# DEVELOPMENT OF A FRAMEWORK FOR AN INTEGRATED TIME-VARYING AGROHYDROLOGICAL FORECAST SYSTEM FOR SOUTHERN AFRICA

## YONAS BEYENE GHILE

Submitted in partial fulfilment of the requirement for the degree of
Doctorate in Philosophy in the Discipline of Hydrology

*School of Bioresources Engineering and Environmental Hydrology*
*University of KwaZulu-Natal, Pietermaritzburg, RSA*

March, 2007

## PREFACE

The research described in this thesis was carried out under the supervision of Professor R.E. Schulze (School of Bioresources Engineering and Environmental Hydrology, University of KwaZulu-Natal). I hereby certify that the research presented in this thesis is my own original work except where specific acknowledgment is made.


_____

Yonas Beyene Ghile


_____

Supervisor: Professor R.E. Schulze


Date: _____

# ABSTRACT

Policy makers, water managers, farmers and many other sectors of the society in southern Africa are confronting increasingly complex decisions as a result of the marked day-to-day, intra-seasonal and inter-annual variability of climate. Hence, forecasts of hydro-climatic variables with lead times of days to seasons ahead are becoming increasingly important to them in making more informed risk-based management decisions. With improved representations of atmospheric processes and advances in computer technology, a major improvement has been made by institutions such as the South African Weather Service, the University of Pretoria and the University of Cape Town in forecasting southern Africa's weather at short lead times and its various climatic statistics for longer time ranges. In spite of these improvements, the operational utility of weather and climate forecasts, especially in agricultural and water management decision making, is still limited. This is so mainly because of a lack of reliability in their accuracy and the fact that they are not suited directly to the requirements of agrohydrological models with respect to their spatial and temporal scales and formats.

As a result, the need has arisen to develop a GIS based framework in which the "translation" of weather and climate forecasts into more tangible agrohydrological forecasts such as streamflows, reservoir levels or crop yields is facilitated for enhanced economic, environmental and societal decision making over southern Africa in general, and in selected catchments in particular. This study focuses on the development of such a framework. As a precursor to describing and evaluating this framework, however, one important objective was to review the potential impacts of climate variability on water resources and agriculture, as well as assessing current approaches to managing climate variability and minimising risks from a hydrological perspective. With the aim of understanding the broad range of forecasting systems, the review was extended to the current state of hydro-climatic forecasting techniques and their potential applications in order to reduce vulnerability in the management of water resources and agricultural systems. This was followed by a brief review of some challenges and approaches to maximising benefits from these hydro-climatic forecasts.

A GIS based framework has been developed to serve as an aid to process all the computations required to translate near real time rainfall fields estimated by remotely sensed tools, as well as daily rainfall forecasts with a range of lead times provided by Numerical Weather Prediction (NWP) models into daily quantitative values which are suitable for application with hydrological or crop models. Another major component of the framework was the development of two methodologies, *viz.* the *Historical Sequence Method* and the *Ensemble Re-ordering Based Method* for the translation of a triplet of categorical monthly and seasonal rainfall forecasts (i.e. Above, Near and Below Normal) into daily quantitative values, as such a triplet of probabilities cannot be applied in its original published form into hydrological/crop models which operate on a daily time step.

The outputs of various near real time observations, of weather and climate models, as well as of downscaling methodologies were evaluated against observations in the Mgeni catchment in KwaZulu-Natal, South Africa, both in terms of rainfall characteristics as well as of streamflows simulated with the daily time step *ACRU* model. A comparative study of rainfall derived from daily reporting raingauges, ground based radars, satellites and merged fields indicated that the raingauge and merged rainfall fields displayed relatively realistic results and they may be used to simulate the "now state" of a catchment at the beginning of a forecast period. The performance of three NWP models, *viz.* the C-CAM, UM and NCEP-MRF, were found to vary from one event to another. However, the C-CAM model showed a general tendency of under-estimation whereas the UM and NCEP-MRF models suffered from significant over-estimation of the summer rainfall over the Mgeni catchment. Ensembles of simulated streamflows with the *ACRU* model using ensembles of rainfalls derived from both the *Historical Sequence Method* and the *Ensemble Re-ordering Based Method* showed reasonably good results for most of the selected months and seasons for which they were tested, which indicates that the two methods of transforming categorical seasonal forecasts into ensembles of daily quantitative rainfall values are useful for various agrohydrological applications in South Africa and possibly elsewhere. The use of the *Ensemble Re-ordering Based Method* was also found to be quite effective in generating the transitional probabilities of rain days and dry days as well as the persistence of dry and wet spells within forecast cycles, all of which are

important in the evaluation and forecasting of streamflows and crop yields, as well as droughts and floods.

Finally, future areas of research which could facilitate the practical implementation of the framework were identified.

*This thesis is dedicated to the memory of my father Beyene Ghile*

# TABLE OF CONTENTS

I

# LIST OF TABLES

# LIST OF FIGURES

# LIST OF SYMBOLS AND ABBREVIATIONS

AAMHS       *ACRU* Agrohydrological Modelling System

ABRESP      Fraction of "saturated" soil water to be redistributed daily from the topsoil into the subsoil when the topsoil is above its drained upper limit

ACRU        Agricultural Catchments Research Unit

ADJIMP      Impervious areas connected directly to a watercourse (fraction)

AGCM        Atmospheric General Circulation Model

ANN         Artificial Neural Networks

AR          AutoRegressive

ARC         Agricultural Research Council

ARIMA       AutoRegressive Integrated Moving Average

AWS         Automatic Weather Station

BCSD        Bias Corrected Spatial Disaggregation

BFRESP      Fraction of "saturated" soil water to be redistributed daily from the subsoil into the intermediate store when the subsoil is above its drained upper limit

BS          Brier Score

CAY         Monthly crop coefficient (fraction)

