Simulation of maize production using a pixel based weather generator: Case of Lake Naivasha catchment.

Musana Segatagara Bernard

March, 2008

Simulation of maize production using a pixel based weather generator: Case of Lake Naivasha catchment

by

Musana Segatagara Bernard

Thesis submitted to the International Institute for Geo-information Science and Earth Observation in partial fulfilment of the requirements for the degree of Master of Science in Geo-information Science and Earth Observation, Specialisation: Water Resources and Environmental Management (Integrated Watershed Modeling and Management)

Thesis Assessment Board

Chairman:	Prof. Dr. Z. Su (WRS/ITC)
External Examinar:	Dr. A. de Wit (Wageningen University)
Primary Supervisor:	Drs. R. Becht (WRS/ITC)
Second Supervisor:	M.Sc. V. Venus (NRS/ITC)

ITC

INTERNATIONAL INSTITUTE FOR GEO-INFORMATION SCIENCE AND EARTH OBSERVATION
ENSCHEDE, THE NETHERLANDS

Abstract

Crop modeling uses different tools to cope with spatial and temporal gaps of weather data. Weather generator are often use. Weather generators were conceived for Meteorological stations and effort for interpolation are not yet popular. Bandari adapted WXGEN using a combination of time-series analysis and spatial analysis with the aim of producing WXGEN input data for every day in the year and for every location inside the Naivasha basin.

The concepts used by Bandhari (2005) to generate daily rain for every day and every pixel were extended. Maximum Temperature, Minimum temperature, and Maximum Solar radiation. The extended rainfall generator for Temperature and Solar Radiation uses the trivariate autoregressive function developed by Richardson (1984). The method generated synthetic data for four weather variables in from aerial input map and pixel by pixel. A simplified WOFOST, a crop growth model developed by the Wageningen University uses the generated data for Simulating Maize production. A simplified version of WOFOST implemented in PCRaster (Spatial modeling language). The crop growth model was spatially distributed. It simulated for each pixel of Weather data.

Results of the simulations show that the stochastic component of the rainfall generator was satisfying. The skewed distribution function has shown some limitation in a non-iterative process. It produces negative and zero values for certain range of data. The weather of average altitudes of the Naivasha catchment was better reproduced. An advantage of using a spatial (raster) approach is that a window of pixels shows variability of climate. Applied to agriculture it can generate information of several years.

After adjustment of the crop growth model using recently observed data, the simulation shows an important range of values between 10 Tonnes per ha and less than 1Tonne per ha. The Yield Simulation of maize was sensitive to Temperature values. Crop failure was simulated for area with average Temperature is below 5 degrees Celsius.

Based on data collected in the Area and methods to extract information from pixel based Yield map, improvement of regression model are suggested. For district level or below, the approach can be used to improved regression models used to predict the water limited yield in the Area.

Acknowledgements

I would like to thank all the organisations and persons who in one way or another contributed to the successful writing of this thesis. I thank my supervisors Drs R Becht and Valentijn Venus, for their arrangements and guidance in Kenya. They inspire the subject and help to understand the contour of the study. The monitoring and the advice of our program director, Ir. Arno van Lieshout, play an important role in the all process of my course in ITC. Many thanks to Partners of ITC in Naivasha, among others I would like to name Dominic Wambuya and his colleagues of Water Management Authority.

I would like to thank also all my classmates for the informal but critical assistance during the thesis writing. My thanks to all Rwandan students of ITC. Thanks to my Brethren of ITC-CF (ITC Christian fellowship) and volunteers who work for patient of the hospital of Enschede (I use to meet on Sunday).

I take the opportunity, to thank my sponsor NUFFIC and ITC's staff for the chance they have given to my candidature for an MSc. I have really appreciated the knowledge gained in the field of Water Management after the last 18 months. I would like to acknowledge the role of my organization which elects me for the multi-year Agreement with NUFFIC. I am the third candidates of ISAR, to join ITC. Financial and moral supports came also from my Parents, brothers and sisters. The assistance needed was diverse but crucial at critical time. I thank you all who contribute to a successful study and stay in Netherlands.

Table of contents

1. Introduction 7
  1.1. Justification....8
  1.2. Research Problem 8
  1.3. Objectives 8
  1.4. Research question 8
  1.5. Hypothesis 8
2. Literature Review....9
  2.1. Rainfed Agriculture and Variability of climate....9
    2.1.1. Importance of Rainfed Agriculture 9
    2.1.2. Rainfed production and Water productivity 10
    2.1.3. Inter-annual and Seasonal variability and Yield prediction....10
    2.1.4. Growing Pattern and Agro-ecological zoning in Kenya 11
  2.2. Spatial Interpolation of Climatic variables....11
  2.3. Stochastic model for weather....12
    2.3.1. General presentation 12
    2.3.2. Inter-annual variability of weather generator....14
    2.3.3. Case study of Interpolated Weather Generator 14
  2.4. Crop modeling 15
    2.4.1. Historical view 15
    2.4.2. Classification of Crop growth models 15
    2.4.3. Characterisics of mechanistics crop growth model....16
    2.4.4. Accounting for the Water Stress 16
3. Methodology 18
  3.1. Study Area 18
  3.2. Plant test....19
    3.2.1. General consideration 19
    3.2.2. Maize in kenya 20
  3.3. Weather data 20
    3.3.1. Rainfall data 20
    3.3.2. Extension to Solar Radiation and Minimum and Maximum Temperature 22
    3.3.3. Other weather data for crop simulation 27
  3.4. The Crop Growth model 27
    3.4.1. Environmental Modeling using PCRaster....28
    3.4.2. Development Stage simulation 28
    3.4.3. Soil Water Balance....29
  3.5. Soil data 30
  3.6. Analysis 30
    Similarity and variability of Weather input data....30
  3.7. Calibration and sensibility analysis 30
    3.7.1. Calibration and Validation of the Crop model....30
  3.8. Assumption 31
4. Weather Generation process: Results and Discussion 32

4.1. Interpolation and Input for Weather Generation ... 32
4.2. Interpolation of Average for Tmin, Tmax and Solar Radiation ... 35
4.3. Characteristics of Synthetic Series ... 36
4.3.1. Rainfall ... 36
4.3.2. Maximum Temperature ... 42
4.3.3. Minimum Temperature ... 45
4.4. Findings ... 47
5. Distributed Crop Growth modeling ... 48
5.1. Adjustment of the Crop model ... 48
5.1.1. Determination of starting date and length of the simulation ... 48
5.1.2. Comparison with the Variation of rainfall in a year ... 49
5.2. Tuning crop parameters ... 50
5.2.1. Optimization curves ... 52
5.3. Simulated Yield ... 52
5.4. Comparison between observed and simulated yield ... 53
5.4.1. Information Retrieval from interviews of farmers ... 53
5.4.2. Information Retrieval from simulated Maize Yield map ... 53
5.4.3. Regression Analysis between Yield, Reliable Rainfall and Topography ... 53
5.4.4. Comparison of the regression models ... 57
6. Conclusion ... 59
7. Bibliography ... 62
8. ANNEXES ... 65

List of figures

Figure 3.2. Rainfall spatial distribution 19
Figure 3.3. Average of mean for each Julian day and Fourier series fitted (Bhandari,2005)................21
Figure 4.1. Variation of Monthly average of Tmax for 12 grids covering the Area..............................34
Figure4.2. Variation of Monthly average of Tmin for 12 grids covering the Area 34
Figure 4.3. Box plot of Variation of simulated annual rainfall For raingauges' locations....................37
Figure 4.4. Map of Average of daily Rainfall for two years (year7 and year8)....................................38
Figure 4.5. Difference Map: Average of daily Rainfall for two years minus the Interpolation of Observed Average Daily rainfall (year7 and year8). 39
Figure 4.6. Map of Standard deviation of daily Rainfall for two years (year7 and year8)....................39
Figure 4.7. Difference Map: Standard deviation of daily Rainfall for two years minus Average interpolated Average Standard deviation of daily Rainfall...39
Figure 4.8. Map of rainy day's probability based on two years average (year7 and year8) 41
Figure 4.9. Difference Map: rainy day's probability average of two years minus the interpolation map of average rainy days (year7 and year 8) 41
Figure 4.10. Variation of simulated Tmax for points of interest 43
Figure 4. 11. Average map of Synthetic Maximum Temperature of two Years...................................44
Figure 4.12. Difference map: Synthetic Maximum Temperature of two Years minus Average Maximum Temperature interpolation 44
Figure 4. 14. Average map of Synthetic Minimum Temperature of two Years....................................46
Figure 4.15. Difference map: Synthetic Minimum Temperature of two Years minus Average Maximum Temperature interpolation 46
Figure 5.1. Calendar of farm Operations in the visited area 48
Figure 5.2. Duration length of maize in field from interview of farmers 49
Figure 5.3. Time series of simulated Average Cumulative rainfall (20 years of data)..........................49
Figure 5.6. Regression of Simulated yield as function of Elevation and reliable rainfall 54
Figure 5.7. Regression of Simulated yield as function of Elevation and reliable rainfall 55
Figure 5.9. Scatter plot of the deviation to observed data of the selected regression vs regression established by Roetter(1993) 57

List of tables

Table 1. Examples of definitions of productivity of water by different stakeholders 10
Table 2.Major features of four stochastic weather generators (Mavromatis & Hansen, 2001)............ 13
Figure 3.1. Map of the study Area 18
Table 4.1. Coordinates of CRU TS 2.1 Grids for seasonal analysis 33
Table4.2. Coefficient of Fourier Polynomial used for Tmax, Tmin and Maximum Solar Radiation... 35
Table 4.3. Characteristics of regressions used for interpolation of Tmax, Tmin and Maximum Solar Radiation. 36
Table 5.1.Parameters used for adjustment and some reference values. 50

1. Introduction

Rain-fed agriculture is important component of the rural economy and poverty reduction strategy at global scale. It is not only a system by default, but can be even more productive in term of value per unit water used. Regionalization of crop and increasing productivity with efficient use of available water are well-known strategies. Unfortunately, tools for understanding variation are inefficient where the gaps of monitoring system are high. *Increasing climate variability and other indications for climate change show that hazards for African farming are increasing in number, extent, and severity* (Olufayo, Stigter, & Baldy, 1998).

Farming is a weather-sensitive occupation and where the agricultural system has enough flexibility, the agricultural `industry' can be highly responsive to weather and climate variables including Africa. In relatively low external input agriculture, as most of the rural agricultural “enterprises”, farm management strategy has to be adapted to climate, top soil and topography . Microclimate and topo-climate should be part of the management strategy often known by farmers but not considered in local or regional planning.

Traditionally, crop modeling uses different tools to cope with the spatial and temporal gaps of weather data. A weather generator is often used for this reason. Weather generators were conceived for Meteorological stations but interpolation efforts are not yet popular.(Daniel S. Wilks, 2008) The need for improving the weather generator is real particularly when spatial interpretation and high resolution of prediction is required. At the same time, evolution of the weather generator and crop model algorithm challenges environmental modellers and Geo-science community to develop comprehensible and accessible tools and assessment methods. Climate change studies have heavily contributed to this renewal of simulation tools.

Crop production sciences, Hydrology and Meteorology recognize the role of GIS and Remote Sensing in describing spatial variation of climate, soil moisture and vegetation response. The aim of adoption of these new techniques is a better description of actual and alternative farming system with a *close to farm* accuracy models .(A.D. Hartkamp, White, & Aguilar, 2000).

Spatial distribution of production potential for agriculture and livestock production (land, climate, topography, labour, vegetation) can measure the exposure of population to food insecurity and also indicate the contribution of farmers to environmental management. These two aspects are crucial in an equity-based IWM project such as IWM in Naivasha (supported by WWF).

1.1. Justification

A personal motivation and need to contribute to agro-meteorological research at the Agricultural Research Institute in Rwanda (ISAR), and awareness of the capabilities of the weather generator, particularly, site specific and distributed ones, especially WXGEN for Naivasha (originally adapted from SWAT algorithm by Lal Muthuwata(2004)) available in a distributed model using ILWIS GIS environment. This format has not yet been adapted and evaluated for crop or hydrological modeling (Bhandari, 2005).

1.2. Research Problem

Understanding the variability of climatic and hydrological factors of the biosphere leaves a gap towards solving food security. For the present study the focus is made on filling in the gaps on describing spatial variation of climate, soil moisture and vegetation response in a long-term scenario. Two approaches are often taken: improving the understanding of present situation using Remote Sensing and precise site observation or combination of statistical information and physical-based model using GIS capabilities.

1.3. Objectives

- To validate the procedure established by Bhandari for generating rainfall on pixel basis.
- To extend the pixel-based weather rainfall generator for Temperature and Solar radiation
- To assess the performance of Synthetic data in simulating Maize Yield
- To assess the relevance and need for irrigation in the studied area.

1.4. Research question

- Is the statistical structure of the observed and synthetic rainfall similar?
- Is the statistical structure of the observed and synthetic Temperature similar?
- Can the WXGEN raster-based implemented by Bhandari (2005) be used to predict productive areas?
- How can this approach be extended and generalized?

1.5. Hypothesis

- WXGEN in ILWIS can capture the spatial variability of climate.
- WXGEN time series can predict crop response amplitudes and variations for each agro-ecological zone.
- Interaction of Soil moisture and yield response and farmer level information can calibrate crop response model at agro-ecological zone level

2. Literature Review

2.1. Rainfed Agriculture and Variability of climate

2.1.1. Importance of Rainfed Agriculture

Rain fed agriculture produces by far the highest proportion (over 60 percent) of food crops in the world. When animal grazing is counted the contribution of rain-fed agriculture to food and commodity production is very high indeed. In sub-Saharan Africa it is estimated that over 90 percent of agricultural production is rain fed. Water Management and food security are highly related even for smallholder farmers. Experience in Tanzania, for example, shows that farmers productivity can be improved through rainwater harvesting and dry spell management (Hatibu,2007. http://www.iwmi.cgiar.org/home/rainwater.htm).

IWMI/UEA Climate Atlas estimates at 7 billion hectares land surface with a potential for rain-fed crop production. 4.7 billion hectares were classified as moderate, high, or very high suitable . Taking the example of Kenya where production potential varies significantly from one region to another, still important region with a potential above 3tonnes per hectare of cereals were identified(Droogers, Seckler, & Makin, 2001). Nevertheless, there is still a gap in term of poverty reduction to be filled: with an area estimated at 582,650 km2, Kenya was still counting 31 %Proportion of undernourishment i.e. approximately Number of 9.7 millions undernourished between 2001 and 2003.

A better characterization of farm in Kenya is now possible using research conducted between 1990's and early 2002: information on production goals (e.g. self-subsistence, market orientation…), the main types of constraints faced, position in the farm developmental cycle and main source of income. Graphic models of the farm system were made using this combined information. The survey highlighted the spatial variation between regions and also variation of fertility and management practices in a certain neighbourhood. Topography plays In an important role in the regional and local variability. (Tittonell, Vanlauwe, Leffelaar, Shepherd, & Giller, 2005).

Although rainfed agriculture produces more less than 40% of the potential of production in Sub-Saharan Africa, the wide range of Crop Productivity observed shown an important opportunity to raise crop yield of rain-fed agriculture. Especially, including water productivity concept extended to rainfed agriculture to agriculture management could improve. Crop failure and yield gaps are the primarily occasioned by either crop season dry spell or end of rainfall and low rainfall utilization efficiency which can create an evapotranspiration deficit.

2.1.2. Rainfed production and Water productivity

The concept of water productivity is applicable irrigated and non irrigated Agriculture. It has been defined as the amount of output produced per unit of water involved in the production, or the value added to water in a given circumstance. (Makarius, 2004).

Table 1. Examples of definitions of productivity of water by different stakeholders

Stakeholder	Useful definition	Scale	Target
Plant physiologists	Dry matter/transpiration	Plant	Productive utilization of light and water resources
Agronomist	Yield/evapotranspiration	Field	Higher yields t/ha
Larger scale farmer	Yield/water supply	Field	Higher yields t/ha
Irrigation engineer	Yield/diverted water	Irrigation scheme	Demand management
Water resources planner	$/total water depletion from the basin	River basin	Optimal allocation of water resources

Source: modified from Bastiaanssen et al. (2003)

Where water is a constraining factor, the importance of this concept is clearer. For years, semi-arid productions were expressed in yield per unit of water, or yield per unit of water evapotranspired, yield per unit of irrigation water supply.

2.1.3. Inter-annual and Seasonal variability and Yield prediction

Several studies in Australia earlier and recently applied to Zimbabwe and Kenya has show the relevance in seasonal prediction. Those predictions are developed for season prediction based identification of similarities with passed years. Particularly Relation between sea surface temperature and atmospheric variables produces encouraging results and implemented in Australian Agriculture forecast Systems and was tested in Zimbabwe(Hammer et al., 2001; Hansen & Indeje, 2004; Venus, 2000).

According to McGregor (McGregor & Nieuwolt, 1997) tropical non-seasonal variations of different frequency affected the tropical climate. Those variations can be classified in ascending order of periodicity as followed:

-Slowly varying 40-50 day tropical oscillation (Madden Julian Oscillation: MJO) in surface pressure, winds and temperatures
-quasi biennal oscillation in equatorial stratospheric winds (QBO)
-El Ninõ southern oscillation (ENSO)

The impacts of climate variability are sensitive in regions affected by the El Nino-Southern Oscillation (ENSO) events. The variability is a challenge for traditional agricultural research and extension (Podesta et al., 2002; Venus, 2000).

2.1.4. Growing Pattern and Agro-ecological zoning in Kenya

FAO define mapping unit, in terms of climate, landform and soils, and land cover. In this method, specific range of potentials and constraints are used to discriminate area(FAO, 1996).At regional and national level, the method suggests thermal zones defined on temperature intervals of 5°C or 2.5°C, and allow more detailed interval of thermal in temperate or subtropical regions. The table 2 based on the Kenya AEZ study, identifies 22 occurring LGP patterns.

Different layer should be compare to define a proper uniform area which approximately homogeneous characteristics for agriculture. In this regard soil map plays an important role, it generally account already for vegetation influence, topography and hydraulic conditions(Bakhsh, Kanwar, & Malone, 2007; Sharma, Mukhopadhyay, & Sidhu, 1998).

2.2. Spatial Interpolation of Climatic variables

According to Downing(1996, pp 215)(Dobesch, Dumolard, & Dyras, 2007)Climatology uses GIS mainly for

(i)interpolation of data and create climatologic fields use in impact studies and climate change scenario. The products can be downscale for local applications according to top-down disaggregating procedures. In this process topography plays often an important role as adjuncts;

(ii) handle climatologic and pollution database.

GIS based spatial interpolation should be confronted to physical process acting at different different time and different space scales. The relative importance of large scale external forcing varies for different meteorological situations. Spatialization develops datasets in form of gridded climate data at different temporal scales and can be taken as an opportunity for distributed modeling, nevertheless it require a carefull quality control process starting from the metadata because of the high spatio-temporal variability of climatic data (Dobesch, Dumolard, & Dyras, 2007).

Although, GIS becomes very important in climatology, interpolation methods still requires independent test in each individual case. Remote sensing data can supplement ground data, but a generalization of existing algorithm has not yet matched expectations. However application of circulation models improves the spatialization of daily values of most climatic variables.

Spatialization develops datasets in form of gridded climate data at different temporal scales and can be taken as an opportunity for distributed modeling, nevertheless it require a carefull quality control process starting from the metadata because of the high spatio-temporal variability of climatic data (Dobesch, Dumolard, & Dyras, 2007). Often Elevation is considered during the interpolation process. Goovaerts (2000) categorizes methods used for accounting for Elevation as followed : straightforward predicting the rainfall as a function of the co-located elevation, Spatial correlation of the residuals taken into account using the three types of geostatistical algorithms :

- SKlm: Simple kriging with varying local means
- KED: Kriging with external drift
- Co-kriging

2.3. Stochastic model for weather

2.3.1. General presentation

From 1980s Scientists have started developing predictive model of climate data. Many reasons contribute to development of stochastic weather generators. The first was the needs for synthetic weather time-series, for hydrological or agricultural applications (risk assessment), recently, new interest in stochastic weather simulation to enable climate change studies was raised (www.rothamsted.bbsrc.ac.uk).

Wilks (2008) define a weather generator as *time-domain stochastic models*. A stochastic process is random process, opposed to deterministic process .According to Doob (1934) It is *defined to be a one parameter set of chance variables:*

$$x(t), -\infty < t < +\infty$$

It is supposed that if $t_1, \ldots, t_n$, *is any finite set of values of t, and* $aj < x < bj, j = 1, \ldots, n$ *any set of intervals, the probability that*

$$aj < x(t_j) < bj, j = 1, \ldots, n \text{ is defined.}$$

Several studies using Richardson's model, have shown the relevance first-order Markov model to simulate rainfall events for point or distributed dataset. WGEN of Richardson is on the most documented weather generator. This approach is widely recognized as shown by implementation of MarkSim by Water and Food challenge program of FAO and CGIAR. MARKSIM is a software package to generate daily weather data for Latin America and Africa. It is based on a stochastic weather generator that uses a third-order Markov process to model daily weather data (Bhandari, 2005; A. D. Hartkamp, White, & Hoogenboom, 2003; Muthuwatta, 2004).

