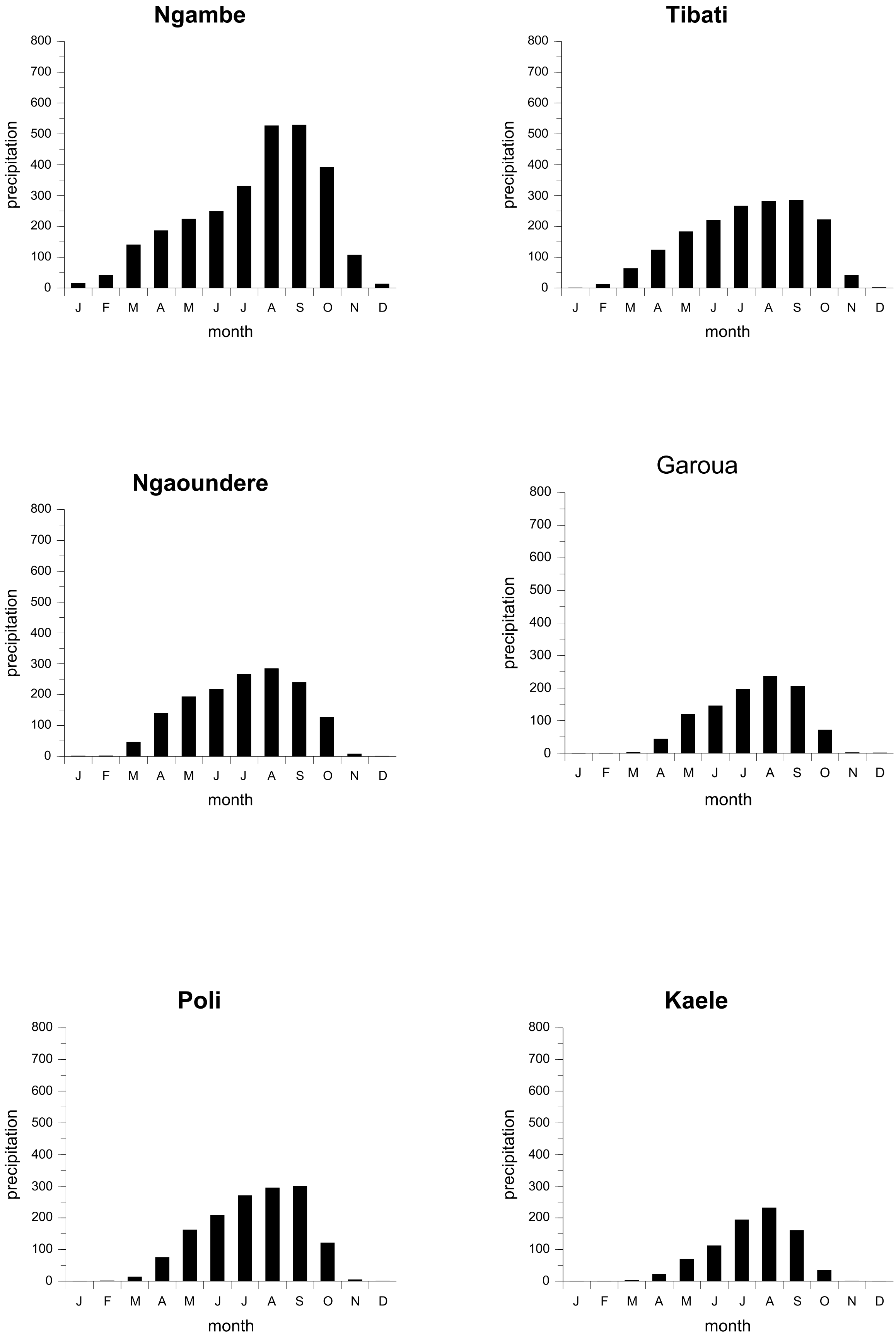


Figure 13: Rainfall regimes with two seasons

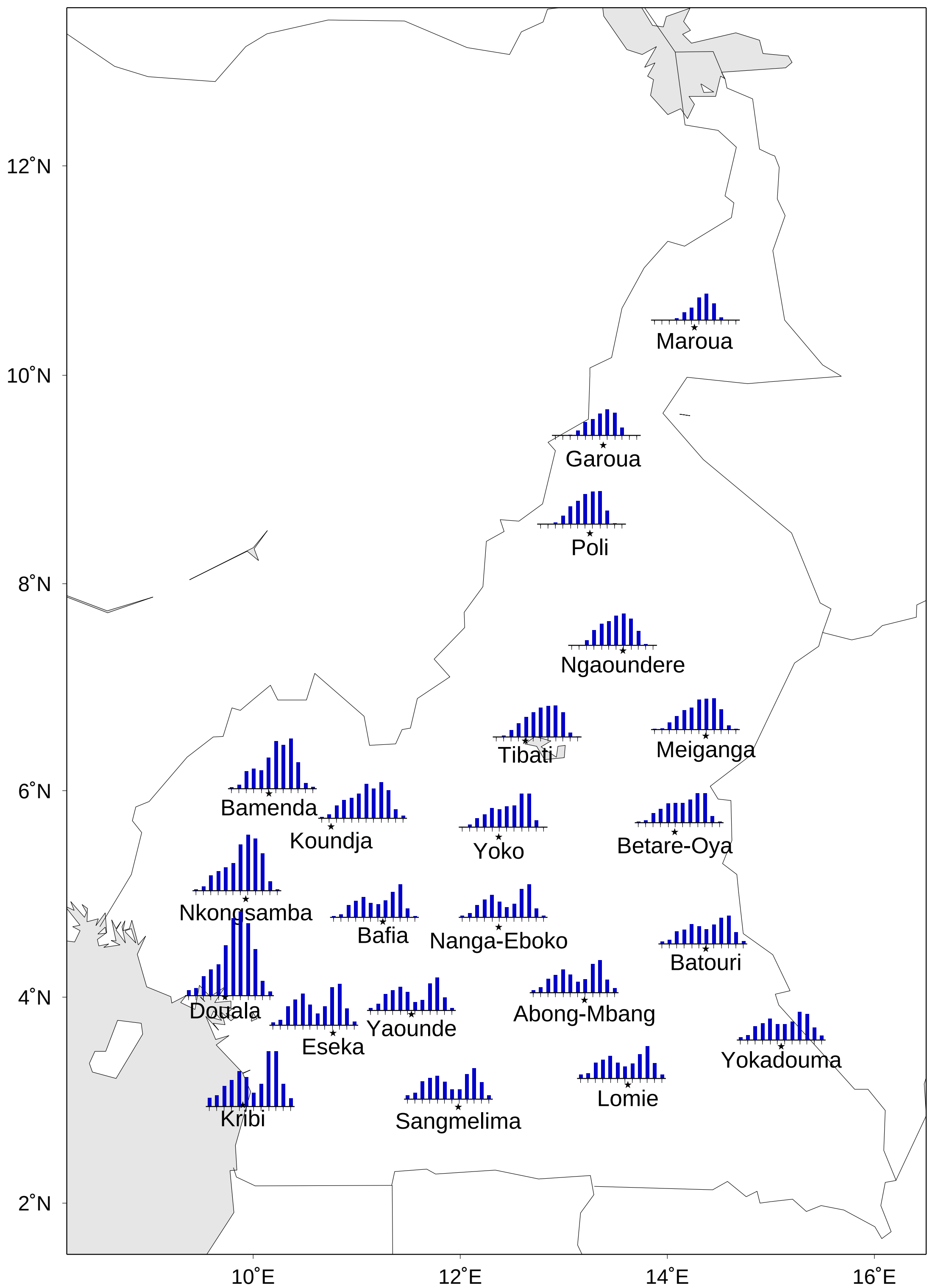


Figure 14: Spatial distribution of rainfall regimes

4 MODELLING DAILY RAINFALL OCCURRENCE

A rainfall generator uses information about the observed weather statistics at the site of interest to parameterise a stochastic tool that can generate sequences of daily rainfall data, all consistent with the statistics of the target site. A typical daily rainfall generator is often rooted in a Markov Chain model. The Markov Chain is a widely-used statistical technique to describe rainfall time-series. For a rainfall generator, we define two states: either the day is classified as dry (no rain) or rainy. The Markov Chain model considers that the likelihood of a particular state on any given day is determined by the states taken in the previous day or sequence of days. Therefore, the likelihood of any given day being rainy is conditioned only by whether the previous day or sequence of days were rainy. For estimating transition probabilities however some previous work has to be carried out before.

4.1 Choice of appropriate order of Markov chains

Following tabulation of frequencies of occurrence, we constructed contengency tables of observed and expected numbers of occurrence in order to test a range of hypotheses about the order of Markov chain. The following hypotheses are formulated for chosen the appropiated order of Markov chain to be used

1. Ho: the weather on day t is independent of weather on day $t-1$
2. Ho: the weather on day t is independent of weather on days $t-2$
3. Ho: the weather on day t is independent of weather on days $t-3$
4. Ho: the weather on day t is independent of on weather on days $t-4$
5. Ho: the weather on day t is independent of on weather on days $t-5$

The χ^2 test is used for this purpose. It compares the observed and the expected data and tell us whether they are significant or not. The first hypothesis is a test of independence of days. Rejection of the null hypothesis implies that the situation on any day is not independent of the situation on previous days (i.e. at least a first order Markov chain). Rejection of the second hypothesis implies that the order is at least two and the

Table 2: The general form of a contingency table

	x_t		
x_{t-n}	w	d	Total
w	f_{11}	f_{01}	$f_{11} + f_{01}$
d	f_{10}	f_{00}	$f_{10} + f_{00}$
Total	$f_{11} + f_{10}$	$f_{01} + f_{00}$	$f_{11} + f_{10} + f_{01} + f_{00}$

rejection of the third hypothesis implies that the order of the chain is at least three. The general form of a contingency table is presented in 2.

where

$$X_t = \begin{cases} 0 & dry \quad at \quad day \quad t \\ 1 & wet \quad at \quad day \quad t \end{cases}$$

Used in this form, with $n = 1$, tests hypothesis 1. Two tables are used in separate tests of hypothesis 2, with $n = 2$. One table involves occurrences when $(t-1)$ is dry, the other when $(t-1)$ is wet. They provide independent tests of the same null hypothesis which can be stated as follows (Feyerherm and Bark 1967): Given the weather dry or wet on day (t), the weather on day t (dry or wet) is independent of the weather on day $(t-1$. In the same way, with $n = 3$, there are four independent tests, one for each combination of weather (dry or wet) on days $(t-1)$, $(t-2)$ to determine if weather on day t is independent of weather on day $(t-3)$ etc.

In these tests, choice of an appropriate significance level for acceptance or rejection of the null hypothesis is the subject of some debate. Gabriel and Neumann (1962), in analysing Tel Aviv data, used 5 % significance level and found that a first order chain was adequate. However, Gates and Tong (1976) point out that if the 10 % level had been adopted, this would not have been the case and using a different method they found that a chain of at least second order should have been used. Feyerherm and Bark (1967) amongst others, adopted a 10 % significance level which, by increasing the likelihood of rejection of null hypothesis, serves as more stringent test of acceptance of a particular order model. In this study, results are presented for 5 % significance levels.

The contingency table analysis shows the rejection of our null hypotheses. Hence, for deciding which order of Markov chain to use, it is necessary to determine how much information is included in any model-given order and compare this with other model orders. The most stringent method consists to weigh calculated χ^2 from contingency table against its critical value using the relation $\frac{\chi^2}{\chi^2_{criticalvalue}}$ for each order. Results are presented in table 3.

Table 3: ratio $\frac{\chi^2}{\chi^2_{critical\ value}}$ calculated from the contingency table

	First order	Second order	Third order	Fourth order	Fifth order
AbongMbang	148.3	120.6	83.9	52.2	33.2
Akonolinga	132.5	99.3	69.4	44.6	28.0
Bafia	123.6	123.7	90.72	57.4	35.5
Bamenda	533.6	358.2	224.0	130.5	74.4
Batouri	125.0	98.0	71.4	49.0	30.3
Bertoua	133.0	108.5	84.0	53.4	32.7
Betare-oya	226.3	161.7	118.1	71.6	44.2
Douala	450.0	323.4	214.6	131.5	79.4
Dschang	449.0	325.1	206.8	124.8	71.4
Edea	262.7	193.3	125.2	77.0	44.9
Eseka	280.5	195.4	127.3	81.1	48.6
Garoua	112.4	41.6	95.1	63.5	21.0
Kaele	167.5	135.3	133.3	92.32	57.5
Koundja	539.2	404.9	273.39	162.1	93.5
Kribi	234.3	144.1	92.6	59.8	37.2
Lomie	67.8	64.4	48.2	34.0	21.9
Maroua	140.0	70. 8	114.4	76.9	19.9
Meiganga	480.0	367.2	243.5	150.9	86.7
Nanga Eboko	148.2	116.2	80.9	49.5	29.6
Ngambe	739.7	514.5	332.9	198.8	116.3
Ngaoundere	626.0	459.8	296.8	176.8	101.2
Nkongsamba	594.9	441.2	290.2	173.4	102.4
Poli	335.6	317.1	222.3	137.9	81.3
Sangmelima	129.6	113.3	79.1	51.4	31.7
Tibati	457.8	357.8	240.2	145.5	84
Yaounde	214.4	160.7	109.5	67.5	41.9
Yokadouma	60.9	60.3	50.9	35.2	23.3
Yoko	255.6	222.2	162.3	101.3	59.6

The analysis of table 3 shows that the amount of information decreases from first-order to fifth order. Thus, for all the orders weighed, the first-order chain presents more information than others; i.e. the first-order chain is satisfactory than other orders to model daily rainfall in Cameroon. The second, third, and greater orders are less likely to be satisfactory. This leads to the conclusion that when the weather is dry, the following day is more likely to be dry than rainy, and vice versa so that the probability that it rains is conditional in the past.

4.2 Temporal distribution of transition probabilities

To carry out a Monte-Carlo simulation of rain occurrence, probability density functions are needed; that is why transition probabilities p_{11} and p_{01} are computed in this study. Having hypothesized that the state (dry or wet) of a day depends on the state of the previous day, we are interested in calculating the probability wet to wet. The transition probabilities are estimated for each month m of the whole series of years in order to represent the seasonal changes. Consider for instance the transition from wet on day $t-1$ to wet on day t and the transition from dry on day $t-1$ to dry on day t. There were $n_{11}(m)$, cases respectively of wet following wet, $n_{00}(m)$ dry following dry in the data. The total number of the year when it rained on day t is $n_{10}(m)+n_{11}(m)$ and the total number of the year where is no rain on day t is $n_{01}(m)+n_{00}(m)$. The required probability is the ratio of the two quantities, noted $p_{11}(m)$:

$$p_{11}(m) = \frac{n_{11}(m)}{n_{01}(m)+n_{11}(m)} \tag{8}$$

Likewise, the dry to wet transition $p_{01}(m)$ is given by:

$$p_{01}(m) = \frac{n_{01}(m)}{n_{01}(m)+n_{00}(m)} \tag{9}$$

It is only necessary to compute the transition probabilities for $p_{01}(m)$ and $p_{11}(m)$ and use those values to derive the other two probabilities, i.e $p_{00}(m) = 1 - p_{01}(m)$ and $p_{10}(m) = 1 - p_{11}(m)$ for dry to dry and wet to dry transition probability respectively.

The temporal distribution $z(m)$ of monthly transition probabilities which means its seasonal variation over the year is described by a double normal function (Schwarz 1980)

$$z(m) = a\left(exp(-(\frac{m-\mu_1}{s_1})^2) + exp(-(\frac{m-\mu_2}{s_2})^2)\right) \tag{10}$$

where

a is a scale parameter,

μ_1 is the first mean belonging to the first normal distribution function,

μ_2 is the second mean belonging to the second normal distribution function

s_1 is the standard deviation of the first normal distribution function

s_2 is the standard deviation of the second normal distribution function.

The idea behind the double normal function is based on the assumption of two different rainy seasons within the year, each one with a peak similar to that of a normal function. But if there is only one rainy season, we expected that this could also be represented by nearly identical values of the two parts of a normal distribution.

Double normal function values are not probabilities needed for the Monte-Carlo simulation of daily rainfall occurrence as previously stated. Such probabilities needed for the Monte-Carlo simulation of the daily rainfall occurrence are computed by transforming the transition probabilities $p_{i1}(m)$ into values $Y_{i1}(m)$ using a logit function i.e. we estimate the $Y_{i1}(m)$ by $z(m)$, and transform the estimations back to probabilities by inverse of the logit function. The logit function used in this context is as follows:

$$Y_{i1}(m) = ln\left(\frac{1}{1-p_{i1}(m)^{0.125}}\right) \qquad i = 0, 1 \tag{11}$$

By this method, we only approximated the logits by the double normal function and not the transition probabilities. Once over our model is ready, the inverse probabilities are calculated by inverting Y_{i1} to

$$p_{i1}(m) = \left(1 - \frac{1}{exp(Y_{i1}(m))}\right)^8 \qquad i = 0, 1 \tag{12}$$

Coefficients of the double normal function are estimated by a minimization of errors. i.e. there is one error function e_i for the set of all five parameters a_i, μ_{i1}, s_{i1}, μ_{i2}, s_{i2}

$$e_i(a_i, \mu_{i1}, s_{i1}, \mu_{i2}, s_{i2}) = \sum_{m=1}^{12}\left(z(m) - Y_{i1}(m)\right)^2 \qquad i = 0, 1 \tag{13}$$

The parameters we are looking for, are the ones which minimize the function e_i. The estimated parameters of $Y_{i1}(m)$ are presented in table 4.

Table 4: Double normal function coefficients for $Y_{i1}(m)$

	$Y_{01}(m)$					$Y_{11}(m)$				
Station	a	μ_1	s_1	μ_2	s_2	a	μ_1	s_1	μ_2	s_2
AbongMbang	2.4	10.1	2.0	4.7	3.9	2.3	10.6	2.8	4.2	5.1
Akonolinga	2.3	10.1	1.9	4.7	4.1	2.2	10.5	2.6	4.3	5.3
Bafia	2.2	9.9	1.9	5.0	4.2	2.1	10.5	2.8	4.1	5.8
Bamenda	2.7	9.2	2.3	5.1	3.7	2.4	9.0	3.4	4.1	6.7
Batouri	2.2	10.1	2.1	4.8	4.3	2.1	10.3	2.9	3.9	5.7
Bertoua	2.2	10.1	2.0	4.8	4.3	2.1	10.4	2.8	3.9	5.9
Betare-oya	2.3	9.7	2.0	5.1	3.9	2.3	10.2	2.9	4.5	4.9
Douala	2.5	9.5	2.7	4.6	4.4	2.2	8.1	11.1	6.2	2.4
Dschang	2.5	9.3	2.6	4.6	3.4	2.5	9.5	3.1	4.4	5.1
Edea	2.5	9.8	2.1	4.9	4.4	2.2	9.4	2.9	4.8	6.3
Eseka	2.6	10.0	1.9	4.9	4.4	2.5	10.0	2.6	4.5	5.5
Garoua	1.7	9.0	2.4	5.4	3.3	1.8	9.1	1.5	5.6	2.9
Kaele	1.6	8.7	2.5	5.7	3.3	1.8	9.2	2.3	5.4	2.9
Koundja	2.5	9.4	2.2	5.1	3.7	2.1	9.5	3.6	3.7	7.9
Kribi	2.4	10.2	2.1	4.4	5.2	2.5	10.0	2.5	4.4	5.9
Lomie	2.4	10.2	2.1	4.6	4.1	2.2	10.3	2.6	3.9	5.3
Maroua	1.7	8.6	2.1	5.6	3.0	1.8	8.7	1.8	5.0	2.8
Meiganga	2.4	9.4	2.3	5.2	3.4	2.3	9.8	3.2	4.5	4.8
Nanga Eboko	2.3	10.0	2.0	4.6	3.8	2.3	10.4	2.5	4.1	4.9
Ngambe	2.7	9.4	1.8	5.5	4.7	2.6	8.5	1.8	6.9	8.6
Ngaoundere	2.0	8.8	2.8	5.1	3.5	1.7	10.0	7.1	4.8	6.4
Nkongsamba	2.5	9.2	2.5	5.1	4.4	2.2	8.4	2.5	6.0	8.8
Poli	2.2	9.0	2.2	5.4	2.9	2.3	9.1	2.1	5.3	2.9
Sangmelima	2.5	10.2	2.0	4.5	3.9	2.3	10.5	2.6	4.1	5.4
Tibati	2.5	9.4	2.1	5.3	3.4	2.6	10.1	3.0	4.9	3.7
Yaounde	2.5	10.1	1.9	4.6	3.8	2.4	10.3	2.4	4.1	5.2
Yokadouma	2.3	10.1	2.3	4.6	4.1	2.1	10.2	3.1	4.0	5.0
Yoko	2. 3	9. 7	2. 0	5. 2	4. 0	2. 1	10. 2	2. 9	4. 3	6. 3

The estimated parameters are then used to compute the monthly $Y_{i1}(m)$ for all the twenty-eight stations. The demonstration of the fit achieved is done by plotting the observed and transformed proportions in figure 24 and figure 25 at the sample stations for wet-dry $Y_{01}(m)$ $i = 0$ and for wet-wet $Y_{11}(m)$ $i = 1$ respectively; the rest of the figures are presented in annexe A for $Y_{01}(m)$ and B for $Y_{11}(m)$ of this study.

Figures 24 for $Y_{01}(m)$ and figure 25 for $Y_{11}(m)$ show that the double normal function provides a good fit to data at many stations despite a few gaps. Gaps can be explained by the fact that sometimes a misbalance exists between estimated and observed values and the associated errors, which should add to zero over the year. This is induced by applying the error minimizing procedure to the $Y_{i1}(m)'s$ and not to probabilities. Monthly variation of $Y_{01}(m)$ and $Y_{11}(m)$ are described above using the double normal function. It is concluded that the double normal function well represents the monthly variation of the transition probabilities. As we are interested in spatial prediction of daily rainfall occurrence processes in Cameroon, we need to estimate the parameters of the double normal function by spatial predictors.