CCA         Canonical Correlation Analysis

C-CAM       Conformal-cubic Atmospheric Model

CHRM        Climate High Resolution Model

CMC         Canadian Meteorological Center

CN          Correct Nulls

COFRU       Coefficient of baseflow response (fraction)

COIAM       Coefficient of initial abstraction (fraction)

COLA        Center for Ocean Land Atmosphere

CoV         Coefficient of Variation

CRED        Centre for Research in the Epidemiology of Disaster

CSI         Critical Success Index

CWG         Conditional Weather Generator

DEAT        Department of Environment Affairs and Tourism

DEPAHO      Thickness of topsoil (m)

DEPBHO      Thickness of subsoil (m)

DISIMP      Impervious areas which are not adjacent to a watercourse (fraction)

DJF         December January February

DWAF        Department of Water Affairs and Forestry

ECMWF       European Centre for Medium Range Weather Forecasts

| | |
|---|---|
| EFS | Ensemble Forecasting System |
| ENSO | El Niňo Southern Oscillation |
| ERM | Ensemble Re-ordering Method |
| ESRI | Environmental Systems Research Institute |
| ET | Ensemble Transform |
| ETKF | Ensemble Transform Kalman Filter |
| FA | False Alarm |
| FAO | Food and Agriculture Organisation of the United Nations |
| FAR | False Alarm Ratio |
| FC1 | Soil water content at drained upper limit for topsoil (m.m$^{-1}$) |
| FC2 | Soil water content at drained upper limit for subsoil (m.m$^{-1}$) |
| FCAO | Farmers' Co-Operatives and Agricultural Organisations |
| FEWS | Food Security Information and Early Warning System |
| FFT | Fast Fourier Transform |
| $F_P$ | Forecast Probability |
| GCM | General Circulation Model |
| GDP | Gross Domestic Product |
| GIS | Geographic Information System |
| $G_K$ | Kriged raingauge rainfall field |
| $G_V$ | Variance fields for raingauges |
| H | Hits |
| HS | Heidke Score |
| ICD | Iterative Constrained Deconvolution |
| IDO | Indian Ocean Dipole |
| IPCC | Intergovernmental Panel on Climate Change |
| IR | Infrared |
| IRUN | Variable to request the exclusion or inclusion of baseflow from the simulation of streamflow |
| ISCW | Institute for Soil, Climate and Water of the ARC |
| ISDR | International Strategy for Disaster Reduction |
| JFM | January February March |
| JMA | Japan Meteorological Agency |
| LAM | Limited Area Model |
| LDA | Linear Discriminant Analysis |
| LEPS | Linear Error in Probability Space |
| Li | Linear Interpolation |
| LTF | Linear Transfer Function |

| | |
|---|---|
| M | Misses |
| MAE | Mean Absolute Error |
| MAP | Mean Annual Precipitation (mm) |
| MAR | Multiple Autoregressive Model |
| MLR | Multiple Linear Regression |
| MRF | Medium Range Forecasting |
| MSG | Meteosat Second Generation |
| MSRR | Multi-Spectral Rain Rate |
| NAO | North Atlantic Oscillation |
| NCEP | National Center for Environmental Prediction |
| NDJ | November December January |
| NWP | Numerical Weather Prediction |
| NWRS | National Water Resource Strategy |
| OND | October November December |
| P(d/w) | Probability that a wet day follows a dry day |
| P(w/d) | Probability that a dry day follows a wet day |
| PC | Primary Catchment |
| PCM | Parallel Climate Model |
| PDO | Pacific Decadal Oscillation |
| PO1 | Soil water content at saturation for topsoil ($m.m^{-1}$) |
| PO2 | Soil water content at saturation for subsoil ($m.m^{-1}$) |
| POD | Probability of Detection |
| PQPF | Probabilistic Quantitative Precipitation Forecasts |
| QC | Quaternary Catchment |
| QCD | Quaternary Catchment Database |
| QDA | Quadratic Discriminant Analysis |
| QFRESP | Catchment's quickflow response (fraction) |
| QPF | Quantitative Precipitation Forecasts |
| $R^2$ | Coefficient of determination |
| RCM | Regional Climate Model |
| $R_G$ | Combined field of raingauge and radar rainfall |
| $R_{GV}$ | Weighted field of $R_V$ and $G_V$ |
| $R_K$ | Kriged radar rainfall field |
| RMSE | Root Mean Square Error |
| ROOTA | Fraction of active root system in the topsoil horizon (fraction) |
| RPS | Ranked Probability Skill |
| $R_V$ | Variance fields for radars |

| | |
|---|---|
| SADC | Southern African Development Community |
| SAWS | South African Weather Service |
| $S_{BSS}$ | Bias Skill Score of satellite field |
| SD | Spatial Disaggregation |
| $S_{FSS}$ | Average Skill Score of $S_{BSS}$ |
| $S_G$ | Combined field of raingauge and satellite rainfall |
| SG | Conditioned Satellite |
| SIMAR | Spatial Interpolation and Mapping of Rainfall |
| SMDDEP | Effective depth of the soil from which stormflow generation takes place (m) |
| SOI | Southern Oscillation Index |
| $S_R$ | Satellite rainfall field |
| $S_{SS}$ | Interpolated $S_{BSS}$ |
| SST | Sea Surface Temperature |
| STOIMP | Surface storage capacity, i.e. initial abstractions (mm) for impervious areas |
| $S_Z$ | Satellite rainfall estimates at positions of raingauges |
| TDS | Total Dissolved Solids |
| TSS | Total Suspended Solids |
| RTSS | Revised True Skill Statistics |
| UKCIP | United Kingdom Climate Impacts Programme |
| UM | Unified Model |
| UNDP | United Nations Development Program |
| UNEP | United Nations Environment Programme |
| USDA | United States Department of Agriculture |
| VEGINT | Land cover's canopy interception loss value (mm. rainday$^{-1}$) |
| VIS | Visible |
| WFP | World Food Programme |
| WMO | World Meteorological Organisation |
| WP1 | Soil water content at permanent wilting point for topsoil (m.m$^{-1}$) |
| WP2 | Soil water content at permanent wilting point for subsoil (m.m$^{-1}$) |
| WWAP | World Water Assessment Programme |