Table 2.Major features of four stochastic weather generators (Mavromatis & Hansen, 2001)

Weather generator	**Precipitation**		**Temperature**		**Solar irradiance**
	Occurence	**Intensity**	**Maximum**	**Minimum**	
WeatherMan	First order Markov Chain	Gamma	Conditional Gaussian	Unconditional Gaussian	Conditional, truncated Gaussian
MARKSIM	Transformed third order Markov Chain	Gamma	Conditional Gaussian	Unconditional Gaussian	Conditional, truncated Gaussian
WM2	Hybrid second order Markov Chain	Mixed exponential	Conditional Gaussian	Conditional Gaussian	Conditional Gaussian of logit-transformed clearness
LARS-WG(LARS)	Altering renewal with semi-empirical distribution	semi-empirical	Conditional Gaussian	Conditional Gaussian	Semi-empirical

Weather generator	Correlation (T_{max},T_{min},R_s)[a]	Parameter intrerpolation
WeatherMan (WM)	Trivariate autoregressive, constant	Mean preserving segment linear
MARKSIM (MS)	Trivariate autoregressive, constant	Mean preserving segment linear
WM2	Trivariate autoregressive, constant	None
LARS-WG(LARS)	Bivariate (T_{max},T_{min})[a] autoregressive	Third order Fourier series

Yang et al. (2005) advocate the use of generalized linear models (GLMs) for the stochastic modeling of daily precipitation and more recently for individual daily weather variables. GLM can provide a general modeling framework for incorporating climate states into parametric stochastic weather generators(C. C. Yang, R. E.; Isham, V. S.; Wheater, H. S., 2005) . Those perceptive are promising because of the simplification and the accessibility of tools utilized. The function GLM in the open source statistical programming software is readily available for fitting such models (e.g., , available at www.r-project.org).

In some Commercial Software package, a simplified weather generator is used and for practical reasons straight forward methods (without parametrization of a gamma function for example) are suggested like WXGEN (Hayhoe, 1998; Schuol & Abbaspour, 2007; Wallis & Griffiths, 1995).

The markov chain of the first order is used for long time already in generating missing data or downscaling rainfall data to lower time scale. Temperature or sun radiation can also be derived using stochastic models. Nevertheless those artificial data have an important inconvenient: in most of the model, they are computed independently and don't reflect natural interactions between climatic variables.

2.3.2. Inter-annual variability of weather generator

Precipitation models which operate on a daily time scale, reproduce lower-frequency variations characterized quantitatively in terms of a variance, which will be analogous to the inter-annual variance of monthly recorded precipitation in the observed data. Reproduction of the observed inter-annual variances by stochastic weather generators can be consider as performance indicator for many applications, and is has been identified as validation test for stochastic weather models. There is well known problem of under representation of low-frequency of weather generator named "overdispersion problem". In stochastic precipitation models widely used (e.g. models using first order Markov chain)(Katz & Parlange, 1998; D. S. Wilks, 1999a).

The procedure used decomposes the inter-annual variability of the monthly precipitation to express different aspects of the daily precipitation occurrence and intensity . This variance is expressed using the relation below (D. S. Wilks, 1999a)

$$\mathrm{Var}[S(T)]=E[N(T)]\sigma^2+\mathrm{Var}[N(T)]\mu^2$$

S(T) denote the sum of T consecutive daily precipitation intensities, i.e., the monthly ($T \approx 30$) total precipitation. Var[S(T)] expresses the realization-to-realization variation in these T-day totals(equivalently the inter-annual variance)…*where N(T) is the number of wet days during the T-day period* (D. S. Wilks, 1999a). $E[N(T)]$ and $\mathrm{Var}[N(T)]$ depend on the precipitation occurrence, and μ and $\sigma 2$ depend on the precipitation intensity models. The respective variance overdispersion can be expressed based on as following(Gregory, Wigley, & Jones, 1993; D. S. Wilks, 1999a):

$$\mathrm{var}iance\_overdispersion = \left[\frac{observed\_\mathrm{var}iance}{\mathrm{mod}eled\_\mathrm{var}iance} - 1\right] \times 100\%$$

2.3.3. Case study of Interpolated Weather Generator

Conventionally, weather generator parameters are built for specific locations for which meteorological data were collected. They are point model. Application of such data to hydrological and crop modeling is therefore problematic. Spatial interpolation of the weather generator parameters is a potential solution to this mismatch between spatial scale of weather generator compare to modeled system(Daniel S. Wilks, 2008).

Wilks (2007) applied for the entire USA , a methods which was mentioned in the literature earlier suggesting to interpolate the weather generator parameters and simulate grid cell by grid cell for high resolution weather generator(Guenni & Hutchinson, 1998) .

Semenov(2007) uses the same approach and accounted for the elevation during the interpolation process . In this second case the objective is the use of LARS-WG as tools for disaggregate climate projection (climate change study) over UK.

2.4. Crop modeling

2.4.1. Historical view

In order to be able to model crop productivity and yields on a global scale we need information on current and previous land cover, land use practices, and changes in land management over time. For the past, this has been done using a combination of census data on cropland inventories and land cover classification from satellite data…(Marko Scholze et al,2005)

The evolution of crop's simulation production from the ex USSR model is not yet accomplished. In general crop model are strong in predicting the biomass production and less in partitioning of photosynthesis's products between different parts of the crop (Monsi, Uchijima, & Oikawa, 1973; Sirotenko, 2001; Wang et al., 2002).Nowadays of the models have evolved mainly by integration of new algorithms for crop management (hybrid selection, planting and harvesting date, irrigation, fertilization, intercropping and multiple cropping, crop phenology, growth and carbon allocation). Although details or resolutions of description of processes and crop types differ. (Marko Scholze et al, 2005). The geoinformation science's possibilities have increased the capabilities of crop model allowing site specific while regional prediction was a standard.

2.4.2. Classification of Crop growth models

Models can be classified in different types: conceptual, physical or mathematical(Acock & Acock, 1991; Dourado-Neto et al., 1998). Conceptual model support construction of hypothesis based on representation of the really World. Pure physical model are rarely for crop modeling. Most of the crop model which will discusses can be consider as mathematical model(Dourado-Neto et al., 1998). Those models represents behavior of crop growth system mathematically. With democratization of computer,They have become popular. They can be classified in a number of classes, but the two main ones are the empirical, and the mechanistic models (Acock & Acock, 1991; Dourado-Neto et al., 1998).
The empirical models are also called statistical models. They determine relationships among variables ignoring the causing factor or reaction process to stress for example. The mechanistic models represent cause-effect relationships among the variables. They describe the plant yield and dry matter

production based on the knowledge of the processes of growth and development (Dourado-Neto et al., 1998).

Simulation of plant yield is used in the management of cropping systems, in the management of the distribution and storage of harvest or giving a notation to commercial farm according to the season (Insurance company, Financial Market…)

2.4.3. Characterisics of mechanistics crop growth model

Among Wageningen crop simulation Software, WOFOST, sucros, are mentioned as mechanistic model. Three modules are common among mechanistic models of this family of software (van Ittersum et al., 2003).

Phenological development

The development stage (DVS) of a crop defines its physiological age and is characterised by the order and rate of appearance of the various organs. Modeling of the development rate is still largely descriptive, and based on temperature and day-length. For many annual species DVS is defined in terms of heat sums (in day-degrees).

The DVS governs the change on assimilation rate of CO_2, Specific Leaf Area or partitioning factor of dry matter produce between plant's organs.

Dry matter partitioning and sink size

The daily increase of dry matter for different (leaves, stems, storage organs and roots) results from the partitioning of daily dry matter production divided according to partitioning coefficients (kg dry matter organ kg^{-1} dry matter crop)

Leaf area development

The green leaf area of plants determines the amount of absorbed light and thus CO_2 assimilation. For a closed canopy, LAI is calculated from leaf dry weight, using the specific leaf area. When the canopy is not closed, the leaf area grows exponentially as a function of the temperature, assuming that temperature is the determining factor for cell division and expansion (sink limitation).

Most of the Wageningen's models consider only leaves photosynthesis as sources of CO_2 assimilation, although or greens part of the plants exposed to sunlight and to biosphere air contribute in reality to it.

2.4.4. Accounting for the Water Stress

Crops model also provide indication of a potential yield using identication on soil-water and nutrients available, all other parameter considered optimum,. water production function (wpf) and crop water production function (cwpf) are used for this purpose. wpf is the relationship between crop yield and water application and cwpf is the relationship between crop yield and evapotranspiration(Al-Jamal, Sammis, Ball, & Smeal, 2000; Brumbelow & Georgakakos, 2007; Kipkorir, Raes, & Massawe, 2002).

Wpf is linear when the soil is in water deficit, because all the applied water is used as evapotranspiration (wpf =cwpf). In the non-deficit case, wpf is non-linear that indicates that the plant did not consume all the supplied water. In case of deficit regime (deficit irrigation or rainfed agriculture) Cwpf is more useful. It is often linear and varies with varieties of crops and climate zones. A general form of Cwpf can be expressed as following(FAO, 2002; Kipkorir, Raes, & Massawe, 2002):

$$1-(Ya/Ym)=Ky(1-(ETa/ETc))$$

Ya : Actual yield, Ym : maximum yield, ETa and Etc: actual and at optimum Evapotranspirations. The reduction of Yield can be estimated qualitatively or using range of yield for give index of stress. The use of CERES-Maize has shown the usefulness of incorporating crop simulation into a decision support system which can determine the best management responses to specific weather conditions.

3. Methodology

Figure 1. 1 illustrates the entire methodology including the analysis. It shows three important periods which are the pre-data collection and refinement of the interview approach, the data collection and data mining to retrieve actual pattern and historical trend of crop yield in the area and modeling part using weather generators and crop model both spatially-distributed or semi-distributed (grid-based). Advances in spatially disaggregated datasets and GIS tools allow further investigation into agricultural production to be characterized and mapped in a much more quantitative manner.(A.D. Hartkamp, White, & Aguilar, 2000).

The Study Area is Lake Naivasha drainage basin which has already been used by other researchers as case study of biophysical characterization using RS and GIS tools. It constitutes a benchmark for spatial approach study and integrated approach innovations outreach.

3.1. Study Area

From the original occupation of Masai people, the lake Naivasha catchment has changed to agriculture and sedentary type of economy. The land uses observed are residential area, Irrigated and non-irrigated agriculture, dairy farming, rangeland, natural vegetation, and out crop rocks. Fishing is also a very important activity in the area.

The catchment of Lake Naivasha located between central province and the province of Rift valley (figure1.2.) is representative of the diversity of climate in Kenya and particularly from the top to bottom of Rift valley landscape. The rainfall varies from ≈ 400 mm to more than 1500mm in this area.

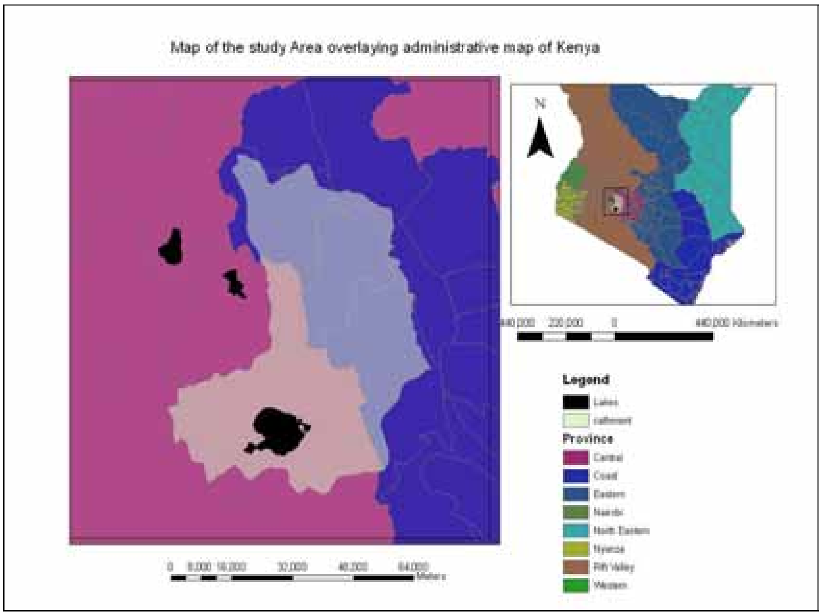


Figure 3.1. Map of the study Area

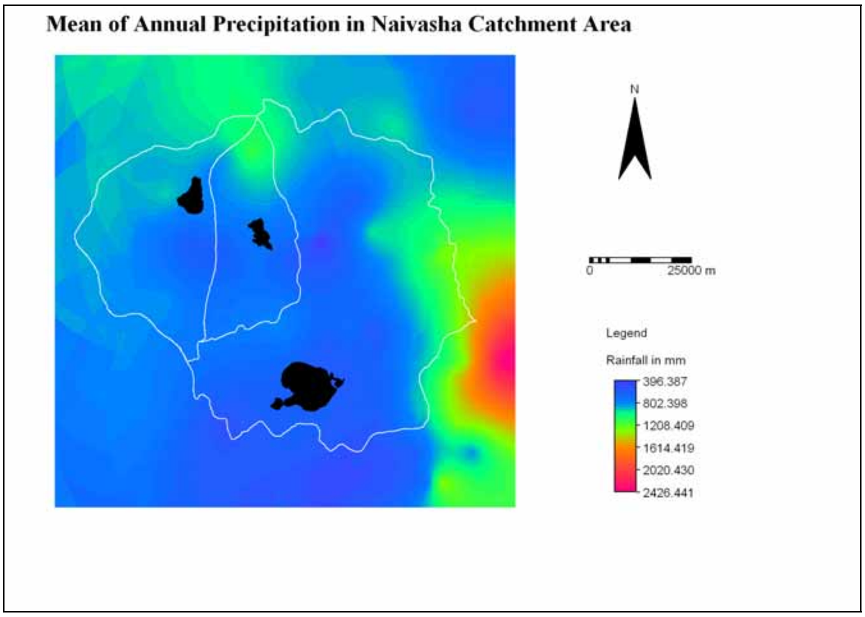


Figure 3.2. Rainfall spatial distribution

The volcanic rocks of the rift valley are the origin of Naivasha catchment soils which varies from moderate to low fertility. Greyish brown to black with clayish loams texture are characteristic for the area with often poor drainage condition.

3.2. Plant test

3.2.1. General consideration

Maize (*Zea mays*) belongs to the family of Gramineae (grass family). It has separated female (ear) and male (tassel) flowers. It is classified as C4 plant meaning that the first stable product of photosynthesis is a molecule with a structure of four atoms of carbon instead of three (FAO, 1992). A supplementary method of CO_2 uptake is used by this group of plants. The extra mechanism allows a faster growth but also creates an increase of demand of input (light, water, nutrients) for dry matter production thus the particular interest of research on ecological condition of the plants highly adopted in America, Africa and Asia(Merwe, Tschauner, Rowan, & Russell, 1999).

The maize plant may be defined as a metabolic system whose end product is mainly starch deposited in specialized organs, the maize kernels. The development of the plant may be divided into two main parts (FAO, 1992):

- The vegetative stage, which is a period of development and differentiation of different tissues until flowering. It starts with a cycle of formation of the first leaves during an upward development with a slow production of dry matter production. This cycle ends with the tissue's differentiation of organs of reproduction which develops in the second cycle as well as leaves organs.
- The reproductive stage starts with the fertilization of the female flowers, where ears and grains are developed. In the first phase, the weight of leaves and other flower parts increase, followed by the second phase of rapid increase of the kernels' weight.

3.2.2. Maize in kenya

Maize is the primary staple food crop in the Kenyan diet with 98 kilograms per capita per year, according to USAID (http:// www.usaidkenya.org / ke.agbuen/ activities/ acdivoca.html, 4th January 2008). The inefficient maize production-marketing system has been identified by this report as contributing to economic stagnation and high levels of poverty. The area under Maize in the whole country stabilize at 1.4 million of ha, and the estimate yield is 5 tonnes per ha (FAO , http://www.fao.org/ag/agL/swlwpnr/reports/y_sf/z_ke/ke.htm, consulted 20 august 2007) .

The first improved Kenyan maize variety was the Kitale Synthetic II (1961), based on inbred lines from the Kenya Flat White complex. From that time a program to widen the genetic base of the Kitale products started using the diversity of center-of-origin material from similar ecological conditions to those of East Africa, close to the Equator with a wide range of altitude. Another variety well known and often used in computer-based simulation is Katumani Composite B (Hansen & Indeje, 2004). This is a fast variety of maize that is fairly short with short cobs. It is a drought-resistant variety, which needs 60-65 days for flowering, 90-120 days for maturation and 250-500mm of rain. The Insect-Resistant Maize for Africa (IRMA) baseline survey identified more than 40 distinct maize varieties in Kenya (Smale, Groote, & Owuor, 2006). According to the interviews of DAOs (District's Agricultural Officer), extension officers and farmers the Katumani, new hybrids adapted to transitional altitude where the temperatures ranges from 12C to 30C and has rainfall requirements ranges from 800-1500mm and high altitude where the temperatures ranges from 8C to 28C and has rainfall requirements ranges from 1500 to 2100mm. A typical example is Hybrid 624 (H624), it performs well in altitude between 1500-2100m, produce more than 10Tonnes per ha in ideal condition with a growing period length between 5 to 7 months.

3.3. Weather data

3.3.1. Rainfall data

Synthetic data of rainfall were available for Naivasha catchment area. They were obtained using the WXGEN rainfall generator (Bhandari, 2005) for 32 rainfall station within the lake Naivasha catchment.

Bhandari (2005) established a procedure to derive aerial rainfall data and the error maps using co-kriging of WXGEN input parameter with elevation allowing for computation of the daily synthetic rainfall and the transition probabilities(Bhandari, 2005).

The Annual variability was optimized by a Fourier series of second or third order for each WXGEN input (skewness, transition probabilities, mean, standard deviation).

$$Fn(x) = a_0 + \sum_{k=1}^{k=n} (a_k \cos(kx) + b_k \sin(kx))$$

Where a_o is a constant and a_k and b_k are constant for i=1,…..,n
The Fourier polynomials are 2π periodic function. The period for Climatic should be the same.

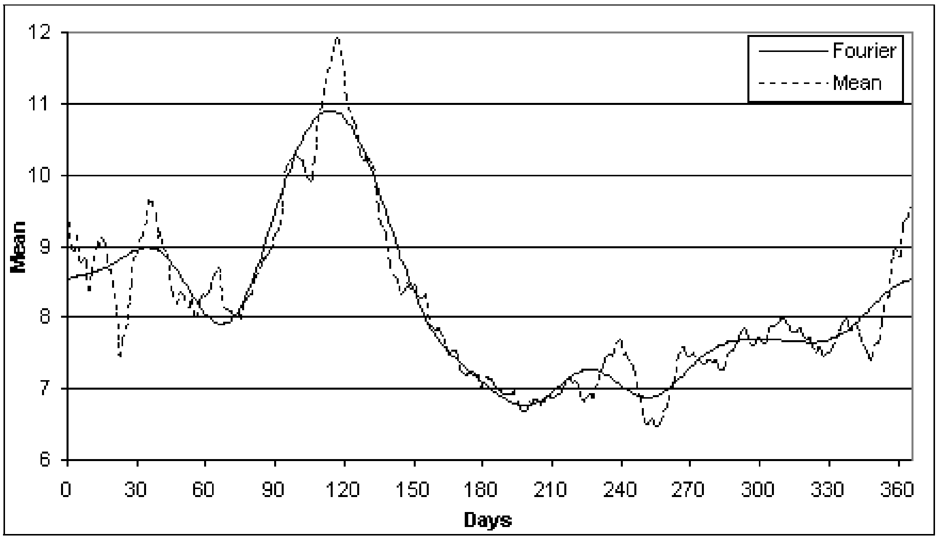


Figure 3.3. Average of mean for each Julian day and Fourier series fitted (Bhandari,2005)

The syntax of WXGEN in ILWIS as an "ILWIS script" uses WXGEN input in a form of specific value per julian day. It first creates the input map for each day (mean of rainfall, standard deviation, skewness, SDN, probability of having a wet day after a dry day, probability of having a wet day after a wet day).

$$R_{day} = \mu_{day} + 2 \cdot \sigma_{day} \cdot \left\{ \frac{\left[\left(SND_{day} - \frac{g_{day}}{6} \right) \cdot \left(\frac{g_{day}}{6} \right) + 1 \right]^3 - 1}{g_{day}} \right\}$$

Where: R_{day} = amount of precipitation on a given day (mm)

μ_{day} = long term mean of daily precipitation (mm) for the day

σ_{day} = long term standard deviation of daily precipitation (mm) for the day

SDN_{day} = standard normal deviation calculated for the day

g_{day} = long term coefficient of skewness for daily precipitation in the day.

Rainfall amount is computed using the randomness of SDN (standard normal deviation calculated for the day), the mean for wet spell, and the probability of having a dry spell. SDN factor of WXGEN is computed by the equation:

$$SDN_{day} = \cos(6.283.rnd_2) * \sqrt{-2 * \ln(rnd_1)}$$

Each input parameter of the rainfall model is function of latitude, longitude and time (y,x and t respectively) but the altitude is also accounted for indirectly (z). The 3 dimensions (with influence of z) is implemented using a combination of an aerial mean of the parameter to Fourier series defining the daily amplitude of the parameter. The value of the parameter is retrieved by the relation:

$$Pxyt=(ft*axy)+axy$$

Where P_{xyt} is the pixel specific daily value of the parameter

f_t is daily value of the Fourier transform for the day t

a_{xy} is mean of daily value of the parameter for a specific pixel (given by the aerial mean)

f_t is taken for a Fourier transform developed by *automatized* fit to time series of deviation to annual mean (dev_t) for all weather station of the area.

$$dev_t=(\text{daily mean for the day t-annual mean})/\text{annual mean}$$

An apparent difference could appear between the syntax of the ILWIS script and Bhandari example file due to development of a full scene of a year of daily data for the purpose of integrated data for crop modeling. In the example file uses only Pwd , the probability of having a wet day after a dry day, but it is explained that for the next day Pww(probability of having a wet day after a wet day) also has to be integrated.