4.3 Spatial distribution of the transition probability parameters: estimation of parameters of $Y_{i1}(m)$

In order to describe the spatial variation of rainfall occurrence, we try to describe its seasonal variation by time invariant parameters and for the latter we need the spatial predictors. With the aim of building simulation models of rainfall occurrence, we want to estimate the parameters of the double normal function by spatially distributed data so that one can use these parameters to simulate rainfall occurrence at a point P within the area of study. The selection of a predictor to be included or to be excluded in this process is done by a stepwise regression analysis between the parameters of $Y_{i1}(m)$ and the predictors. Parameters of $Y_{i1}(m)$ are dependent variable whereas the latitude, the longitude, the altitude, the relief orientation, the rain exposition are the independent variables known as predictors. Our objective is to find the best predictors well describing the rainfall occurrence process anywhere in the country; as outcome of stepwise regression, 5 % significance level is chosen to decide about the inclusion or exclusion of a independent variable. By best predictors, we mean them to

- be easy to observe
- explain a good deal of the (spatial) variance of the parameters.

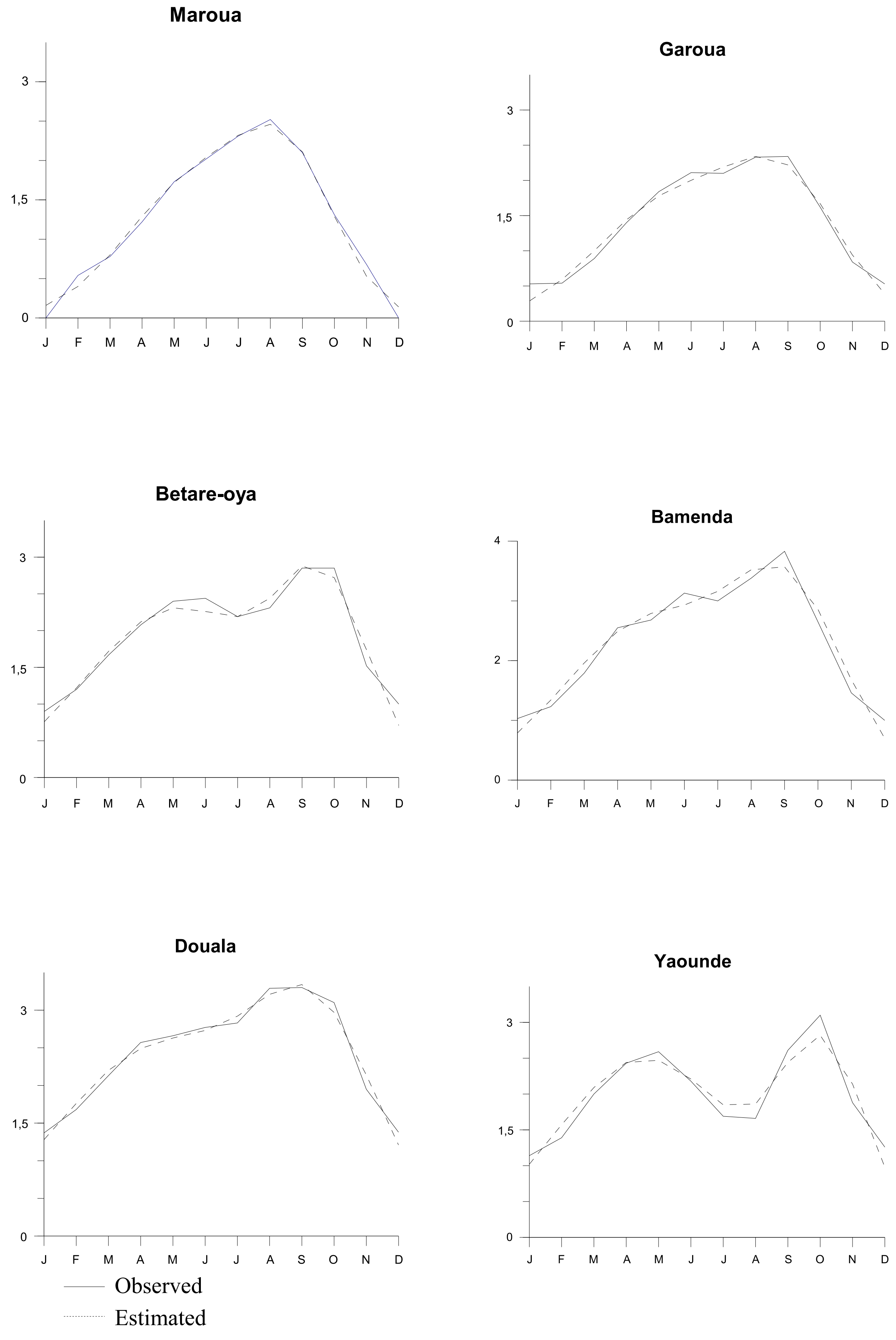


Figure 15: Temporal distribution of the logits $Y_{01}(m)$ of wet/dry transition probability $p_{01}(m)$ using the double normal function

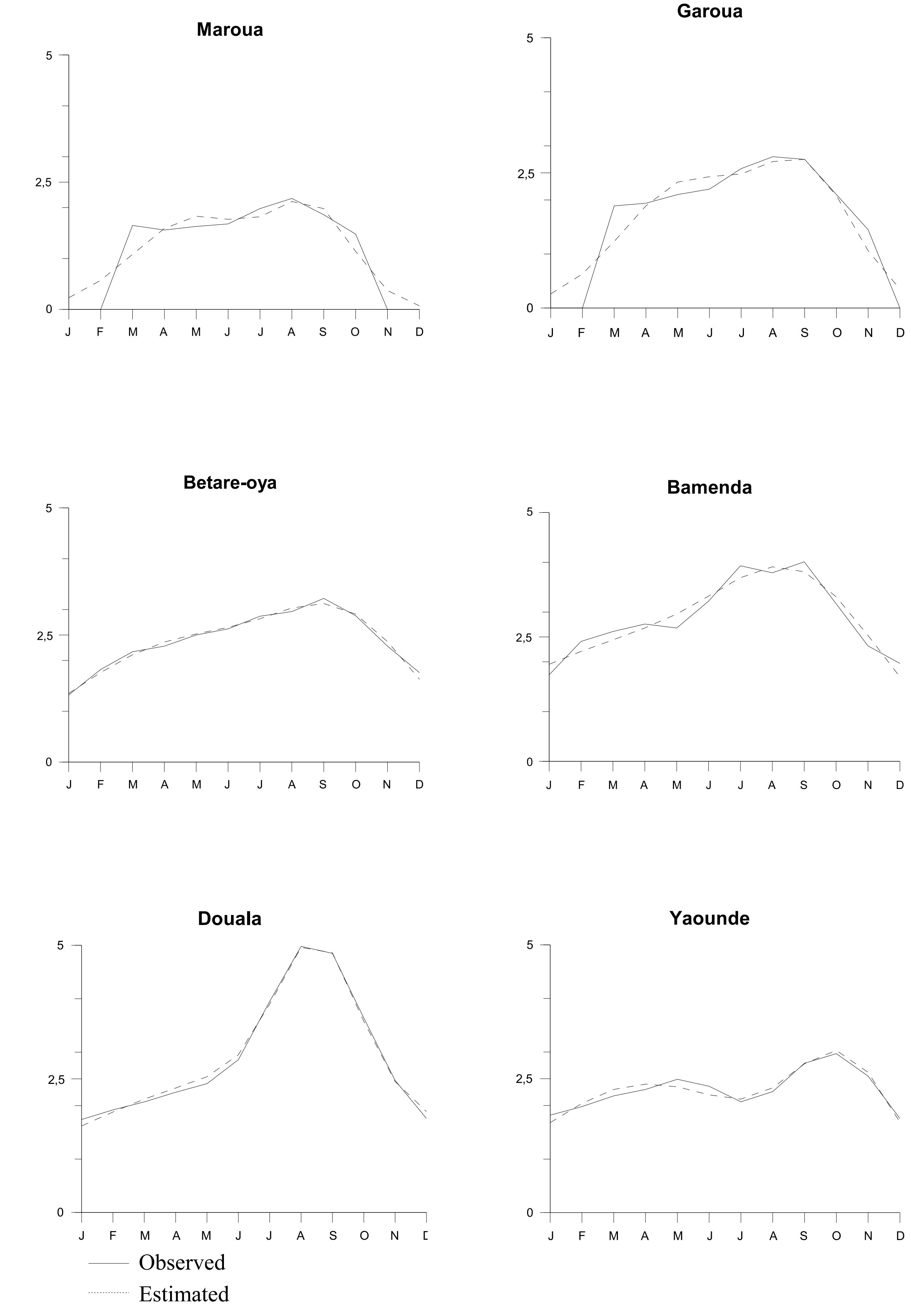


Figure 16: Temporal distribution of the logits $Y_{11}(m)$ of wet/wet transition probability $p_{11}(m)$ using the double normal function

Stepwise regression analyses are applied to parameters of $Y_{i1}(m)$ and spatially distributed data. The latitude, the longitude, the altitude, the relief orientation, and the rain exposition are independent variables or predictors (see the section about data description). We got them from topographic maps and from field works. These are geographical factors thought to have an influence on rainfall distributions in Cameroon. As independent spatial variable are used, they include latitude φ, longitude λ, altitude ω, relief orientation ϑ and rain exposition κ.

The dependent variables are parameters a_i, μ_{i1}, s_{i1}, μ_{i2}, s_{i2} of the double normal function used to describe the seasonal variation of $Y_{i1}(m)$ with $i = 0, 1$;
a_i is a constant which allows the superposition of the double normal function on the observed data curve.
Parameter μ_{i1} represents the first peak i.e. first maximum or "little"rainy season.
Parameter s_{i1} is the width of the first season; it the standard deviation.
Parameter μ_{i2} represents the second peak i.e. second maximum or second rainy season.
Parameter s_{i2} is the width of the second rainy season; it the standard deviation.

Estimated parameters of $Y_{i1}(m)$, independent variables selected for the estimation of coefficients, unstandardized coefficient B , standardized Beta coefficient as well as the t-value and significance level are summarized in table 5.

Table 5: Summary of regression coefficients for estimating $Y_{i1}(m)$ parameters

parameters	variable	coefficients (B)	Beta coefficients	t	significance value
a_0	$constant$	-38.86		5.81	.000
	$\cos(\varphi)$	43.00	.631	6.33	.000
	λ	$-.06$	$-.368$	-3.69	.001
	ω	.00018	$-.256$	2.82	.010
μ_{01}	$constant$	11.78		70.91	.000
	$\ln(\varphi)$	-1.22	$-.871$	-11.66	.000
	ϑ	$-.08$	$-.193$	-2.54	.018
	κ	$-.0019$	$-.194$	-2.48	.021
s_{01}	$constant$	2.02		27.90	.000
	κ	.0028	.467	2.69	.012
μ_{02}	$constant$	3.61		21.19	.000
	$\ln(\varphi)$	.084	.848	8.15	.000
s_{02}	$constant$	6.26		19.59	.000
	φ^2	$-.93$	$-.758$	-6.79	.000
	ω	$-.00034$	$-.257$	-2.30	.000
a_1	$constant$	-31.07		-3.50	.002
	$\cos(\varphi)$	33.45	.592	3.75	.001
μ_{11}	$constant$	9.95		44.08	.000
	ω	$-.0007$	$-.360$	2.30	.000
	κ	$-.01$	$-.700$	-4.47	.030
s_{11}	$constant$	2.02		6.23	.000
	ω	.0018	.482	2.80	.009
μ_{12}	4.639				
s_{12}	$constant$	-148.46		-2.13	.043
	$\cos(\varphi)$	161.31	.362	-3.09	.005
	λ	$-.52$	$-.479$	-2.34	.028

From table 5 for $i = 0$, the t-values show that $\cos(\varphi)$, λ and, ω are the best predictors for describing parameter a_0. The regression model of parameter a_0 with cosine of latitude, the longitude and the altitude produced $R^2 = .806$. The latitude has a bigger Beta coefficient than longitude and altitude; i.e the latitude has strength of influence on the dependent variable a_0 than other independent variables used to estimate a_0.

The t-value suggests that $\ln(\varphi)$, ϑ, κ are the best predictors to be used when estimating parameter μ_{01}. The regression model of parameter μ_{01} with the logarithm of latitude, the relief orientation, and the rain exposition produced $R^2 = .947$. All the Beta coefficients are negative. The latitude has a higher Beta coefficient than relief orientation and relief exposition; that means the latitude is the variable which influences strongly the dependent variable μ_{01}.

The t-value suggests that only κ can be used to estimate parameter s_{01}. The Beta coefficient is not too high. Nevertheless, the rain exposition has an influence on the dependent variable s_{01}. The regression model of parameter s_{01} with rain exposition produced $R^2 = .218$.

According to the t-values, $\ln(\varphi)$ is the best predictor to be used when estimating parameter μ_{02}. The Beta coefficient is comparatively high; the logarithm of the latitude has a strong influence on the dependent variable μ_{02}. The regression model of parameter μ_{02} with the logaritm of latitude produced $R^2 = .718$.

The t-value suggests that φ^2 and ω are the best predictors to be used for representing parameter s_{02}. The predictor φ^2 has higher Beta coefficient than ω; the square of latitude has a strength of influence on the dependent variable s_{02} than altitude. The regression model of parameter s_{02} with square of latitude and altitude produced $R^2 = .694$.

For $i = 1$, the t-value shows the cosine of latitude $\cos(\varphi)$ to be the best variable for estimating parameter a_1. It has a strength of influence on the dependent variable a_1 because of the high value of Beta coefficient. The regression model of parameter a_1 with the cosine of latitude produced $R^2 = .351$.

The t-value suggests that the altitude ω and the rain exposition κ are the best variables for predicting parameter μ_{11}. The rain exposition has comparatively higher Beta coefficient than the altitude; the rain exposition has stronger influence on parameter μ_{11} than altitude. The regression model produced $R^2 = .454$.

The altitude ω is the best variable for estimating parameter s_{11} as suggested by t-value. It has a positive regression weight represented by the Beta coefficient. The altitude has a strength of influence on the dependent variable s_{11}. The regression model with altitude produced $R^2 = .232$.

There is no linear relation between parameter μ_{12} and all the predictors. We consider parameter μ_{12} to be constant for all stations, its value being the mean value of all μ_{12}.

The cosine of latitude $\cos(\varphi)$ and longitude λ are found to be the best variables for predicting parameter s_{12} as suggested by the t-value. The longitude has comparatively higher Beta coefficient than cosine of the latitude; this means that the longitude has a stronger influence on parameter s_{12} than the cosine of the latitude. The regression model produced $R^2 = .511$.

In general for $i = 0$ and $i = 1$, the t-value suggests that latitude φ, longitude λ, rain exposition κ, relief orientation and altitude ω are the best predictors well describing $Y_{i1}(m)$.

Estimated parameters of $Y_{i1}(m)$ are:

for $i = 0$, we get spatial parameters of $Y_{01}(m)$ as follows:

$$\begin{aligned}
a_0 &= -38.86 + 43.00\cos(\varphi) - .06\lambda + .00018\omega \\
\mu_{01} &= 11.78 - 1.22\ln(\varphi) - .08\vartheta - .0019\kappa \\
s_{01} &= 2.02 + .0028\kappa \\
\mu_{02} &= 3.61 + .084\ln(\varphi) \\
s_{02} &= 6.26 - .93\varphi^2 - .00034\omega
\end{aligned} \tag{14}$$

for $i = 1$ we get spatial parameters of $Y_{11}(m)$ as follows:

$$\begin{aligned}
a_1 &= -31.07 + 33.45\cos(\varphi) \\
\mu_{11} &= 9.95 - 0.0007\kappa \\
s_{11} &= 2.023 + .0018\omega \\
\mu_{12} &= c = 4.639 \\
s_{12} &= -148.46 + 161.31\cos(\varphi) - 0.524\lambda
\end{aligned} \tag{15}$$

By this way, equations (14) and (15) are used for describing the spatial distribution of logit $Y_{i1}(m)$ as follows:

$$Y_{i1}(m) = a_i\left(exp(-(\frac{m-\mu_{i1}}{s_{i1}})^2) + exp(-(\frac{m-\mu_{i2}}{s_{i2}})^2)\right) \qquad i = 0, 1 \tag{16}$$

Resulting values of $Y_{i1}(m)$ at some stations have been plotted in figure 17 for $i = 0$ and in figure 18 for $i = 1$. It is noted that the model fits well at some stations whereas at others there is a gap between observation and estimation. As previously stated, the discrepancy is explained by the fact that the peaks don't look like bell shaped normal distribution but rather resemble shield volcanoes; because of such a form, they cannot be well represented by a normal function. Sometimes a misbalance exists between estimated and observed values and the associated errors, which should be added to zero over the year. This is induced by the fact that the error minimizing procedure is applied to $Y_{i1}(m)$ and not to probabilities. The bias occurred when inverting $Y_{i1}(m)$ to $Y_{i1}(m)^{-1}$ is carried out in order to get probabilities again.

First-order Markov chain model is used to model the daily rainfall series. The double normal function is used to describe the temporal variation of the parameters. Stepwise regression analysis is used to estimate parameters of the double normal function from the spatial predictors. Estimated parameters are used to build input models which will be latter used for simulating daily rainfall occurence process using the Monte Carlo simulation methods. The next step consits to study the temporal and spatial distribution of daily rainfall amount; our goal is to develop an independent model for simulating the rainfall amounts on a wet day.

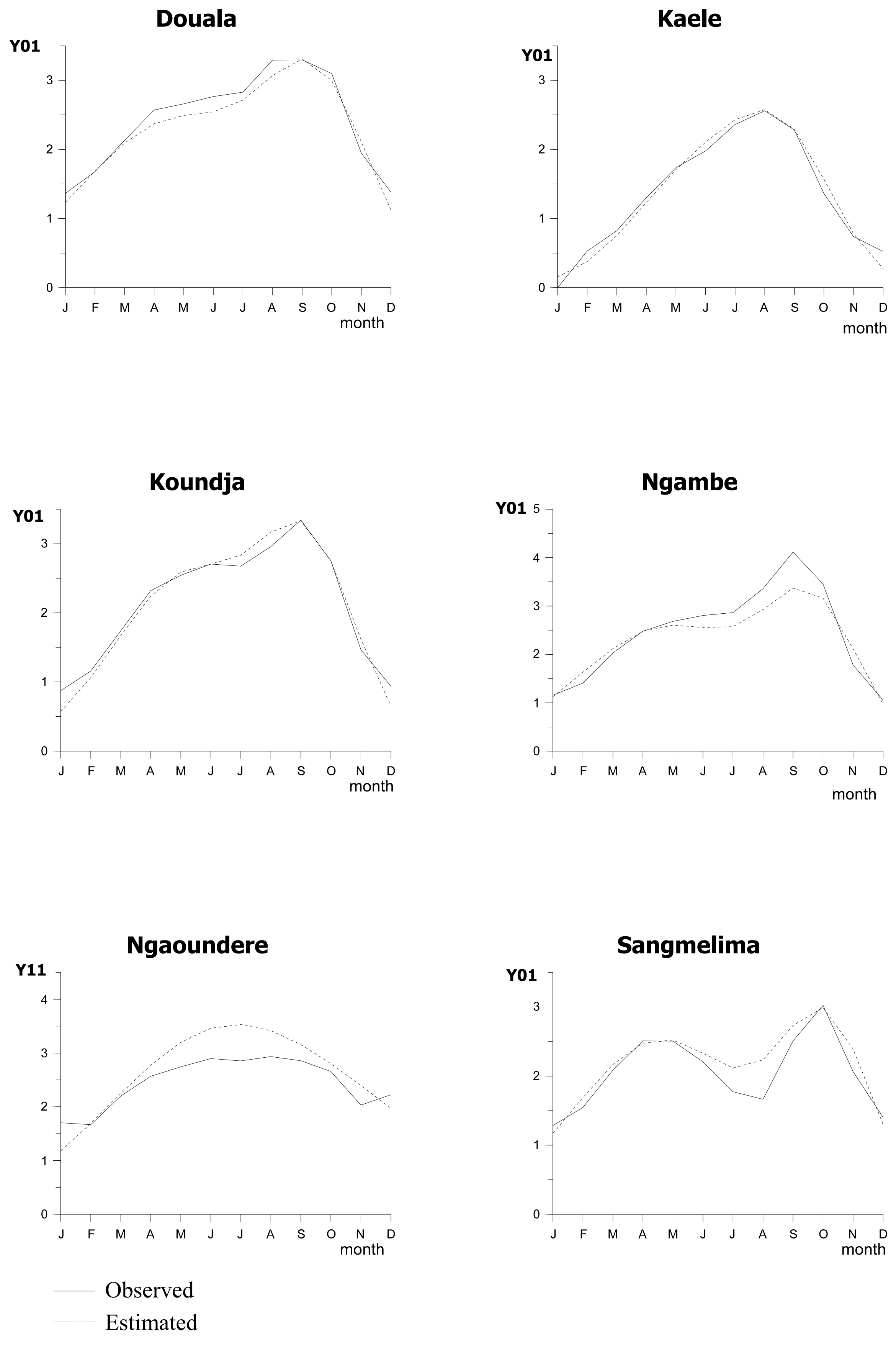


Figure 17: Spatial distribution of of the monthly wet-dry logits $Y_{i1}(m)$ with $i = 0$ using the double normal function