# 1.   INTRODUCTION

Uncertainty about hydro-climatic conditions in the immediate future (today), as well as the near (up to one week) and more distant futures (up to one season) remains a fundamental problem challenging decision makers in fields such as water resources, agriculture and many other water sensitive sectors in southern Africa, where the climate is highly variable both temporally and spatially. Many critical agricultural and water management decisions that depend on climatic conditions must be made in advance, based on climate information and assumptions, before the actual impacts of the climate materialise (Hansen, 2002). Hence, water resource managers and agriculturalists in southern Africa need to be advised of likely climatic and hydrological conditions well in advance by producing skilful hydro-climatic forecasts that have the potential to reduce risk in both the near and long terms, and to provide valuable support to meet the increasing and competing demands for limited water resources.

In South Africa, several institutions such as the South African Weather Service (SAWS), the University of Pretoria and the University of Cape Town have been actively involved in providing short (1 - 3 day) and medium (4 - 14 day), as well as long (up to 6 month) term rainfall forecasts across a range of space scales, which could potentially be applied to benefit agriculture, water resources and many other climate sensitive sectors. The scientific application of these forecasts in operational decision-making in water resources management could potentially save affected water related industries millions of Rand annually and may also spare hardship as well as loss of life (Schulze, 2002). Research into, and targeted application of, weather and climate forecasts in the management of agricultural and water resources decisions is, however, not simple, and remains an area which requires significant research effort.

## 1.1   Rationale for the Study

A Water Research Commission (WRC) project K5/1646 titled "Applications of Rainfall Forecasts for Agriculturally Related Decision Making in Selected Catchments" is currently (2005 - 2010) being undertaken by a consortium of six institutions, *viz.* the University of KwaZulu-Natal (UKZN) as the lead organisation,

the University of the Free State (UFS), the University of Pretoria (UP), the Agricultural Research Council (ARC), the University of Cape Town (UCT) and the South African Weather Service (SAWS). The project aims to address the gaps that exist between the weather and climate forecasts and the links (via agrohydrology models) into targeted applications in agricultural and water related decision making. The rationale for the project includes the following:

- The climate in southern Africa is generally highly variable both in time and space. Climate variability is likely to be amplified with the climate change in the future.

- Climate variability has a profound impact on agriculture and water resource management.

- Many agricultural decisions are made, based on climate information (short, medium and long term) and on assumptions relating to weather and climate in the near future.

- Forecasting with different lead times from one day, through multiple days, one month and one season has the potential to reduce risk in the long term and to provide valuable support to meet the increasing and competing demands for limited water resources.

- However, gaps exist between the weather and climate forecasts, both in linking them to the agrohydrology models, and in application of forecasted information for targeted agricultural and water related decision-making.

## 1.2 Limitations in the Use of Weather/Climate Forecasts

There are many scientific problems that need to be addressed in the use of weather/climate forecasts for applications with agrohydrological models. Some of the problems identified in this study are:

- the skill level of the current weather and climate forecasts, which is highly dependent on the geographic locations and seasonality,
- the mismatch in scales between the output from weather/climate models and the spatial scale at which hydrological models are applied,
- climate products that are not in the form that can be used directly in agrohydrological models, and

- the lack of information transfer on how end users can apply the various forecast types.

## 1.3 Objective and Specific Tasks of the Study

The overall objective of this study is to develop, test and operationalise the translation of weather (i.e. short to medium term) and climate (i.e. long term) forecasts into integrated, time-varying and self-correcting agrohydrological forecasts for enhanced economic, environmental and societal decision making over South Africa in general, and in selected catchments in particular, in order to reduce risks to the agricultural and water sectors which are associated with vagaries of day-to-day to seasonal climate variability. This study forms part of a wider ranging WRC funded research project which encompasses three major components, *viz.*

- the collation of weather and climate forecasts, including their downscaling to appropriate spatial and temporal scales,
- the "translation" of the weather and climate forecasts into agrohydrological forecasts, and
- the applications of the forecasts to targeted end users.

This study addresses in particular the second component, focusing on development of a framework for "translating" weather and climate forecasts into agrohydrological forecasts. In this research it is hypothesised that, if short, medium and long lead time daily to seasonal rainfall forecasts can be made with a degree of confidence, these forecasts can be transformed into short to long term agrohydrological forecasts and that such information would be beneficial to farmers, water managers as well as catchment and disaster risk managers in southern Africa in order to make operational decisions from one day up to several months ahead of time. In this context, agrohydrological forecasts encompass variables that are most required in planning and management/operation of water and agricultural sectors. These forecasts include, but are not limited to:

- streamflows,
- soil water contents of the top- and subsoil,

- irrigation requirements,
- reservoir levels, and
- crop yields.

Within the study, some new methodologies have been developed and tested to translate available near real time values of rainfall from radar/satellite images as well as daily and multiple day forecasts through to seasonal rainfall forecasts, into quantitative values which can be input into the daily time step *ACRU* agrohydrological simulation model as well as into other daily models. This facilitates the generation of agrohydrological forecasts suitable for use in operational decisions of water resources systems such as water releases from dams, water rationing, flood warnings or hydropower generation projects, as well as for agricultural decisions such as crop selection, planting decision, pest control, fertilizer application, irrigation scheduling, or implementation of soil and water conservation programmes. This study has, furthermore, made a contribution to some new knowledge of forecasting on issues related to model verification strategies and uncertainty analysis by estimating errors that cascade through the translation of weather/climate forecasts into streamflow forecasts.