3.3.2. Extension to Solar Radiation and Minimum and Maximum Temperature

The realizations of radiation and minimum and maximum temperature followed the methods established by Richardson (1984), which generate randomly residuals of the three variables to create their correlation and cross-correlation in time (see § …).

Richardson (1984) was followed by replacing the current case mean of tmax, tmin and solar radiation by aerial mean daily average of each of these parameters and corresponding mean coefficient of variation of daily values. Aerial means were obtained by fitting a linear regression model .The regression estimation used monthly data, to fit an empirical relation. A first order regression was chosen. The interpolated parameters were obtained using elevation as an explicative variable following a general form equation:

$$Yi=a*Zi+b$$

Y_i is the parameter interpolated

Z_i is the elevation of particular grid i: is an identifier for the grid

a: slope determine of the linear equation

b: the intercept of the linear equation

The harmonics were replaced by the Fourier series of the 3 means (tmax, tmin and Solar radiation) and the corresponding coefficient of variation. Time series of 12 grids of monthly weather data for 102 years (from 1901 to 2002), covering the area, provided by CRU (Centre Research Unit, http://www.cru.uea.ac.uk/) were used for the establishment of those periodic functions.

3.3.2.1. Cross correlation and Auto correlation Matrix

For generating daily values of t_{max}, t_{min}, and Maximum Solar radiation, WGEN is based on the following equation:

$$\chi_i \begin{pmatrix} Solar\_Radiation_i \\ T\max_i \\ T\min_i \end{pmatrix} = A * \begin{pmatrix} Solar\_Radiation_{i-1} \\ T\max_{i-1} \\ T\min_{i-1} \end{pmatrix} + B * \begin{pmatrix} Rand1 \\ Rand2 \\ Rand3 \end{pmatrix}$$

Maximum Solar Radiation replaces Solar radiation used in the WGEN algorithm. *Rand1*, *Rand2* and *Rand3* are three independant random number between 0 and 1. A and B are 3 X 3 matrices. Defined by the following equations:

$$A = M_1 * M_0^{-1}$$
$$B * B^T = M_0 - M_1 * M_0^{-1} * M_1^T$$

M_0 and M_1 are matrices of correlation and cross correlation of solar radiation, Tmin and Tmax, for lagtime equal to 0 and lagtime equal to one day. M_0 and M_1 are defined as follows:

$$M_0 = \begin{pmatrix} \rho_0(SR,SR) & \rho_0(SR,T\max) & \rho_0(SR,T\min) \\ \rho_0(T\max,SR) & \rho_0(T\max,T\max) & \rho_0(T\max,T\min) \\ \rho_0(T\min,SR) & \rho_0(T\min,T\max) & \rho_0(T\min,T\min) \end{pmatrix}$$

$$M_1 = \begin{pmatrix} \rho_1(SR,SR) & \rho_1(SR,T\max) & \rho_1(SR,T\min) \\ \rho_1(T\max,SR) & \rho_1(T\max,T\max) & \rho_1(T\max,T\min) \\ \rho_1(T\min,SR) & \rho_1(T\min,T\max) & \rho_1(T\min,T\min) \end{pmatrix}$$

ρ_0: is the correlation of two variables without lagtime

ρ_1: is the correlation of two variables with lagtime of one day.

Solving matrix A is a straight forward process while solving matrix require a software to solve the equation which can have different solutions.

Steps followed for the computation are:

step1: Compute average values of Maximum Solar radiation , Tmax , Tmin for each day of the year. Create a table with a time series of the three climate variables

step2: Generate the Matrix M_0. For this a computation of the cross correlation and the correlation have to be done.

step3: Generate the Matrix M_1. For this a computation of the cross correlation and the correlation have to be done

step4: Compute the inverse of matrices M_0 and M_1

Step5: compute matrix A using the relation

A spreadsheet solution step by step was chosen to solve matrix B. Because it is easy to produce inputs of the equation used. The principle used is to minimize an objective function which links two parts of the equal sign. Different values of B can be a solution to the equation. Additionally, three function which forced the cell (1,2), (1,3), (2,3) were added. They were implemented in the spreadsheet's solver as constraining functions (constraint) .

The steps used are described below:

Step 1: Compute the matrix M_1^T; the transpose of matrix M_1

Step 2: Compute the matrix S_1 using matrices M_0 M_1 M_1^T and M_0^{-1}

$$S_1=M_0-M_1*M_0^{-1}*M_1^T$$

Step 3: create a system of two matrices B' and B'^T of 3X3, B' and B'^T. Both should be linked by formula.

Step 4: Compute S_2 using matrix B' and B'^T

$S_2=B * B^T$

Step 5: Create an objective function linking S_1 and S_2. The objective function can be defined as follows:

$F(S_1,S_2)=(S_1(1,1)-S_2(1,1))^2+(S_1(1,2)-S_2(1,2))^2+(S_1(1,3)-S_2(1,3))^2+(S_1(2,1)-S_2(2,1))^2+(S_1(2,2)-S_2(2,2))^2+(S_1(2,3)-S_2(2,3))^2+(S_1(3,1)-S2(3,1))^2+(S_1(3,2)-S_2(3,2))^2+(S_1(3,3)-S_2(3,3))^2$

Step 6: Solve $F(S_1,S_2)=0$, by changing B' (in Excel is available in Tools, Solver Add in)

Step 7: Verify the results of the computation. S_1 is static in spreadsheet and doesn't change with the computation. S_2 is the dynamic part which change and should be equal to S_1 at the end of the process. The final B' is the value of matrix B.

Additionally three function which force the cell (1,2), (1,3), (2,3) were added. Directly implemented in the solver during the computation of the matrix B by adding 3 constraints

B'(1,2)=0
B'(1,3)=0
B'(2,3)=0

Mean Solar radiation is used in the Richardson methods, instead of the data of the weather station installed at doodeman farm, for the determination of correlation and cross correlation.

Using Matlab software this process can be reduced by computing from step 2, S_1. It is decomposed using an algorithm called *Cholesky decomposition* (function chol in Matlab). The solution of this decomposition is the matrix B^T. Finally B can be computed by transposing B^T.

3.3.2.2. Linking WGEN Residuals and daily mean

The residuals are correction terms which affect the variation to mean of Tmin, Tmax or Solar radiation. WGEN defines these effects by the following relation:

$t_i (j) = m_i (j)* [\chi_i(j)*c_i (j) +1]$

i : a given day in a julian calendar
j : identifier for the variables (j=1 for Tmax, j=2 for Tmin, j=3 for Solar radiation)
$t_i (j)$: final value of the variable j generated for the day i
$c_i (j)$: Coefficent of Variation of the variable j generated for the day i
$m_i (j)$: mean value of the variable j generated for the day i
$\chi_i(j)$: residuals of the variable j generated for the day i

The seasonal variation of the coefficient of variation and mean of Tmin, Tmax or Solar radiation are modeled by an harmonic for dry and wet condition. In WGEN the harmonic is a trigonometric function. For the extension of WXGEN, Fourier series were selected and the dry and wet periods using a correction factor after computation of t_i.

It is strongly recommended to build the harmonic with long time series of observed data. This was possible for weather station established in the catchment. Assuming regional uniformity of seasonal variation of Tmin, Tmax, Solar radiation grid weather data of CRU were used. CRU grid weather data are monthly data produced from 1900 to 2002 at a resolution of 0.5°. Twelve (12) grids cover the study area. Finally, the different variables were obtained using the equation below.

$Tmax_i = m_i(Tmax_i)* [\chi_i(Tmax)*c_i (Tmax) +1]$

$Tmin_i = m_i(Tmin_i)* [\chi_i(Tmin)*c_i (Tmin) +1]$

Maximum Solar radiation$_i$ = m_i(Maximum Solar radiation$_i$)* [χ_i(Maximum Solar radiation)*c_i (Maximum Solar radiation) +1]

$m_i(Tmax_i)$, c_i (Tmax), $m_i(Tmin_i)$, c_i (Tmin), m_i(Maximum Solar radiation$_i$) and c_i (Maximum Solar radiation) are mean and coeffient of variation for Tmin, Tmax and Maximum Solar radiation. For generation pixel by pixel , it has chosen to use each of those term as a daily map with a specific value for each pixel.

The fourier is a normalized variation of the mean combined to map the annual mean of Tmax, Tmin or Solar radiation. The normalized variation of mean is defined as follows:

dev_i(Tmax) =(daily mean$_i$ (Tmax) -annual mean(Tmax))/annual mean(Tmax)

dev_i(Tmin) =(daily mean$_i$ (Tmin) -annual mean(Tmin))/annual mean(Tmin)

dev_i(Maximum Solar Radiation) =(daily mean$_i$ (Maximum Solar Radiation) -annual mean(Maximum Solar Radiation))/annual mean(Maximum Solar Radiation)

dev_i(CV_Tmax) =(daily mean$_i$ (CV_Tmax) -annual mean(CV_Tmax))/annual mean(CV_Tmax)

dev_i(CV_Tmin) =(daily mean$_i$ (CV_Tmin) -annual mean(CV_Tmin))/annual mean(CV_Tmin)

dev_i(CV_Maximum Solar Radiation) =(daily mean$_i$ (CV_Maximum Solar Radiation) -annual mean(CV_Maximum Solar Radiation))/annual mean(CV_Maximum Solar Radiation)

Six Fourier series were fitted to the series of deviations to mean. The methods of determination of a Fourier using spreadsheet iterative computation was preferred in stead of FFT in SPSS software package (Mentioned in the thesis of Bhandari(2005)).

$dev_i(y) \approx f_y(i)$

3.3.2.3. Linking Fourier series and aerial map of daily mean in GIS environment

4^{th} order of Fourier series was selected according to degree of fitting judged visually. Therefore 9 parameters had to be determined for each of the fitted harmonic. The general equation of the Fourier series becomes:

$$f(i)=a +b*(\cos(da_i))+c*(\sin(da_i))+ d*(\cos(2* da_i))+e*(\sin(2* da_i)) + f*(\cos(3* da_i)) +g*(\sin(3* da_i)) + h*(\cos(4* da_i)) +j*(\sin(4* da_i))$$

a, b, c, d, e, f, g, h and j are coefficient to be determined.
da_i is the day angle for the day i, in radian computed using the relation:

$$da_i =2*\pi*i /365$$

A simple multiplication of the Fourier value for the day and mean was applied to link Fourier series and aerial map of daily mean in a GIS environment. It is a straightforward method used for rainfall generation (Bhandari,2005). Pixel by pixel a specific value was computed according WXGEN algorithm for each realization (not necessarily equal to the value of the next pixel or the next day).

3.3.3. Other weather data for crop simulation

i. **Relative humidity** was computed using the temperature, rainfall, and periodic (trigonometric) function. The general form of the regression model is presented below:

$$RHi=a*Pi+b*(Tmaxi+Tmini)+F(i)$$

$$F(i)=aF*(\cos(dai))+bF*(\sin(dai))+ cF*(\cos(2*dai))+dF*(\sin(2*dai))$$

Where , i days of the year expressed in julian day
RHi , for the day i
Pi, Tmaxi, Tmini are precipitation, Tmax, Tmin generated for the day i
dai is the day angle for the day i
a, b, aF, bF, cF and dF are coefficient determined by the solution of the regression analysis using observed data of RHi , Pi, Tmaxi, Tmini. data for kedong, mutobio, North Kinangop hospital and Doodeman rainfall stations (2000-2002 and 2007).

ii. **Wind speed** is tabular value kept constant in space and time.

3.4. The Crop Growth model

The selected model for crop-modeling is WOFOST because of its capabilities to integrate different types of observation from the field and multiple type of simulation possible (potential yield , water limited yield). For potential yield, WOFOST uses temperature dependant growth stage to simulate the plant growth. Presently WOFOST has been incorporated in the Crop Growth Monitoring System (CGMS) software used by the European Union (Wit, Boogaard, & Diepen, 2005). For this thesis a model implemented in PCRaster in an experimental phase (Prof. Victor Jetten, personal communication) was chosen.

Potential and water limited production was simulated by the simplified WOFOST algorithm implemented in PCRaster language and using Nutshell software (http://www.itc.nl/lisem/nutshell/) as a compiler and shell for the computer code. The potential production is a function of temperature, solar radiation and crop characteristics.

3.4.1. Environmental Modeling using PCRaster

The crop growth is a spatially distributed one. It can input for each time a specific value for each grid of the spatial domain considered. It has been built using PCRaster Environmental Modeling language which is a computer language with dynamic modeling capabilities (iterative spatio-temporal environmental models) .

PCRaster is GIS system developed by the University of Utrecht. It has dynamic and iterative computation capacities. The language is conceived for research in the environmental sciences, and allows an easy development of diverse form of distributed model.

3.4.2. Development Stage simulation

D is the development stage: stepwise linear function: it is set at 0 at emergence of the plant, 1 at flowering and 2 at maturity and based on the concept of thermal time(TSUM).

TDbase, base temperature below which no development occurs.

TDmax1 maximum effective temperature above which a the phenological development

TDmax2 maximum temperature the development stops

D, *TDbase, TDmax1* and *TDmax2* are related to temperature and sum of Temperature par the following equations, *dTe* indicates the daily effective temperature.
(http://www.geog.uu.nl/landdegradation/teaching/mlcm/MLCM2005_2.pdf)

$$
\begin{array}{ll}
dT_e = 0 & for\ T < TD_{base} \\
dT_e = T - TD_{base} & for\ TD_{base} < T < TD_{\max 1} \\
dT_e = TD_{\max} - TD_{base} & for\ TD_{\max 1} < T < TD_{\max 2} \\
dT_e = 0 & for\ T > TD_{\max 2} \\
T_e = T_e + dT_e &
\end{array}
$$

The other crop parameters are *Ce,* the efficiency assimilates to biomass conversion, *Cm* maintenance respiration coefficient, *Rootmax* max root depth in cm, *DRoot* root increase cm per day. 6 development's stage dependant crop parameters are defined as well: SLA specific leaf , AMD CO_2 assimilation rate, and partitioning parameters prt**,** plv**,** pst and pso, to root, leaves, stem and storage organs**.**

3.4.3. Soil Water Balance

The daily soilwater balance is calculated by WOFOST. The moisture availability in the soil affects the growth of the crop. The maximum production rate is when the soil is at field capacity. Also the percolation beyond the root zone is calculated.

The evapotranspiration is estimated using the Penman Monteith equation. The evapotranspiration is constrained by water availability and modeled by water retention curve presented in the following equation.

$$\boxed{\theta = \theta_p * (\exp(-\gamma * \ln(h)^2)}$$

θ, soil moisture
θ_p, porosity
γ, texture dependent parameter
h positive matrix potential in cm

The lower boundary percolation is estimated by the following equation

$$\boxed{Perc = K_s \left(\frac{\theta - \theta_{fc}}{\theta_p - \theta_{fc}}\right)^2}$$

where θ_{fc} is the field capacity if $\theta > \theta_{fc}$,

Ks is the saturated hydraulic conductivity in cm/d.

Therefore 3 soil parameters are required for the computation of the two outgoing terms in accordance to the equations mentioned above and used in WOFOST, namely θ_p, γ and K_s. In the simplified model, two (2) more parameters are needed: albedo and soil depth. The volume of the soil exploited by roots is dependent on the development stages (function of DVS). The incoming term is the effective rainfall from which interception is deducted.

3.5. Soil data

The data retrieve from KENSOTER database was used to set soil texture, and average organic matter of soil unit.

3.6. Analysis

Similarity and variability of Weather input data

Mean and variance test was applied to evaluate simulated temperature and solar radiation data, variation of input data for both variable were investigated and Inter-annual variability was evaluated (A. D. Hartkamp, White, & Hoogenboom, 2003).

The sensibility analysis was used in the interpretation of the interpolation of parameters in the catchment area. One can discuss the boundary of the parameters and limits of realistic value.

Validation of the weather generator was based on

- variability of monthly and annual means
- distribution of length of dry and wet periods
- normality of variables of WGEN model SRAD, TMAX, TMIN

3.7. Calibration and sensibility analysis

3.7.1. Calibration and Validation of the Crop model

Calibration should be understood as adjustment of model parameter to match a set of real data. Normally, for a typical crop simulation study to calibrate a crop model is a long time investment in data collection (Leaf Area Index, Biomass, weight of storage organs through the growing periods), (Roetter, 1993). The nature of the study and the time frame didn't allow such a plan. Documents retrieved during the data collection, were used to create a set of training data for the generic maize similar to variety H635.

The calibration consisted of modification of crop-specific parameters using simulated Crop parameter. Parameters dependant or not on Development stage were fine tune. Biomass production and LAI indicated in the literature for Maize in this area comforted the calibration. (Wu, Yu, Lu, & Hengsdijk, 2006; H. S. Yang et al., 2004).The validation used point information from fieldwork campaign and yield measurement survey results from station. (Doodeman Farm, naivasha catchment area)

Performance Criteria of Yield estimation

The calibrated results were compared to yield information retreived from farmers. Regional statistics and potential production map were used for spatial variability comparison. The *goodness-of-fit* simulated yield was compared to the observed data and evaluated using Mean-absolute error (MAE) (Hansen & Indeje, 2004).

A comparison of the performance of distributed yield computed from direct relation rainfall-DTM-yield and rainfall-yield was made.

3.8. Assumption

The first assumption is the uniform use of input for agriculture over the agro-topoclimate unit, what is almost impossible to find in reality, but which can be acceptable if the size of farms and farming system are homogenous over the lumped area. This allows assumption of the same type of fertility management and the same source/sink source for plant water balance namely irrigation or mulching. The latter practice is used for cash crops in rainfed agriculture but systematic uses in large scale are rare.

Soil variability is related to Geology, climate and land use and the factors considered for mapping soil units and agro-climatic classification. Nevertheless within a soil unit, soil parameters vary. In the current study only aggregated value collected by previous soil survey in the study area will be considered. The calibration of those data will be also done for lumped agro-topoclimate.

Assumption related to economic aspects concern mainly the social parameter and variables assumed to be constant and steady over time and over space. This can be a reasonable assumption in the case of a small area, but depends highly on local dynamics, which are beyond the scope of the current study. Another factor which can be important and is not feasible in the current case is the impact of road vicinity which is a source of externalities and can influence the cost (benefit for farmers) of crop products and the productivity and opportunities for alternative practices

4. Weather Generation process: Results and Discussion

Eight raingauge stations with long sequences of measured data were selected to represent a range of climates. The proximity to study area and availability of long daily rainfall data limited the choice of observed time series.

For each weather variable, statistical tests were applied to compare “inter-annual” (among years), and seasonal (among months) variations. Weather generator outputs were analyzed considering the variations of mean or standard deviation. Solar irradiance was excluded from statistical analysis.

4.1. Interpolation and Input for Weather Generation

Grid data of CRU TS 2.1 offered by CRU (Climatic Research Unit) were used for the analysis of seasonal variation of Tmax and Tmin. 12 grids covering the study area were selected. Each grid represents a square cell of 0.5°. The Bounding box created by those 12 grids together is

Longitude : 36° East to 37.5° East
Latitude : 1° South to 1° North

A cell is referenced following an origin located in the South Pole at the Reference meridian. A cell is named using a system of two coordinates. The table 4.1. presents reference of each grid used for the seasonal analysis. The coordinates of the grid indicate the distance to longitude and the latitude to the reference point. The reference point (South Pole at the Reference meridian) has the coordianate (1,1) . The last grid (with the higher longitude 180° East and the higher latitude 90° North) has as reference the coordinate (720,360). For each grid monthly data of Tmin and Tmax are provided form 1901 to 2002.

Table 4.1. Coordinates of CRU TS 2.1 Grids for seasonal analysis

Coordinate of the cell refereeing to Longitude.	Coordinate of the cell refereeing to Latitude.	Symbols used in the presentation of results
433	179	(433,179)
433	180	(433,180)
433	181	(433,181)
433	182	(433,182)
434	179	(434,179)
434	180	(434,180)
434	181	(434,181)
434	182	(434,182)
435	179	(435,179)
435	180	(435,180)
435	181	(435,181)
435	182	(435,182)

The Figures 4.1. and 4.2. represent the variation of average value of Monthly Tmax and Tmin for the the selected 12 grids of CRU TS 2.1. Average values were computed using the 102 Monthly average of Tmin and Tmax (related to 102 years).