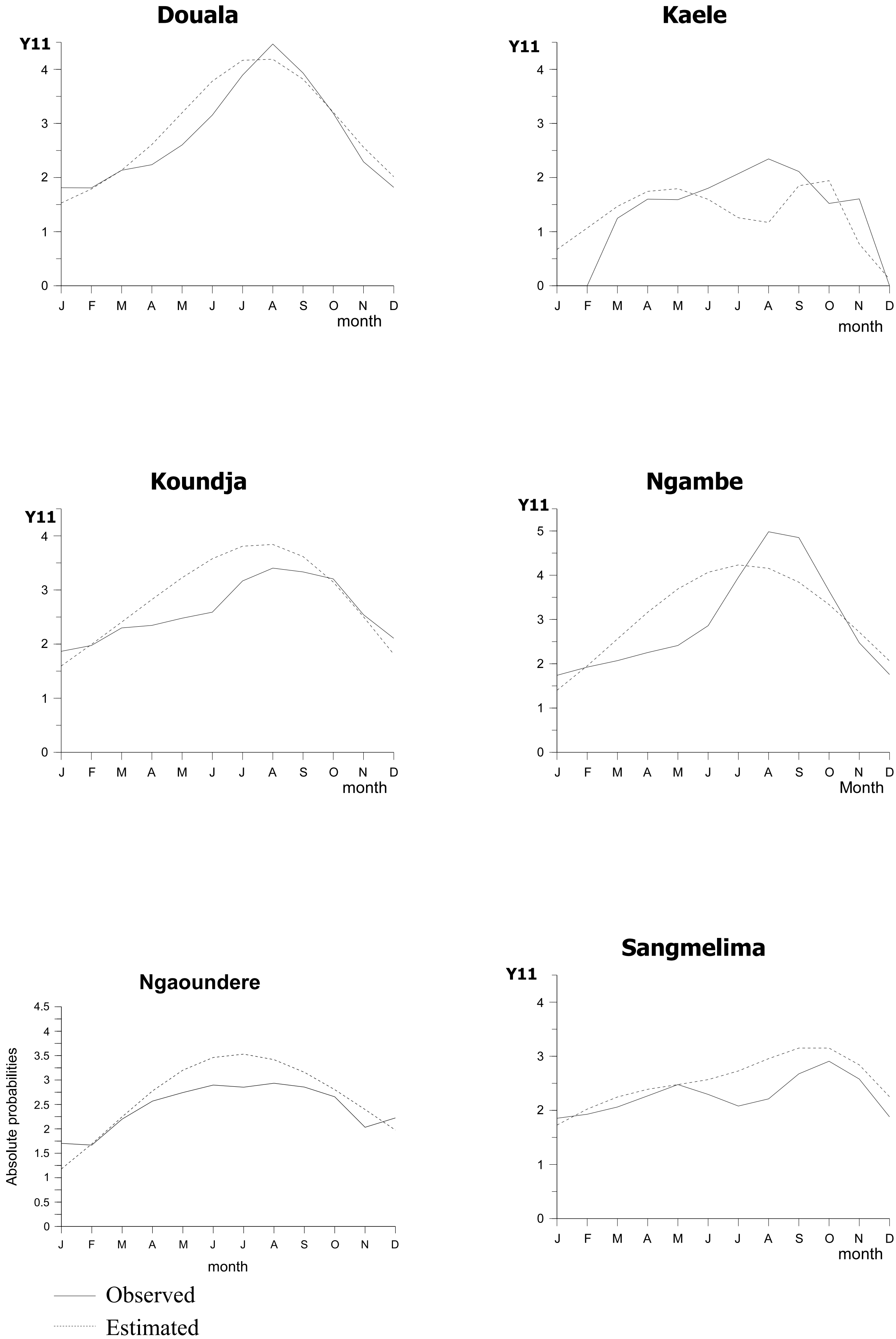


Figure 18: Spatial distribution of the monthly wet-wet logit $Y_{i1}(m)$ with $i = 1$ using the double normal function

5 DISTRIBUTION OF RAINFALL AMOUNTS

5.1 Fitting of a suitable theoretical distribution function

The idea to fit a suitable theoretical distribution function follows from the fact that we need a "good" distribution which has to be simple and heuristically convincing. "Simple" means that the chosen distribution is simply measured by a less number of parameters. To date, there is no general rule for chosing the type of distribution function (Sevruk and Geiger 1981) when fitting curves to a set of data. There is also no universally accepted method concerning the selection of a distribution function and according statistical tests used to this purpose (Kite 1977). In this study, to find out a "good"distribution, we first of all represent the distribution of daily rain amount r on wet days graphically, with data grouped in intervall of 10^{th} of mm; this allows us to have a general idea on how the curves look. For this purpose, only few stations are used. The chosen station represents a typical rainfall regime.

From figure 19, it can be seen that the distribution is unimodal and asymetric, which is in agreement with the Weibull distribution. We then formulate the null hypothesis that "Weibull distribution function fits the daily rain amount on wet days". Weibull distribution has an advantage that it is integrable i.e. there is possibilities getting probability which useful in Monte Carlo simulation methods. Its probability density function is as follows:

$$w(r) = \frac{g}{b}\left[\frac{r}{b}\right]^{g-1} exp\left[-\left[\frac{r}{b}\right]^{g}\right] \tag{17}$$

where $\quad g > 0, \quad b > 0, \quad r > 0$.

For Monte Carlo simulation of rain amount r using a random number $R \in [0,1]$ equally distributed in the interval [0,1] as generated by pseudo-random number generator, we take the probability distribution $D(r)$ belonging to $w(r)$

$$D(r) = \int_0^r w(z)\,dz = 1 - exp\left[-\left[\frac{r}{b}\right]\right]^{g} \tag{18}$$

The substitution of $D(r)$ with any random number R gives

$$R = 1 - exp\left[-\left[\frac{r}{b}\right]\right]^{g} \tag{19}$$

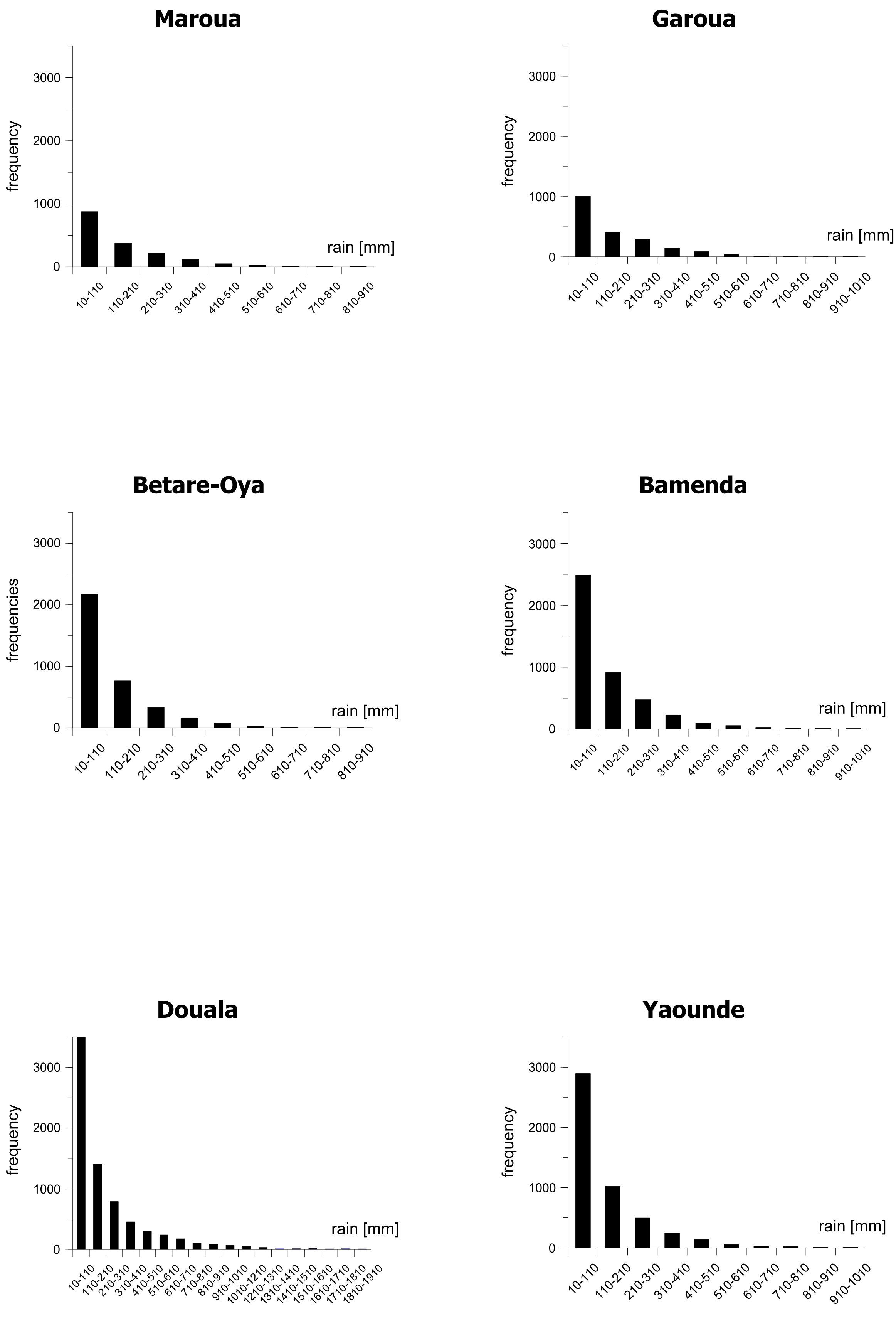


Figure 19: Distribution of observed daily rainfall frequency in wet days (1951-1993)

To produce the corresponding rain amount r, one has to solve the equation (19) for r

$$r = b * \left(ln(\frac{1}{1-R}) \right)^{\frac{1}{g}} \tag{20}$$

To simulate rain amounts that follow the Weibull distribution is to generate random numbers R, each random number then inserted in the equation (20) gives rain amount r.

The use of Weibull distribution in this study follows some major steps. These steps are carried out for every month m. At the first step, we adapted a Weibull distribution to the observed frequencies, using a random search algorithm for the parameter $g(m)$ and $b(m)$ as in the equations (21) and (22). Data are grouped in intervals i.e. classes j of 10^{th} of mm. The equation (21) is used to compute an error ε from expected frequencies E_j as follows:

$$\varepsilon(g(m), b(m)) = \sum_j \left(E_j(m) - O_j(m)\right)^2 \tag{21}$$

where

j is the intervals i.e. classes of daily rainfall amounts in 10^{th} of mm

O_j is the number of observed frequency belonging to class j

E_j is the number of expected frequency belonging to class j

The parameters $g(m)$ and $b(m)$ we are looking for, are the ones which minimize the error ε

$$\varepsilon(g(m), b(m)) = Minimum \tag{22}$$

we then decide on the goodness of fit χ^2 statistics by calculating χ^2 statistics

$$\chi^2(m) = \sum_j^n \frac{\left(E_j(m) - O_j(m)\right)^2}{E_j(m)} \tag{23}$$

Equation (23) yields the χ^2 contribution. We compare it to the critical value at 5 % significance level and the result is presented on table 6 for all the 28 stations under study.

Table 6: χ^2 - test of Weibull distribution function on daily rainfall

Stations		J	F	M	A	M	J	J	A	S	O	N	D
AbongMbang	χ^2	.344	.719	3.561	7.627	4.103	7.607	2.645	4.811	7.546	9.880	**16.14**	6.335
	df	3	4	6	6	7	6	6	6	9	8	4	4
	cv	7.81	9.49	12.59	12.59	14.07	12.59	12.59	12.59	16.92	15.51	9.49	9.49
Akonolinga	χ^2	.010	.017	11.979	5.087	4.335	3.588	2.197	7.192	6.209	5.636	1.689	.807
	df	2	2	6	5	6	5	5	6	7	7	4	3
	cv	5.99	5.99	12.59	11.07	12.59	11.07	11.07	12.59	14.07	14.07	9.49	7.81
Bafia	χ^2	.008	1.997	1.466	.991	3.064	.919	3.577	6.396	5.587	6.072	.415	.005
	df	1	3	5	5	6	5	6	5	7	7	4	1
	cv	3.84	7.81	11.07	11.07	12.59	11.07	12.59	11.07	14.07	14.07	9.49	3.84
Bamenda	χ^2	.005	.308	7.026	6.695	6.088	2.664	**20.12**	8.375	8.117	5.098	.009	.003
	df	1	3	4	6	5	6	7	7	7	5	2	1
	cv	3.84	7.81	9.49	12.59	11.07	12.59	14.07	14.07	14.07	11.07	5.99	3.84
Batouri	χ^2	2.709	.153	6.171	3.988	1.777	5.656	2.857	7.853	3.346	12.676	2.184	.275
	df	3	3	5	6	6	6	6	8	6	7	5	3
	cv	7.81	7.81	11.07	12.59	12.59	12.59	12.59	15.51	12.59	14.07	11.07	7.81
Bertoua	χ^2	.393	2.321	3.407	2.490	3.055	3.923	3.486	2.902	4.201	4.241	4.673	.105
	df	3	3	5	5	6	6	6	7	7	7	6	3
	cv	7.81	7.81	11.07	11.07	12.59	12.59	12.59	14.07	14.07	14.07	12.59	7.81
Betare-oya	χ^2	.047	.015	6.364	1.755	2.644	**20.09**	1.735	3.084	8.044	6.702	.018	.0004
	df	1	2	4	5	5	3	6	7	6	5	3	1
	cv	3.84	5.99	9.49	11.07	11.07	7.81	12.59	14.07	12.59	11.07	7.81	3.84
Douala	χ^2	2.909	.298	5.381	2.258	6.509	8.350	**32.54**	**35.65**	14.667	**17.660**	3.391	.036
	df	4	4	7	7	8	12	19	13	14	9	7	3
	cv	9.49	9.49	14.07	14.07	15.51	21.03	30.14	22.36	23.68	16.92	14.07	7.81
Dschang	χ^2	.002	3.903	.381	5.553	1.512	**19.54**	5.536	1.240	5.546	**13.73**	.004	.0004
	df	1	3	4	5	4	4	5	6	6	4	2	1
	cv	3.84	7.81	9.49	11.07	9.49	9.49	11.07	12.59	12.59	9.49	5.99	3.84
Edea	χ^2	.948	.466	3.254	4.595	7.202	9.335	12.163	4.874	2.508	10.391	.688	8.961
	df	3	3	6	7	7	9	8	8	10	8	6	3
	cv	7.81	7.81	12.59	14.07	14.07	16.92	15.51	15.51	18.31	15.51	12.59	7.81
Eseka	χ^2	.876	2.243	**35.50**	**17.85**	12.404	5.193	5.035	1.461	4.097	**18.20**	4.329	1.059
	df	4	4	5	8	8	6	5	5	7	8	5	3
	cv	9.49	9.49	11.07	15.51	15.51	12.59	11.07	11.07	14.07	15.51	11.07	7.81
Garoua	χ^2	0	0	.0004	2.279	9.005	11.407	7.743	15.954	10.781	8.259	0	0
	df	0	0	1	4	6	6	8	9	7	5	0	0
	cv	0	0	3.84	9.49	12.59	12.59	15.51	16.92	14.07	11.07	0	0
kaele	χ^2	0	0	0	.003	1.196	2.123	2.768	4.951	5.238	3.247	0	0
	df	0	0	0	2	5	5	8	7	5	4	0	0
	cv	0	0	0	5.99	11.07	11.07	15.51	14.07	11.07	9.49	0	0
Koundja	χ^2	.006	.002	4.621	7.592	1.958	5.272	2.056	5.612	**28.66**	2.365	4.457	.0004
	df	1	2	5	5	5	6	7	7	8	5	4	1
	cv	3.84	5.99	11.07	11.07	11.07	12.59	14.07	14.07	15.51	11.07	9.49	3.84

to continue. . .

continuation...													
Stations		J	F	M	A	M	J	J	A	S	O	N	D
Kribi	χ^2	2.501	6.609	4.502	6.823	1.326	3.437	7.419	**16.37**	8.929	17.201	7.389	2.951
	df	5	7	8	9	10	10	7	8	11	10	9	5
	cv	11.07	14.07	15.51	16.92	18.31	18.31	14.07	15.51	19.68	18.31	16.92	11.07
Lomie	χ^2	2.559	2.537	2.752	.891	2.746	8.745	4.022	6.546	1.139	3.423	3.672	.177
	df	3	3	6	5	6	6	7	7	7	7	6	2
	cv	7.81	7.81	12.59	11.07	12.59	12.59	14.07	14.07	14.07	14.07	12.59	5.99
Maroua	χ^2	0	0	0	.206	2.840	7.697	3.219	2.906	5.096	4.044	0	0
	df	0	0	0	2	4	5	7	7	6	3	0	0
	cv	0	0	0	5.99	9.49	11.07	14.07	14.07	12.59	7.81	0	0
Meiganga	χ^2	.001	.002	.066	1.286	10.270	1.485	3.806	4.529	**16.60**	2.841	2.953	.0003
	df	1	1	3	4	5	5	7	7	7	5	3	1
	cv	3.84	3.84	7.81	9.49	11.07	11.07	14.07	14.07	14.07	11.07	7.81	3.84
Nanga-Eboko	χ^2	.031	.288	7.661	2.892	3.424	.339	4.718	3.796	16.101	9.159	7.983	.024
	df	2	4	6	6	6	5	5	6	10	8	5	2
	cv	5.99	9.49	12.59	12.59	12.59	11.07	11.07	12.59	18.31	15.51	11.07	5.99
Ngambe	χ^2	.010	.841	3.677	5.655	4.940	6.931	**95.67**	**49.38**	9.829	3.055	7.048	.089
	df	2	3	6	6	6	7	5	6	9	8	5	2
	cv	5.99	7.81	12.59	12.59	12.59	14.07	11.07	12.59	16.92	15.51	11.07	5.99
Ngaoundere	χ^2	0	0	.248	2.195	1.063	2.013	**15.92**	1.306	2.749	3.182	.030	0
	df	0	0	3	5	5	5	7	7	6	5	1	0
	cv	0	0	7.81	11.07	11.07	11.07	14.07	14.07	12.59	11.07	3.84	0
Nkongsamba	χ^2	.042	.383	2.310	1.645	3.084	2.358	**19.83**	**21.01**	11.847	6.227	.798	.011
	df	2	3	5	5	5	5	9	9	8	7	4	2
	cv	5.99	7.81	11.07	11.07	11.07	11.07	16.92	16.92	15.51	14.07	9.49	5.99
Poli	χ^2	0	0	.055	1.212	14.02	5.881	7.868	12.821	12.796	7.881	.002	0
	df	0	0	1	5	7	7	8	8	7	6	1	0
	cv	0	0	3.84	11.07	14.07	14.07	15.51	15.51	14.07	12.59	3.84	0
Sangmelima	χ^2	3.921	.0621	4.871	3.591	.793	8.540	1.543	7.273	3.894	10.786	1.555	.063
	df	4	4	6	6	6	5	5	6	7	8	5	2
	cv	9.49	9.49	12.59	12.59	12.59	11.07	11.07	12.59	14.07	15.51	11.07	5.99
Tibati	χ^2	0	.005	3.362	1.685	2.159	2.806	.623	2.342	6.882	10.017	.435	0
	df	0	2	4	3	5	6	7	7	7	5	3	0
	cv	0	5.99	9.49	7.81	11.07	12.59	14.07	14.07	14.07	11.07	7.81	0
Yaounde	χ^2	.443	1.501	7.217	7.833	.910	.519	2.415	4.982	3.733	12.009	.967	.0020
	df	3	3	6	6	6	6	4	5	7	8	5	2
	cv	7.81	7.81	12.59	12.59	12.59	12.59	9.49	11.07	14.07	15.51	11.07	5.99
Yokadouma	χ^2	**10.69**	.662	3.733	5.476	3.641	.947	**16.24**	3.385	15.170	2.755	1.372	.114
	df	4	3	5	5	6	6	8	6	9	7	5	4
	cv	9.49	7.81	11.07	11.07	12.59	12.59	15.51	12.59	16.92	14.07	11.07	9.49
Yoko	χ^2	.0002	.242	1.018	5.437	4.770	3.452	4.586	1.143	3.029	4.906	.586	.008
	df	1	3	4	5	5	5	6	6	7	7	4	1
	cv	3.84	7.81	9.49	11.07	11.07	11.07	12.59	12.59	14.07	14.07	9.49	3.84

The high χ^2 contribution means the rejection of the null hypothesis at 5% level at some months at some stations i.e. November in Abong-Mbang, July in Bamenda, June in Betare-Oya, July, August and October in Douala, May and October in Dschang, October in Eseka, September in Koundja, September in Lomie, July and August in Ngambe and Nkongsamba. The numbers in fat in the table indicate the high χ^2 contribution against the critical value (see table 6). It is noted that the high χ^2 contribution mostly occurs in the heart of the rainy season. This might be explained by the fact that these months experience continous rain over several days each of them with the biggest possible daily rain amount. Accordingly, all these days have biggest rainfall amounts of nearly the same order of magnitude which cause a lump in the frequency distribution in figure 20. The Weibull distribution predicts much fewer big rain amount the higher the rain amount is. Thus we can expect high χ^2 contribution at high rain amounts, especially if we enlarge the span of big rainfall class, in order to get an expected frequency higher than 5.