Climate variability affects many socio-economic activities in southern Africa at the present time, yet it receives relatively little attention of the public when compared to the awareness of people to future climate change impacts. This study therefore commences in Chapter 2 with a brief review of literature on southern Africa's climate variability and its potential impacts on water resources and agriculture. In the chapter the two most frequently occurring hydrological hazards, *viz.* droughts and floods are also discussed. The economic, social and environmental impacts of these hazards are described within the context of southern Africa. The chapter ends with a review of the current approaches to managing climate variability and minimising risks from a hydrological perspective.

In Chapter 3 the role of hydro-climatic forecasting in decisions to modify the vulnerability of human beings and properties to the adverse of impacts of climate variability is described briefly. In this chapter, an attempt has been made to review the types of agrohydrological forecasts and their potential applications to modify

vulnerability in water resources and agricultural water management, followed by a brief review of the current state of climate and hydrological forecasting techniques, as well as their qualities and limitations. Challenges and approaches to maximise benefits from hydro-climatic forecasts are presented in Chapter 4. Key elements of the review are the techniques developed to improve forecast quality, the impediments in communicating the information and the application of forecasts to modifying decisions, (i.e. forecast value).

Chapter 5 commences with the motivation behind the development of a GIS based framework for the agrohydrological forecasting system, followed by a brief description on how to use the outputs of the weather/climate models imbedded within the framework. This is followed in Chapter 6 by some general background on the test catchment, *viz.* the Mgeni catchment and the set up of the *ACRU* model for streamflow simulations. The evaluation of rainfall information from a network of daily reporting gauges, radars, satellite images and merged fields is discussed in Chapter 7, followed in Chapter 8 by an evaluation of the short and medium term rainfall forecasts from three selected Numerical Weather Prediction (NWP) models. Chapters 9 and 10 are devoted, respectively, for the verification of two temporal downscaling methodologies, *viz.* the *Historical Sequence Method* and the *Ensemble Re-ordering Based Method*, which are developed in this study to translate the categorical monthly and seasonal (i.e. 3 months) rainfall forecasts into daily quantitative rainfall values.

A summary and conclusions, followed by recommendations for further research, are presented in Chapter 11.

# 2. IMPACTS OF CLIMATE VARIABILITY ON WATER RESOURCE AND AGRICULTURAL SYSTEMS, WITH A FOCUS ON SOUTHERN AFRICA

Climate drives virtually all biophysical processes (e.g. runoff, droughts, floods, fire, frost, snowmelt, erosion, photosynthesis), and climate variability has, therefore, always been a major determinant of the socio-economic activities of humankind (Jose *et al.*, 1996; Sauchyn, 2000). How people dress, how they build their homes, when they go on vacation, what kind of crops they grow, as well as many other human activities are dependent on climate variability (Kinuthia, 1999). Before end users use any climate information from hydro-climatic forecasts, they must be made aware of climate risks and their impacts on them. In this chapter a review of the impacts of climatic variability on two of the most sensitive climate determined sectors, *viz.* water resources and agriculture, is presented with reference to southern Africa. Following that, current approaches to manage climate variability and minimising risks from hydrological perspective, with a focus on the two most frequently occurring hydrological hazards in southern Africa, (i.e. floods and droughts) are briefly described, in order to demonstrate the importance of integrated time-varying agrohydrological forecasting as a tool to reduce vulnerability to climate hazards in these two sectors, and to manage the risk of droughts and floods through anticipatory actions.

## 2.1 Short Term Variability of Climate in Southern Africa

Climate is a function of global wind patterns, amount and variability of precipitation and rates of evaporation (Van Zyl, 2003). Climate has never been stable for any extended period of time (Kabat and Bates, 2002). Natural external causes of climate variability include variations in the amount of energy emitted by the sun, changes in the distance between the earth and the sun and presence of volcanic pollution in the upper atmosphere. Internal variations of the climate system also produce fluctuations through the feedback processes that connect various components of the climate system. These variations arise when the more rapidly varying atmospheric conditions "force" the so-called slow components of the system such as internal variations in the ocean or biosphere. The El Niño - Southern Oscillation (ENSO) phenomenon is a good example of slow internal climate variability (Kabat and Bates, 2002).

6

There is sometimes confusion between the two terms climate variability and climate change. However, distinguishing the long term climate change from the shorter term climate variability is the point of departure in understanding the way in which weather patterns are changing over time, and in reducing the impacts that extreme events have on life and property. Climate variability can be defined as any deviation from a long-term expected value, calculated typically from 30 years of time series data (Schulze, 2003). Unlike climate change, climate variability is an entirely natural phenomenon, is reversible and is non-permanent (IPCC, 2001; Schulze, 2003). Although, climate variability is an inherent feature of the natural climate system, it may be exacerbated as a result of global warming (IPCC, 2001; Kabat and Bates, 2002; Schulze, 2003). According to Schulze (2003) climate variability has time scales ranging from

- diurnal (within the course of a day, e.g. time of occurrence of convective thunderstorms), to
- daily (i.e. variations from one day to the next), to
- intra-seasonal (e.g. monthly climate variations), to
- inter-annual (e.g. year-to-year variability), and
- decadal (e.g. consecutive wet years or dry years).

Climate change, on the other hand, may be defined in a contemporary context as a change which is attributed directly or indirectly to human activities, through which the composition of the global atmosphere is altered and which is, in addition to natural climate variability, observed over comparable time periods (Kabat and Bates, 2002). Climate change is irreversible and permanent, and occurs where a trend over time (either positive or negative) is superimposed over naturally occurring variability. The time scale of climate change is decades to centuries and the trend is more likely to occur in steps than linearly over time (Schulze, 2003). Even though climate variability is likely to be amplified with climate change, this chapter concentrates on present day shorter term climate variability that often creates a need for natural resource management strategies, or natural hazard mitigation, over southern Africa. Southern Africa, as used in the context of this study, is defined to include Swaziland and Lesotho in addition to the Republic of South Africa.