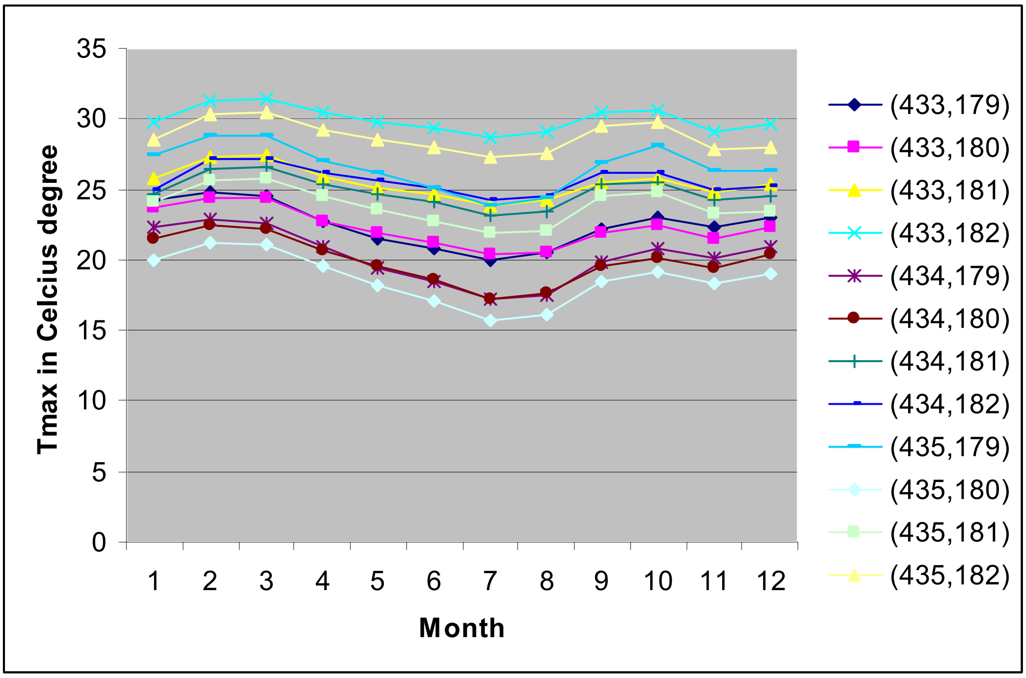


Figure 4.1. Variation of Monthly average of Tmax for 12 grids covering the Area

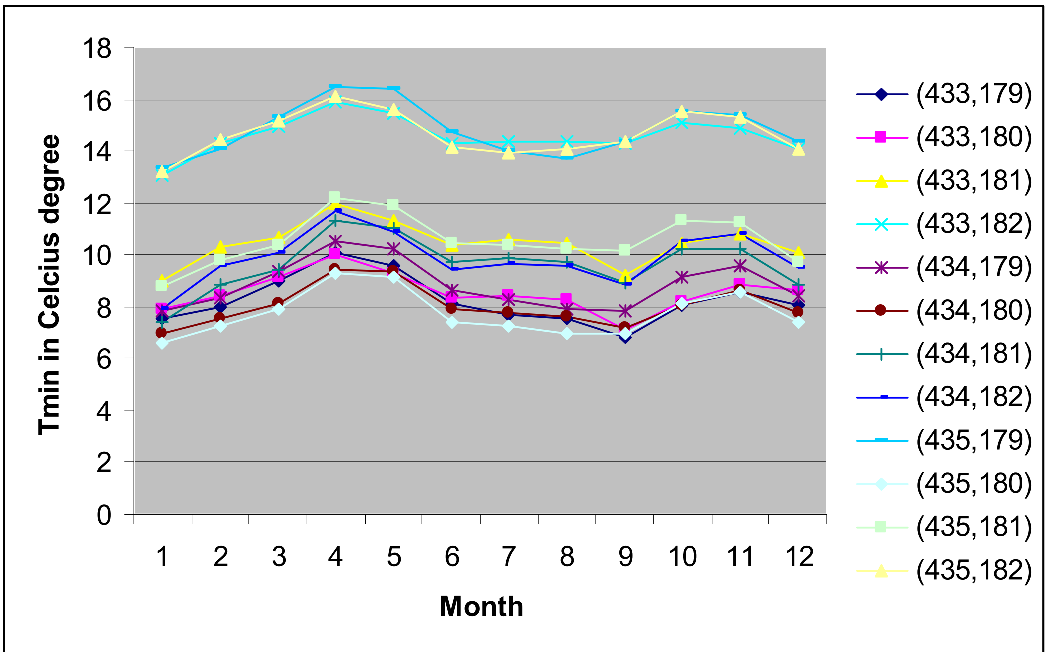


Figure4.2. Variation of Monthly average of Tmin for 12 grids covering the Area

The time series of monthly value of Tmin and Tmax for the 102 years were averaged. The tendency is approximately followed for most the year. A Fourier fitted to normalized time series could acceptably represent the seasonal variation of Tmax and Tmin. (Annex 2)

The Fourier of Maximum Solar radiation uses as proxy the seasonal variation of Mean Solar Radiation. The daily Maximum Solar radiation was only available from 2007. Mean Solar Radiation was recorded from 2000 to 2003 in the Weather Station located at Nakuru (Synoptic Weather Station).

Ideally daily could be preferred for Fourier series for Weather variables. Applications of harmonic to seasonal variation suggest using long period variation at large scale and Short period variation for local study.

Table4.2. Coefficient of Fourier Polynomial used for Tmax, Tmin and Maximum Solar Radiation.

	Variables					
Coefficient	**Tmin**	**CV Tmin**	**Tmax**	**CV Tmax**	**Maximum Solar Radiation**	**CV Maximum Solar radiation**
a	-0.00021256	-0.0002157	-0.004198219	-0.004198219	0.001407359	
b	0.040099132	0.04781353	0.183881824	0.183881824	0.082089376	
c	-0.024680654	-0.0248615	0.197427351	0.197427351	0.012738306	
d	-0.066643291	-0.0724497	0.21996234	0.21996234	0.118868951	
e	-0.054437733	-0.052836	-0.044864044	-0.044864044	0.056296371	
f	-0.02917172	-0.0324986	0.174404506	0.174404506	0.059426858	
g	-0.009772391	-0.0092802	-0.154391919	-0.154391919	0.031203184	
h	0.020230521	0.0218348	-0.000218911	-0.000218911	-0.050743081	
i	-0.018746947	-0.0202912	-0.198854536	-0.198854536	0.132836667	

The functions used are 4th and 6th order Fourier. The higher is the order more precise can be the function. The criteria for selecting a higher order is the relative complexity of the seasonal variation curve to be modeled.

4.2. Interpolation of Average for Tmin, Tmax and Solar Radiation

Aerial surface of Tmin, Tmax and Solar radiation were produced using elevation as unique variable in a first-order polynomial function (Maximum Solar radiation uses Tmax which is defined by a first-order polynomial of elevation). First-order regression was generated for Tmin and Tmax and Maximum Solar Radiation. Monthly of mean Tmin and Tmax and Maximum Solar Radiation of were plotted with the elevation for Station of North Kinangop, Mutobio, Kedong and Doodeman Farm. Additionally value of Solar Radiation was constrained to the interval of value observed at Doodeman farm. The maximum and the minimum of the average rainfall could not have lower or higher value than the interval of observed data.

The table 4.2.presents regression curve used with the R^2 expressing the performance of the regression to modeled the change of Tmin, Tmax and Maximum Solar Radiation over the area.

Table 4.3. Characteristics of regressions used for interpolation of Tmax, Tmin and Maximum Solar Radiation.

Variable dependant Y	**Variable independent X**	**R^2**	**Standard error**	**Significance of F**
Tmin	Elevation	0.689165	1.540429	0.000
CV Tmin	Elevation	0.005860749	0.136974072	0.626
Tmax	Elevation	0.953843	1.019076	0.000
CV Tmax	Elevation	0.422899	0.158815	0.004723
Maximum Solar Radiation	Tmax	-	-	-
CV Maximum Solar radiation	-	-	-	-

Mean value of Tmax and Tmin are well modeled by the regression. Results of the regression analysis and Fourier fitting for Coefficients of variation show a weak statistical meaning of those functions. Netherveless the purpose was to capture a general trend in time of Coefficients of variation not to modeled not established empirical models. The noise in time and in space of the Coefficients of variation is too high. The regression and Fourier of Coefficients of variation are just numerical adjustment.

4.3. Characteristics of Synthetic Series

4.3.1. Rainfall

For two years daily rainfall maps were generated. The results were grouped by average of annual rainfall. Yearly rainy day, daily rainfall intensity and standard deviation of daily rainfall are presented. Subsequently, records for the station of simulated rainfall were retrieved in the GIS environment for comparison with the simulated ones . Additionally realizations for point of interest were fulfilled. Those maps and points information were used in the performance assessment. Statistical test were applied only for points (specific locations) attributes. Daily rainfall intensity is measured by dividing the total amount of rainfall by yearly rainy day (number of rainy day events in a year).

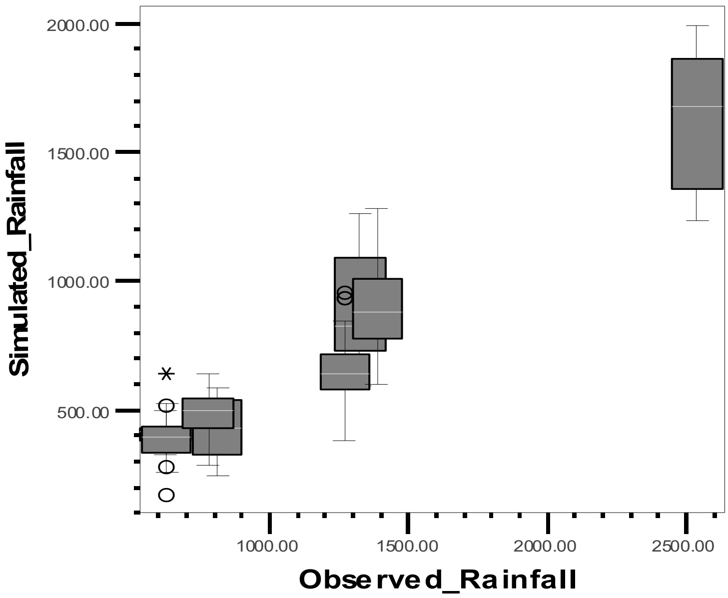


Figure 4.3. Box plot of Variation of simulated annual rainfall For raingauges' locations

The figure 4.3. presents the variation of simulated annual rainfall in a period of 19 years for the raingauges' location. The realizations shows that annual rainfall is underestimated by the Weather generator. The error increases gradually from the dry areas to wet areas. This underestimation is due to diminution of rainy day events in the simulated data. Underestimation of rainy days events is not the consequence of the simulation of rainfall occurrence events but the effect of the skewed distribution function used by WXGEN. The function creates negative value for given ranges of input parameter. In the workflow developed to adapt WXGEN in a GIS environment, those negative values are excluded and replaced by no rainfall events (0mm of rainfall for the day).

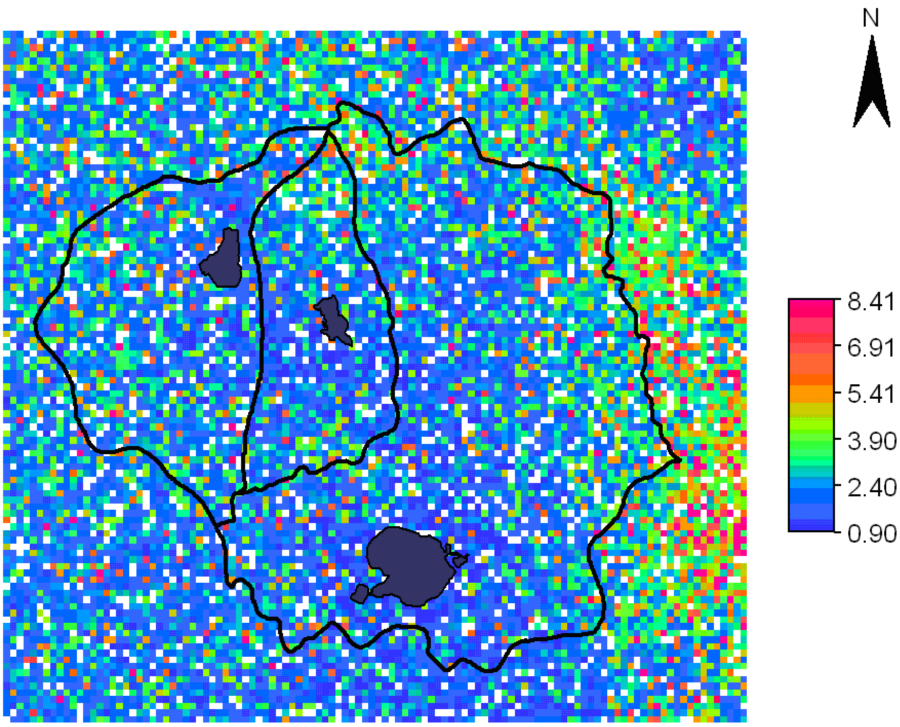


Figure 4.4. Map of Average of daily Rainfall for two years (year7 and year8)

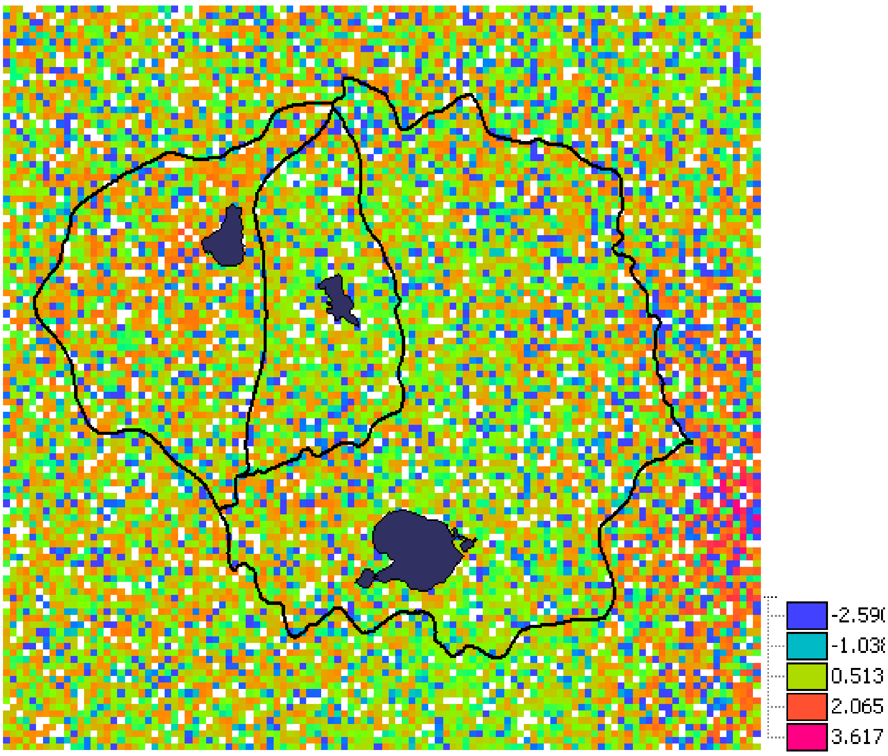


Figure 4.5. Difference Map: Average of daily Rainfall for two years minus the Interpolation of Observed Average Daily rainfall (year7 and year8).

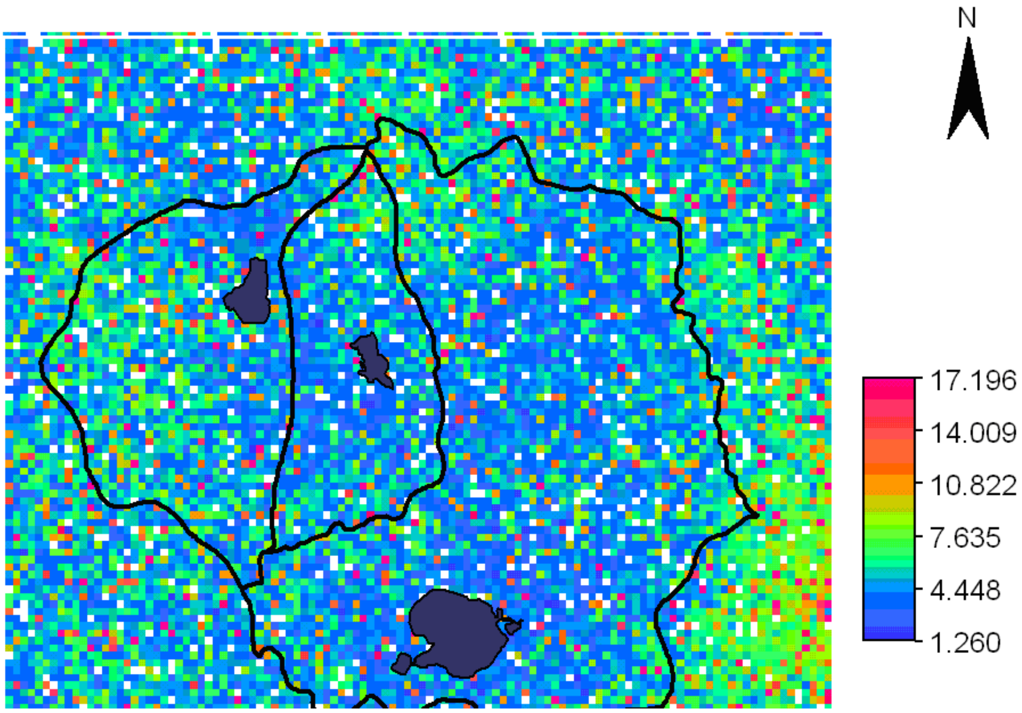


Figure 4.6. Map of Standard deviation of daily Rainfall for two years (year7 and year8)

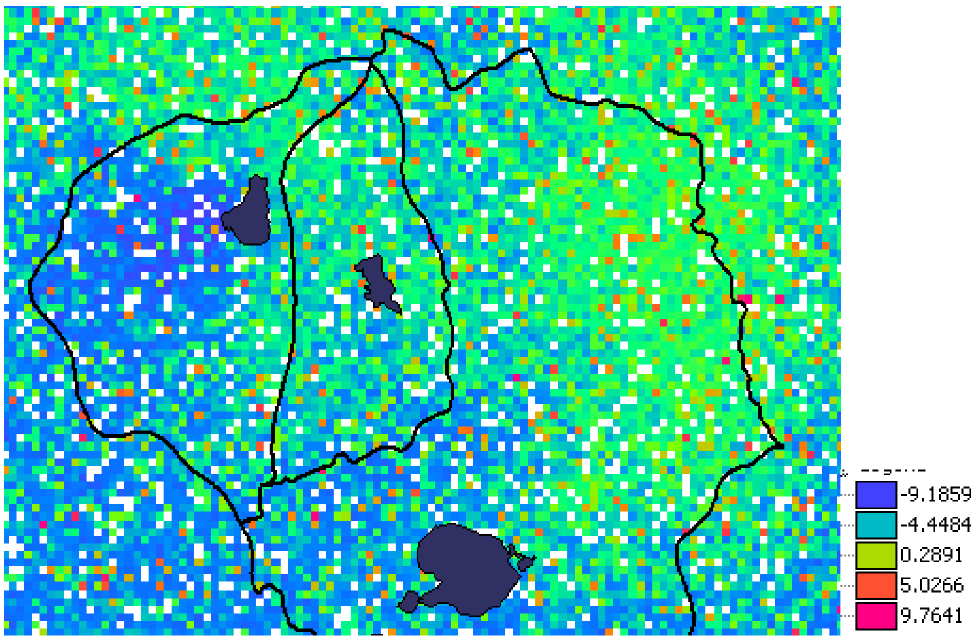


Figure 4.7. Difference Map: Standard deviation of daily Rainfall for two years minus Average interpolated Average Standard deviation of daily Rainfall

The Figure 3.6. presents the map of standard deviation observed on the simulated annual rainfall in a period of 2 years. The difference map represents a result of a map Algebra subtracting observed values to Simulated. Mean Annual rainfall used as the observed value is retrieved from the average of observed daily rainfall interpolated by krigging. Values presented suggest similar spatial variation compared to observed . The generator tends to overestimate rainfall for the highest area of the Lake Naivasha catchment. Those points are not monitored by a weather station, and the errors assumed for different interpolation at those locations are important. The difference maps with the observed average shows a gradual decrease of the difference and abrupt change of the signed (the difference become positive showing increase of the simulated values compared the observed) for the highest location.

The Figure3.6 and 3.7.present maps of the average standard deviation of simulated daily rainfall in for 10 years for the raingauges' location and the difference map between the average annual standard deviation of simulated daily rainfall for 19 years and the observed ones (the average annual standard deviation of observed daily rainfall was taken from interpolated map used in weather generator). The values obtained are acceptable. The difference with observed standard deviation is for the area of interest below 10% of the probability of rainy day. Such difference was observed for example between two weather generator using different markov chain order.(Katz & Parlange, 1998)

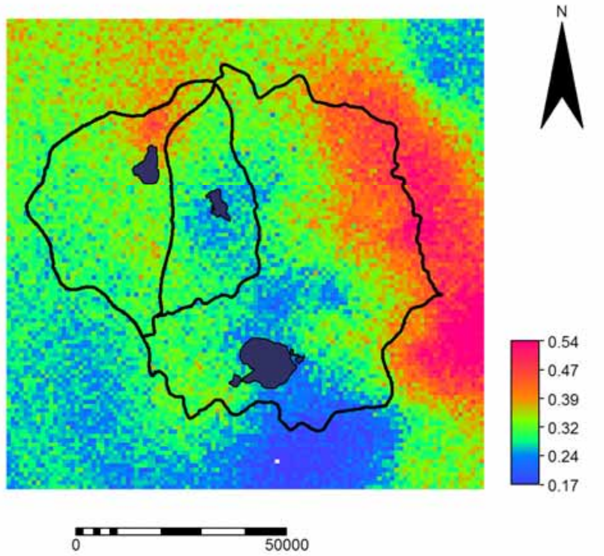


Figure 4.8. Map of rainy day's probability based on two years average (year7 and year8)

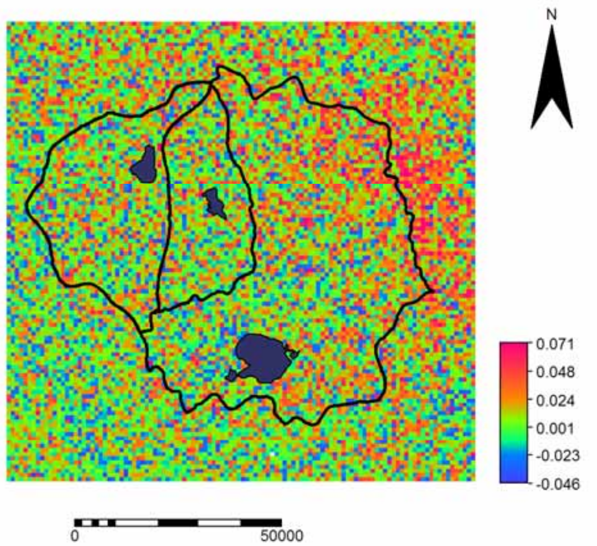


Figure 4.9. Difference Map: rainy day's probability average of two years minus the interpolation map of average rainy days (year7 and year 8)

The Figure4.9 presents the variation of simulated number of rainy day days in a period of 19 years for the raingauges' location. The variation shows a good performance of the weather generator for the variable. The number of rainy day events simulated using the transition probability Pwd and Pww. It estimates theoretical occurrence of rainfall for a day. The amount of rainfall for the a rainy day is simulated using the skewed distribution function of WXGEN. This component of WXGEN can for the certain value compute a value close to 0, even negative (for value used in the area). This apparent weakness requires a deeper investigation which was not tackled in the definition of the research questions. Problems with WXGEN in some area (far from USA or Canada for example) was also pointed out by Wallis & Griffiths (1995). Including Fourier polynomial is part of effort to improve this weather generator. It is still relevant to assess this tool which is used as component in EPIC (Erosion Productivity-Impact Calculator) and SWAT (Soil and Water Assessment Tool)(Priya & Shibasaki, 2001; Schuol & Abbaspour, 2007).