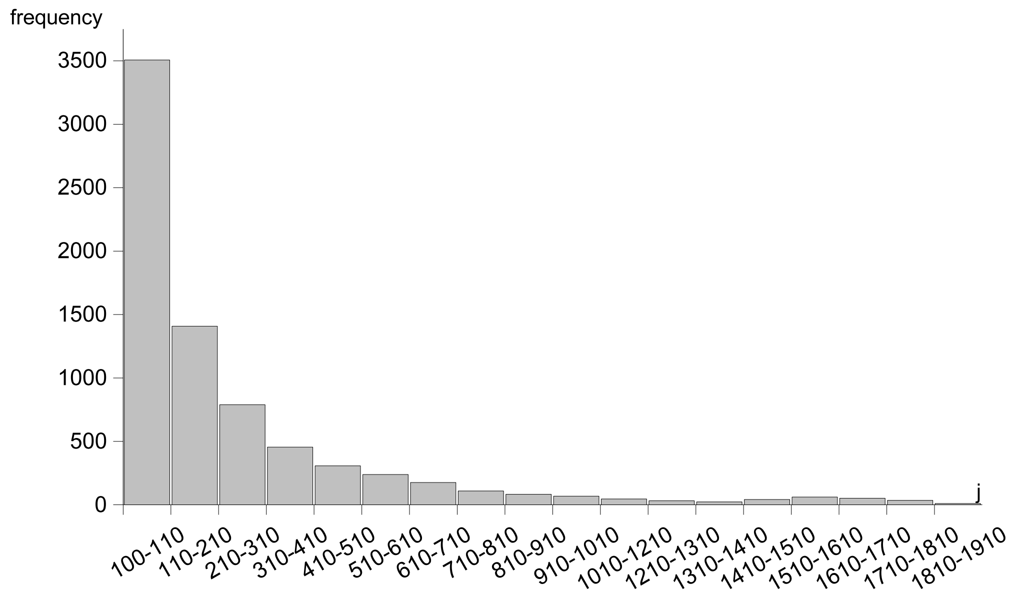


Figure 20: Frequency distribution of daily rain amounts in Douala

To correct the misfit of Weibull distribution, we represent the lumped part of the curve by overlay of normal distribution. The new function $v(r)$ to be used for computing the new expected frequencies includes four parameters, g and b of Weibull distribution and the mean and standard deviation of the overlay normal distribution

$$v(r) = \frac{\frac{g}{b}\left[\frac{r}{b}\right]^{g-1} * exp\left[-\left[\frac{r}{b}\right]^{g}\right] + \frac{1}{b} * exp\left[-\left[\frac{r-a}{s}\right]^{2}\right]}{\left(1 + \frac{1}{b} * \sqrt{2\pi} * cf\right)} \qquad (24)$$

cf is a correction factor to be used for the probability distribution property and is defined as follows

$$cf = \int_0^{2a} exp\Big[- \Big[\frac{r-a}{s}\Big]^2\Big]\, dr \tag{25}$$

The newly computed χ^2 contributions fitting four parameters (two of Weibull distribution and two of the overlay normal function) is presented in table 7. This is done for every month m troughout the years; we do this without argument m for the sake of simplicity.

Table 7: χ^2 - test of Weibull distribution function on daily rainfall after adding the normal function

Stations		J	F	M	A	M	J	J	A	S	O	N	D
AbongMbang	χ^2	.018	1.971	1.764	7.939	7.258	10.007	1.250	3.567	10.528	9.506	**15.29**	5.952
	df	3	4	6	6	7	7	6	6	10	7	6	5
	cv	7.81	9.49	12.59	12.59	14.07	14.07	12.59	12.59	18.31	14.07	12.59	11.07
Akonolinga	χ^2	.046	.0008	9.922	4.391	1.584	3.376	1.965	2.374	1.735	3.267	.977	.599
	df	2	2	6	5	7	6	5	6	7	7	5	3
	cv	5.99	5.99	12.59	11.07	14.07	12.59	11.07	12.59	14.07	14.07	11.07	7.81
Bafia	χ^2	1.497	1.082	1.215	2.848	4.557	10.115	.426	8.083	4.663	8.807	.006	.005
	df	2	3	5	5	6	6	5	6	7	7	3	1
	cv	5.99	7.81	11.07	11.07	12.59	12.59	11.07	12.59	14.07	14.07	7.81	3.84
Bamenda	χ^2	.005	.093	1.991	4.099	5.275	2.858	10.720	11.495	8.364	5.235	.016	.000
	df	1	3	4	5	5	6	8	7	7	5	2	1
	cv	3.84	7.81	9.49	11.07	11.07	12.59	15.51	14.07	14.07	11.07	5.99	3.84
Batouri	χ^2	.0968	.136	4.615	2.511	3.542	3.590	4.653	8.971	5.521	14.061	.028	.078
	df	3	3	5	6	6	6	7	8	7	7	4	3
	cv	7.81	7.81	11.07	12.59	12.59	12.59	14.07	15.51	14.07	14.07	9.49	7.81
Bertoua	χ^2	2.299	.006	6.764	.659	7.985	1.869	5.937	3.972	1.511	7.874	5.919	.655
	df	3	3	5	5	6	6	6	8	7	7	5	3
	cv	7.81	7.81	11.07	11.07	12.59	12.59	12.59	15.51	14.07	14.07	11.07	7.81
Betare-Oya	χ^2	0	.041	4.450	.449	8.677	5.678	5.189	4.934	7.016	5.614	.067	.0001
	df	0	2	4	4	5	5	6	7	6	7	3	1
	cv	0	5.99	9.49	9.49	11.07	11.07	12.59	14.07	12.59	14.07	7.81	3.84
Douala	χ^2	3.772	.175	**20.70**	4.424	6.758	8.138	**52.94**	15.310	7.443	10.646	2.737	.522
	df	4	4	9	8	10	13	21	16	14	9	7	3
	cv	9.49	9.49	16.92	15.51	18.31	22.36	32.67	26.30	23.68	16.92	14.07	7.81
Dschang	χ^2	.0007	6.189	1.394	4.489	1.959	7.432	5.494	1.766	7.698	6.399	3.735	.0003
	df	1	3	4	6	5	5	5	6	6	6	3	1
	cv	3.84	7.81	9.49	12.59	11.07	11.07	11.07	12.59	12.59	12.59	7.81	3.84
Edea	χ^2	1.931	.271	1.532	4.404	5.698	4.707	12.279	4.074	1.578	5.856	2.491	4.829
	df	3	3	6	7	7	9	8	8	10	8	6	3
	cv	7.81	7.81	12.59	14.07	14.07	16.92	15.51	15.51	18.31	15.51	12.59	7.81
to continue...													

continuation...													
Stations		J	F	M	A	M	J	J	A	S	O	N	D
Eseka	χ^2	1.983	1.776	8.808	**16.27**	9.443	3.014	6.884	1.931	4.046	**22.07**	4.088	.243
	df	4	4	6	7	8	6	5	6	7	8	6	3
	cv	9.49	9.49	12.59	14.07	15.51	12.59	11.07	12.59	14.07	15.51	12.59	7.81
Garoua	χ^2	0	0	.0004	.593	10.791	**13.13**	10.413	10.928	12.011	5.305	0	0
	df	0	0	1	4	6	6	8	8	7	4	0	0
	cv	0	0	3.81	9.49	12.59	12.59	15.51	15.51	14.07	9.49	0	0
Kaele	χ^2	0	0	0	.003	.816	2.535	3.130	3.681	1.407	.834	0	0
	df	0	0	0	2	5	5	8	7	6	3	0	0
	cv	0	0	0	5.99	11.07	11.07	15.51	14.07	12.59	7.81	0	0
Koundja	χ^2	.003	.040	1.779	10.189	.520	4.356	1.905	7.126	**21.80**	6.411	6.193	.0004
	df	1	2	5	6	6	6	7	7	8	6	4	1
	cv	3.84	5.99	11.07	12.59	12.59	12.59	14.07	14.07	15.51	12.07	9.49	3.84
Kribi	χ^2	4.405	8.908	5.439	2.767	3.469	5.035	6.125	12.078	6.632	11.660	9.392	3.514
	df	5	7	8	9	10	11	7	8	12	11	8	5
	cv	11.07	14.07	15.51	16.92	18.31	19.68	14.07	15.51	21.03	19.68	15.51	11.07
Lomie	χ^2	.774	.029	1.960	2.156	3.135	**19.61**	3.386	10.875	3.283	.755	.124	.012
	df	3	3	6	6	6	7	7	7	7	7	5	2
	cv	7.81	7.81	12.59	12.59	12.59	14.07	14.07	14.07	14.07	14.07	11.07	5.99
Maroua	χ^2	0	0	0	.031	1.050	11.047	2.959	2.483	1.467	1.671	0	0
	df	0	0	0	2	4	5	7	7	6	3	0	0
	cv	0	0	0	5.99	9.49	11.07	14.07	14.07	12.59	7.81	0	0
Meiganga	χ^2	.003	.001	.212	2.619	2.839	.103	4.253	14.322	9.115	2.676	1.612	.0003
	df	1	1	3	4	6	5	7	9	7	5	3	1
	cv	3.84	3.84	7.81	9.49	12.59	11.07	14.07	16.92	14.07	12.59	7.81	3.84
Nanga-Eboko	χ^2	.003	3.147	5.501	6.706	1.980	1.843	.591	7.247	15.699	13.629	8.978	.0002
	df	2	3	6	6	6	5	5	6	10	9	5	2
	cv	5.99	7.81	12.59	12.59	12.59	11.07	11.07	12.59	18.31	16.92	11.07	5.99
Ngambe	χ^2	.010	1.343	3.752	6.645	8.511	7.680	**16.41**	**26.39**	5.852	2.148	3.736	.081
	df	2	3	6	6	7	7	8	12	9	8	5	2
	cv	5.99	7.81	12.07	12.07	14.07	14.07	15.51	21.03	16.92	15.51	9.49	5.99
Ngaoundere	χ^2	0	0	.199	2.903	.403	4.155	11.986	.913	.725	10.615	.0004	0
	df	0	0	3	5	5	6	7	7	6	5	1	0
	cv	0	0	7.81	11.07	11.07	12.59	14.07	14.07	12.59	11.07	3.84	0
Nkongsamba	χ^2	.0003	.098	2.812	.151	4.223	3.106	**22.46**	**26.12**	16.952	8.831	.546	.024
	df	2	3	5	5	6	6	10	11	9	7	4	2
	cv	5.99	7.81	11.07	11.07	12.59	12.59	18.31	19.68	16.92	14.07	9.49	5.99
Poli	χ^2	0	0	.034	1.192	13.456	4.770	12.949	**17.94**	11.660	9.252	.002	0
	df	0	0	2	5	8	7	9	9	7	6	1	0
	cv	0	0	5.99	11.07	15.51	14.07	16.92	16.92	14.07	12.59	3.84	0
Sangmelima	χ^2	6.088	.427	5.642	3.293	2.843	5.432	.555	10.831	7.046	13.258	1.520	.0004
	df	4	4	6	6	7	6	5	5	7	7	5	2
	cv	9.49	9.49	12.59	12.59	14.07	12.59	11.07	11.07	14.07	14.07	11.07	5.99
	χ^2	0	1.158	2.273	.703	6.790	2.827	1.912	5.182	5.024	**16.25**	.003	0
	df	0	2	4	4	6	6	7	7	7	7	3	0
to continue...													

continuation...													
Stations		J	F	M	A	M	J	J	A	S	O	N	D
Tibati	cv	0	5.99	9.49	9.49	12.59	12.59	14.07	14.07	14.07	14.07	7.81	0
Yaounde	χ^2	.025	.964	6.353	2.369	.548	1.078	2.821	3.230	2.479	15.078	1.219	.001
	df	3	4	6	6	6	6	4	5	7	8	5	2
	cv	7.81	9.49	12.59	12.59	12.59	12.59	9.49	11.07	14.07	15.51	11.07	5.99
Yokadouma	χ^2	8.154	.097	2.421	.681	15.251	2.183	10.412	8.457	13.409	3.591	1.541	.106
	df	4	4	5	6	8	6	8	7	9	7	5	3
	cv	9.49	9.49	11.07	12.59	15.51	12.59	15.51	14.07	16.92	14.07	11.07	7.81
Yoko	χ^2	.0006	.223	.157	5.949	6.401	.350	4.761	3.542	3.099	6.242	3.655	.0004
	df	1	3	4	5	6	5	6	6	7	7	4	1
	cv	3.84	7.81	9.49	11.07	12.59	11.07	12.59	12.59	14.07	14.07	9.49	3.84

The addition of an overlay normal distribution to the Weibull distribution shows a better fit, characterized by a reduction of χ^2 contribution at some stations. The better fit is the proof for our hypothesis the long rain periods overlapping several days to cause the misfit of the Weibull distribution. The equation (24) would be a better distribution for simulating daily rainfall amounts; but we are not able to use it for simulation because we cannot integrate it as we did in equation (20). We are not able to invert it in order to use Monte Carlo simulation of daily rainfall amounts.

5.2 The temporal distribution of parameters

In Cameroon an understanding of the seasonal variation of precipitation throughout the year is important for its influence on the economic activities of the population and in the construction of models. Several approaches have been used to deal with the seasonality. In this study, we assume that parameters $g(m)$ and $b(m)$ of Weibull distribution vary as step functions for each month and we use the double normal function (see, equation (10)) to describe the seasonal variation of parameters. We use a double normal function here because it describes the two rainy seasons and their interaction with big and small dry season. However, if there is only one rainy season, we expected that this could also be represented by near identical values of the two part of the normal distribution.

The parameters $g(m)$ and $b(m)$ of Weibull distribution vary from month to month; therefore, double normal function (see equation (10)) is used to describe the monthly variation of $g(m)$ and $b(m)$. Its coefficients are estimated by a minimization of errors e_g and e_b

$$e_g(a_g, \mu_{g1}, s_{g1}, \mu_{g2}, s_{g2}) = \sum_{m=1}^{12} \big(z(m) - g(m)\big)^2 \tag{26}$$

$$e_b(a_b, \mu_{b1}, s_{b1}, \mu_{b2}, s_{b2}) = \sum_{m=1}^{12} \big(z(m) - b(m)\big)^2 \tag{27}$$

where

$g(m)$ is the shape parameter for a month m

$b(m)$ is the scale parameter for a month m

The parameters we are looking for, are the ones which minimize the error functions e_g and e_b.

We have applied the above techniques to the data. Table 8 presents estimated coefficients of double normal function for $g(m)$ and $b(m)$. It is shown that coefficients vary from station to station.

Table 8: Estimated coefficients of the double normal function

	$g(m)$					$b(m)$				
Station	a_g	μ_{g1}	s_{g1}	μ_{g2}	s_{g2}	a_b	μ_{b1}	s_{b1}	μ_{b2}	s_{b2}
AbongMbang	0.9	2.7	4.2	10.8	5.1	139.0	3.6	4.8	10.7	3.1
Akonolinga	0.9	3.5	3.6	11.3	5.7	143.0	3.5	4.0	10.0	3.1
Bafia	1.0	3.1	3.1	9.8	4.5	148.0	3.8	4.1	10.3	2.5
Bamenda	0.9	2.7	4.2	10.8	5.1	133.0	3.3	4.6	10.6	4.0
Batouri	0.9	2.7	4.9	11.6	5.7	120.0	3.2	5.5	12.2	6.0
Bertoua	1.0	2.5	3.6	11.1	5.8	129.0	1.6	5.3	10.0	4.9
Betare-oya	1.0	1.0	3.2	9.2	6.1	129.0	1.5	.9	8.2	9.5
Douala	0.9	3.3	4.6	11.2	3.7	125.0	6.4	9.1	8.0	2.4
Dschang	0.9	2.7	4.3	11.0	5.2	139.0	3.6	4.8	10.7	3.1
Edea	1.0	3.4	3.6	11.0	4.0	143.0	3.5	4.5	10.6	3.2
Eseka	0.9	3.9	4.2	10.8	2.8	142.0	3.5	4.6	10.6	3.1
Garoua	0.9	2.5	3.0	9.5	5.5	145.0	1.6	.3	7.0	11.9
Kaele	1.1	1.0	6.3	11.6	4.7	133.0	3.3	4.6	10.6	4.2
Koundja	0.9	1.0	6.0	10.3	6.7	151.0	2.6	.3	7.2	7.0
Kribi	0.8	1.1	4.8	10.5	5.8	143.0	3.5	4.5	10.6	3.2
Lomie	1.1	3.2	5.4	11.9	3.0	122.0	2.9	4.1	10.7	5.3
Maroua	1.0	2.7	1.5	8.7	6.6	132.0	3.2	4.5	10.6	4.2
Meiganga	0.9	2.9	5.0	11.9	5.4	114.0	7.5	.4	6.6	8.9
Nanga-Eboko	0.9	1.5	4.8	11.0	2.9	133.0	3.3	4.6	10.6	4.0
Ngambe	1.0	1.0	6.5	9.1	4.3	146.0	3.2	2.8	9.4	4.0
Ngaoundere	1.1	3.5	6.1	12.9	4.0	144.0	3.1	3.5	10.1	4.0
Nkongsamba	1.0	.8	5.5	10.6	6.6	113.0	3.6	5.5	9.2	4.5
Poli	1.0	1.0	3.1	8.8	6.1	118.0	3.3	6.0	12.6	6.5
Sangmelima	.9	2.3	4.4	11.0	3.5	129.0	2.2	3.6	10.2	5.7
Tibati	1.0	3.6	3.1	9.7	4.0	141.0	2.4	.5	7.1	5.3
Yaounde	.9	2.1	4.4	11.0	6.5	122.0	2.2	4.3	9.6	4.8
Yokadouma	.9	3.0	3.1	9.6	5.3	119.0	2.9	4.5	9.8	4.5
Yoko	1.0	1.0	4.6	10.0	5.3	138.0	3.6	4.9	10.7	3.1

5.3 Spatial distribution of parameters from predictors

In this part of the study, our main concern is the spatial distribution of different seasonal course of rain amounts. This is why we try to describe the seasonal variation by time invariant parameters. It's in order to describe its spatial variation and for the latter, we need the spatial predictors.

Like in the case of probability estimation, stepwise regression analysis is also applied to parameters of $g(m)$, $b(m)$ and spatially distributed data. The square of the latitude φ^2, logarithm of the latitude $\ln(\varphi)$, and its cosine $\cos(\varphi)$, the longitude λ, the altitude ω, the relief orientation ϑ, the rain exposition κ, and yearly rain amount θ are independent variables or predictors whereas parameters $a_g, \mu_{g1}, s_{g1}, \mu_{g2}, s_{g2}$ of g and $a_b, \mu_{b1}, s_{b1}, \mu_{b2}, s_{b2}$ of b of the double normal function used to describe the seasonal variation of $g(m)$ and $b(m)$ are dependent variables.

As outcome of stepwise regression, 5 % significance level is chosen to decide about the exclusion of an independent variable. Estimated parameters of $g(m)$ and $b(m)$, independent variables selected for the estimation of coefficient, unstandardized coefficient B, Beta coefficient as well as the t-values and significance value are summarized in table 9.

Table 9: Summary of regression coefficients for estimating $g(m)$ and $b(m)$ parameters

parameters	variable	coefficients (B)	Beta coefficient	t	5 % significance value
a_g	*constant*	.795		10.92	.000
	φ^2	.069	.399	2.22	.036
μ_{g1}	*constant*				
s_{g1}	*constant*				
μ_{g2}	*constant*				
s_{g2}	*constant*	4.42		11.04	.000
	λ	.011	.401	2.23	.035
a_b	*constant*				
μ_{b1}	*constant*	1.85		2.82	.009
	θ	.0008	.414	2.32	.028
s_{b1}	*constant*	1.77		2.80	.008
	θ	.0012	.397	2.20	.036
μ_{b2}	*constant*				
s_{b2}	*constant*	-2.35		11.04	.000
	λ	.588	.451	2.57	.016

From table 9 for g, the t-values show that only the square of latitude φ^2 and rain exposition κ are included in the equations whereas the logarithm of latitude, the cosine of latitude, the longitude, the altitude, the relief orientation and the yearly rain amount are excluded. The square of latitude φ^2 is the only spatial predictor which is able to be used for predicting parameter a_g. On the other hand, the rain exposition κ appears to be a good spatial predictor for approximating parameter s_{g2}.

There are no independent variables to be used for predicting parameters μ_{g1}, s_{g1}, and μ_{g2}. We then consider μ_{g1}, s_{g1}, and μ_{g2} to be constant and we estimate them by calculating their mean value. Parameters of $g(m)$ are summarized in equations (28) below

$$\begin{aligned} \mathrm{a}_g &= .795 + .0697\varphi^2 \\ \mu_{g1} &= 2.346 \\ \mathrm{s}_{g1} &= 4.325 \\ \mu_{g2} &= 10.596 \\ \mathrm{s}_{g2} &= 4.424 + .0109\kappa \end{aligned} \tag{28}$$

For b, the t-values show that only longitude and yearly rain amounts are included in the equations whereas other independent variables are excluded. The yearly mean rainfall amount θ is found to be a good predictor to be used for approximating parameters μ_{b1} and s_{b1} of $b(m)$. On the other hand, the longitude λ is found to be a good predictor to be used for approximating parameters s_{b2}. No dependent variable is found to be used for estimating parameters a_b and μ_{b2}. We estimated a_b and μ_{b2} by computing their mean value. Parameters of $g(m)$ are summarized in equations (29) below

$$\begin{aligned} \mathrm{a}_b &= 133.321 \\ \mu_{b1} &= 1.85 + .0008\theta \\ \mathrm{s}_{b1} &= 1.77 + .0012\theta \\ \mu_{b2} &= 9.768 \\ \mathrm{s}_{b2} &= -2.35 + .588\lambda \end{aligned} \tag{29}$$

Equations (28) and (29) are then used to spatially describe the seasonal variation of parameter $g(m)$ and $b(m)$ as follows

$$g(m) = a_g \left(exp(-(\frac{m-\mu_{g1}}{s_{g1}})^2) + exp(-(\frac{m-\mu_{g2}}{s_{g2}})^2) \right) \tag{30}$$

$$b(m) = a_b \left(exp(-(\frac{m-\mu_{b1}}{s_{b1}})^2) + exp(-(\frac{m-\mu_{b2}}{s_{b2}})^2) \right) \tag{31}$$

$g(m)$ is fixed by five parameters namely $a_g, \mu_{g1}, s_{g1}, \mu_{g2}, s_{g2}$.
$b(m)$ is fixed by five parameters namely $a_b, \mu_{b1}, s_{b1}, \mu_{b2}, s_{b2}$.

Thus, the Weibull distribution is used to model the daily rainfall amounts, the double normal function is used for modelling the monthly variation of Weibull parameters over the year. By stepwise regression analysis we estimate the parameters of double normal function from the spatial predictors describing the conditions of location. Estimated coefficients are then used to develop models for simulating daily rainfall amounts r on a wet day, using a pseudo random number generator and Monte Carlo simulation methods.