Southern Africa is sandwiched between sub-tropical high pressure cells to the west and east (the Atlantic Ocean and Indian Ocean anticyclones), by virtue of which it becomes prone to frequent droughts and an uneven rainfall distribution in time and in space (Tyson, 1990; Kinuthia, 1999; Dyson and Van Heerden, 2002; Dyson *et al*., 2002; Hallowes, 2002). Variations in the positions and intensity of the two high pressure systems play an important role in the rainfall distribution over southern Africa (Dyson and Van Heerden, 2002). The mid-latitude westerly circulation, extending northward to these high pressure systems, controls the weather of southern Africa to a large extent. During the summer months, the influence of the westerly circulation is diminished as the high pressure systems migrate southwards and invade southern Africa in the form of tropical cyclones or easterly waves (Taljaard, 1994; Kinuthia, 1999; Dyson and Van Heerden, 2002). Invading tropical weather systems are often associated with heavy rainfall and flooding (Dyson and Van Heerden, 2002).

Generally, six main climate regions may be identified in southern Africa, *viz.* the hot humid and high summer rainfall eastern coast region, the southwestern winter rainfall region, the all year rainfall south coast, the dry western coast, the semi-arid Karoo with late summer rains and the sub-humid northeastern central summer rainfall territory (Schulze, 1997; Van Zyl, 2003). However, extensive research has shown that large parts of southern Africa experience amongst the most highly variable rainfall worldwide (Haines *et al.,* 1988; Schulze, 1997; Schulze, 2003). Over the western interior the annual average rainfall varies between 100 and 200 mm. Over the remainder of the interior the annual average rainfall fluctuates between 200 and 700 mm with only the elevated mountain areas receiving more than 1 000 mm of rain annually (Taljaad, 1996; Schulze, 1997). The rainfall over many areas is often concentrated within a short period of time and displays a high inter-seasonal and inter-annual variability (Tyson, 1990; Schulze, 1997). Consequently, southern Africa experiences exceptionally high coefficients of variation in runoff (Schulze, 1997). Research conducted by Schulze (1997) has indicated that the coefficient of variation (CoV%) of annual rainfall over parts of southern Africa is in excess of 40%, which is high by world standards. The ratio of the conversion of rainfall to runoff over approximately half of South Africa is less than 5%, which is low in comparison to the world mean of 35%.

Year - to - year climate variations over southern Africa are strongly influenced by interactions between the atmosphere and the underlying ocean and land surfaces (Landman and Klopper, 1998; Hansen, 2002). Although the atmospheric conditions may fluctuate quite rapidly, surface characteristics such as surface temperature and soil moisture change more slowly and are capable of influencing climate over longer periods. The ENSO, the largest source of natural variability in the global climate system, is responsible for a large portion of the inter-annual rainfall variability over southern Africa (Landman and Klopper, 1998). It is an anomalous large-scale ocean-atmosphere system associated with an irregular cycle of warming and cooling of sea surface temperatures (SSTs) in the tropical Pacific Ocean (Mason, 1990; Mason *et al.,* 1996; Kabat and Bates, 2002). The variability in seasonal rainfall over southern Africa is also related to sea surface temperature variations over the Atlantic and Indian Oceans (Mason, 1990). The associations between rainfall and sea surface temperature vary over the summer rainfall season. According to Landman and Klopper (1998), the Arabian Sea area and the equatorial Pacific Ocean show significant associations with December rainfall. Similarly, the central equatorial Indian Ocean is highly associated with February and March rainfall, while January rainfall has a poor association with sea surface temperatures.

Uncertainty about future climatic conditions creates a major risk not only to natural resources such as water, agriculture, forestry or fisheries, but also to other climate sensitive sectors such as traffic, energy, city planning and environmental protection. Agriculture and water resources are, however, considered as the most dependent of all human activities on weather and climate (Hansen, 2002; Maini *et al.,* 2004). The marked intra-seasonal and inter-annual variabilities of climate over southern Africa result in a high-risk environment for water resources and agriculture decision takers because these variabilities affect the major drivers of the hydrological system and certain processes within it (Kunz, 1993; Schulze, 1997). In turn, changes in the hydrological system may impact soil moisture, extreme events (floods and droughts), reservoir storage, groundwater recharge, water quality and rainfed agriculture (Kunz, 1993). It is the variability of rainfall and runoff from season to season, and within a season, rather than averaged amounts,

that cause many of the complexities and uncertainties in the management of agricultural and water systems in southern Africa (Schulze, 2003).

The two main climatic elements that affect the hydrological cycle are rainfall variability (spatial and temporal distribution) and temperature fluctuations. In this study, however, more emphasis is given to rainfall variability, since the amplitude of rainfall fluctuations are more evident than those of temperature fluctuations.

## 2.2    Potential Impacts of Climate Variability on Water Resources

Water is considered to be one of the most critical sectors associated with climate variability and change (Perks, 2001). The economic well-being of southern Africa is critically dependent upon the available water resources and the variability in rainfall patterns imposes stresses on the natural system, which can result in major social and economic dislocations. In southern Africa, summer season convective storms are the major source of most of water supply, both directly to crops and indirectly by way of farm dams, larger reservoirs and streams, all of which are also utilised to meet demands for water (NWRS, 2002). The following sections contain a review of how the supply and demand, availability, quality as well as the spatial and temporal distribution of southern Africa's water resources are affected by climate variability.