The maps of Figure4.10. and 3.11.present the average of simulated probability of rainy days for a period of 2 years, and the difference map between probability of rainy days for a period of 2 years and the observed ones (average of observed number of rainy day events in year was taken from interpolated rainy days used in weather generator). No particular spatial trend can be observed on occurrence of rainy day. The performance is acceptable according to literature which validated similar range for USA. Detailed study are rare in this field for sub-Saharan Africa.(Katz & Parlange, 1998)

The created map has an important advantage The information of surrounding pixels, can give information on possible values which can be taken by a given pixel. For example in a window of 10*10 pixels of one realization map contains information on 100 possible values for the day of the year considered. The assumption is that neighbouring pixels has similar climatic condition. This randomness is transferred to annual statistics. Neighbouring pixel simulate annual statistics which can be considered as proxy.

4.3.2. Maximum Temperature

Realization of 2 years, generated daily Tmax maps were aggregated in average of daily maps, and standard deviation of daily value. Information retrieval for points was conducted for locations where weather stations were installed. Records of simulated Tmax were compared to the observed ones. Those maps and point information were used in the performance assessment.

The figures4.12.presents for point of interest, the variation of Tmax of daily values simulated for 2 years.

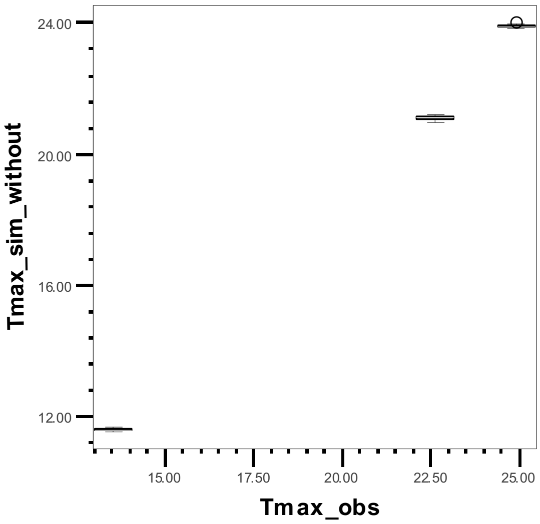


Figure 4.10. Variation of simulated Tmax for points of interest

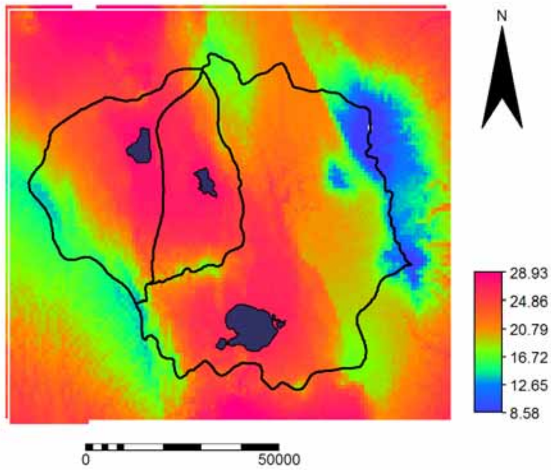


Figure 4. 11. Average map of Synthetic Maximum Temperature of two Years (degree celcius)

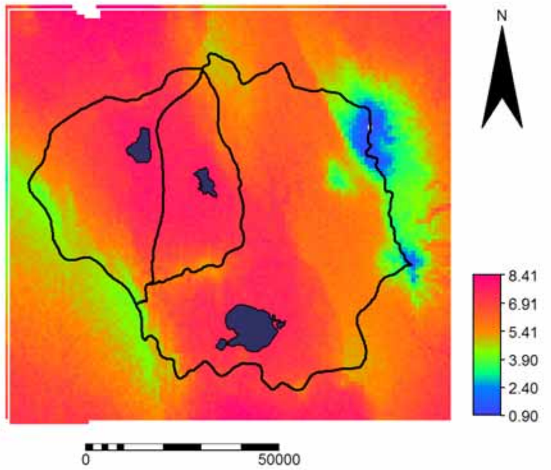


Figure 4.12. Difference map: Average Maximum Temperature interpolation minus Synthetic Maximum Temperature of two Years(degree celcius)

The deviation to average are govern by the effect of the residuals defined by Richardson (1984). The residuals dependant on the determination of three random numbers. The value of those random numbers are stated to be between 0 and 1 according to Richardson for WGEN (Parlange & Katz, 2000; Richardson & Wright, 1984). Wilks(1999b) suggested to determine the range of the random number experimentally. The procedure suggests fitting extreme values to maximum and minimum read from an inverted normal distribution of the studied variable.

The Figures 4.13. and 4.14. present maps of the average of Tmax simulated for 19 years location of interest and the difference maps between the average of Tmax simulated for 2 years and the average of observed Tmax and its Coefficient of Variation (used by the weather generator).

4.3.3. Minimum Temperature

Generated daily Tmin maps (2 years) were aggregated in average of daily maps, and standard deviation of daily value. Records of simulated of simulated Tmin were were compared to observation weather station . Maps and point information were used in the performance assessment.

The Figure 4.13 presents for point of interest, the variation of Tmin of daily values simulated for 19 years.

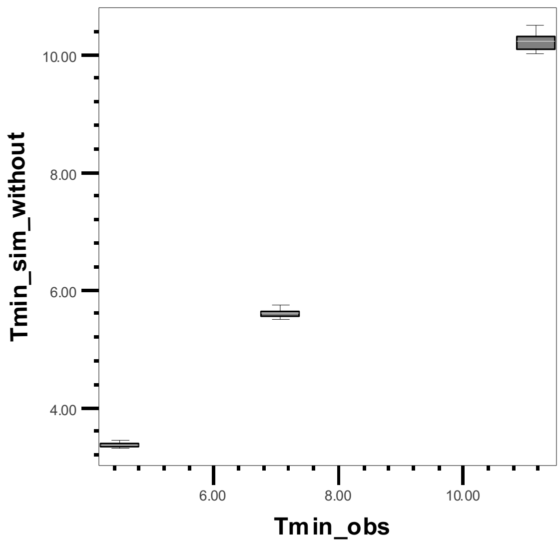


Figure 4.13. Variation of simulated Tmin for point of interest

The Figures 4.14. and 4.15. present maps of the average of Tmin . Results are based on 19 years of realizations for location of interest. The difference map between the average of Tmin and its Coefficient of Variation of daily values simulated for 2 years and the average of observed Tmin and its Coefficient of Variation (used by the weather generator). Tmin varies more than Tmax in daily basis and is more sensitive to rainy day or dry day occurrence, therefore the error caused by under representation of rainy day is more visible for this variable.

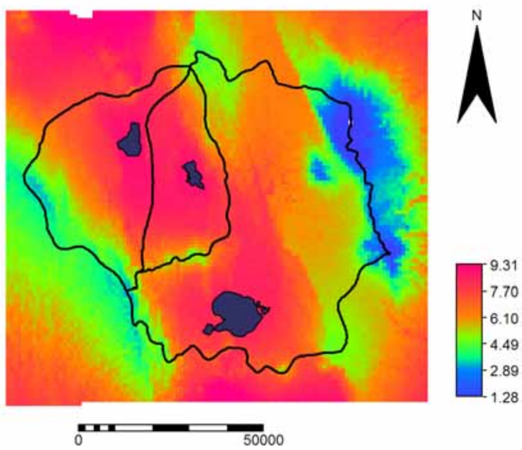


Figure 4. 14. Average map of Synthetic Minimum Temperature of two Years

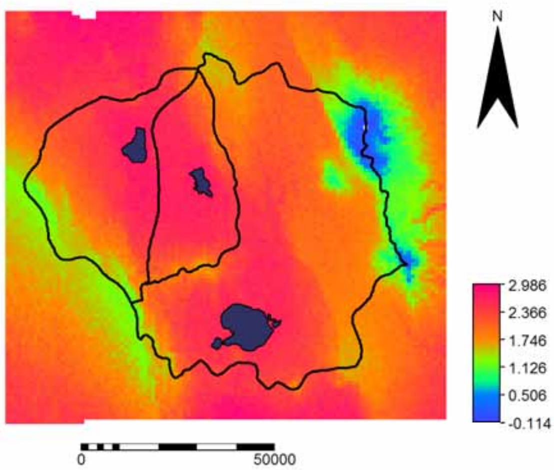


Figure 4.15. Difference map: Synthetic Minimum Temperature of two Years minus Average Maximum Temperature interpolation

The number of rainy day events affected the annual rainfall total. It governs at the certain extend the deviations to mean (by the correction coefficient). The underestimation of number of rainy days increases the error and the annual average of Tmin and Tmax.

Statistical assessment of performance of representing inter-annual or seasonal variation is difficult for Tmax and Tmin. The period of observation is short for available data. Nevertheless a comparison with the variation of grid of 0.5 ° (CRU TS 2.1 from the Climatic Research Unit mentioned in §3.3.2.) could give an insight on the performance of the weather generator.

The first order regression as defined in this procedure used to generate inputs for the Weather generator, and changed the value of annual average at the location of weather stations. A solution to this error is to use a higher order of regression, to increase the number of variables used by the regression or use more detailed or sensitive methods. Many of those methods require more observation point and preferably for a long periods. Nevertheless, a mathematical solution should not contradict physical relationship as the raise of temperature with altitude. Thus including more observation points and validate the interpolation (Latitude and longitude for example) should be suggested to improve the Weather Generator (Richardson & Wright, 1984; Daniel S. Wilks, 2008) .

4.4. Findings

The method for extending Rainfall to trivariate regression developed for WGEN has been applied to reproduce the correlation and cross-correlation among Maximum temperature, Minimum Temperature an Maximum Solar Radiation . The inputs for non-precipitation variables are presented. 6 variables are required: mean and Coefficient of variation Maximum temperature, Minimum Temperature and Maximum Solar Radiation. The interpolations of those variables are better for mean of Maximum temperature, Minimum Temperature and Maximum Solar Radiation compared to Coefficient of variation of Maximum temperature and Minimum Temperature.

The seasonal variations of Maximum temperature and Minimum Temperature were modeled by Fourier series. Long term Regional climate database developed by The Climatic Research Unit were used to determine seasonal variation of Maximum temperature and Minimum Temperature.

The quality of the data generated was constrained by the quality of the interpolation methods used. Averages of Maximum temperature, Minimum Temperature and Maximum Solar Radiation are based on short observation periods. They were developed mainly to complete the weather generator with complete set of climatic variables required by the crop model used.

The precipitation component of the weather generator uses the procedure developed by Bhandari. It is able to represent successfully the probability of rainy days. It represents the amount of rainfall with an increasing absolute error toward lower or higher altitudes. At locations with average altitude the error on precipitation is lower. Theoretically, including a iterative process should successfully correct the total annual rainfall.

5. Distributed Crop Growth modeling

5.1. Adjustment of the Crop model

5.1.1. Determination of starting date and length of the simulation

The Figure 5.1 presents the crop calendar observed in the Agro ecological zone UH2, according to the Farm Management Handbook survey (Jaetzold, Schmidt, Hornetz, & Shisanya, 2006). The survey shows a growing period length for maize is between 182 to 252 days. The simplified crop model simulated for the emergence and these periods should be deducted. The starting date and the length of the simulation are critical for the model. From the first day of the simulation the development stage, the water requirement is simulated.

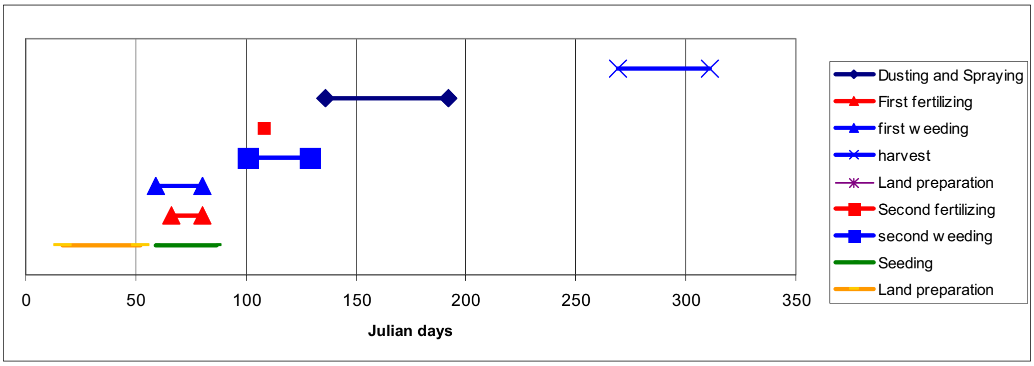


Figure 5.1. Calendar of farm Operations in the visited area

The interview of farmers visited in different focal area shows that the growing periods for maize varies between 183 to 334 days. The values presented are taken from average of date estimated by farmers. As shown by the figure 5.2. , even at sub-location level the date planting date and the date of harvest vary sensibly.

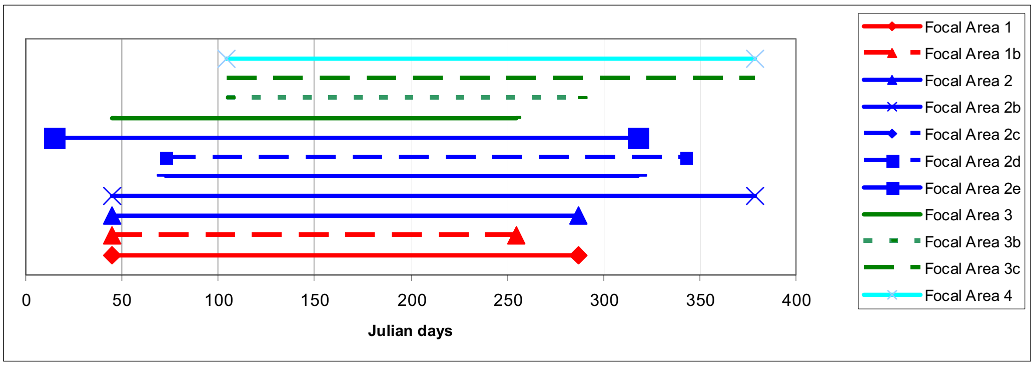


Figure 5.2. Duration length of maize in field from interview of farmers

From Figure 5.1 and 5.2. , a date of 75 julian day was selected as average date representing a large area in the upper part of the Naivasha Catchment. The simulation had to run at least for 185 days. The minimum periods of running time was used for the results presented in the following sections.

5.1.2. Comparison with the Variation of rainfall in a year

A simple comparison of cumulative curves of rainfall was made to assess the similarity with the growing periods determined by farmers. The weather generator produce series which recreate wet periods and dry periods.

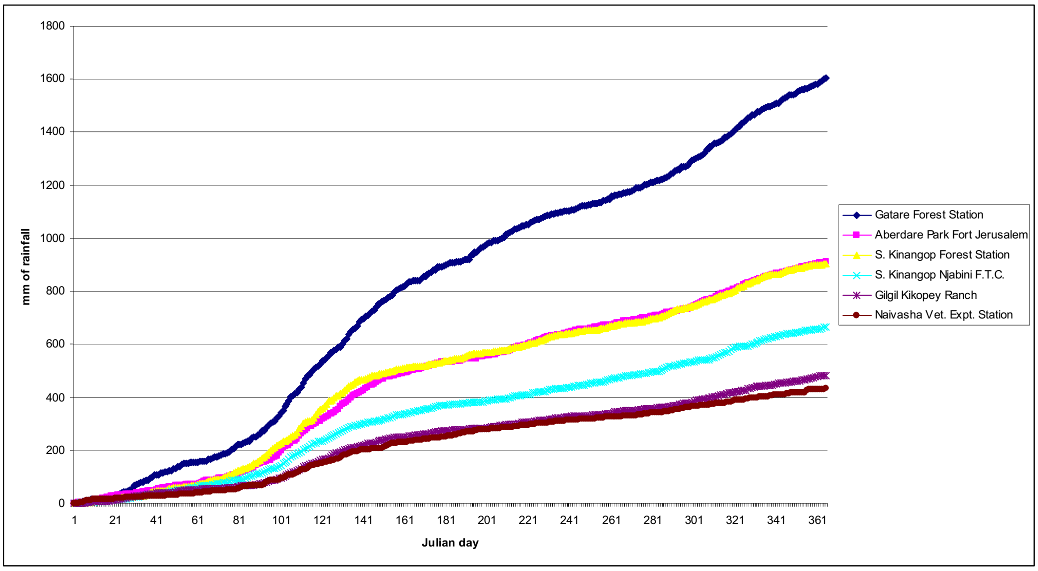


Figure 5.3. Time series of simulated Average Cumulative rainfall (19 years of simulated data)

The pattern of the cumulative curve shows an change between day 70 and 120 for most of the location represented.

5.2. Tuning crop parameters

The crop parameter dependant and non dependant to Crop development stage were fine-tuned to reproduce the growing periods and the best average yield. The latest recorded with a good rainfall pattern was selected as climatic datasets for the calibration. The most sensitive crop parameters for maize were fine tuned. The lengths of growing stage were checked using the Development stage output variables of the model. It varies from 0 to 2 (at 2 the crop is at harvesting time). The meanings of important parameters were introduced in the section 3.4.2.

Table 5.1.Parameters used for adjustment and some reference values.

	Symbol chosen for crop parameter			For Zimbabwe(PSn adapted by Venus,2002)	For Botswana (Mweso,2003)
	PSn	WOFOST	PCRaster WOFOST		
m^2/kg	SLA	SLATB	SLA	14-21	14-35
℃	TO℃	TBASE	TDbase	10	10
℃d	Tleaf	TSUM1	Te_ant	1000	1000
℃d	Tsum	TSUM2	Te_mat	1500	1600

For tropical area the often mentioned calibration reference is the work of Roetter(1993, pp. 36-38) in Kenya, who applied the curve fitting approach to adapt with reasonable limits weak o unknown parameters or relations. Essentially Crop varieties used in the area have not datasets to be used as reference, but from literature crop parameters data as to be developed and assessed. For the variety 613 (close to the variety 614 used in the upper part of the Study area) condition of Kitale it is mentioned that TDbase should be between 0-8°C and TDmax2 inferior to 30°C. Trials at 2000m at Muguga (closest agricultural station of KARI-Kenya Agricultural Research Institute) has shown that with mean air temperature of 16.5, 18.3, and 23°C, temperature sums were 825, 810 and 846 d°C. This value was pre-validated by observed LAI (Roetter, 1993; Vinocur & Ritchie, 2001).

Different settings of parameter were combined step by step. Discrete values of parameter were selected in trial and error process. All combinations of chosen settings were compared and the best solution selected. The Figure 5.2. shows the value selected for Specific Leaf Area (SLA) and for partitioning factor of plants organs above ground. The first was is a default set provided with the computer code the second is adapted from partitioning factor used for Muguga as explained by Roetter(1993).