6 SIMULATION PROCEDURES AND VALIDATION OF THE RESULTS

Simulation procedures use models constructed in chapter 4 and 5 for generating daily rainfall occurrence and daily rainfall amounts on wet days.

6.1 Simulating rainfall occurrence

Be r the rain amount falling on day t. Then rainfall occurrence probability $\nu(t+1)$ on the next day is modelled by first order Markov chain; with $m = 1, ..., 12$, monthly transition probabilities $p_{01}(m)$ and $p_{11}(m)$ depend from r:

$$\nu(t+1) = \begin{cases} p_{01}(m) & \text{if} \quad r \leq 1mm \\ p_{11}(m) & \text{if} \quad r > 1mm \end{cases} . \tag{32}$$

Transition probabilities $p_{i1}(m)$ were transformed into logits $Y_{i1}(m)$ (see equation (12)) approximated by $z(m)$ (see equation (10)). This method only approximates logits; Therefore, transition probability need for Monte Carlo simulation is determined for month m by inverting the logit functions $Y_{i1}(m)$ again to probability as follows

$$p_{i1}(m) = \left(1 - exp(-Y_{i1}(m))\right)^8 \qquad i = 0, 1 \tag{33}$$

Now knowing the transition probability wet-dry and wet-wet (see equation (33)), we can easily derive the dry-wet $p_{10}(m)$ and dry-dry $p_{00}(m)$ transition probabilities from the latters $p_{i0}(m) = 1 - p_{i1}(m)$, $(i = 0, 1)$.

As proposed by Schwarz (1980), the function $Y_{i1}(m)$ describing the seasonal distribution of rainfall is chosen to be the double normal functions each fixed by five parameters $a_i, \mu_{i1}, s_{i1}, \mu_{i2}, s_{i2} \quad (i = 0, 1)$.

$a_i, \mu_{i1}, s_{i1}, \mu_{i2}, s_{i2}$ represent the spatial distribution of rainfall regime in Cameroon. Accordingly, they are estimated by variables describing the condition of location. Thus the models for simulating daily rainfall occurrence wherever in Cameroon by the help of spatially distributed data are sumarized by equation (16).

6.2 Simulating daily rainfall amounts

Be $W(r)$, the Weibull density distribution of daily rain amount r

$$W(r) = \frac{g(m)}{b(m)} \left[\frac{r}{b(m)}\right]^{g(m)-1} * exp\left[-\left[\frac{r}{b(m)}\right]^{g(m)}\right] \tag{34}$$

The functions $g(m)$ and $b(m)$ (see equations (30) and (31)) describing the seasonal variation of Weibull parameters are chosen to be the double normal function (Schwarz 1980) each fixed by five parameters.

The parameters represent the spatial distribution of rainfall regime in Cameroon. Accordingly, they are estimated by variables describing the condition of location.

For Monte Carlo simulation of rain amount r using a random number $R \in [0, 1]$ equally distributed in the interval [0,1] as generated by pseudo-random number generator, we take the probability distribution $D(r)$ (see equation (18)) belonging to $w(r)$ (equation (34)). The substitution of $D(r)$ with any random number R yields the equation (19) and from it, we can derive the rainfall amount r as follows:

$$r = b(m) * \big(-lnR\big)^{\frac{1}{g(m)}} \tag{35}$$

We then use the random number R to produce Weibull distributed rainfall r.

Models for simulating daily rainfall occurrence on the one hand and models for independently simulating daily rainfall amounts on wet day wherever in Cameroon by the help of spatially distributed data have been constructed. At this step of the study, we can now apply the Monte-Carlo simulation methods and pseudo random number generator for generating the daily rainfall series.

If computers were able to produce true random numbers, you could not use them. No computer can work randomly. The main problem of generating uniformly distributed random numbers in [0,1] is that computers can only produce pseudo-random numbers; therefore for this study, we use pseudo-random numbers. Constructing pseudo-random numbers is an art. A lot of methods exist for this purpose. For our study, we used a multiplicative congruential method with sequence repetition of not less than $2^{31} - 1$.

Conditions are combined to "let it rain" and to determine "if it rain", how much rain?. Regrettably, we were not able to use the observation data that are independent from the ones we used for calibration, i.e estimating all the parameters.

6.3 Validation of the outcome of rainfall series

The purpose of the stochastic generator is to produce data which are statistically similar to the observed series. In other words, the statistics including means, variances, relative frequency of occurrence, correlations, and lag correlations between variables derived from synthetic data should be statistically insignificantly different from those derived from the observed data. In validating our generator, time series of 100 years are generated. The synthetic series are analysed and resultant statistics are compared with those derived from the observed series. The time series from 1951-1993 are used for calibrating the models whereas time series from 1994-2000 available at some stations (Douala, Kaele, Koundja, Ngambe, Ngaoundere and Sangmelima) are used to test the observed against the estimated series.

6.3.1 Distribution of daily precipitation amounts

Equation (35) is used to produce Weibull distributed rain amounts r. The distribution of daily rain amounts is tested by comparing the frequency distribution of observed and estimated daily rainfall amounts. Generated and observed daily rainfall totals are grouped into classes of 10 mm intervals, beginning with 1 mm and tested using χ^2 . The frequency classes and the results of the test are presented in table 10. For carrying out χ^2 tests, classes with expected frequencies less than 5 are joined to a new combined class. By combining two classes to one we have one degree of freedom less; this explains the reduction of degrees of freedom in the table.

Table 10: Test of distribution functions of observed and generated rainfall amounts in 10^{th} of mm

Stations	χ^2	degrees of freedom	critical value
Douala	53.87	10	18.31
Kaele	5.79	4	9.49
Koundja	70.60	6	12.59
Ngambe	19.58	6	12.59
Ngaoundere	7.91	6	11.07
Sangmelima	11.00	5	11.07

The analysis of table 10 shows the high χ^2 contribution in Douala, Koundja and

Ngambe. This suggests a difference between the frequency distribution of the generated and the observed daily rainfall amount series. In others words, the frequency distribution of the generated daily rainfall amount is not resembling the ones of the observed series. In Kaele, Ngaoundere, and in Sangmelima, the low value of χ^2 contribution indicates the resemblance between the two frequency distributions. The generator produces good estimate of daily rainfall amount in the northern area (Kaele) where the climate is generally dry, and in the southern plateau (Sangmelima) which is characterized by four seasons. Weak rainfall amounts are generated by the generator, at the stations situated on highlands (Koundja) and in the coastal area (Douala and Ngambe); the latter regions experience high rainfall amount during the rainy season.

6.3.2 Test of the seasonal distribution of wet-day series

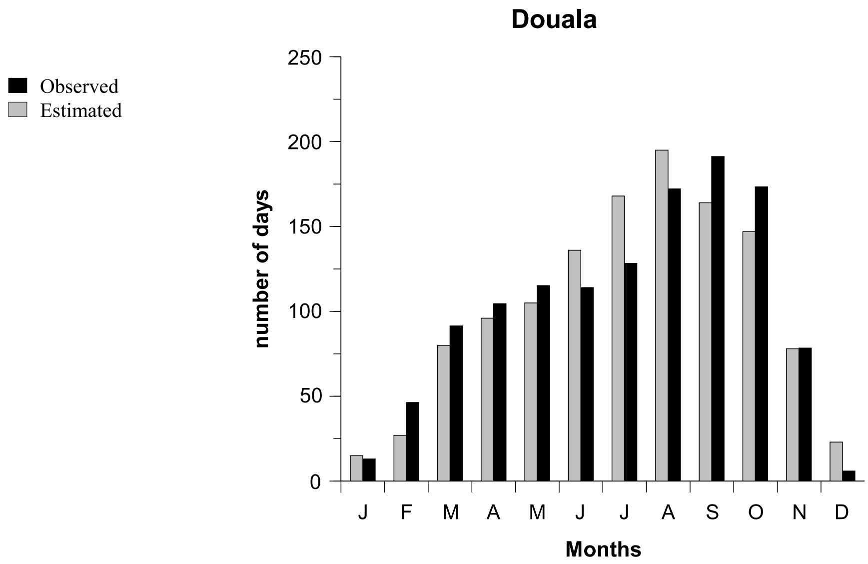
The rainfall occurrence probability $\nu(t+1)$ on day $t+1$ is modelled by first-order two-state Markov chain with $m = 1, \ldots, 12$ monthly transition probabilities $p_{01}(m)$ and $p_{11}(m)$ depending from r as defined in equation (32). Using the latter equation, we simulate rainfall occurrence series. From the simulated series, we then compute the frequency of wet days for each month. This frequency distributions of observed and simulated series are compared using χ^2 test and the results are presented at 5 % significance level. Data are divided into classes j, with 10^{th} of mm interval between classes. Chi square test requests expected frequencies above 5. To deal with this problem, the classes with expected frequencies less than 5 are joined to a new combined class and counts of both classes are added. Thus, by combining two classes to one we have one degree of freedom less. Results of statistical tests comparing the observed data for six sites with synthetic data generated by our simulator for the seasonal distribution of wet series are presented in table 11.

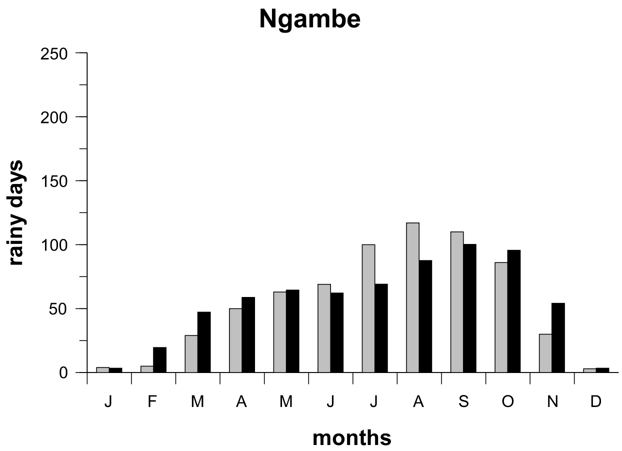
The analysis of table 11 shows that the χ^2 contribution is very high in Douala and Ngambe; i.e. the distribution frequency of the generated series do not resemble the distribution frequency of the observed series in Douala and Ngambe. The low χ^2 contribution suggests that the statistics of the generated series and the ones of the observed series are considered to be the same; this is the case in Kaele, Ngaoundere, Koundja, and Sangmelima. Overall the generator seems to be able to generate rain series for inland stations with statistics resembling the observations. For the coastal areas (Douala and Ngambe), it seems that other factors than the chosen predictors may well explain the behaviour of daily rainfall; Cameroon's coastal climate is of an equatorial type and is influenced by the meteorological equator, being the meeting point between the anticyclone of Azores (North Atlantic) and that of Saint Helen (South Atlantic). This climate results from the combined effect of convergence of the tropical oceanic low-pressure zone and the inter-tropical front within the continent. South-westerly monsoon winds predominate, modified by land sea breezes causing humidity values to almost saturation point. Wind speeds exceptionally reach values of 18 $m\ sec^{-1}$ (April, 1993) with average values recorded over a period of 10 years (1983 - 1993) varying between 0.5-2.5 $m\ sec^{-1}$. Thus, the dynamics of the monsoon circulation, the surface wind and the sea surface temperatures are factors that mostly influence rainfall in the coastal area and should be taken into account as predictor variables when simulating rainfall in this area. Another reason is that of the split of rain over many days in this area (see figure 20). The monthly distribution of observed and simulated wet-days series is plotted in figure 21.

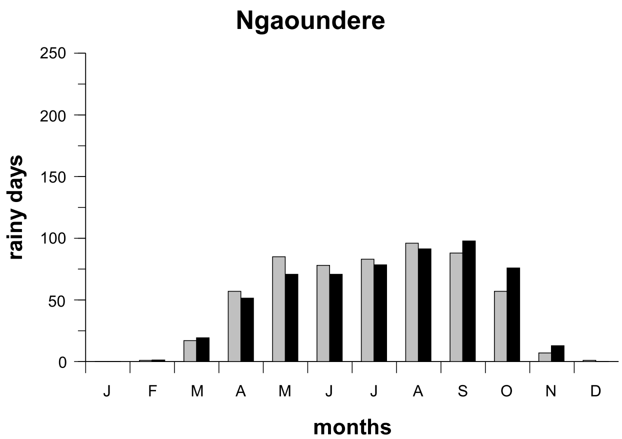
In general, the simulated series present the same configuration as the observed series. The maximum of wet days occurs during the rainy season. In Douala, Ngambe and Ngaoundere, the maximum wet days occurs in August for the observed series whereas it occurs in September for the estimated series. At Kaele and Koundja, the maximum of both observed and estimated occurs in the month of August. Sangmelima presents the maximum of both observed and estimated in the month of October.

6.3.3 Yearly mean and variance of monthly rainfall

Monthly rainfall of a certain month for various years are tested. i.e. we make tests for each month of a yearly rainfall of month m with $m = 1, 2, \ldots, 12$. Yearly means μ are tested using the t-test whereas the F-test is used to test the variance σ^2. Table 12 gives the values of the F-test statistic and t-test statistic comparing synthetic and observed series. The significance of F values indicate that all values are above .05 over all months in Kaele, Koundja, Ngambe, Ngaoundere, Sangmelima; this suggests that there is no difference between the variance of observed and simulated series at 5 % significance level. In Douala, the significance of F values are above .05 in January, February, March, April May, June, July, September and December whereas in August, October, and November these values are less than .05. This suggests that there is no difference between the variance


Observed
Estimated
Douala
250
200
150
100
50
0
number of days
J F M A M J J A S O N D
Months


Ngambe
250
200
150
100
50
0
rainy days
J F M A M J J A S O N D
months


Ngaoundere
250
200
150
100
50
0
rainy days
J F M A M J J A S O N D
months

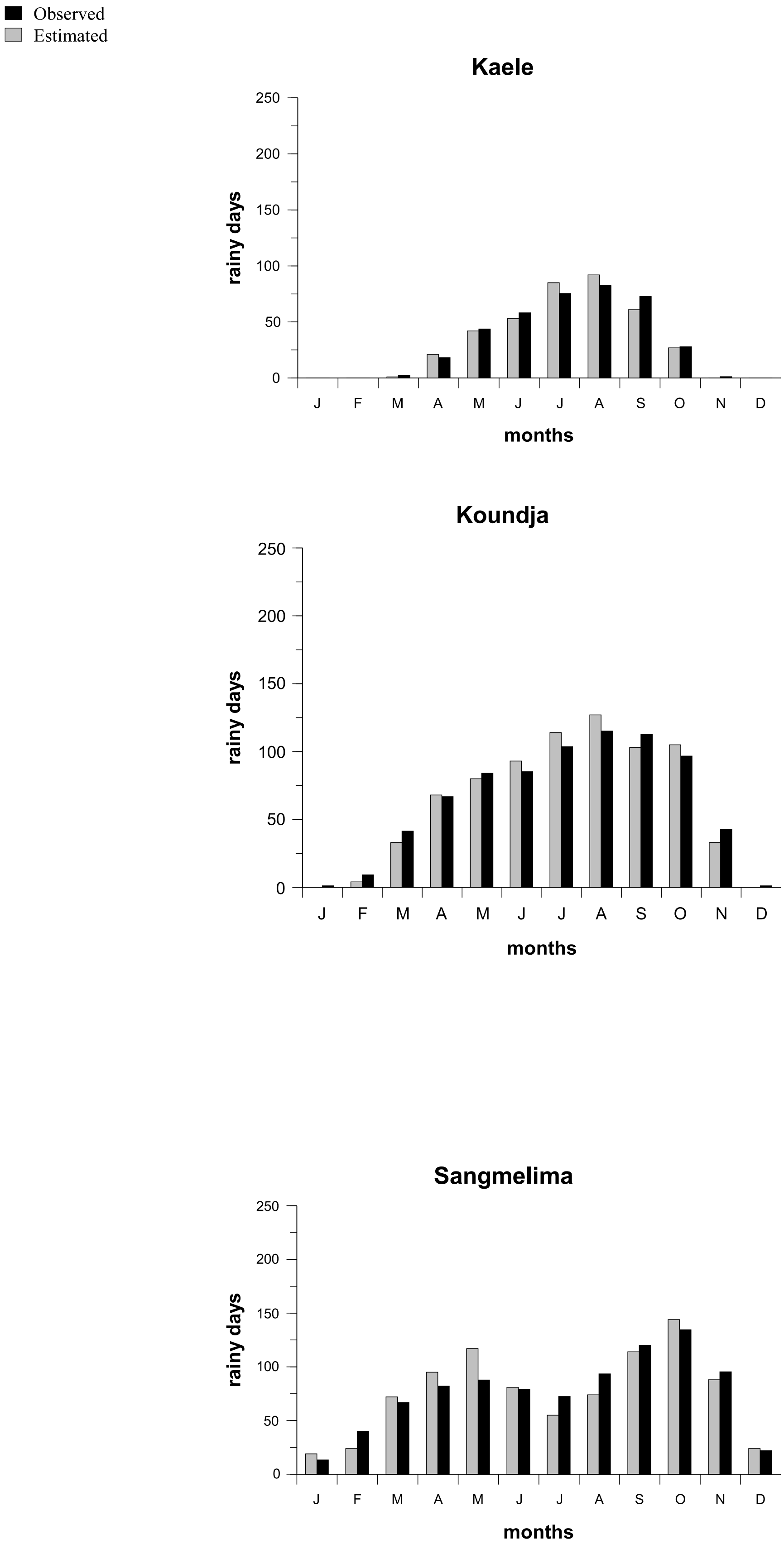


Figure 21: Monthly distribution of observed and estimated absolute rainy days

of observed and simulated series at 5 % significance level during the months of January, February, March, April May, June, July, September and December. The values under .05 indicate that there is a difference between the variance of observed and simulated series in August, October, and November.

The test of the mean rainfall indicates that the significance value is above .05 in Kaele, Koundja, Ngambe, Sangmelima for all the 12 months suggesting that there is no difference between the mean of observed and simulated rainfall series. In Douala, the significance is less than .05 in August. In Ngaoundere the significance is less than .05 in December. The significance value under .05 suggests a difference between the mean of observed and simulated data in the later months.

The generator is considering to simulate the wet days series for inland stations with statistics resembling the observations. This is not the case in the coastal areas represented by Douala and Ngambe stations. That can be interpreted by the fact in the coastal area south-westerly monsoon winds predominate, modified by land sea breezes causing humidity values to almost the saturation point, bringing thus the region more rain than the inland stations. There is sometimes continuous rain over several days especially during wet months; by splitting over few days, all these days have high rainfall amounts of nearly the same order of magnitude. Thus, when fitting a distribution to the daily rainfall amounts, the chosen distribution could predict much fewer big rain amounts, the higher the rain is. The effect of these methodological factors and errors is primarly reflected in the occurrence process and in the smaller amounts of rainfall. The generator produces good estimate of daily rainfall amounts in the northern area where the climate is generally dry, and in the southern plateau which is characterized by four seasons. The underestimation of rainfall amount in coastal and highland region can be explained by the fact that these regions are under the influence of monsoon winds, see breeze which have to be taking in account as predictors when estimating the rainfall in coastal region. The result also suggests that the far distance from the ocean, we have good estimation; we have then to include the distance from the sea as another predictor.

The result of the present study exhibits no difference between the monthly mean of rainfall, excepted Douala which presented a difference between the mean of the generated and observed rainfall series in August, which corresponds to its wettest period of the year.

The situation is the same when testing the monthly variance; the values under .05 indicate that there is a between the variance of observed and simulated series in August, October, and November in Douala.

6.3.4 Outcome of probabilities

Equation (33) is used to compute the wet-dry and wet-wet transition probabilities and the results are presented in table 13. The validation here consists in graphically compare the estimated probabilities against the observed ones. This comparison is presented in figure 22 and figure 23 for wet-dry and wet-wet transition probabilities respectively.

The models reproduced well the seasonal variation of wet-dry and wet-wet transition probabilities at all the stations tested in spite of the gap existing between the observed and the estimated probabilities. It is also noted that there is a discrepancy between the month where the maximum (peak) and the minimum of estimated probabilities occurred against the observed ones. The difference can be explained by the fact that the peaks do not look like bell shaped normal distribution but rather resemble shield volcanoes or matterhorus; because of such form, they cannot be well represented by a normal function.