### 2.2.1    Impacts on Water Supply and Demand

Southern Africa experiences a high inter-annual variability of rainfall and runoff, thereby placing significant demands on surface and groundwater resources (Perks, 2001). Over the past century, complex water related infrastructures have been built to provide clean water for drinking and for industries, for disposal of wastes, to facilitate transportation, generate electricity, irrigate crops and to reduce the risks of floods and droughts (Garbrecht and Schneider, 2004).  However, water demand and use have also been increased markedly as a result of a growing population, higher standards of living, industrial development, supplemental agricultural irrigation, and growing power generation in southern Africa (Van Jaarsveld and Chown, 2001; NWRS, 2002).  The demand for water must be met by the limited available water. Most reservoirs have substantial conservation

storage capacities and dependable yields to bridge persistent dry periods. However, the dependable water supplies from rainfed aquifers and withdrawals from uncontrolled streamflows can be substantially reduced during persistent dry periods (NWRS, 2002).  It is believed that municipalities, industries and agriculture which rely on these sources for all or part of their water supply may be vulnerable to shortages during persistent dry periods. This is particularly pertinent for the arid and semi-arid regions of southern Africa, where demands frequently exceed supply. During the 1980s and 1990s, for example, urban areas and industry were severely affected by drought. Most water dependent industries in southern Africa were forced to reduce their activities after water reservoirs fell to critically low levels (UNEP, 1997).

Changes in hydrological variables (e.g. precipitation and evapotranspiration) can affect water availability in soils, rivers, dams and aquifers, with implications for water supplies for domestic, industrial and agricultural uses, as well as for ecological requirements (IPCC, 1996; Perks, 2001). Climate variability can also affect the complex water related infrastructure and systems, including reservoir operations, hydro-electric generation and navigation (Gleick, 2000). Research conducted in the USA, for example, has indicated that the additional costs imposed by climate variability are considerably larger than the additional costs imposed by industrial changes, or changes in agricultural water demands (Gleick, 2000). According to Smithers *et al*. (2001), the February 2000 rainfall event in some parts of the Sabie catchment South Africa exceeded the 200 year return period event and resulted in disastrous flooding and severe damage to infrastructure. More recent research conducted in Portugal by Cunha *et al*. (2005) has also indicated that relatively small changes in the inflow to reservoirs due to climate variability may cause large changes in the reliability of water yields from those reservoirs.

### 2.2.2   Impacts on Water Quality

Water quality could deteriorate from impacts of climate variability, either directly or indirectly. The indirect impacts of climate variability may result of an increase or decrease in runoff. Both types of hydrological extremes have been shown in studies to negatively affect the quality of water. During low flow events, increased

concentrations of pollutants, toxins, bacterial contaminants and nuisance algae are common, whereas heavy flow events have been shown to increase soil erosion, chemical leaching, transport of livestock wastes and nutrients into the source water systems (Cunha *et al.,* 2005). An assessment of water quality conducted by Faniran *et al.* (2001) on the Isinuka springs in the Eastern Cape indicated that the measured water quality variables (e.g. TDS, turbidity, $NH_4^+$, $SO_4^{2-}$, $Cl^-$) were found to show seasonal fluctuations depending on the magnitude of flow events. Schulz *et al.* (2001) reached a similar conclusion in the Western Cape orchard areas, where pesticides, total phosphates and total suspended solids were measured in the Lourens River at the beginning of April 1999 prior to the first rainfall of the season and in the middle of April during high discharge periods. Pre-runoff season samples indicated only contamination with total endosulfan (α, ß, sulphate) at levels up to 0.06 ug/l. Runoff during the first rainfall event resulted in an increase in the levels of endosulfan, chloropyrifos and azinphos-methyl.

Certain direct impacts may result from an increase of water temperature as a consequence of climate change, causing a decrease in the actual level of dissolved oxygen in water and interfering with chemical and biological processes occurring in water bodies, such as eutrophication processes (Cunha *et al.,* 2005).

### 2.2.3   Impacts on Groundwater Recharge

Groundwater is an important source of water supply for domestic, industrial and agricultural use, particularly in the rural and the more arid areas of southern Africa (Basson, 1997). Impacts of climate variability on groundwater are poorly understood, and the relationships between climate variables and groundwater are more complicated than those of surface water, as the influence of climate variability on groundwater levels cannot be detected immediately (Chen *et al.,* 2004). In addition, groundwater responses to climatic fluctuations are complicated by factors such as proximity to rivers, times of heavy pumping, degree of saturation, as well as the effective porosity and permeability of the soil. These factors are responsible for a delay in response of the groundwater table to rain (Soveri and Ahlberg, 1989; Van Kleef, 2003).

However, the impact of climate variability on groundwater levels can be investigated by analysing the relationship between historical climate records and groundwater level fluctuations (Chen *et al.,* 2004). Groundwater resources are related to climate variability and change through hydrological processes such as precipitation and evapotranspiration, and through interaction with surface water. With increased evapotranspiration and decreased precipitation, the impact of climate variability could result in declining groundwater levels, which would cause some wells to become dry or be less productive owing to the loss of available drawdown (Rutulis, 1989; Van Kleef, 2003; Chen *et al.,* 2004). During the 1982/83 El Niño season over southern Africa, for example, recharge to groundwater zone was less than 20% of the median recharge, since recharge takes place only following either individual large events or sustained rainfalls (Schulze, 2003). Research conducted in Finland by Soveri and Ahlberg (1989) has shown that groundwater storage changes from season to season as a consequence of different precipitation and evapotranspiration patterns. Similarly, studies conducted in Canada (e.g. Rutulis, 1989; Chen *et al.,* 2002, Chen *et al.,* 2004) have shown that under natural conditions, annual groundwater level fluctuation and long-term trends depend on net groundwater recharge, which is a function of precipitation and evapotranspiration. Long lasting severe dry weather conditions may also alter hydraulic properties of an aquifer and could alter recharge rates for aquifer systems (Chen *et al.,* 2004).

### 2.2.4   Impacts on Irrigation Demand

Agriculture of any kind is strongly influenced by the availability and quality of water (Olesen and Bindi, 2002). Irrigation is the largest consumer of water resources, particularly in semi-arid regions of southern Africa (Basson, 1997). Irrigated agriculture is less vulnerable to climate variability and change than rainfed agriculture, provided that there is a sufficient supply of water. The amount of water required to irrigate a crop depends principally on crop evapotranspiration (Perks, 2001). The demand for water for irrigation use tends to rise in a warmer or drier climate conditions (Schulze, 1997), thereby increasing the competition for available water between agriculture and other users (Arnell, 1999). Peak irrigation

13

demands are also predicted to rise as a result of more severe heat waves that affect crop evapotranspiration (Olesen and Bindi, 2002).