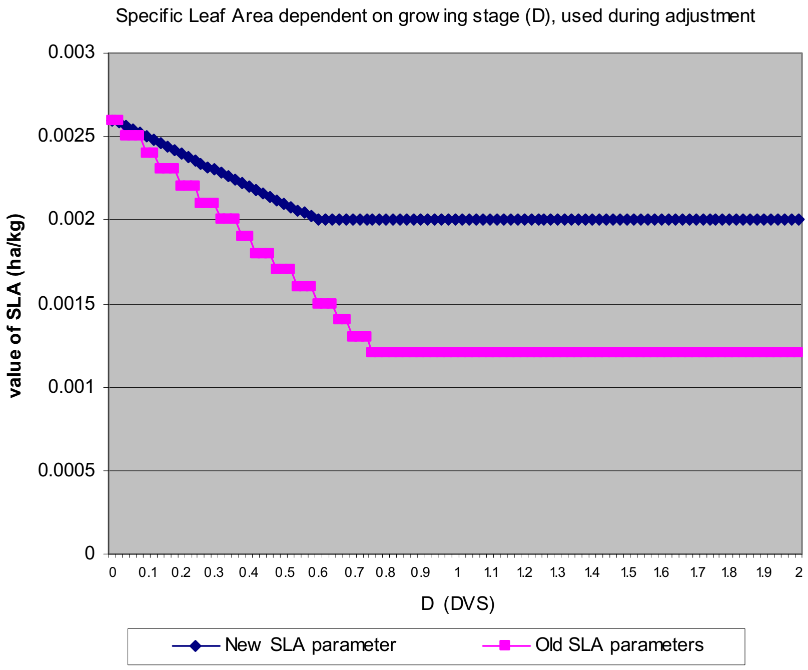


Figure 5.4. Settings of Specific Leaf Area used for the adjustment of the crop model

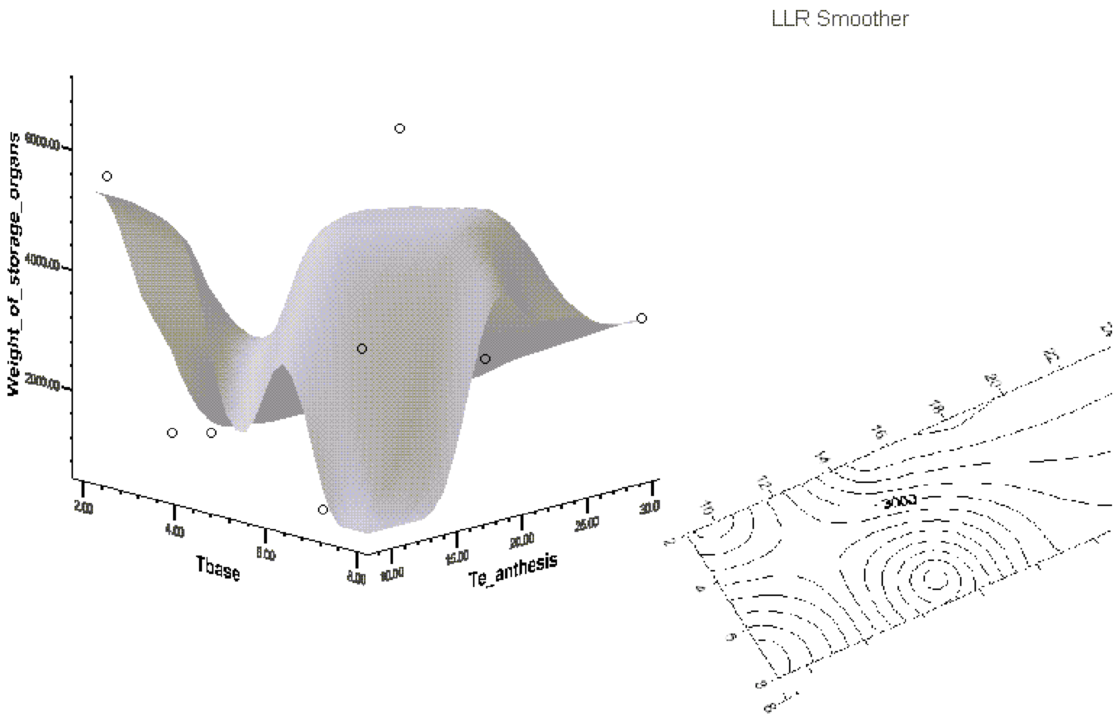


Figure 5.5. Representation of variation of Yield as function of variable Tbase and Te_anthesis simulated with data of Doodeman farm.

In annexes two different partitioning factors tested during the adjustment of the crop parameters are presented. The first was is a default set provided with the computer code the second is adapted from partitioning factor used for Kitale.

5.2.1. Optimization curves

To simplify the visualization of the adjustment process and the interpretation of results, optimization curves were plotted. The x-axis of the Figure 5.5. represents different levels of Tbase , the y-axis of the Figure a scale of Te_anthesis and the z-axis shows the yield observed for each x and y values.

For the case represented by the Figure 4.5.the value is Te_anthesis fix and equal to 35.

The same procedure was repeated for each settings of Specific Leaf Area (SLA), and each partitioning factor settings. The new Specific Leaf Area (adapted for calibration of Muguga) was better results. The setting of partitioning factor used for Kitale was also preferred to the default. The example of optimization curves used the selected settings of partitioning factor and Specific Leaf Area.

After visual interpretation of curve the value of non-dependant crop parameters has been set to

Tbase = 8,

Te_anthesis=18,

Te_anthesis=30,

The yield are results of simulations under climatic conditions of Doodeman farm (latitude:, longitude:).

5.3. Simulated Yield

Distributed simulation used the grid weather data for Rainfall, Tmax, Tmin, Solar Radiation and Relative humidity. Lumped data (not spatially distributed) applied for wind speed. The crop growth simulation is specific for each grid. For each grid cell soil, land use, a unique weather condition can be applied.

A successful weather generator or distributed weather datasets could be improved the understanding of agricultural condition in the area. Nevertheless the use of distributed model suppose a definition of optimum resolution for economic reason (computations and time required by the simulation) and for the nature of the study (higher is space scale, less is the need for resolution).

The simulation has produced a large number of quasi crop failures for the upper catchment. Different reason can explain this quasi failure situation. One reason could be a particular the lower temperature.

An other reason the solar radiation which was weighted according to observations of Nakuru and doodeman farm. The regression reduces gradually Maximum Solar Radiation with the elevation.

The results of each simulation was investigated. An overlay of Simulated Yield , reliable rainfall, and Elevation was prepared. And matrix from pixel values constructed for analysis.

5.4. Comparison between observed and simulated yield

The comparison of observed and simulated was not direct read from map of simulated yield. It used regression built with a map yield. The procedure can be synthetized in three main part:

i. Establish correlations of the simulated yield with rainfall and Elevation.
ii. Estimate Statistically performance of each correlation
iii. Compare the best predictor with an alternative method.

5.4.1. Information Retrieval from interviews of farmers

Interviews of 30 farmers were in 4 different areas were grouped. The areas selected by the Agricultural Officer are used for the baseline survey of the Service of Ministry of Agriculture at District and Division level. The district agricultural office has divided areas to according to differences and agro-climatic conditions et population density.

A map of yield information retrieved from farmers' interviews is presented in annex. The map presented three: the average yield, the maximum yield and the yield gap of the worst record.

5.4.2. Information Retrieval from simulated Maize Yield map

A yield surface with harvest and crop failure was obtained earlier. To be able to discussed point information from the interview of farmers in different sub location, two approaches can be taken:

The first is to read the simulated yield maps within the boundary of yield map. The second is operation create a matrix with layers of interest by crossing operation in GIS environment. The layers selected are Reliable rainfall, and Elevation.

5.4.3. Regression Analysis between Yield, Reliable Rainfall and Topography

A relationship of Maize Yield, as simulated following the steps explained in the previous section, with reliable rainfall and Topography was analysed. Three regressions were investigated:

- Simulated Maize Yield explained by reliable Rainfall

- Simulated Maize Yield explained by Elevation

- Simulated Maize Yield explained by reliable Rainfall and Elevation

The Figures 5.6., 5.7 and 5.8present the three regressions. In Figure 5.6., the x-axis represents different levels Rainfall and the y-axis Maize Yield. The x-axis of the Figure 5.7.represents different altitudes, the y-axis Maize Yield. The x-axis of the Figure 5.8. represents reliable rainfall and the y-axis is elevation. Maize Yield is represented by the z-axis.

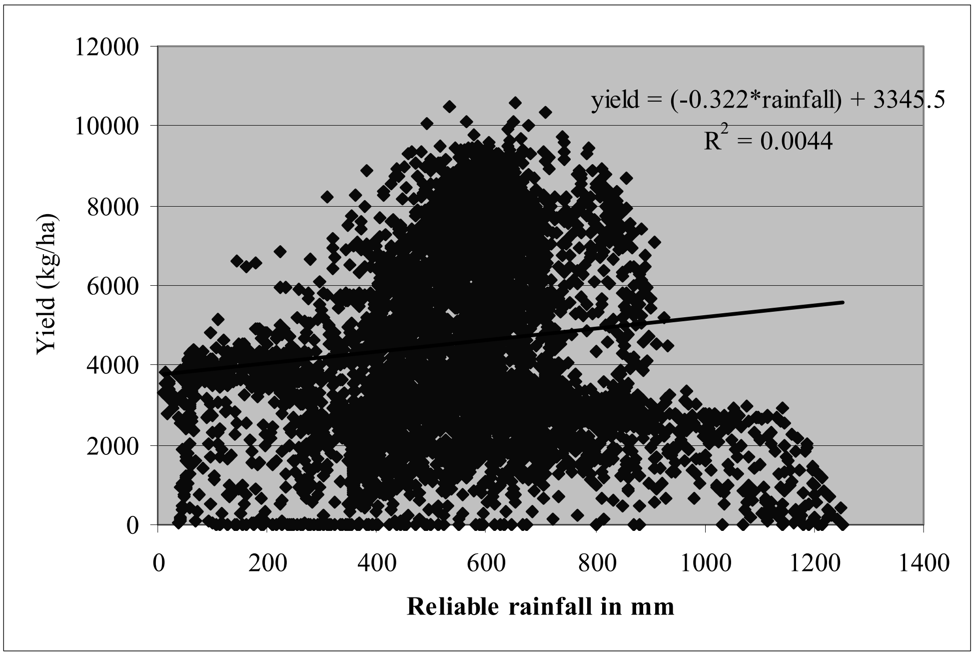


Figure 5.6. Regression of Simulated yield as function of Elevation and reliable rainfall

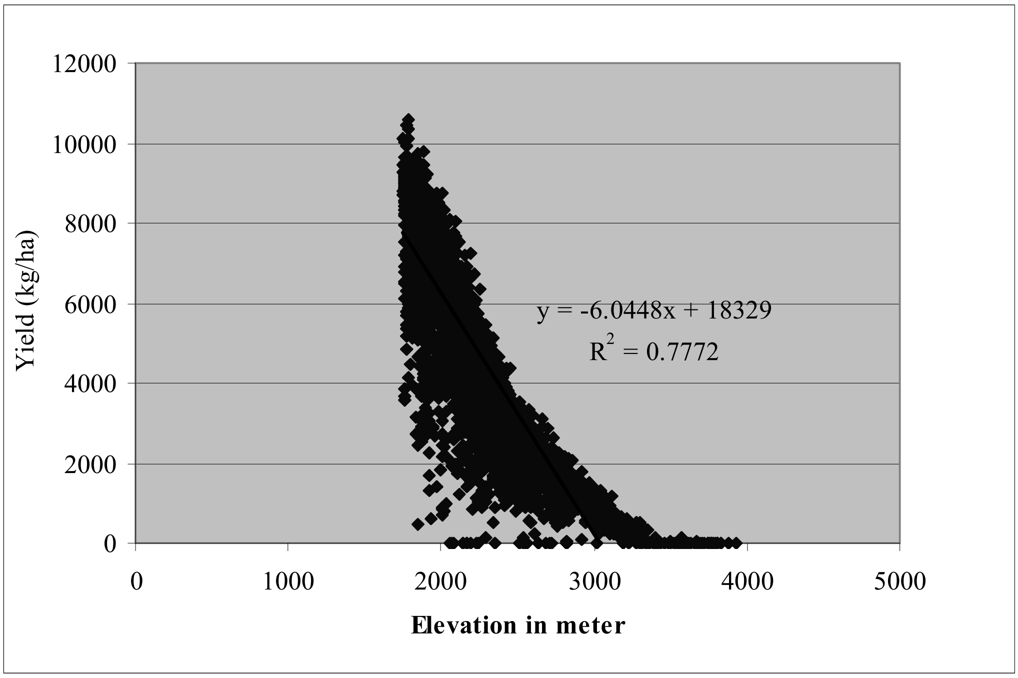


Figure 5.7. Regression of Simulated yield as function of Elevation and reliable rainfall

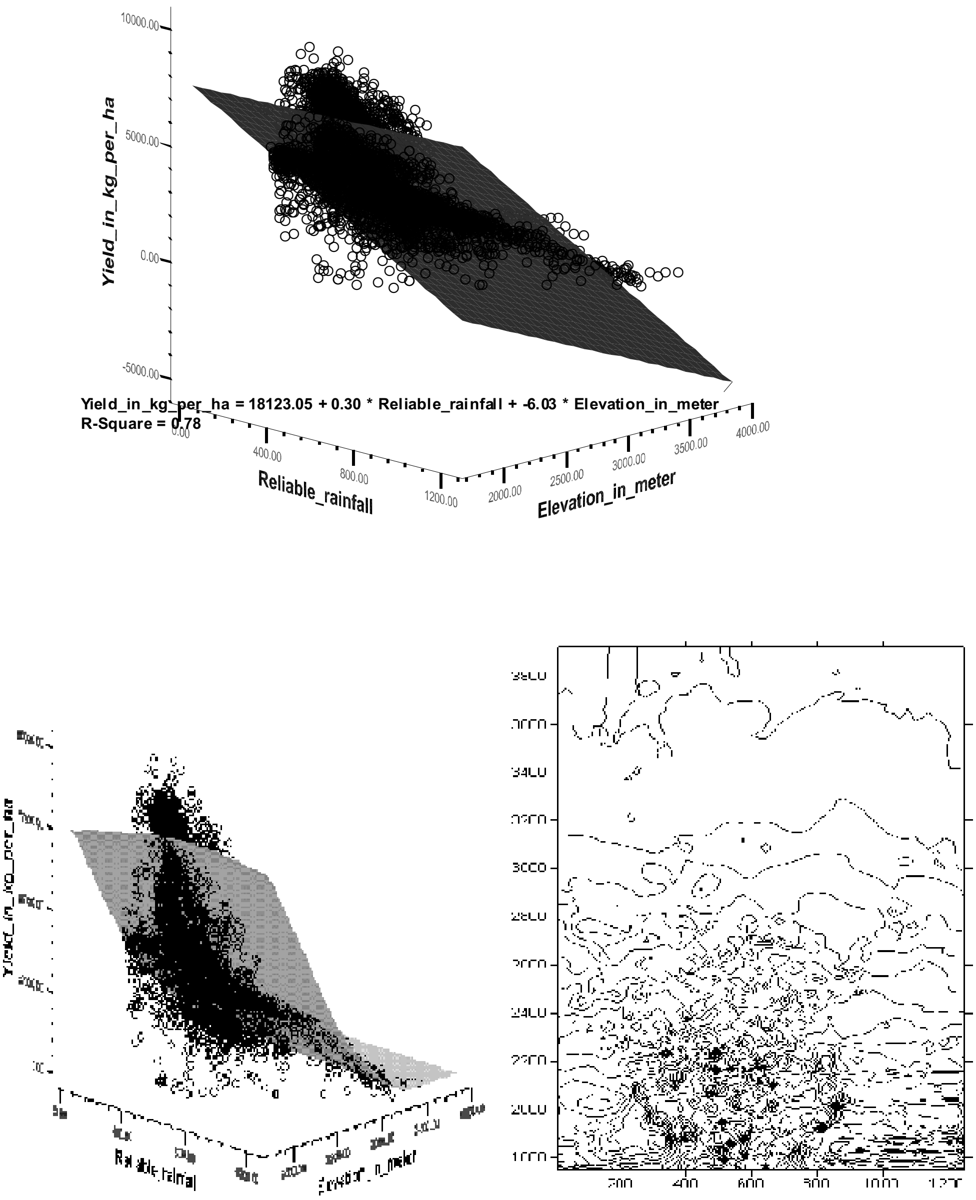


Figure 5.8. Regression of Simulated yield as function of Elevation and reliable rainfall fitted with a regression and a smoother

The regression shows a better agreement with two explicative variables instead of one used for options 1 and 2.

5.4.4. Comparison of the regression models

Two regression methods where compared to observed data. The first explains Yield by reliable Rainfall and Elevation and was developed using the distributed simulation presented in the previous sections. The second is the regression of Maize yield with reliable rainfall presented by Roetter(1993). The second regression used a comprehensible datasets of experimental data and survey in the former district of Muranga (in the central province), neighboring location to the study Area. The predictor of Roetter(1993) is a regression which computes the Ratio between potential yields and the water limited yield. The potential yield was estimate at 11Tonnes for the Area.

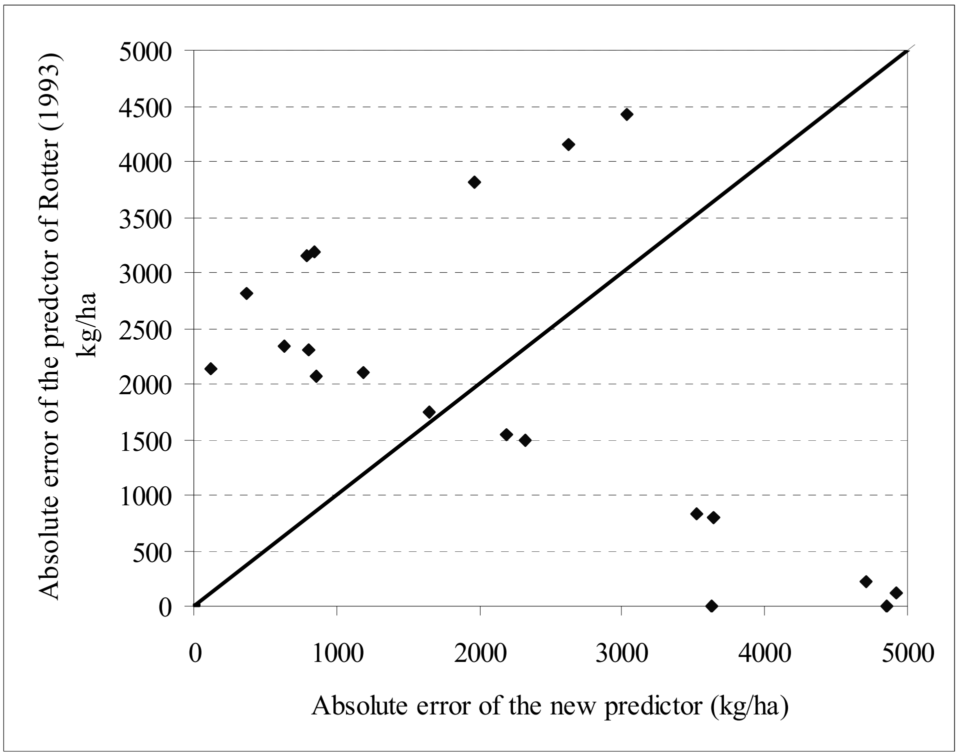


Figure 5.9. Scatter plot of the deviation to observed data of the selected regression vs regression established by Roetter(1993)

The Figure 5.9. presents the comparison of the deviation to observation by both regression. In the x-axis represent the absolute deviation to observed data by the first regression (first explaining Yield by reliable Rainfall and Elevation) and the y-axis the absolute of the second regression (by Roetter(1993)). The absolute errors are computed using the difference to mean yield. Each point represents a farm visited. The regression of Roetter(1993) gives a the ratio between the potential yield and water limited yield.

The graphs shows in more than 53 % of case the new regression is better in predicting the average yield while in less than 47 % of the case the regression suggested by Roetter(1993) was better. The assessment has to be continued, to compare observed Ratio of potential yield vs water limited yield. More important should be the effort of improving the weather generator, which can empower the understanding of spatial variation of Maize yield.

Findings

The simulation presents large number of crop failures events using the generated weather data for relatively cold area (temperature below 6 degree celcius). There is statistical evidence on the importance including elevation for yield estimation. The simulated yield is sensitive to Elevation. The sensitivity results from the combine effects of Weather generator parameters interpolated using Elevation. The best regression has used simulated data, precipitation and Elevation. It has found to be a better predictor of yield for the study area compared to regression developed by Roetter(1993).

The results of the chapter 4 answer the second research question positively. The weather generator implemented by Bhandari extended to Temperature (Maximum and Minimum), to Solar radiation can predict productive area. It can also be used to investigate the average planting periods. The determination of length of the growing periods requires a lot of information. The calibration made should be improved by field observations.

The third research asks if the methods can be out-scaled or up-scaled. Although for the weather generator can recreate statistics of observed (annual statistics mainly). The linkage with the crop model requires improvement on understanding the reason which lead to crop failure. If the crop failures are better explained, the adapted weather generator could have a level applicability and utility similar to non-distributed ones.

6. Conclusion

The availability of reliable spatial climatic data is important for the planning of agriculture, natural resources and environmental management. Especially in developing countries the spatial and temporal coverage of meteorological data has many gaps.

This thesis shows the usefulness of a spatially distributed weather generator (WXGEN) to realize meteorological time series used as an input to a crop growth model (WOFOST) implemented using PCRaster.

The approach to spatially and temporally interpolate the input parameters for the rain generator could also be applied for the input parameters for the generators of maximum and minimum temperature and solar radiation. The relative humidity was reliably estimated based on regression with the precipitation, and minimum and maximum temperature.

The statistical structure of the generated meteorological daily data in a year is well reproduced.

The lack of reliable meteorological data from a high elevation station may produce a considerable error in the estimation of the parameters for areas over 2800 m of altitude.

In the current implementation of WXGEN under ILWIS the long term variations in weather pattern (series of drought/wet years) is not well reproduced.

The optimum crop yields could be estimated fitting well the yields predicted by the Farm management handbook of Kenya, and field data collected by interviewing agricultural officers and farmers. However, the yield is also influenced by management factors (soil preparation, fertilizers, crop-calendar, weeding etc) not captured by the water limited simulation using WOFOST.

An advantage of using a spatial (raster) approach is that a window of pixels also shows the variability of yields. The yield estimates of a 10*10 window is a proxy of 100 years of point simulations.

For simulated maize variety a linear regression between elevation and mean and variance of the yield could be established. The relation is an improvement over methods commonly used in Kenya. Therefore, it allows better agricultural planning without going into the complexity of generating meteorological time series and crop modeling. However, it is not known whether this relation holds true outside the research area.

No formal sensitivity analysis was carried, but based on experience of the model calibration and field evidence; temperature seems to be more a limiting factor than water availability.

Considering the well known attitude that small rural farmer in Africa are inclined to reduce risk instead of maximizing long term yield, the risk of crop failure is a critical factor, and can be analyzed by the suite WXGEN- WOFOST.

Recommendation

It would be interesting to explore the usefulness of the spatial implementation of WXGEN in combination with hydrological models. For long term planning the multi year variability should be integrated. Here the method already developed by Lal Muthawatta could be used.

Since the series are generated for every individual pixel no spatial correlation exists for short periods (days). A next step in the weather is to build in the short duration spatial dependency, resulting in daily rainfall fields.