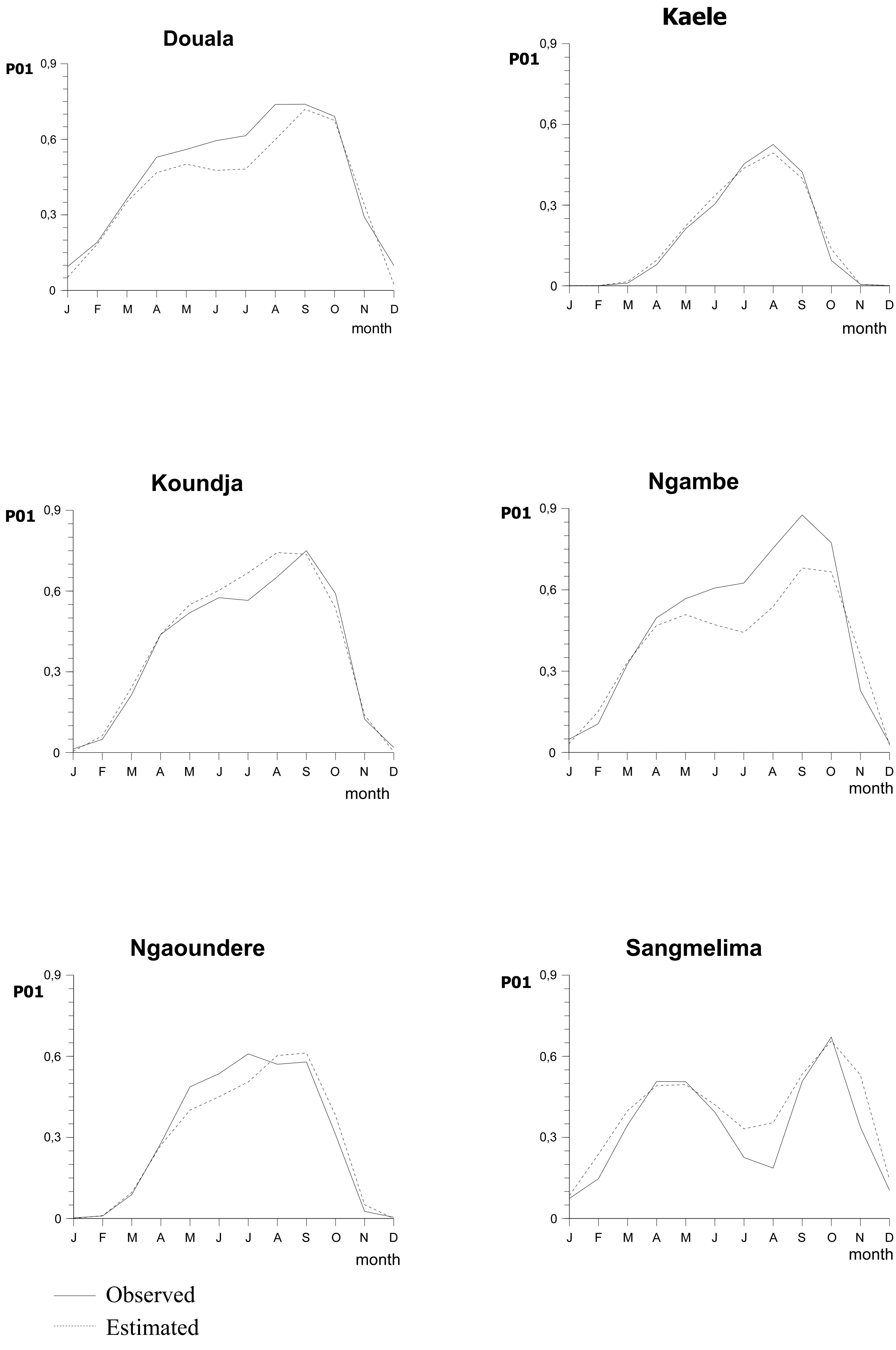


Figure 22: Observed and estimated wet/dry probabilities

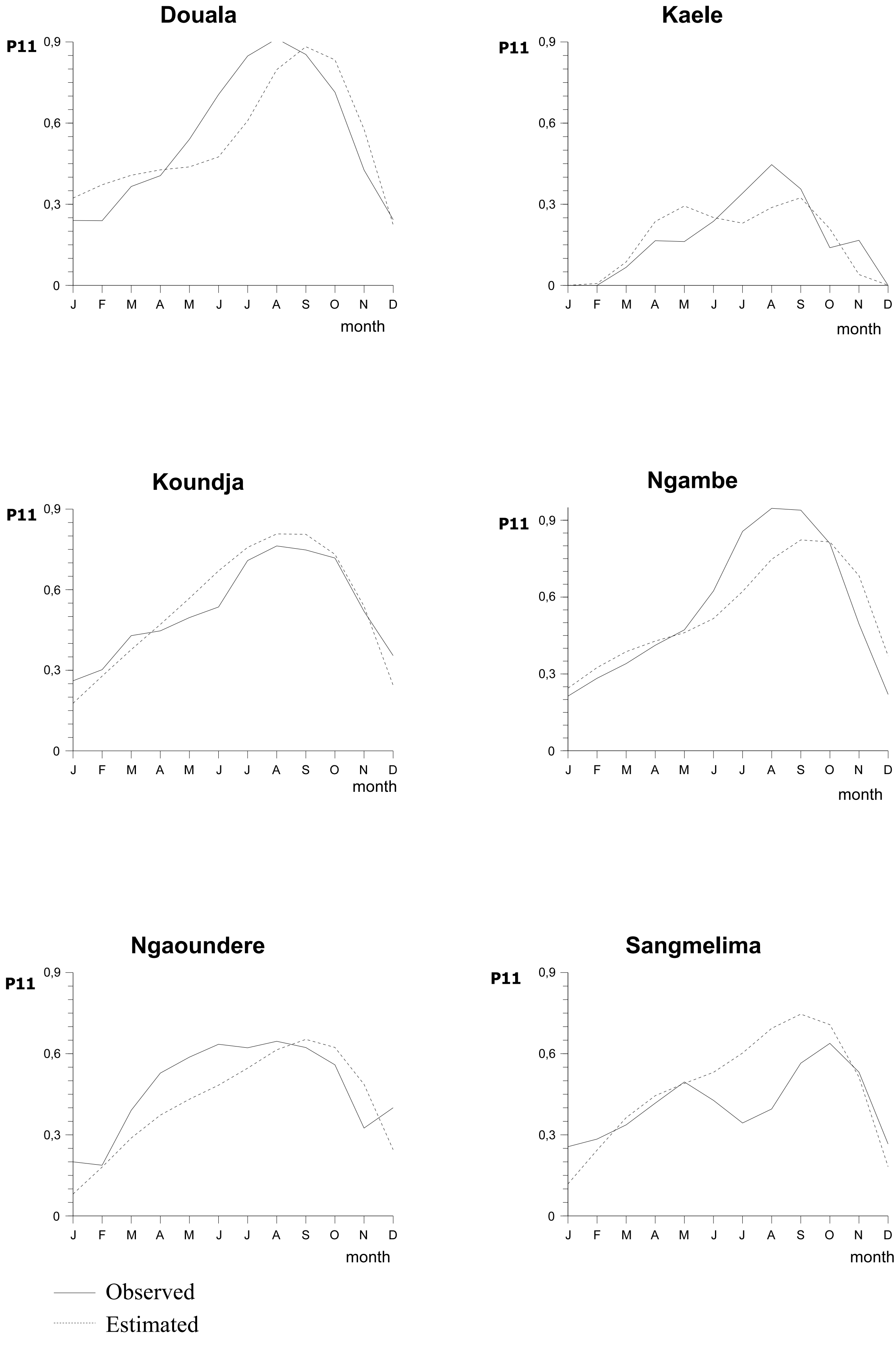


Figure 23: Observed and estimated wet/wet probabilities

Table 11: Test of seasonal distribution of wet day series

		J	F	M	A	M	J	J	A	S	O	N	D	df	χ^2	CV
Kaele	observed	0	0	1	21	42	53	85	92	61	27	0	0	6	5,39	12,59
	estimated	0	0	3	17	44	58	75	82	73	28	1	0			
Ngaoundere	observed	0	1	17	57	85	78	83	96	88	57	7	1	8	12,65	15,51
	estimated	0	2	18	51	71	71	78	91	98	76	13	0			
Koundja	observed	0	4	33	68	80	93	114	127	103	105	33	0	8	6,38	15,51
	estimated	1	9	34	67	84	85	104	115	113	97	28	1			
Sangmelima	observed	19	24	72	95	98	81	55	84	114	144	88	24	11	19,42	19,68
	estimated	13	40	67	82	88	79	72	93	120	134	95	22			
Ngambe	observed	3	5	29	50	63	69	100	117	110	86	30	3	9	71,22	16,92
	estimated	4	16	38	48	54	52	69	73	85	86	52	4			
Douala	observed	15	27	80	96	105	136	168	195	164	147	78	23	11	30,96	19,68
	estimated	13	36	91	81	115	114	128	172	165	155	75	16			

Table 12: Values of F-test statistic and t-test statistic for the monthly rainfall amount

			J	F	M	A	M	J	J	A	S	O	N	D
Douala	σ^2	F	0.588	1.315	0.364	1.971	2.604	2.859	3.092	15.915	0.419	7.976	11.341	2.470
		Sig	0.444	0.252	0.546	0.161	0.107	0.091	0.079	0.00	0.518	0.005	0.001	0.119
	μ	t	-0.548	0.228	-0.531	0.440	0.540	-1.681	1.068	2.917	-0.429	1.237	1.354	1.948
		Sig	0.584	0.820	0.595	0.660	0.589	0.093	0.285	0.004	0.668	0.216	0.176	0.054
Kaele	σ^2	F	-	-	-	0.001	0.185	2.423	0.481	2.100	1.932	0.074	-	-
		Sig	-	-	-	0.973	0.668	0.120	0.488	0.148	0.165	0.786	-	-
	μ	t	-	-	-0.376	0.199	-0.043	-1.850	-0.213	-0.677	-1.246	-1.108	-	-
		Sig	-	-	0.709	0.843	0.965	0.071	0.831	0.499	0.213	0.269	-	-
Koundja	σ^2	F	-	0.597	1.618	1.870	2.165	3.277	2.459	1.860	1.887	2.975	2.899	-
		Sig	-	0.441	0.204	0.172	0.141	0.070	0.117	0.173	0.170	0.085	0.089	-
	μ	t	-	0.058	-0.915	-0.677	-0.447	-0.582	-0.685	0.530	-0.358	-0.628	-0.749	-
		Sig	-	0.954	0.360	0.498	0.655	0.561	0.494	0.596	0.720	0.530	0.454	-
Ngambe	σ^2	F	0.987	0.985	1.311	0.486	0.606	0.792	3.538	2.069	0.314	0.076	0.036	1.046
		Sig	0.323	0.321	0.252	0.486	0.437	0.373	0.060	0.151	0.575	0.783	0.850	0.306
	μ	t	-0.664	-0.334	-1.623	-0.032	0.262	-0.970	-1.502	0.834	-0.670	-0.086	-0.252	-0.087
		Sig	0.508	0.738	0.105	0.975	0.793	0.332	0.133	0.404	0.503	0.932	0.801	0.931
Ngaoundere	σ^2	F	-	-	2.906	0.862	2.472	1.646	2.450	3.548	1.730	0.064	0.096	-
		Sig	-	-	0.089	0.353	0.116	0.200	0.118	0.060	0.189	0.800	0.758	-
	μ	t	-	-0.432	-1.094	-0.117	-1.237	-0.964	-1.286	-1.165	-1.718	-0.362	-0.282	-30.60
		Sig	-	0.671	0.272	0.907	0.216	0.335	0.199	0.244	0.086	0.717	0.778	0.020
Sangmelima	σ^2	F	0.069	0.227	0.209	1.309	0.162	0.466	0.076	0.682	0.167	1.189	0.423	1.355
		Sig	0.793	0.634	0.648	0.253	0.723	0.495	0.783	0.409	0.683	0.276	0.515	0.245
	μ	t	-0.435	0.320	-0.063	-0.742	-0.097	-0.942	0.091	0.678	-0.573	-0.095	0.122	-0.422
		Sig	0.663	0.749	0.950	0.458	0.923	0.346	0.928	0.498	0.577	0.924	0.903	0.673

Table 13: Observed and estimated monthly p_{01} and p_{11}

			J	F	M	A	M	J	J	A	S	O	N	D
Douala	p_{01}	observed	.095	.192	.365	.529	.560	.595	.615	.739	.739	.691	.292	.099
		estimated	.053	.184	.356	.468	.501	.477	.482	.598	.719	.675	.341	.023
	p_{11}	observed	.240	.239	.366	.406	.540	.706	.848	.912	.854	.714	.426	.243
		estimated	.323	.372	.407	.427	.438	.475	.609	.798	.882	.834	.577	.223
Kaele	p_{01}	observed	0	.0008	.009	.079	.213	.304	.453	.525	.423	.094	.006	.0008
		estimated	0	.0005	.015	.095	.223	.336	.437	.494	.339	.136	.005	0
	p_{11}	observed	0	0	.067	.165	.162	.237	.341	.446	.356	.139	.167	0
		estimated	0	.007	.087	.236	.294	.251	.229	.288	.324	.209	.040	.0007
Koundja	p_{01}	observed	.013	.049	.214	.438	.519	.575	.565	.653	.750	.591	.123	.019
		estimated	.005	.062	.242	.439	.549	.602	.668	.743	.737	.537	.139	.003
	p_{11}	observed	.261	.302	.429	.447	.496	.536	.709	.763	.748	.718	.520	.355
		etimated	.177	.278	.377	.471	5685	.670	.758	.808	.806	.732	.539	.244
Ngambe	p_{01}	observed	.049	.105	.324	.496	.567	.607	.625	.753	.877	.774	.228	.032
		estimated	.032	.151	.332	.468	.508	.471	.443	.537	.680	.666	.358	0.027
	p_{11}	observed	.213	.283	.340	.411	.472	.624	.856	.946	.939	.809	.494	.220
		estimated	.244	.325	.387	.428	.461	.516	.622	.748	.823	.815	.683	.369
Ngaoundere	p_{01}	oberved	.0002	.009	.089	.279	.487	.536	.609	.571	.579	.311	.026	.005
		estimated	.0002	.0106	.097	.271	.401	.430	.505	.602	.612	.382	.051	.0002
	p_{11}	observed	.020	.188	.391	.528	.587	.635	.622	.646	.623	.559	.325	.400
		estimated	.081	.181	.288	.373	.432	.484	.547	.614	.653	.623	.487	.245
Sangmelima	p_{01}	observed	.074	.147	.346	.507	.506	.393	.226	.187	.507	.671	.338	.105
		estimated	.084	.238	.399	.492	.495	.421	.332	.355	.534	.657	.531	.144
	p_{11}	observed	.256	.284	.337	.417	.495	.428	.344	.396	.565	.638	.532	.266
		estimated	.118	.244	.364	.445	.490	.531	.602	.694	.746	.707	.513	.182

7 CONCLUSION

The Markov chain model is used on daily precipitation in 28 sites in Cameroon. A two-state first-order model was found to be applicable to daily rainfall occurrence of the rain. Transition probability between dry or wet and wet days are calculated and fitted with double normal function to describe its seasonal variation. Weibull distribution function is fitted to daily rainfall amount; the seasonal variation of its parameters is also described by the double normal function. Parameters of double normal function are estimated by minimizing errors and are approximated from spatially distributed predictors such as latitude, longitude, altitude, relief exposition, using a stepwise regression analysis. As the outcome of regression analysis, 5 % significance level was used to decide of the inclusion or the exclusion of a predictor.

The results reveal that latitude (cosine, square and logarithm of it), longitude, altitude, rain exposition and relief exposition are found to be "good" variables to describe the dry-wet and wet-wet transitions. The outcome of the simulation reveals that our simulation models present many deficiency in representing the observed and simulated probability, in spite of the fact that in both series the seasonal variation are considering to be preserved. The deficiency can be explained by the quality of our data; in fact, it has been well documented that the time of observation is a significant factor influencing the occurrence process as defined by first-order Markov and in the distribution of rainfall depth; this is attributed to the diurnal variation of rainfall and to the evaporation of small rainfall amounts for gauges read in the afternoon. In Cameroon gauges are read twice a day (morning and evening); during the gap between the first and second read the latter problem could happen.

This study also reveals that square of latitude, rain exposition, longitude and the mean of yearly rainfall are found to be "good" variables to describe the parameters of Weibull distribution function fitting to daily rainfall amounts. New methods to improve the derivation of spatial predictors are suggested; these include the use of digital topographic analysis techniques with which it is possible to derive all necessary physiographic outputs (surface gradients, mean elevation, land cover, distances to ocean or streams) rapidly and accurately.

The generator is considering to simulate the wet days series for inland stations with statistics resembling the observations. This is not the case in the coastal areas represented by Douala and Ngambe stations. In the coastal area south-westerly monsoon winds predominate, modified by land sea breezes causing humidity values to almost the saturation point, bringing thus the region more rain than the inland stations. There is sometimes continuous rain over several days especially during wet months; by splitting over few days, all these days have high rainfall amounts of nearly the same order of magnitude. Thus, when fitting a distribution to the daily rainfall amounts, the chosen distribution could predict much fewer big rain amounts, the higher the rain is. The effect of these methodological factors and errors is primarily reflected in the occurrence process and in the smaller amounts of rainfall.

The study shows that the frequency distribution of the generated daily rainfall amounts is not resembling the ones of the observed series in Douala, Koundja and Ngambe whereas simulated and observed daily rainfall series of Kaele, Ngaoundere, and Sangmelima indicate the resemblance between the two frequency distributions. The generator produces good estimate of daily rainfall amounts in the northern area where the climate is generally dry, and in the southern plateau which is characterized by four seasons. The underestimation of rainfall amount in coastal and highland region can be explained by the fact that these regions are under the influence of monsoon winds, see breeze which have to be taken into account as predictors when estimating the rainfall in coastal region. The result also suggests that the far distance from the ocean, we have good estimation; we have then to include the distance from the sea as another predictor.

The result of the present study exhibits no difference between the monthly mean of rainfall, excepted Douala which presented a difference between the mean of the generated and observed rainfall series in August, which corresponds to its wettest period of the year. The situation is the same when testing the monthly variance; the values under .05 indicate that there is a between the variance of observed and simulated series in August, October, and November in Douala.

The improving of our simulation models using time invariants predictors means the search of predictors using new techniques for deriving physiographic parameters of the area of study such as the digital topographic analysis techniques, and/or remote sensing. These techniques are tools providing relevant landscape data. As Cameroon is situated in the Gulf of Guinea, it will be interesting to investigate the influence of the West Africa monsoon and of the squall lines on the rainfall series in the country, and then use both also as predictors in the possible future projects. Improving our simulation models also suggests the use of daily rainfall series measured automatically; Automated stations will measure exactly the quantity of rain which has really fallen. This is costly but would provide good quality data, and accordingly better results .

8 ACKNOWLEDGMENTS

Working in the pleasant atmosphere of the Department of Physical Geography at the Intitute of Geography, University of Hamburg has been an excellent opportunity for me.

I would especially like to thank Prof. Dr. R. Schwarz for making this thesis possible and for his expert advises. His profound knowledge, enthousiastic support and continuous guidance were very helpful and highly motivating.

I' am grateful to Prof. Dr. Hans von Storch for the examination of this work.

I'm also greateful to Pr. Maurice Tsalefac Department of Geography, University of Yaounde I, Cameroon who I learned much of the research diverse ways.

Many thanks go to Marion Dohr, Anne Gnauck for their valuable helps in all domains especially during my integration within the staff of the department of Physical Geography, University of Hamburg.

I greatly appreciate the assistance I received from the staff of the Institute of Geography of the University of Hamburg especially from Prof. Jurgen Lafrenz.

Most of all, I want to thank M. Ajoumissi Jean for his encouragement; thanks to your effort, my dream to be Doctor in Physical Geography has been realized.

Most of all, I want to thank my lovely wife Nguendia Liliane Hortence for her love and encouragement.

I will like to dedicate this thesis to all the members of my family especially in the memories of my **late grandfather Djoumessi Tatsangue and late mother Ngoufack Madeleine** who passed away in July 2002 and 11^{th} Febraury 2004 respectively. Wherever you are, keep eyes on your children.

Last but not the least, to my friends especially Edouard Penlap for their continuous support.

References

Abramowitz, M., and I.A. Stegun, 1984: *Pocketbook of mathematical functions.* Frankfurt: Verlag Harri Deutsch.

Agnew, J.L., and R.C. Knapp, 1978: *Linear algebra with applications* (Third ed.). Pacific Grove: Brooks/Cole Publishing Compagny.

Akaike, K., 1974: A new look at the statistical model identification. *IEEE Transactions on Automatic control*, **19**, 716–723.

Anderson, T.W., and L.A. Goodman, 1957: Statistical inference about Markov chains. *Annals, Mathematical Statistics*, **28**, 89–109.

Anyadike, R.N.C., 1993: Patterns and variations of rainfall over banjul, gambia. *Singapore J. Trop. Geogr.*, **14**, 1–14.

Bardossy, A., 1998: Generating precipitation time series using simulated annealing. *Water Resources Research*, **34**, 1247–1259.

Berensten, P.B.M., A.A. Dijkhuizen, and A.J. Oskam, 1992: A dynamic model for cost benefit analyses of food-and-mouth disease control strategies. *Preventive Veterinary Medecine*, **12**, 229–243.

Besson, L., 1924: On the probability of rain. *Monthly Weather Review*, **52**, 308.

Bhat, U.N., 1984: *Elements of Applied stochastic Processes.* John Wiley Sons, Second ed.New York, Chichester, Brisbane, Toronto, Singapore.

Bratley, P., B.L. Fox, and L.E. Schrage, 1987: *A guide to simulation.* Springer-Verlag.

Bruhn, J.A., W.E. Fry, and G.W. Fick, 1980: Simulation of daily weather using theoritical probability distributions. *Journal of Applied Meteorology*, **19**, 1029–1036.

Buckle, C., 1996: *Weather and Climate in Africa.* Addison Wesley Longman. Malaysia.

Buishand, T.A., 1977: Stochastic modeling of daily rainfall sequences. *Communications, Agricultural University, Wageningen*, **77 (3)**, 211.

Buishand, T.A., 1978: Some remarks on the use of daily rainfall models. *Journal of Hydrology*, **36**, 295–308.

Burlando, P., R. Rosso, L. Cadavid, and J. Salas, 1993: Forecasting of short term rainfall using ARMA models. *Journal of Hydrology*, **144**, 193–211.

Chatfield, C., 1996: *The analysis of Time Series: An Introduction* (5th ed.). Chapman and Hall, London.

Chin, E.H., and J.F. Miller, 1980: On the conditional distribution of daily precipitation amounts. *Mon. Wea. Rev.*, **108**, 1462–1464.

Cochran, W.G., 1938: An extension of gold's method of examining the apparent persistence of one type of weather. *Quart. J. Roy. Meteor.Soc.*, **64**, 631–634.

Cox, D.R., and V. Isham, 1994: Stochastic models of precipitation. In: *Statistic for environment*, V.Barnet and K. Turkman (Eds.), 2, pp. 3–18. Wiley, New York: Water issues.

Daly, C., D.L. Phillips, and R.P. Nielson, 1994: A statistical topographic model for mapping climatological precipitation over mountainous terrain. *Journal of Applied Meteorology*, **33**, 140–158.

Dubrovsky, M., 1997: Creating daily weather series with use of the weather generator. *Environmetric*, **8**, 409–424.