## 2.3    Potential Impacts on Rainfed Agriculture

Agricultural output is highly sensitive to year-to-year climate variability. One reason for climate variability so often being devastating to rainfed agriculture is that farmers do not know what to expect in the upcoming growing season. Thus, farmers and other decision makers in agriculture tend to be unprepared for the climate conditions that do occur, and make decisions based on their understanding of general climate patterns for their regions (Jones *et al.*, 2000). The impact of climate variability is felt by farmers wanting to optimise agricultural production, by agribusiness managers wanting to best manage their inventories or commodity marketing strategies, and by governments wanting to best manage taxation and to enact drought relief policies (Hammer and Nicholls, 1996).

Droughts and heavy storms are the major climatic threats to agricultural production, and they affect the livelihoods of millions of people annually around the world (FAO, 2004). Rainfall variability, which includes erratic and unpredictable seasonal rainfall as well as the occurrences of floods and droughts, contributes to the risks in managing water resources and agricultural production across southern Africa, particularly in marginal rainfed agricultural areas characterised by low and variable precipitation. Extended periods of drought or, alternatively, unusually high rainfall with associated flooding, can have devastating effects on the already marginal levels of production, placing subsistence farming at increased risk. The drought over southern Africa during the 1982/1983 El Niño caused estimated damages about US$ 1 billion (Moura *et al.,* 1992; cited by Landman *et al.,* 2001). The impact of this event was made worse by the major drought experienced during 1991/1992, which exacerbated social and economic problems, and also reduced the overall livelihood security of the society (Kinuthia, 1999; Klopper, 1999; Landman *et al.,* 2001; FAO, 2004). The loss in GDP in South Africa during the 1991/1992 drought was estimated to be at about 1.8%, equivalent US$ 500 million (FAO, 2004). In contrast to this, major flooding events in southern Africa that lasted several days during 1984, 1988 and 2000 equally caused significant loss of life and of livestock. Large areas of agricultural land were submerged, as

were livestock and farming implements. Hundreds of hectares of fertile alluvial soil were reduced to bare rock beds (Landman *et al.,* 2001; FAO, 2004).

Within a more complex environment and through sophisticated production systems, people, livestock, crops and wildlife are competing for increasingly scarce resources. Over time, pressures on land and intensification of agricultural practices will lead to greater susceptibility of farmers to future droughts and floods, resulting in possible further degradation of resources and loss of productivity. Climate uncertainty often leads to conservative strategies which then sacrifice some productivity in order to reduce the risk of losses in poor years (Jones *et al.*, 2000). In southern Africa, the majority of the land area is currently used for agriculture, but this is diminishing with the expansion of afforestation, urban and industrial areas. The combined effects of a growing demand for, and the shrinking area of, agricultural land are forcing agriculture into more marginal areas with ever greater sensitivity to climate variability (Schulze *et al*., 1993). Climate variability could impact agriculture in many ways, including the following:

- planning of farm operations (e.g. land preparation, planting decision, crop selection, fertilizer applications, pests/disease control operations, firebreak burning and infield machinery operation) could become more difficult (Ogallo, *et al.,* 2000; Schulze, 2005),
- yields in rainfed and irrigated agriculture could become more variable than in average years (Kinuthia, 1999; Ogallo *et al.,* 2000),
- pests and vector-borne could spread into areas where they were previously unknown (Olesen and Bindi, 2002),
- the cost of production and marketing could be affected (Ogallo *et al.,* 2000),
- biological diversity would be reduced (Olesen and Bindi, 2002),
- livestock production, such as milk output (volume / unit time) and live-weight gain over a year may face the problem of poor and variable rangeland productivity and desertification processes (Kinuthia, 1999),
- afforestation activities could be negatively affected by deficient rainfall and runoff (Kinuthia, 1999; Glieck, 2000),
- distributions and quantities of fish and sea-foods could change markedly, and

- decimation of wildlife populations could occur through recurrent droughts (Kinuthia, 1999).

In summary, climate variability has both direct and indirect impacts on many water resources operations such as water releases from dams, water rationing for domestic, irrigation and industrial uses, flood warning and hydro-electric generation, and on many agricultural activities such as land preparation, sowing and harvesting times, crop selection, consistency in yield, cost of production, irrigation needs, transportation, storage, pest and disease control and marketing. Many researchers (e.g. Rowlston, 2003; Schulze, 2003) have asserted that the climate in southern Africa varies highly within a year, and from year to year, even without the additional uncertainties brought about by climate change. Understanding the dynamics of climate variability over southern Africa and assessing the impacts of climate extremes on water resources and agriculture provide a base for exploring opportunities in assisting decision makers to better manage their resources through appropriate use of timely and skilful hydro-climatic forecasts. In addition to the implications of climate variability, however, full knowledge of the magnitude and timing of the climate hazards, as well as other factors of vulnerability such as social, technical and financial coping capacities of exposed communities are also needed in order to reduce the loss of human life and economic disruption associated with extreme climatic fluctuations. A brief review of hydrological hazards, followed by an overview of strategies to risk management from a more of hydrological perspective, is presented in the sub-section which follows.

## 2.4 Managing Climate Variability Impacts from a Hydrological Perspective

Climate variability impacts the lives of human beings in many ways, some positive and many negative (Bogardi *et al.,* 2005). The impacts of inter-seasonal and inter-annual climate variability, especially of the more extreme events, are often devastating, especially in less developed regions where technological adaptations are still minimal (Ogallo *et al.,* 2000; Tychon *et al.,* 2003).