A very promising field of research is combine stochastic and earth observation approaches. The idea to improve the estimate interpolated WXGEN input statistics using the information from MSG time series. For example, the probability of a rainy day is 50% and MSG images show only clouds in 25%, obviously our estimate of 50% is wrong.

Good quality ground observations are equally important to improve the quality of the realizations.

The occurrence of dry spells is critical in many processes. The statistical properties of this should be investigated and examined how well reproduced by WXGEN.

It could be investigated whether the estimate of relative humidity can be approved using saturated and unsaturated vapor prerssure instead of the simple empirical relation used in this study.

Also the estimate of wind speed, a constant in the current study, should be deepened.

To better calibrate and /or validate the crop model better crops statistics are necessary. Here a simple GPS used by the extension officers could already greatly improve data quality.

Similar to the use of MSG time series to improve calibration of the weather generator, the combined model (generator & crop simulation) could be improved by confronting the output (dry matter production) with NDVI time series from eg. MSG and SPOT vegetation.

WOFOST calculates the water balance components such as soil moisture contend and actual evaporation and deep percolation. Combining this approach with hydrological models and satellite derived energy balance approaches seem obvious.

7. Bibliography

Acock, B., & Acock, M. C. (1991). Potential for Using Long-Term Field Research Data to Develop and Validate Crop Simulators. *Agron J, 83*(1), 56-61.

Al-Jamal, M. S., Sammis, T. W., Ball, S., & Smeal, D. (2000). Computing the crop water production function for onion. *Agricultural Water Management, 46*(1), 29-41.

Bakhsh, A., Kanwar, R. S., & Malone, R. W. (2007). Role of landscape and hydrologic attributes in developing and interpreting yield clusters. *Geoderma, 140*(3), 235-246.

Bastiaanssen, W. G. M., Dam, J. C. v., & Malik, R. S. (2003). Introduction. *Water Productivity of Irrigated Crops in Sirsa District, India. Integration of Remote Sensing, Crop and Soil Models and Geographical Information Systems*, 11-20.

Batjes, N., & Gicheru, P. (2004). Soil data derived from SOTER for studies of carbon stocks and change in Kenya (ver. 1.0; GEFSOC Project). (Report 2004/01, ISRIC - World Soil Information, Wageningen), 31.

Bhandari, A. K. (2005). *A study to assess accurate spatial rainfall data in lake Naivasha basin, Kenya.* ITC, Enschede.

Brumbelow, K., & Georgakakos, A. (2007). Determining crop-water production functions using yield-irrigation gradient algorithms. *Agricultural Water Management, 87*(2), 151-161.

Dobesch, H. e., Dumolard, P. e., & Dyras, I. e. (2007). *Spatial interpolation for climate data : the use of GIS in climatology and meteorology*: ISTE.

Dourado-Neto, D., Teruel, D. A., Reichardt, K., Nielsen, D. R., Frizzone, J. A., & Bacchi, O. O. S. (1998). Principles of crop modeling and simulation: I. uses of mathematical models in agricultural science. *Scientia Agricola, 55*, 46-50.

Droogers, P., Seckler, D., & Makin, I. (2001). Estimating the potential of rain-fed agriculture. *Working Paper, 20*(IWMI, Colombo, Sri Lanka).

FAO. (1992). Maize in human nutrition. *FAO Food and Nutrition Series, 25*(FAO, Rome (Italy)
).

FAO. (1996). AGRO-ECOLOGICAL ZONING Guidelines. *FAO Soils Bulletin 73*(Rome).

FAO. (2002). Deficit irrigation practices *Water Report, 22*(Rome).

Gregory, J. M., Wigley, T. M. L., & Jones, P. D. (1993). Application of Markov-Models to Area-Average Daily Precipitation Series and Interannual Variability in Seasonal Totals. *Climate Dynamics, 8*(6), 299-310.

Guenni, L., & Hutchinson, M. F. (1998). Spatial interpolation of the parameters of a rainfall model from ground-based data. *Journal of Hydrology, 212-213*, 335-347.

Hammer, G. L., Hansen, J. W., Phillips, J. G., Mjelde, J. W., Hill, H., Love, A., et al. (2001). Advances in application of climate prediction in agriculture. *Agricultural Systems, 70*(2-3), 515-553.

Hansen, J. W., & Indeje, M. (2004). Linking dynamic seasonal climate forecasts with crop simulation for maize yield prediction in semi-arid Kenya. *Agricultural and Forest Meteorology, 125*(1-2), 143-157.

Hartkamp, A. D., White, J. W., & Aguilar, A. R. (2000). Maize Production Environments Revisited: A GIS-based Approach. (Mexico, D.F.: CIMMYT.).

Hartkamp, A. D., White, J. W., & Hoogenboom, G. (2003). Comparison of three weather generators for crop modeling: a case study for subtropical environments. *Agricultural Systems, 76*(2), 539-560.

Hayhoe, H. N. (1998). Relationship between weather variables in observed and WXGEN generated data series. *Agricultural and Forest Meteorology, 90*(3), 203-214.

Jaetzold, R., Schmidt, H., Hornetz, B., & Shisanya, C. (2006). Farm management handbook of Kenya *2B: Central Kenya, Subpart B2 Central Province*(2nd Edition), 491.

Katz, R. W., & Parlange, M. B. (1998). Overdispersion phenomenon in stochastic modeling of precipitation. *Journal of Climate, 11*(4), 591-601.

Kipkorir, E. C., Raes, D., & Massawe, B. (2002). Seasonal water production functions and yield response factors for maize and onion in Perkerra, Kenya. *Agricultural Water Management, 56*(3), 229-240.

Mavromatis, T., & Hansen, J. W. (2001). Interannual variability characteristics and simulated crop response of four stochastic weather generators. *Agricultural and Forest Meteorology, 109*(4), 283-296.

McGregor, G. R., & Nieuwolt, S. (1997). *Tropical climatology : an introduction to the climates of low latitudes* (Second edition ed.). Chichester etc.: Wiley & Sons.

Merwe, N. J. v. d., Tschauner, H., Rowan, F. S., & Russell, K. M. (1999). C4 Plants and the Development of Human Societies. In *C4 Plant Biology* (pp. 509-549). San Diego: Academic Press.

Monsi, M., Uchijima, Z., & Oikawa, T. (1973). Structure of Foliage Canopies and Photosynthesis. *Annual Review of Ecology and Systematics, 4*, 301-327.

Muthuwatta, L. P. (2004). *Long term rainfall - runoff : lakelevel modeling of the lake Naivasha basin, Kenya.* ITC, Enschede.

Olufayo, A. A., Stigter, C. J., & Baldy, C. (1998). On needs and deeds in agrometeorology in tropical Africa. *Agricultural and Forest Meteorology, 92*(4), 227-240.

Parlange, M. B., & Katz, R. W. (2000). An Extended Version of the Richardson Model for Simulating Daily Weather Variables. *Journal of Applied Meteorology, 39*(5), 610-622.

Podesta, G., Letson, D., Messina, C., Royce, F., Ferreyra, R. A., Jones, J., et al. (2002). Use of ENSO-related climate information in agricultural decision making in Argentina: a pilot experience. *Agricultural Systems, 74*(3), 371-392.

Priya, S., & Shibasaki, R. (2001). National spatial crop yield simulation using GIS-based crop production model. *Ecological Modeling, 136*(2-3), 113-129.

Richardson, C. W., & Wright, D. A. (1984). WGEN: A Model for Generating Daily Weather Variables. (United States Department of Agriculture, Agriculture Research Service, ARS-8), 86.

Roetter, R. (1993). Simulation of the biophysical limitation to maize production under rainfed conditions in Kenya : evaluation and application of the model WOFOST. 261 p.

Schuol, J., & Abbaspour, K. C. (2007). Using monthly weather statistics to generate daily data in a SWAT model application to West Africa. *Ecological Modeling, 201*(3-4), 301-311.

Semenov, M. A. (2007). Development of high-resolution UKCIP02-based climate change scenarios in the UK. *Agricultural and Forest Meteorology, 144*(1-2), 127-138.

Sharma, B. D., Mukhopadhyay, S. S., & Sidhu, P. S. (1998). Microtopographic controls on soil formation in the Punjab region, India. *Geoderma, 81*(3-4), 357-368.

Sirotenko, O. D. (2001). Crop Modeling: Advances and Problems. *Agron J, 93*(3), 650a-653.

Smale, M., Groote, H. D., & Owuor, G. (2006). BIODIVERSITY OF MAIZE ON FARMS IN KENYA. *Promising Crop Biotechnologies for Smallholder Farmers in East Africa: Bananas and Maize*(IFPRI).

Tittonell, P., Vanlauwe, B., Leffelaar, P. A., Shepherd, K. D., & Giller, K. E. (2005). Exploring diversity in soil fertility management of smallholder farms in western Kenya - II. Within-farm variability in resource allocation, nutrient flows and soil fertility status. *Agriculture Ecosystems & Environment, 110*(3-4), 166-184.

van Ittersum, M. K., Leffelaar, P. A., van Keulen, H., Kropff, M. J., Bastiaans, L., & Goudriaan, J. (2003). On approaches and applications of the Wageningen crop models. *European Journal of Agronomy, 18*(3-4), 201-234.

Venus, V. (2000). *Towards a framework for maize yield forecasting : a comparison of two promising methods : the case of Mashona land, west province, Zimbabwe.* ITC, Enschede.

Vinocur, M. G., & Ritchie, J. T. (2001). Maize Leaf Development Biases Caused by Air-Apex Temperature Differences. *Agron J, 93*(4), 767-772.

Wallis, T. W. R., & Griffiths, J. F. (1995). An assessment of the weather generator (WXGEN) used in the erosion/productivity impact calculator (EPIC). *Agricultural and Forest Meteorology, 73*(1-2), 115-133.

Wang, E., Robertson, M. J., Hammer, G. L., Carberry, P. S., Holzworth, D., Meinke, H., et al. (2002). Development of a generic crop model template in the cropping system model APSIM. *European Journal of Agronomy, 18*(1-2), 121-140.

Wilks, D. S. (1999a). Interannual variability and extreme-value characteristics of several stochastic daily precipitation models. *Agricultural and Forest Meteorology, 93*(3), 153-169.

Wilks, D. S. (1999b). Simultaneous stochastic simulation of daily precipitation, temperature and solar radiation at multiple sites in complex terrain. *Agricultural and Forest Meteorology, 96*(1-3), 85-101.

Wilks, D. S. (2008). High-resolution spatial interpolation of weather generator parameters using local weighted regressions. *Agricultural and Forest Meteorology, 148*(1), 111-120.

Wit, A. J. W. d., Boogaard, H. L., & Diepen, C. A. v. (2005). Spatial resolution of precipitation and radiation: The effect on regional crop yield forecasts. *Agricultural and Forest Meteorology, 135*(1-4), 156-168.

Wu, D., Yu, Q., Lu, C., & Hengsdijk, H. (2006). Quantifying production potentials of winter wheat in the North China Plain. *European Journal of Agronomy, 24*(3), 226-235.

Yang, C. C., R. E.; Isham, V. S.; Wheater, H. S. (2005). Spatial-temporal rainfall simulation using generalized linear models. *Water Resour. Res., 41*, 1 to 13.

Yang, H. S., Dobermann, A., Lindquist, J. L., Walters, D. T., Arkebauer, T. J., & Cassman, K. G. (2004). Hybrid-maize--a maize simulation model that combines two crop modeling approaches. *Field Crops Research, 87*(2-3), 131-154.

8. ANNEXES

Station data used for the rainfall generator (modified from Bhandari)-Part A

St_Id	St_Name	Elev. (m)	Pwd	Pww	mean	sd	skew	rainy days
9036002	Naivasha D. O.	1962	0.21	0.54	607.19	165.08	-0.44	0.31
9036011	Kedong Valley, Maai Mahiu	1869	0.14	0.56	709.06	262.71	0.11	0.24
9036025	N. Kinangop Forest Station	2661	0.24	0.60	1116.96	213.21	0.59	0.37
9036034	Gilgil Railway Station	2005	0.18	0.53	577.97	189.85	-0.73	0.27
9036059	Kangari Farm, Naivasha	2209	0.15	0.45	564.37	230.14	-0.39	0.22
9036061	Kirita Forest Station	2403	0.20	0.58	1322.69	364.62	0.84	0.32
9036062	Naivasha Kongoni Farm	2002	0.21	0.55	621.18	174.46	-0.75	0.32
9036065	Naivasha Nanga Gerri	2383	0.18	0.54	715.33	252.16	0.21	0.28
9036072	Mweiga Estate	1912	0.21	0.47	853.92	227.26	0.78	0.28
9036073	Naivasha K.C.C. Ltd.	1899	0.19	0.47	583.61	164.09	0.49	0.26
9036076	Technology Farm, Nakuru	1913	0.23	0.55	854.53	223.85	-0.89	0.34
9036081	Naivasha Vet. Expt. Station	1925	0.19	0.49	689.34	207.12	0.47	0.27
9036085	Karameno Shopping Centr, N/Moru	2060	0.16	0.50	597.50	318.98	0.37	0.24
9036109	Naivasha Marula Estate	1890	0.15	0.46	641.04	153.19	0.32	0.22
9036129	Chokerereia F.C. Society	2230	0.18	0.50	717.55	207.55	-0.12	0.26
9036147	Elementaita, Soysambu Estate	1819	0.18	0.43	721.31	165.35	-0.03	0.24
9036150	Gilgil Kikopey Ranch	1846	0.19	0.50	606.99	202.31	-0.23	0.27
9036151	Subukia Pyrethrum Nursery	2083	0.22	0.61	1021.25	294.81	-0.47	0.36
9036152	S. Kinangop Njabini F.T.C.	2535	0.24	0.63	1153.02	343.91	1.03	0.40
9036162	Kijabe Railway Station	2119	0.12	0.52	660.12	269.21	-0.22	0.20

Station data used for the rainfall generator (modified from Bhandari)-Part B

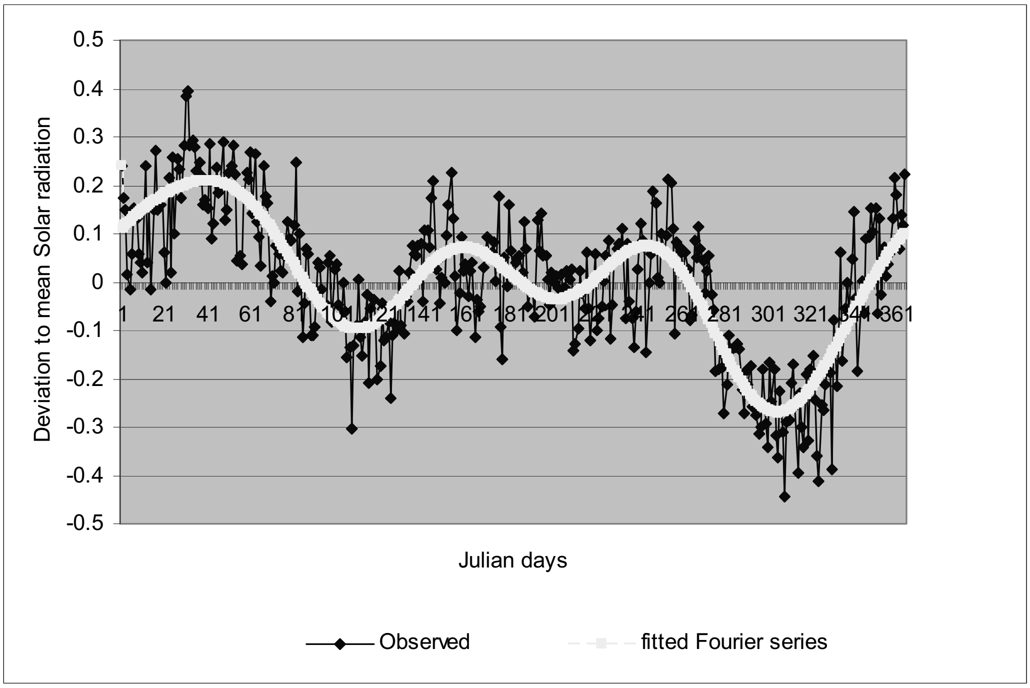
St_Id	St_Name	Elev. (m)	Pwd	Pww	mean	sd	skew	rainy days
9036164	S. Kinangop Forest Station	2531	0.24	0.68	1299.49	392.10	-0.28	0.43
9036174	Aberdare Park Fort Jerusalem	3173	0.21	0.77	1261.90	779.76	-0.89	0.47
9036179	Naivasha Korongo Farm	1888	0.17	0.49	631.15	175.42	0.51	0.26
9036183	Naivasha Karati Scheme	2546	0.20	0.59	782.31	300.94	-0.38	0.33
9036188	Kinangop Sasumua Dam	2517	0.28	0.73	1438.86	331.85	0.54	0.51
9036198	New Gakoe Farm (Nakuru)	1936	0.24	0.58	865.76	308.80	-1.12	0.36
9036214	Naivasha Longonot Farm	1897	0.14	0.50	614.48	177.72	-0.05	0.21
9036227	Elementaita Nderit Ranger Post	1828	0.21	0.53	740.43	183.16	0.42	0.31
9036236	Nakuru Lanet Police Post	1873	0.24	0.60	794.12	184.01	0.30	0.37
9036241	Geta Forest Station	2605	0.24	0.63	1081.97	312.31	-0.68	0.39
9036243	Dundori Forest Station	2237	0.21	0.50	1130.29	251.14	0.83	0.30
9036244	Kieni Forest Station	2536	0.21	0.61	1321.07	442.19	-0.05	0.35
9036252	Menengai Forest Station	2226	0.24	0.57	923.27	235.19	0.17	0.36
9036253	Thome Farmers No. 2	2385	0.18	0.53	811.08	321.44	-0.55	0.28
9036257	Eastern Rift Saw Mill Ltd.	2336	0.21	0.49	828.56	267.43	0.23	0.29
9036259	Gatare Forest Station	2602	0.32	0.79	2428.49	514.40	0.18	0.60
9036261	Nakuru Meteorological Station	1919	0.31	0.65	891.67	203.45	-0.46	0.47
9036262	Olaragwai Farm Naivasha	2070	0.19	0.49	674.49	149.43	0.90	0.27
9036264	N. Kinangop Mawingo Scheme	2407	0.23	0.64	846.14	314.89	-0.34	0.39
9036272	Mutubia Gate (A.N. Park)	2643	0.34	0.70	1296.67	456.67	-1.35	0.54

Station data used for the rainfall generator (modified from Bhandari)-Part C

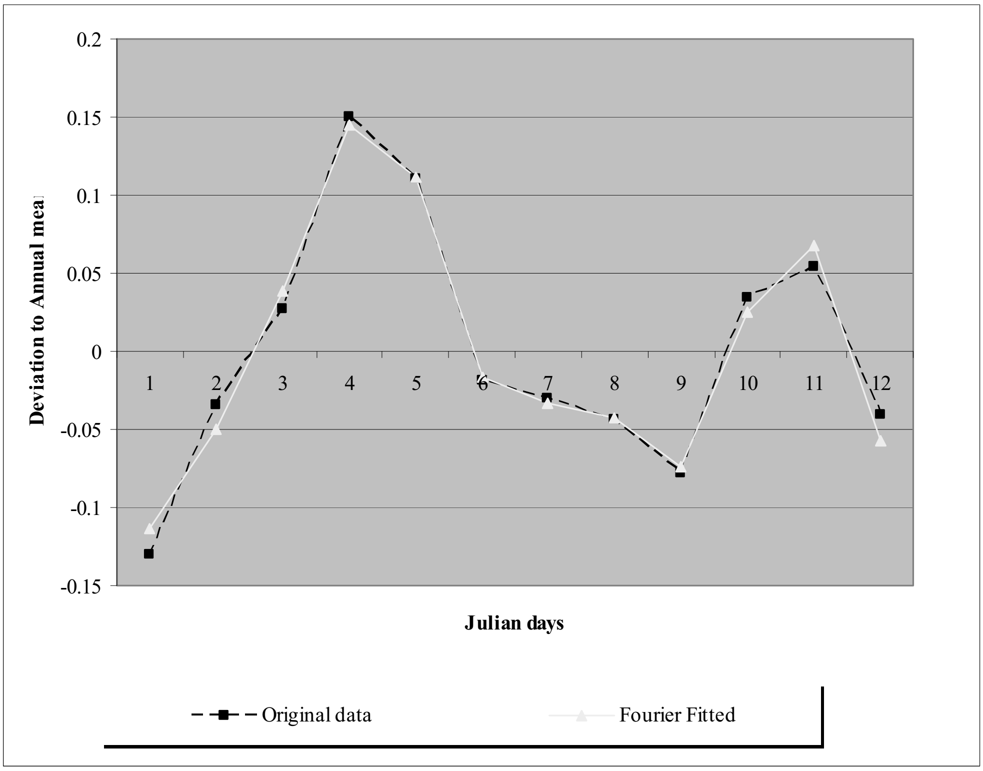
St_Id	St_Name	Elev. (m)	Pwd	Pww	mean	sd	skew	rainy days
9036277	Magura River	3009	0.02	0.00	118.28	36.70	0.56	0.02
9036278	Riunge Hill	3167	0.01	0.00	92.53	24.56	-0.86	0.02
9036279	Culvert Camp	2619	0.02	0.09	40.35	24.78	-0.45	0.02
9036280	Chania River Aberd. Nat. Park	2990	0.02	0.00	104.40	25.99	-0.95	0.02
9036281	Naivasha W.D.D.	2031	0.21	0.51	619.24	176.42	-0.97	0.30
9036285	Longonot Akira Ranch	1694	0.11	0.42	467.60	238.36	-0.72	0.16
9036289	Wanjohi Chief's Camp	2438	0.26	0.64	871.30	217.20	0.78	0.42
9036290	Malewa Farmer Coop. Soc.	2323	0.18	0.60	595.31	256.32	-0.94	0.31
9036294	Ngecha New Farmers Coop.	2154	0.16	0.64	394.72	261.56	-0.03	0.31
9036296	Kurase Hill Aberdare Park	3347	0.02	0.29	60.07	49.04	-1.72	0.03
9036307	Kangui Secondary School	2542	0.17	0.62	862.39	299.52	-0.02	0.32
9036308	Ngethu Water Supply	1798	0.21	0.63	1286.37	409.16	0.08	0.37
9036309	Miti Mingi Farm	1908	0.24	0.55	672.12	279.05	-1.84	0.35
9036310	Kamirithu Fancy Farm	1929	0.23	0.50	932.60	845.97	2.01	0.31
9036312	Chamata Gate	2822	0.27	0.65	968.77	272.34	0.02	0.43
9036313	Chebuswa Hill	3229	0.01	0.57	49.40			0.02
9036317	Sakutiek C.C. Outpost	2695	0.20	0.55	516.69	289.47	-0.35	0.31
9036319	Mugunda Primary School	2337	0.15	0.49	540.77	285.00	-0.04	0.22
9036320	Naishi Ranger's Post	1790	0.21	0.53	603.24	211.78	-0.30	0.31
9036322	Crescent Island	1881	0.16	0.44	537.23	148.99	0.23	0.22
9036323	Kianganye Farm Ichichi	2286	0.28	0.73	2337.28	553.66	0.21	0.50
9036331	Olchoro Agri. Office	2571	0.19	0.43	766.72	263.13	-0.94	0.25
9036336	Tumaini N.Y.S. Camp	2499	0.20	0.58	941.14	412.71	0.28	0.32
9036337	Sururu Forest Station	2551	0.20	0.51	834.53	287.33	-0.24	0.29
9036342	Ndunyu Njeru	2177	0.19	0.57	582.70	407.62	-0.19	0.31
9036343	Olkaria Geothermal Station	2031	0.19	0.44	627.30	141.51	-0.95	0.25

Annex 2. Example of Fourier used for mean Tmax, Tmin and Maximum Solar Radiation.