Feyerherm, A.M., and L.D. Bark, 1967: Goodness of fit of a markov chain model for sequences of wet and dry days. *Journal of Applied Meteorology*, **6**, 770–773.

Foufoula-Georgiou, E., and D.P. Lettenmaier, 1987: A markov renewal model for rainfall occurrences. *Water Resources Research*, **23**, 875–884.

Gabriel, K.R., and J. Neumann, 1962: A markov chain model for daily rainfall occurrence at tel aviv. *Quart. J. Roy. Meteor.Soc.*, **88**, 90–95.

Gates, P., and H. Tong, 1976: On markov chain modeling to some weather data. *Journal of Applied Meteorology*, **15**, 1145–1151.

Geng, S., F.T.W. Penning de Vries, and I. Supit, 1986: A simple method for generating daily rainfall data. *Agricultural and forest Meteorology*, **36**, 363–76.

Gilman, D.L., M.F. J, and J.M. Jr, 1963: On the power spectrum of "red noise". *Journal of Atmospheric Sciences*, **20**, 182–184.

Gold, E., 1929: Note on the frequency of occurrence of sequences in a series of events of two types. *Quart. J. Roy. Meteor.Soc.*, **36**, 363–376.

Gregory, J.M., T.M.L. Wigley, and P.D. Jones, 1993: Application of markov models to area average daily rainfall series and interannual variability in seasonal totals.

Climate dynamics, **8**, 299–310.

Guttorp, P., 1995: *Stochastic modelling of scientific data*. Chapman and Hall. London.

Haan, C.T., M.D. Allen, and J.O. Street, 1976: A markov chain model for daily precipitation. *Water Resources Research*, **12**, 443–449.

Hanson, C.L., K.A. Cumming, D.A. Woolhiser, and C.W. Richardson, 1994: Microcomputer program for daily weather simulation in contiguous United States. US Department of Agriculture.Agricultural Research Service ARS-114.

Harrison, M., and P. Waylen, 2000: A note concerning the proper choice for markov model order for daily precipitation in the humid tropics: A case study in costa rica. *jclimatology*, **20**, 1861–1872.

Harrison, P.A., R.E. Butterfield, and T.E. Downing, 1995: Climate change and agriculture in Europe. Assessment of impacts and adaptations. University Oxford, Environmental change unit, United Kingdom.

Hirsch, R.M., J.R. Slack, and R.A. Smith, 1982: Techniques of trends analysis for monthly weather quality data. *Water Resources Research*, **18**, 107–121.

Hutchinson, 1995: Stochastic space-time model from ground based data. *Agriculture and Forest Meteorology*, **73**, 237–264.

Jackson, I.J., 1981: Dependence of wet and dry days in the tropics. *Archives for Meteorology, Geophysics, and Bioclimatology*, **B29**, 167–179.

Jones, P.G., and P.K. Thornton, 1997: Spatial and temporal variability of rainfall related to a third order markov model. *Agric. and For. Meteorol.*, **86**, 127–138.

Katz, R.W., 1977: Precipitation as time dependent process. *Journal of Applied Meteorology*, **16**, 671–676.

Katz, R.W., 1981: On some criteria for estimating the order of markov chain. *Technometrics*, **23**, 243–249.

Katz, R.W., 1985: Probabilistic models. In: *Probability, Statistics and decision making in the atmospheric sciences*, A. H. Murphy and K. R. W. (Eds.), pp. 261–288. Boulder, CO: Westview press.

Katz, R.W., 1995: Generalization of chain-dependent processes: Application to hourly precipitation. *Water Resources Research*, **31**, 1331–1341.

Katz, R.W., and P.M. B, 1998: Overdispersion phenomenon in stochastic modeling of precipitation. *Journal of Climate*, **11**, 591–601.

Kite, W.G., 1977: Frequency and risk analyses in hydrology. In: *Water Resource Publication* (2nd ed.). Fort Collins, Colorado: Water Resource Publication, US.

Longley, R.W., 1953: The length of dry and wet periods. *Quart. J. Roy. Meteor.Soc.*, **79**, 520–527.

Lund, I.A., and D.D. Grantham, 1977: Persistence, runs and recurrence of precipitation. *Journal of Applied Meteorology*, **16**, 346–358.

Mirza, M.Q., R.A. Warrick, N.J. Ericksen, and G.J. Kenny, 1998: Trends and persistence in precipitation in the Ganges, Brahmaputra and Meghna River basins. *Hydrological Science Journal*, **43**, 845–858.

Montgomery, B.A., J.W. Lloyd, and A. Jensen, 1994: Use of markov chain model to evaluate employment policies for a vetenary school faculty. *Journal of Medical Education*, **21**, 28–32.

Newnham, E.V., 1916: The persistence of wet and dry weather. *Quarternary Journal of Royal Meteorology Society*, **42**, 153–162.

ORSTOM, 1973: *Précipitations journalières de l'origine des stations à 1965. République du Cameroun.* Paris: ORSTOM.

Parthasarathy, B., and O.N. Dhar, 1975: Trends analysis of annual indian rainfall. *Hydrological Science Bulletin*, **20**, 257–260.

Pedgley, D.E., and T.N. Krishnamurti, 1976: Structure and behaviour of a monsoon cyclone over West Africa. *Monthly Weather Review*, **104**, 149–167.

Phillips, D.L., J. Dolp, and D. Marks, 1992: A comparison of geostatistical procedures for spatial analysis of precipitation in mountainous terrain. *Agriculture and Forest Meteorology*, **58**, 119–144.

Press, W.H., B.P. Flannery, S.A. Teukolsky, and W.T. Vetterling, 1986: *Numerical recipes, the art of scientific computing.* Cambridge University Press, London.

Racsko, P., L. Sleidl, and M.A. Semenov, 1991: A serial approach to local stochastic weather models. *Ecological modelling*, **57**, 27–41.

Richardson, C.W., 1981: Stochastic simulation of daily precipitation, temperature, and solar radiation. *Water Resources Research*, **17**, 182–190.

Richardson, C.W., and D.A. Wright, 1984: "WGEN: A model for generating daily weather variables". US Department of Agriculture, Agricultural Research Service ARS-8.

Roldan, J., and D.A. Woolhiser, 1982: Stochastic daily precipitation models. A comparison of occurrence processes. *Water Resource Research*, **18**, 1451–1459.

Running, S.W., R.R. Newmani, and R.D. Hungerford, 1987: Extrapolation of synoptic meteorological data in mountainous terrain and its use for simulating forest evapotranspiration and photosynthesis. *Canadian Journal Forest Resource*, **17**, 472–483.

Schwarz, B., 1978: Estimating the dimension of a model. *Annals of Statistics*, **6**, 461–464.

Schwarz, R., 1980: Das Weidepotentiel der trockenen Tropen, ermittelt mit Hilfe dynamischer Systemsimulation. In: *BARTH, H.-K./WILHELMY, H. (Hrsg.): Trockengebiete. Natur und Mensch im Ariden Lebensraum*, 80, pp. 303–320. Tübinger Geographische Studien.

Semenov, M.A., and E.M. Barrow, 1997: Use of stochastic weather generator in the development of climate change scenarios. *Climate Change*, **35**, 397–414.

Semenov, M.A., and R.J. Brook, 1999: Spatial interpolation of LARS-WG stochastic weather generator in Great Britain. *Climat Resource*, **10**, 95–107.

Sevruk, B., and H. Geiger, 1981: Selection of distribution types for extrems of precipitation. Operational Hydrolical Report 15, WMO.

Singh, S.V., and R.H. Kripalini, 1986: Analysis of persistence in daily monsoon rainfall over India. *Journal of Climatology*, **6**, 625–39.

Stander, J., 1989: Markov chain analysis and specialization in criminal careers. *The British Journal of Criminology*, **29**, 317–335.

Stern, R.D., 1980a: Analysis of daily rainfall at Samuru, Nigeria, using a two-part model. *Archive of Meteorology, Geophysics, and Bioclimatology*, **B28**, 123–135.

Stern, R.D., 1980b: The calculation of probability distributions for models of daily precipitation. *Archive of Meteorology, Geophysics, and Bioclimatology*, **B28**, 137–147.

Stern, R.D., and R. Coe, 1984: A model fitting analysis of daily rainfall data. *Journal of Royal Statistics*, **Serie A147**, 1–34.

Stern, R.D., M.D. Dennett, and I.C. Dale, 1982: Analysing daily rainfall measurements to give agronomically useful results. II. A modelling approach. *Experimental Agriculture*, **18**, 237–253.

Stewart, W.J., 1994: *Introduction to the numerical solution of Markov Chains*. Princeton University Press, Princeton.

Suchel, J.B., 1988: *Les climats du Cameroun*. Ph. D. thesis, Univ.de Saint Etienne.

Thom, H.C.S., 1958: A note on the gamma distribution. *Monthly Weather Review*, **86**, 117–122.

Todorovic, P., and D.A. Woolhiser, 1975: A stochastic model of n-day precipitation. *Journal of Applied Meteorology*, **14**, 17–24.

Tsuji, G.Y., G. Hoogenboom, and P.K. Thornton, 1998: *Understanding options for agricultural production*. Kluwer Academic: Dordrecht.

UNEP, 2002: Africa environment outlook. United Nation Environmental Program.

Wilby, R.L., 1994: Stochastic weather type simulation for regional climate change impact assessment. *Water Resources Research*, **30**, 3395–403.

Wilks, D.S., 1989: Conditioning stochastic daily precipitation models on total monthly precipitation. *Water Resources Research*, **25**, 1429–39.

Wilks, D.S., 1992: Adapting stochastic weather generation algorithms for climate change studies. *Climate Change*, **22**, 67–84.

Wilks, D.S., 1995: *Statistical methods in the atmospheric sciences*. San Diego, CA: Academic Press.

Wilks, D.S., 1998: Multisite generalizations of a daily stochastic precipitation generation model. *Journal of Hydrology*, **210**, 178–191.

Wilks, D.S., 1999: Interannual variability and extreme value characteristics of several stochastic daily precipitation models. *Agriculture and Forest Meteorology*, **93**, 153–169.

Wilks, D.S., and R.L. Wilby, 1999: The weather generation game: a review of stochastic weather models. *Progress in Physical Geography*, **23**, 329–357.

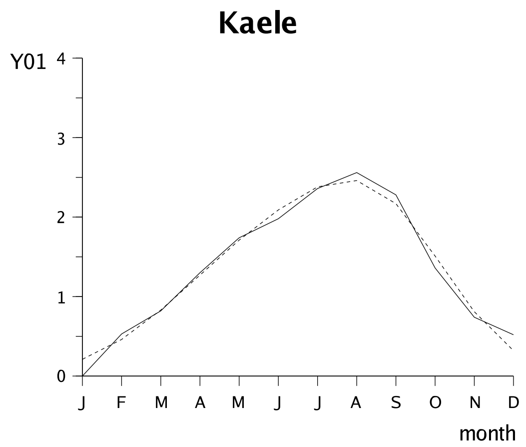
Williams, C.B., 1952: Sequences of wet and dry days considered in relation to the logarithmic series. *Quarternary Journal of Royal Meteorology Society*, **78**, 91–96.

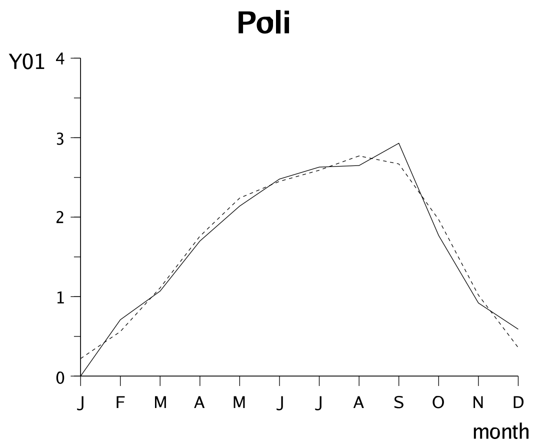
Woolhiser, D.A., and G.G.S. Pegram, 1979: Maximum likelihood estimation of Fourier coefficients to describe seasonal variation of parameters in stochastic daily precipitation models. *Journal of Applied Meteorology*, **18**, 34–42.

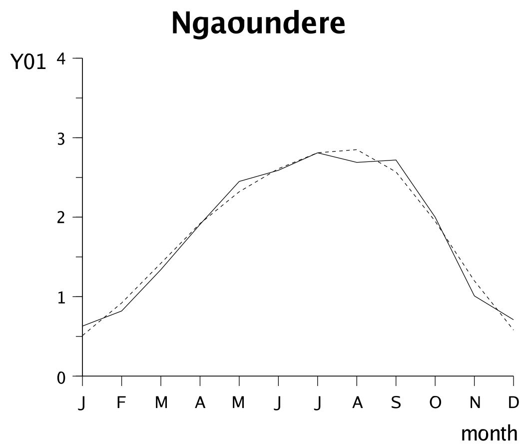
Woolhiser, D.A., and J. Roldan, 1982: Stochastic daily precipitation models. a comparison of amounts. *Water Resources Research*, **18**, 1461–1468.

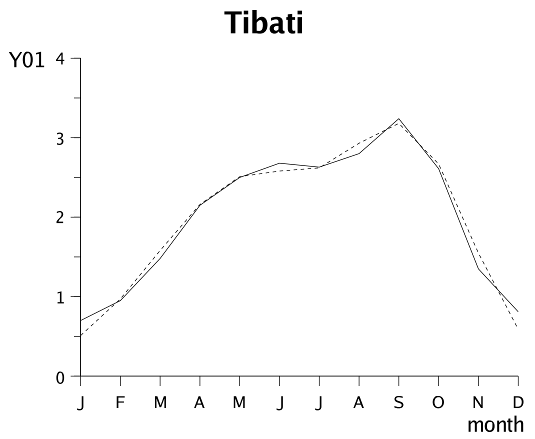
9 ANNEXE -A-

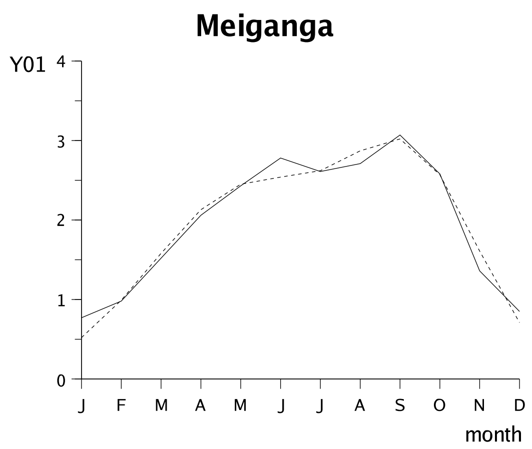
This annexe shows the absolute wet-dry probability using the double normal function for the rest of stations.