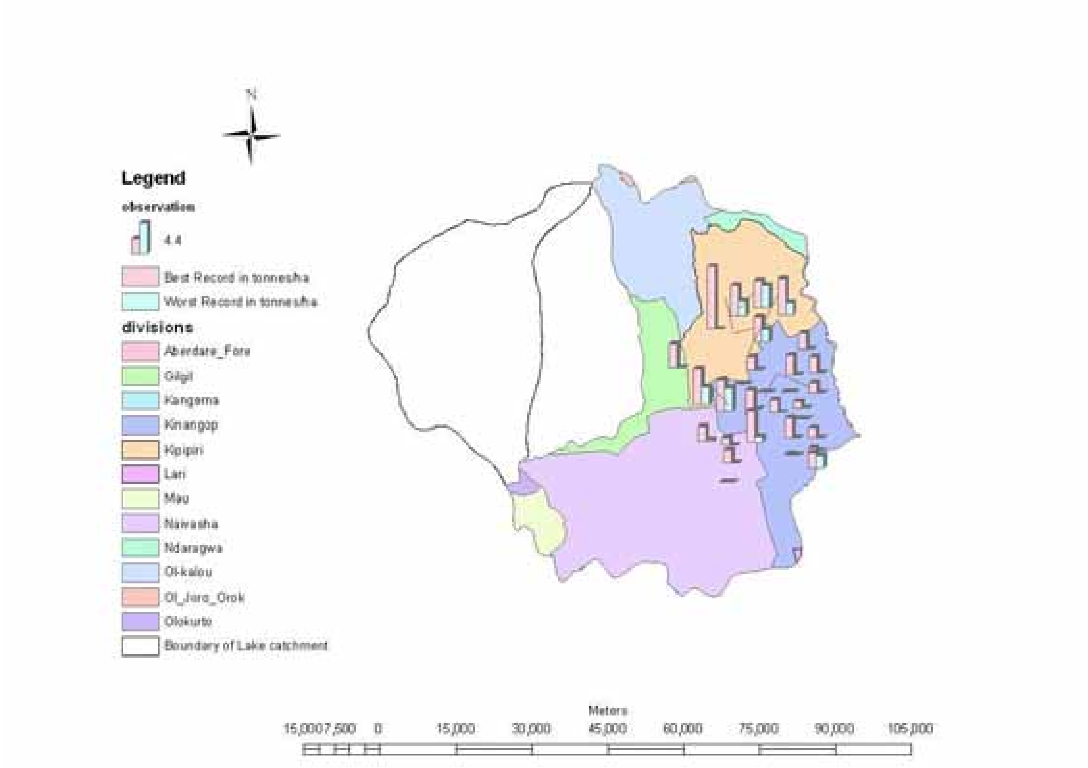
A.Fourier Series used for Mean of Solar Radiation (from Daily Data of Nakuru Synoptic Station)



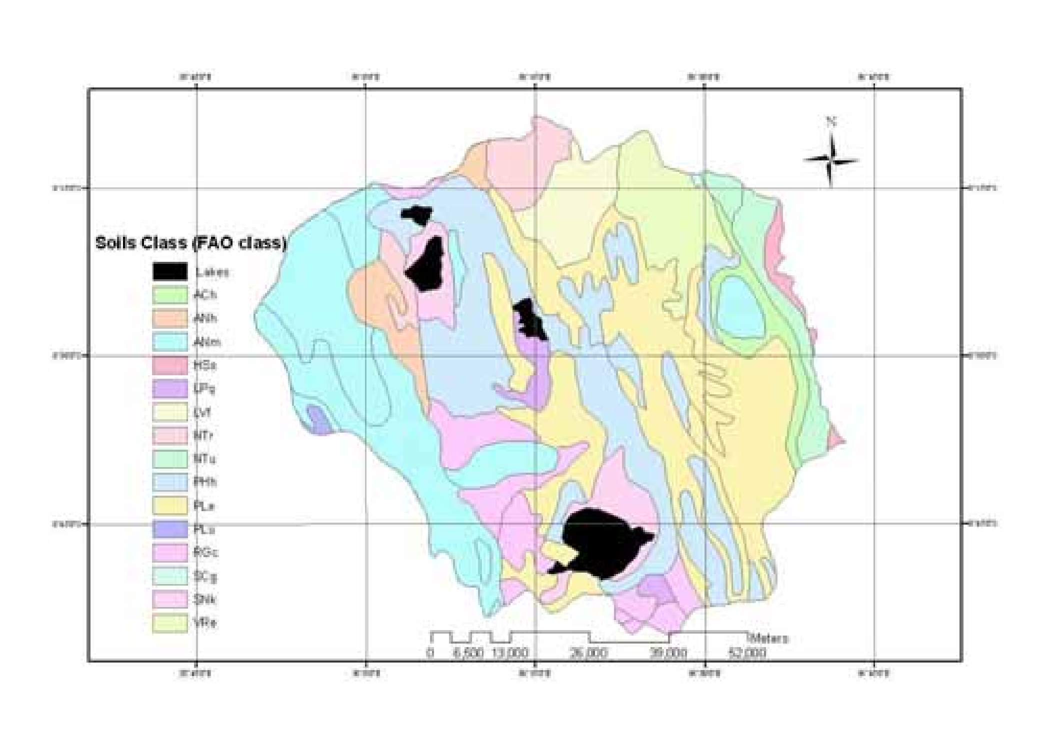
B. Fourier series used for Mean of Tmax (from monthly data, CRU T.S.2)



Annex 3. Yield Map from farmer interview



Annex 4. Soil map of the study Area from ISRIC(Batjes & Gicheru, 2004).



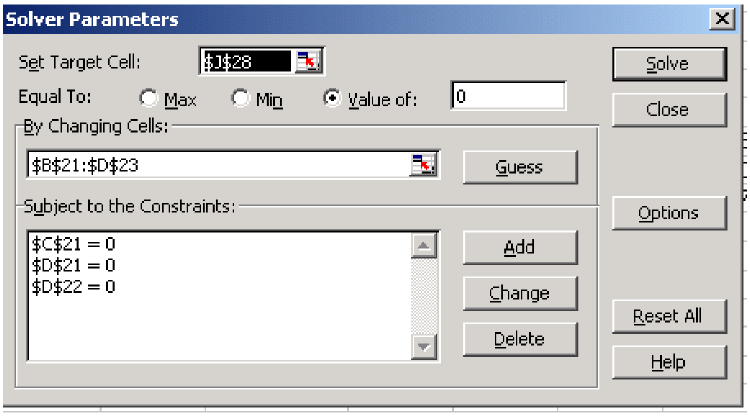
Annexe 5. Input matrix for WGEN using data of Naivasha Area

A. Using Excel

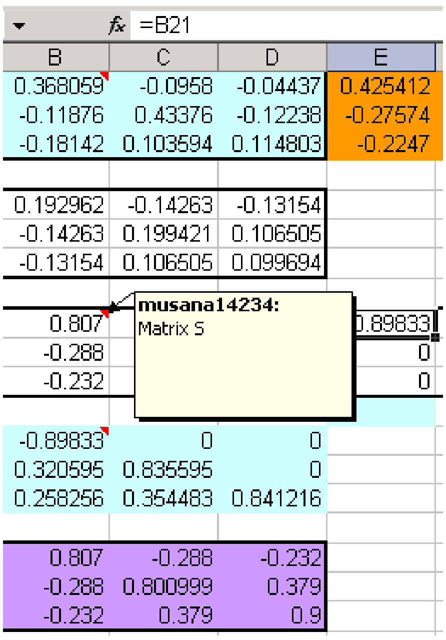
$$M_0 = \begin{bmatrix} 1 & -0.43 & -0.36 \\ -0.43 & 1 & 0.48 \\ 0.36 & 0.48 & 1 \end{bmatrix} \quad M_1 = \begin{bmatrix} 0.42 & -0.27 & -0.22 \\ -0.26 & 0.42 & 0.13 \\ 0.26 & 0.23 & 0.23 \end{bmatrix}$$

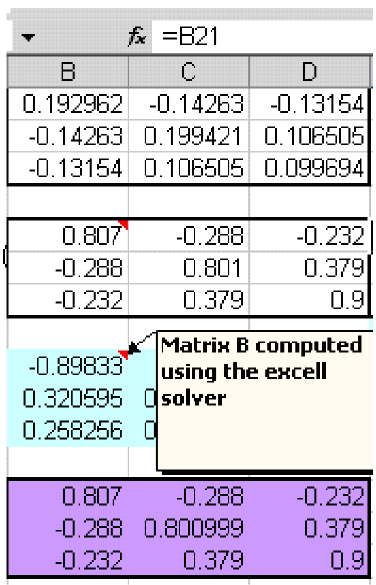
$$A = \begin{bmatrix} 0.36 & -0.09 & -0.04 \\ -0.11 & 0.43 & 0.12 \\ 0.18 & 0.10 & 0.11 \end{bmatrix} \quad B = \begin{bmatrix} -0.89 & 0 & 0 \\ 0.32 & 0.83 & 0 \\ 0.25 & 0.35 & 0.84 \end{bmatrix}$$

Windows of Excel representing the Objectives Function (0 cell J28), the 3constraints



Windows of Excel representing the Static Matrix S used to solve Matrix B. The dynamic Matrix S' is equal to S at the end of the process.





B. using Mathlab (*Cholesky decomposition)*

```
To get started, select MATLAB Help or Demos from the Help menu.
>> M=[0.807043347443903 -0.287797883141581 -0.231672996441832;-0.287797883141581
0.800589045223451 0.380086493498991;-0.231672996441832 0.380086493498991
0.898792657608992;]
M =
    0.8070   -0.2878   -0.2317
   -0.2878    0.8006    0.3801
   -0.2317    0.3801    0.8988
>> Lt = chol( M )
Lt =
    0.8984   -0.3204   -0.2579
         0    0.8354    0.3561
         0         0    0.8399
>> L = Lt'
L =
    0.8984         0         0
   -0.3204    0.8354         0
   -0.2579    0.3561    0.8399
>>
```

Annex6. Procedures for Import of Series of Map to PCRaster

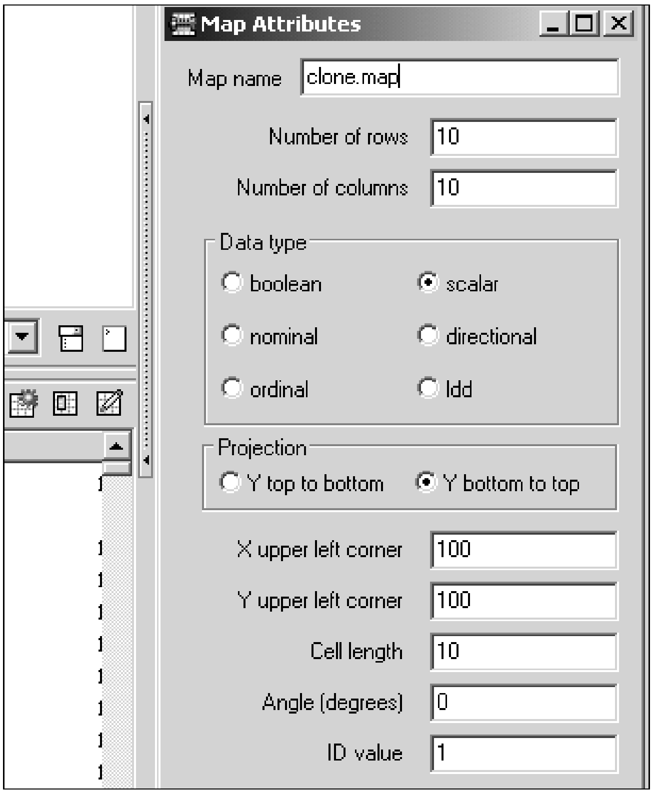
1. Create a batch file

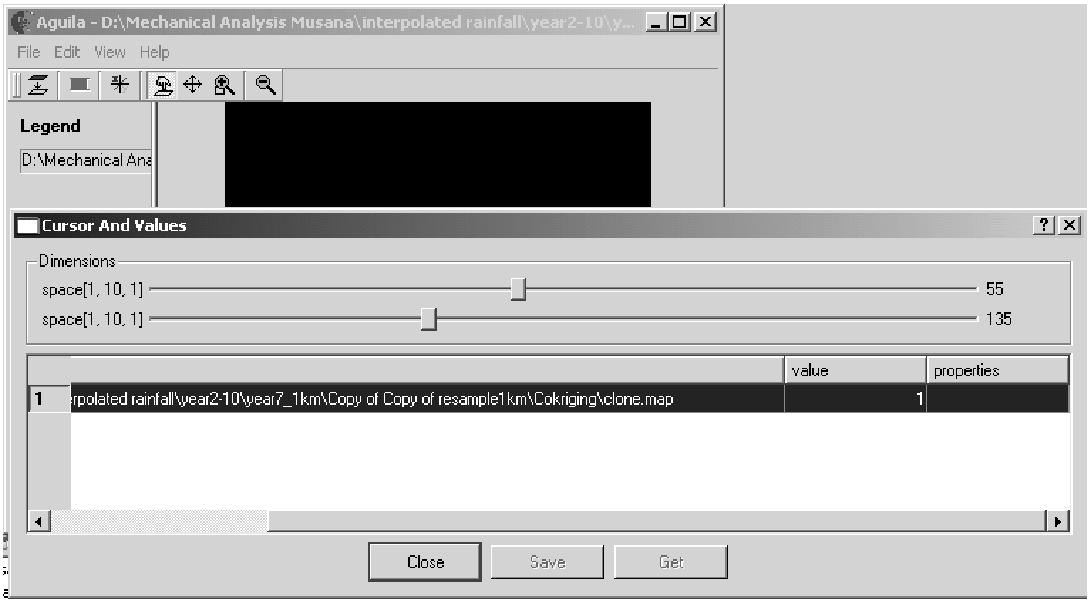
By using the application asc2map. The general form of the syntax of asc2map is provided in the help file of PCRaster (http://pcraster.geo.uu.nl/documentation/pcrman/r14839.htm),

asc2map --clone mapclone.map -S -m mv -v 4 AscFile1.asc Result1.map

Mapclone is grid map in pcraster (not in a ascii format). It has the bounding box (x and y limits) of the map you are importing. In nutshell create a mapclone is even simply by using mapedit application. In mapedit using create map and specify unique value. The results can be display in Nutshell

To use mapedit and Aguila in Nutshell are explained in the Nutshell web page (http://www.itc.nl/lisem/nutshell/)

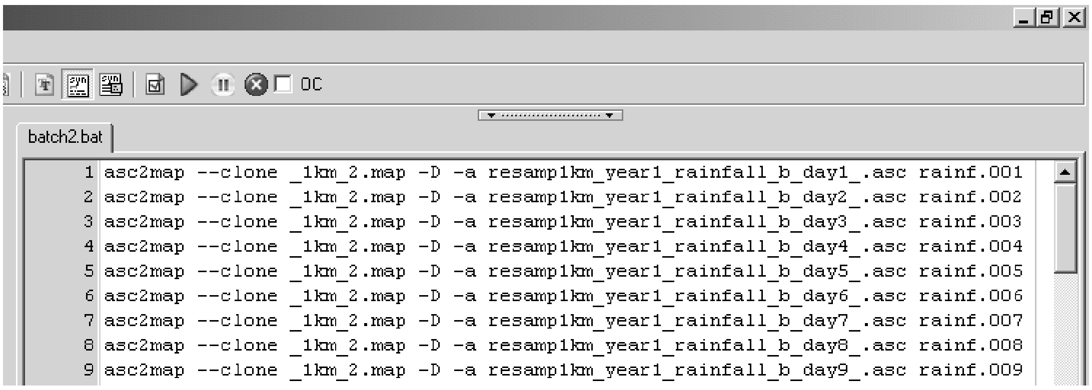




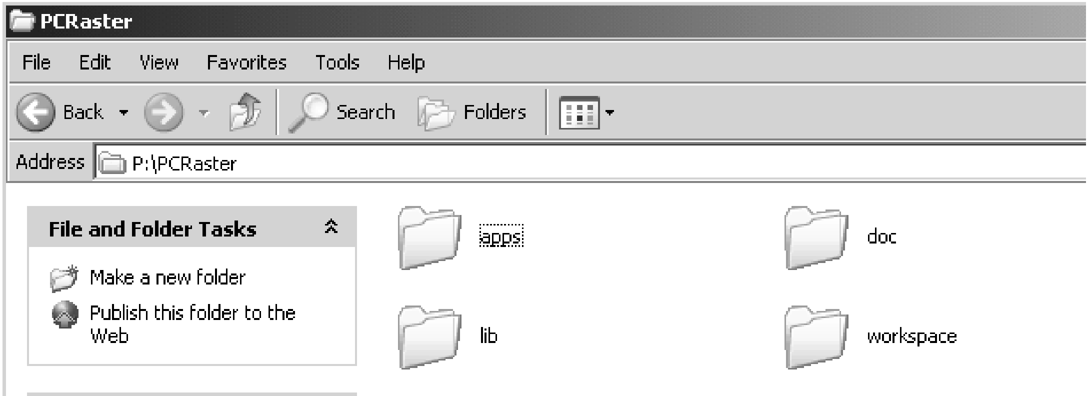
2. For the case used in the dissertation the following bounding box were applied :

Ncols = 114
Nrows = 112
Lower left x coordinate = 141950
Lower left y coordinate = 9884000
Cellsize = 1000.00

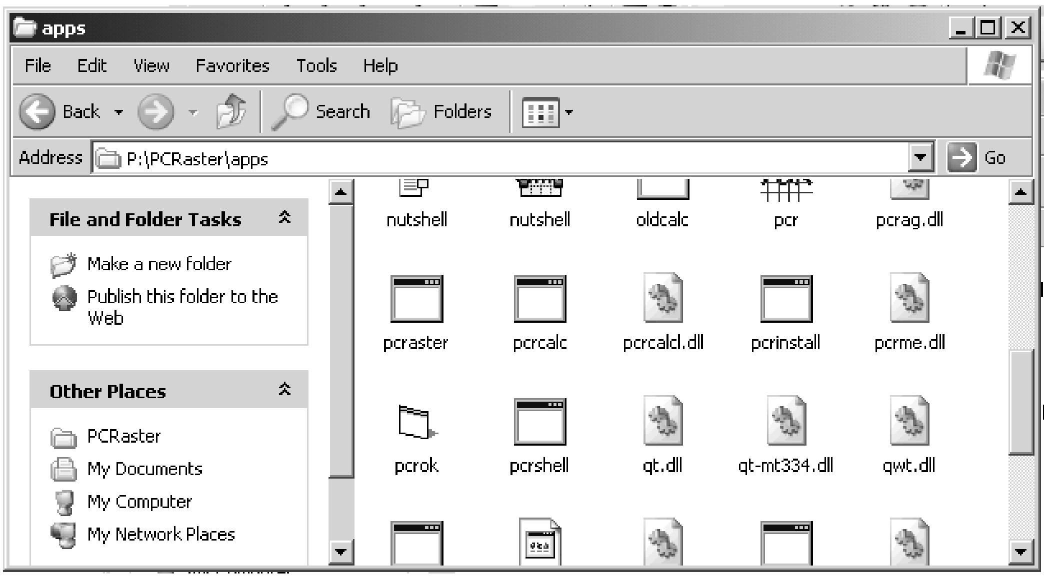
For each map imported write one line in the batch file. No comma or semi column are used. As presented below.



3. Go to apps file



4. Go to pcrshell



5. Open pcrshell and check if the working directory is the same (the location of the input of your model written).

6. Run the batch file by typing is name after the sign”>”.

A notification on the results of importing will be given.