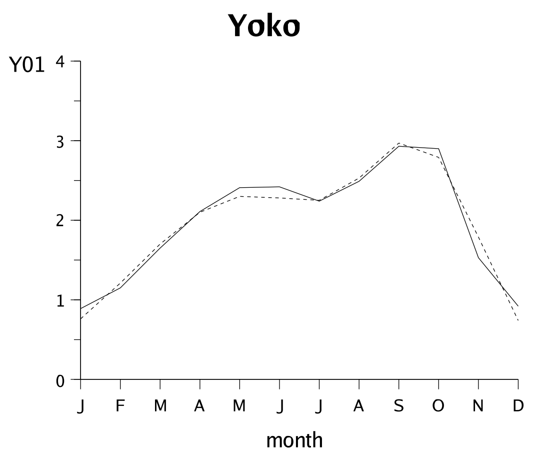




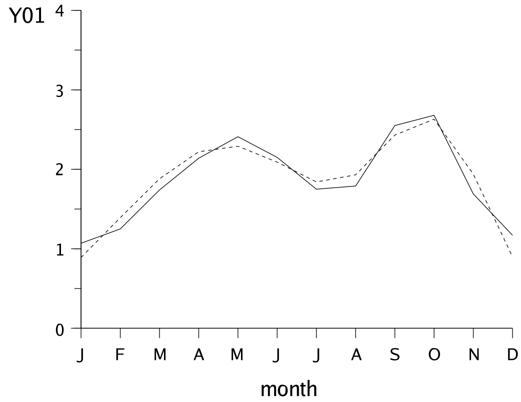




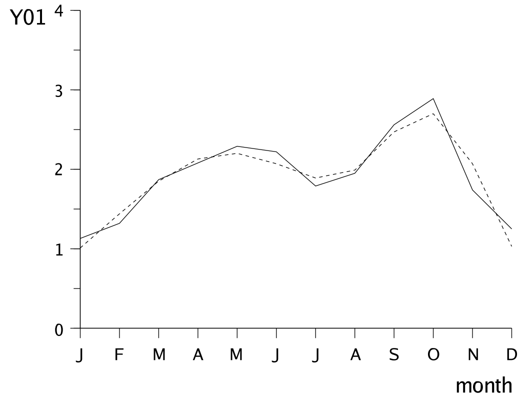




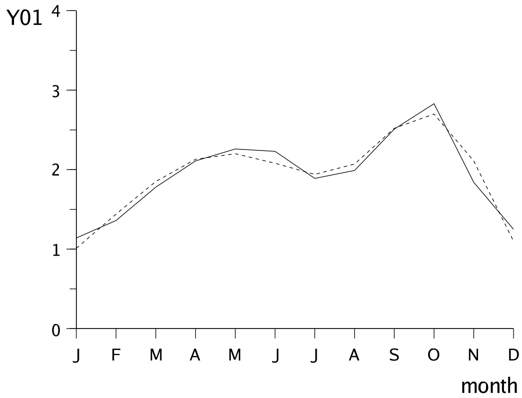
Nanga-Eboko



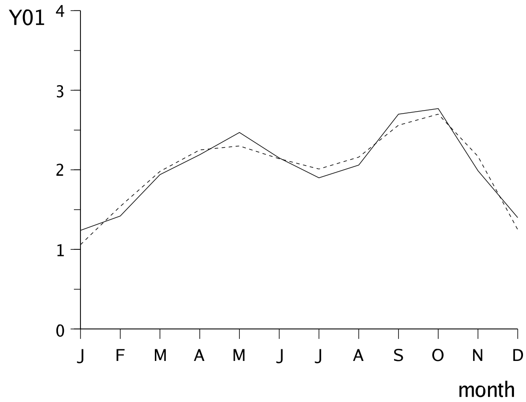
Bertoua



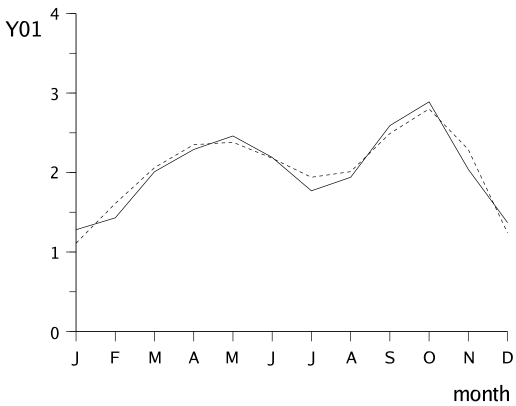
Batouri



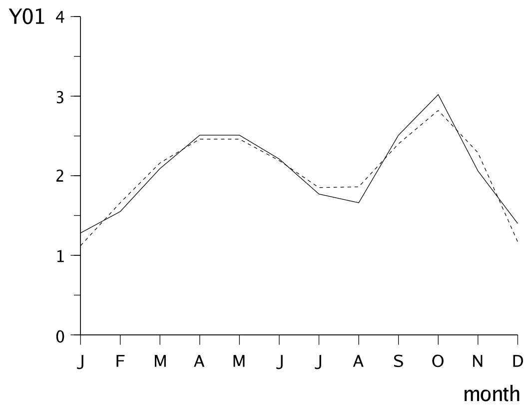
Yokadouma



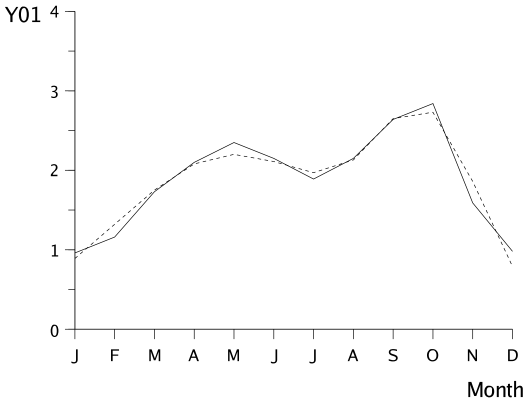
Lomie



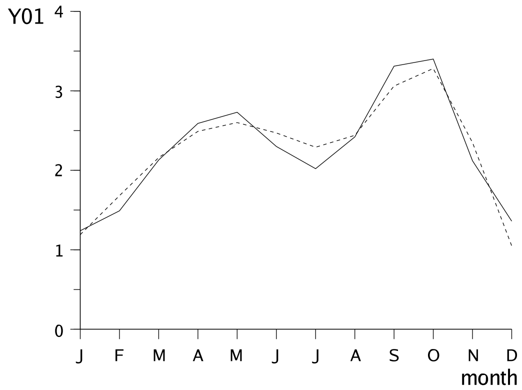
Sangmelima



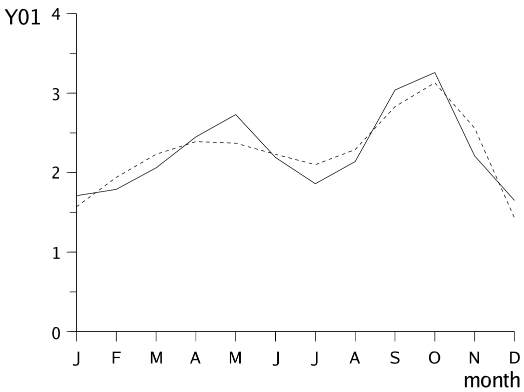
Akonolinga



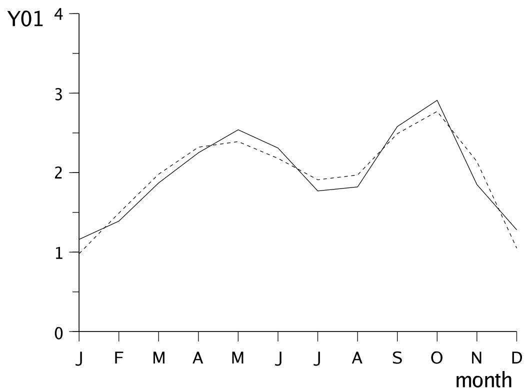
Eseka



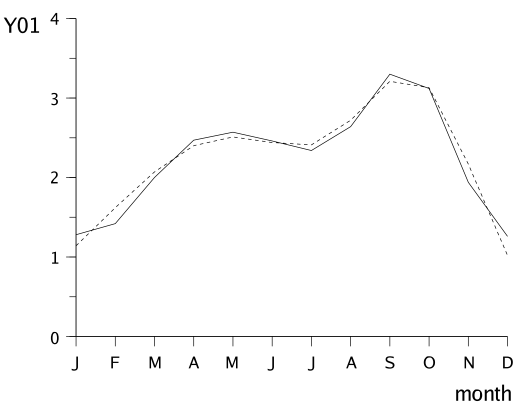
Kribi



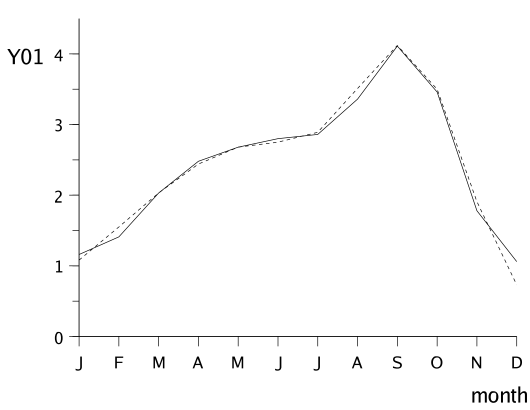
Abong-Mbang



Edea



Ngambe



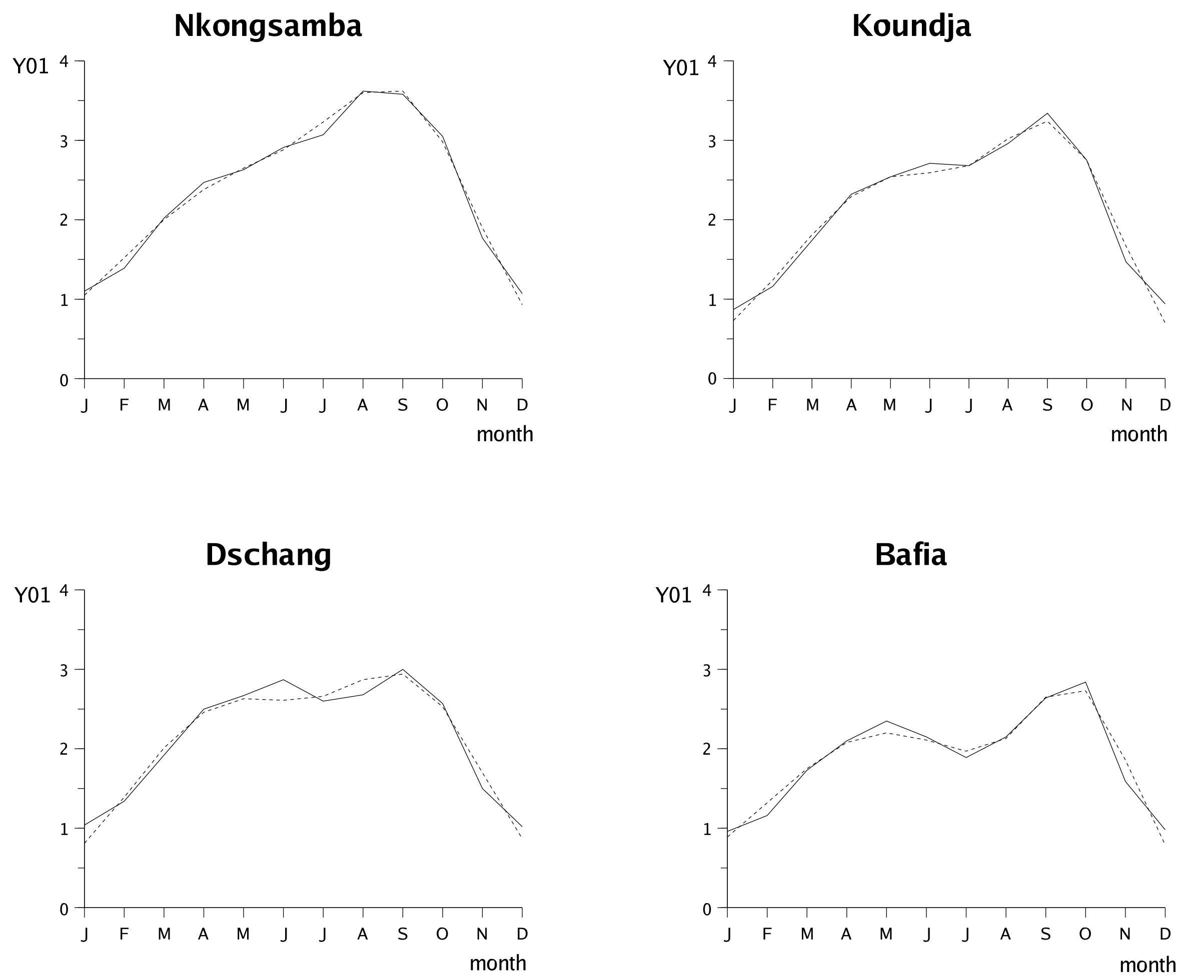
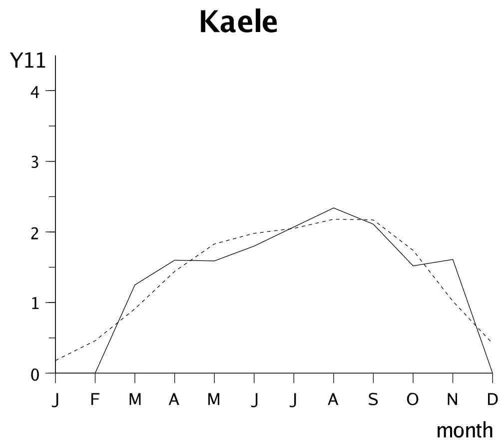
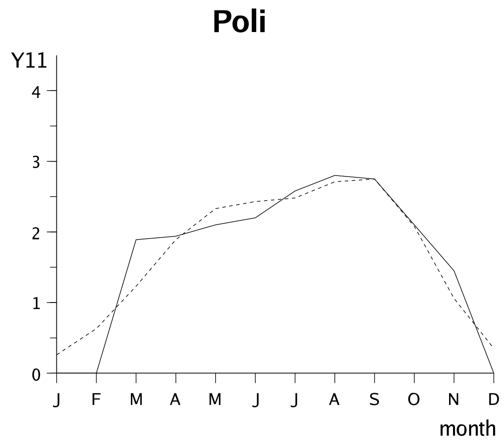


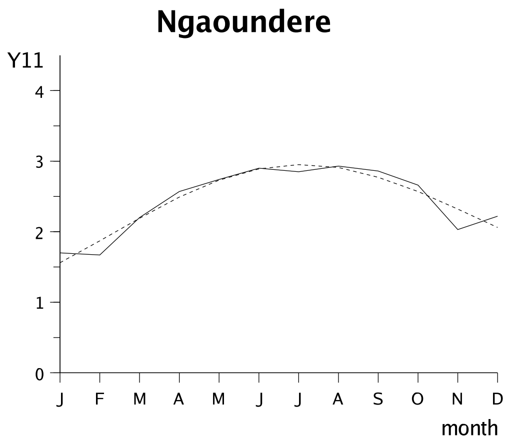
Figure 24: Transformed wet-dry probability using from the double normal function

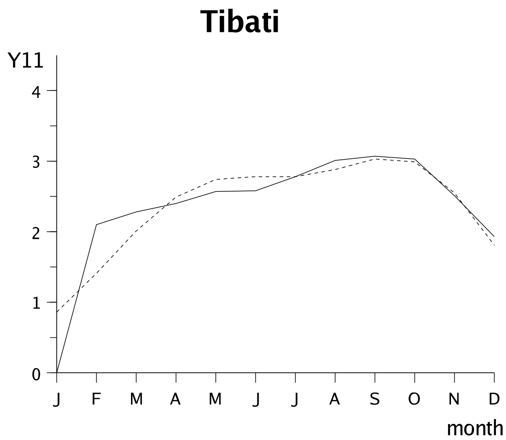
10 ANNEXE -B-

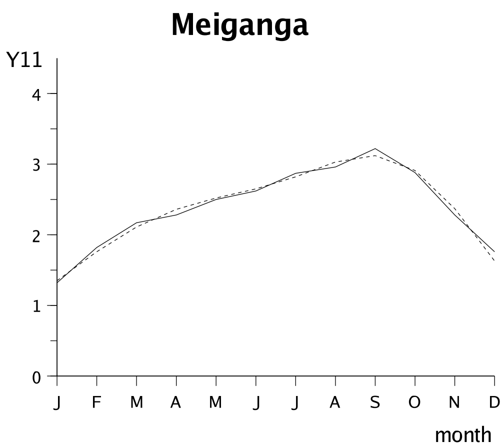
This annexe shows the absolute wet-wet probability using the double normal function for the rest of stations.

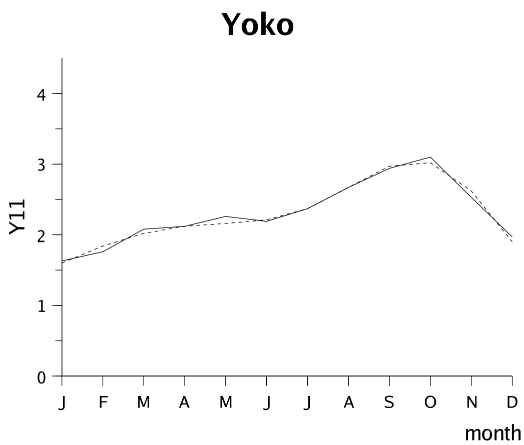




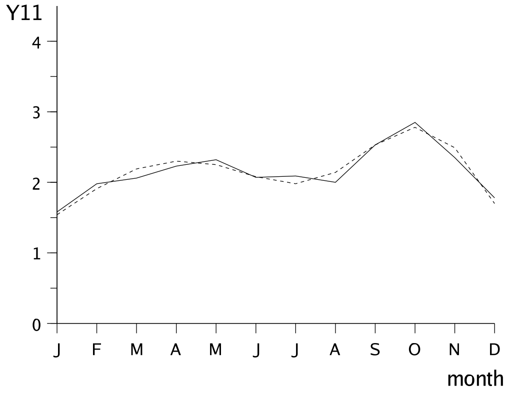




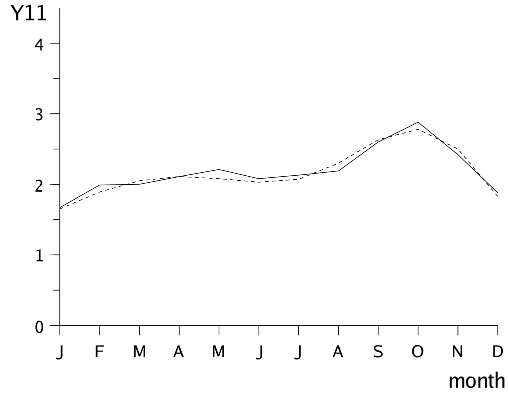




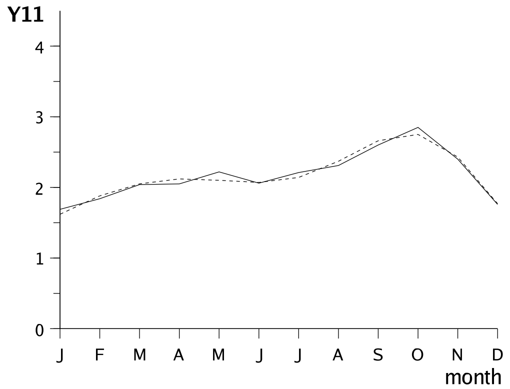
Nanga-Eboko



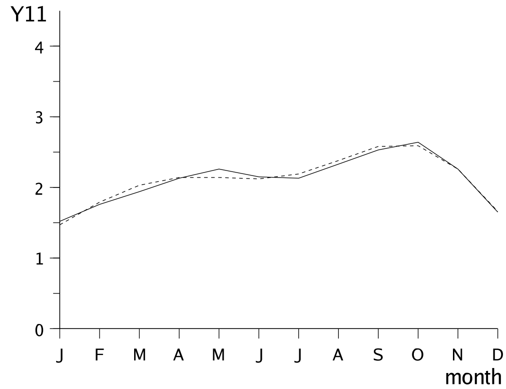
Bertoua



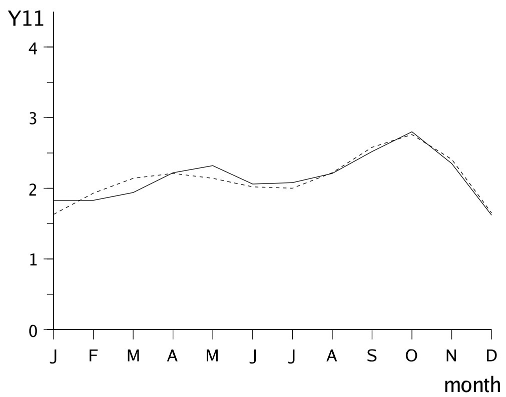
Batouri



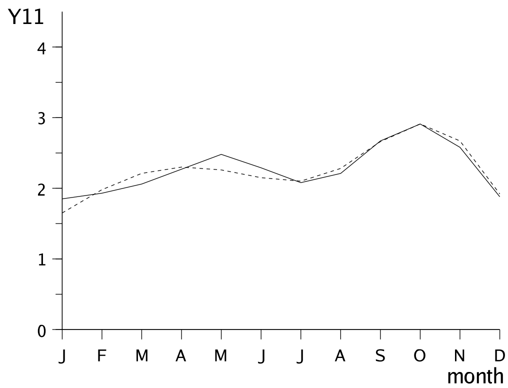
Yokadouma



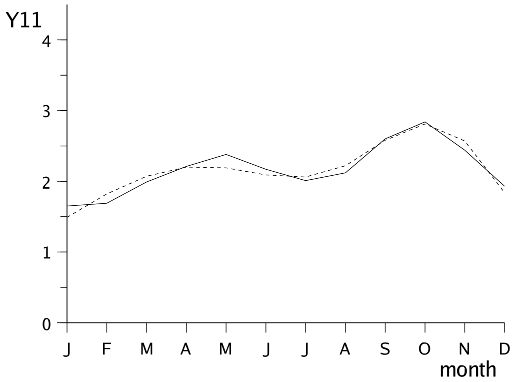
Lomie



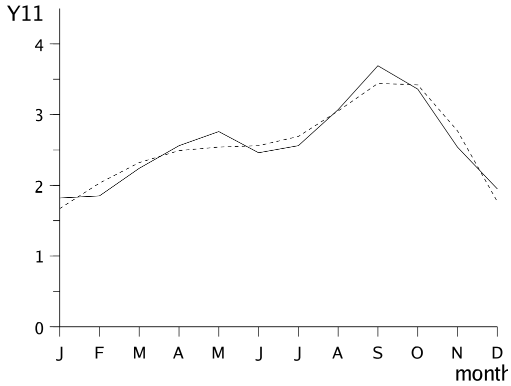
Sangmelima



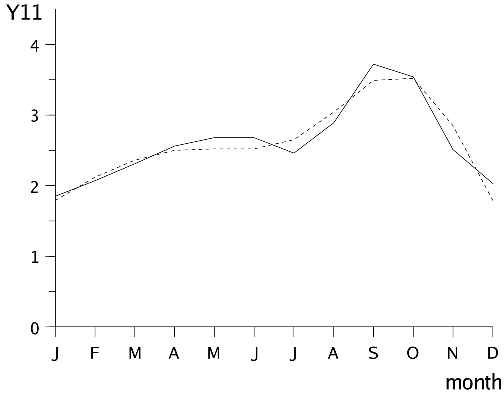
Akonolinga



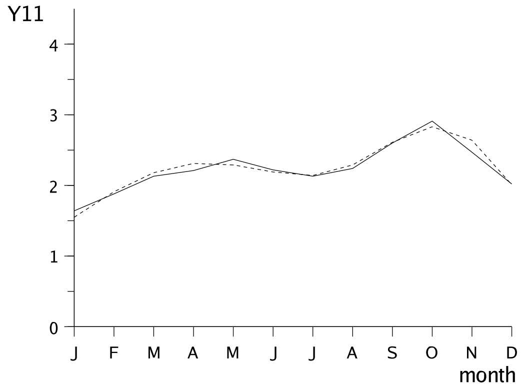
Eseka



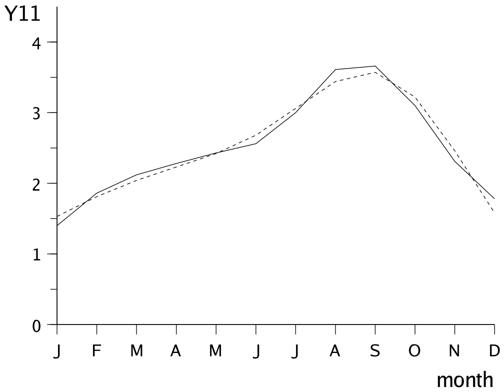
Kribi



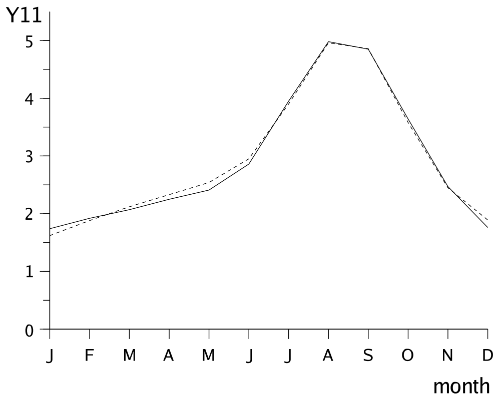
Abong-Mbang



Edea



Ngambe



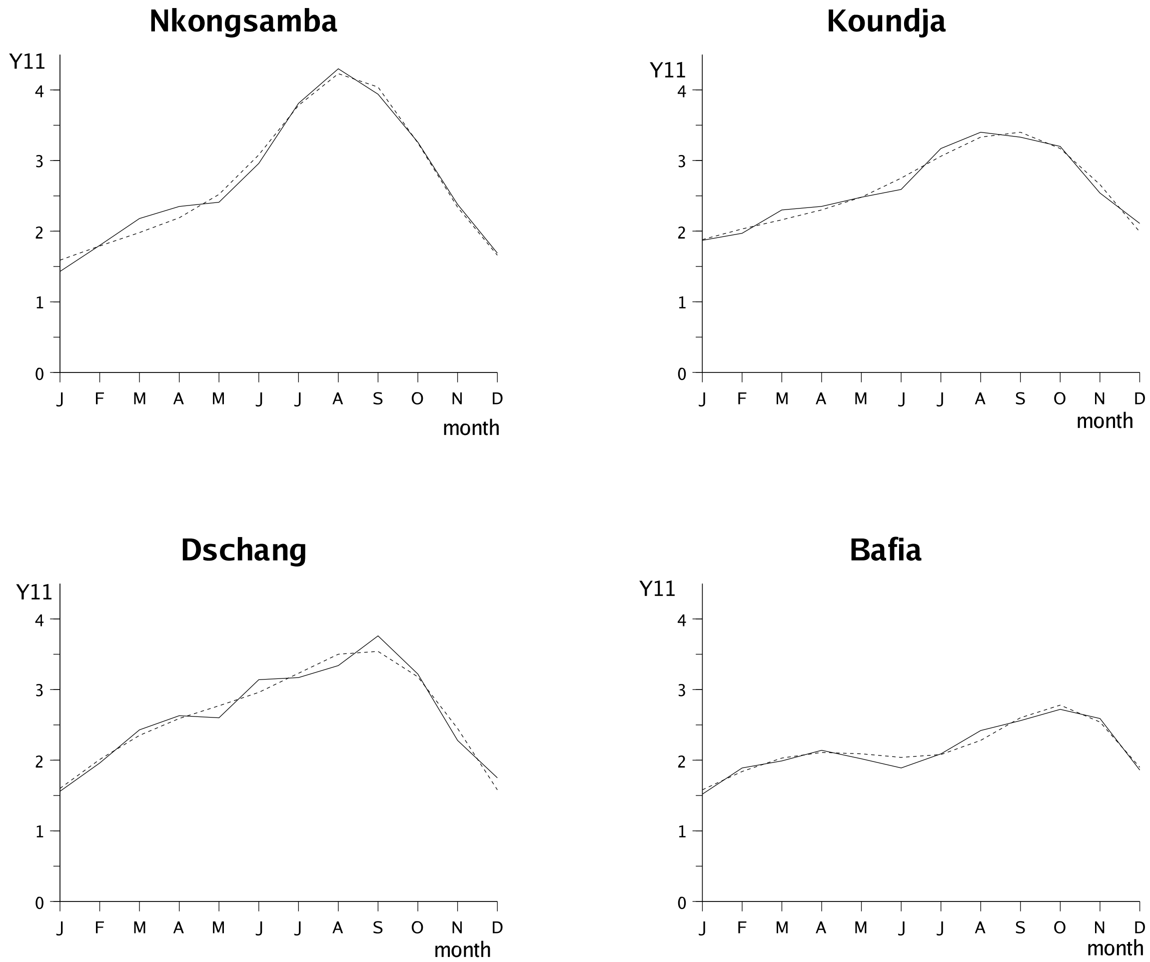


Figure 25: Transformed wet-wet probability using from the double normal function