Responses to climate variability can be of two broad types (Schulze, 2001; Fischer *et al.,* 2002). The first employs adaptive measures to reduce the impacts of risks

and maximise the benefits and opportunities of climate variability by assessing hazards, vulnerability, and impacts of disasters. The second employs mitigation measures by avoiding hazards and modifying vulnerability to climate hazards (Schulze, 2001). The ultimate objective of climate risk management is to minimise the probability that damage will occur from an event to as close to zero as possible. However, climate variability, which usually cannot be predicted or accounted for, can lead to uncertainty (Hallowes, 2002; UKCIP, 2003). Uncertainty may be construed as a component of risk, if its consequences have an impact on human activities (Hallowes, 2002; UKCIP, 2003). Hydro-climatic forecasting is an integral part of climatic risk management methods that reduces vulnerability to climate hazards in agriculture and water resources management operations (Schulze, 2001). Advance warning of future expected meteorological events (e.g. related to rainfall, temperature, humidity, wind, tornadoes) and agrohydrological variables (e.g. soil moisture status, streamflow, reservoir levels and crop yield) at time scales of days to several months would, therefore, be extremely important in order to enable agricultural and water resource operations to take maximum advantage of any future expected climate anomalies. However, it is essential to first know all the elements contributing to risk and disaster. This section has, therefore, been divided into three sub-sections in order to adequately address the risks of climate variability. The first two sub-sections review the concepts of hazard, disaster, vulnerability and risk within a context of hydrological risk management, while the third sub-section contains a review of the current state of risk management techniques used to prevent and reduce climate related risks.

### 2.4.1 Hazard and Disaster in the Context of Climate Variability

A hazard may be defined as the probability of occurrence of a damaging event in a given period of time (Tychon *et al.,* 2003). Such events could be naturally occurring (e.g. a cyclone, hurricane, earthquake, flood, drought, wildfire), or be human induced (e.g. deforestation), or be accidental (e.g. a dam failure, contamination of water). The event has to have the potential to create damage (Hallowes, 2002; Tychon *et al.,* 2003).

All hazards do not necessarily produce disasters. Disaster is an extreme form of hazard realisation which causes widespread human, material, economic or environmental losses that exceed the ability of the affected community to cope when using its own resources (Hallowes, 2002). Databases from the Centre for Research in the Epidemiology of Disaster, CRED (2002) reveal that more than 2 200 major and minor water related disasters occurred in the world during the period 1991-2000. Of these, floods accounted for half of the total disasters, water-borne and vector disease outbreaks accounted for 28% and drought accounted for 11% of the total disasters. Of these disasters, 35% occurred in Asia, 29% in Africa, 20% in the Americas and 13% in Europe and the rest of Oceania. According to the first report of World Water Assessment Programme, WWAP (2003), the economic losses from water disasters are currently equivalent to 20% of new investment needs in water related infrastructures. Although disasters are associated with a wide range of hazard types, hydro-meteorological events account for a very significant part of disaster loss each year. According to Kishore (2002), hurricanes, flooding, drought, hail storms and storm driven wave action account over 70% of economic losses from all disasters, with a far higher relative incidence experienced in the developing countries.

Figure 2.1 compares reported hydro-meteorological and geophysical disasters worldwide for the period 1991-2000.
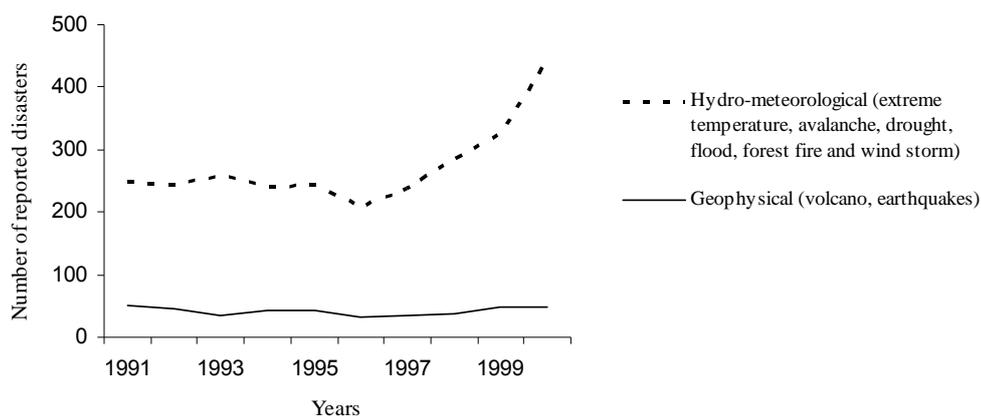


Figure 2.1   Reported hydro-meteorological and geophysical disasters worldwide from 1991-2000 (after CRED, 2002)

While the number of reported disasters associated with geophysical events such as volcanic eruptions and earthquakes remained remarkably constant, those associated with hydro-meteorological events such as floods, drought, forest fires and storms have demonstrated a curve of exponential growth. The number of reported hydro-meteorological disasters in 2001 was approximately double the figure reported in 1996 (Kishore, 2002). This triggers an increased attention in the scientific community to place more emphasis on understanding the cause and extent of hydro-meteorological hazards for a better predictability of hydro-climatic fluctuations at the spatial scales of impacts and at which decisions are made.

Of the hydro-meteorological hazards, droughts and floods are normally recurring events that affect the livelihoods of millions of people in southern Africa. Flooding has resulted in significant damage to infrastructures such as dams, roads, bridges and natural vegetation, in addition to the loss of life. The agricultural sector has also been affected by variations in onset, timing and amount of rainfalls as well as temperature fluctuations, with devastating effects on the economy of the region. These two hydrological hazards are reviewed in more detail below.

### 2.4.1.1 Droughts

Drought is a slow-onset, creeping natural hazard that occurs in virtually all regions of the world. It results in serous economic, social, and environmental impacts (Wilhite, 2000; Wilhite *et al.,* 2000; FAO, 2004). Drought occurs in high as well as in low rainfall areas and is a temporary aberration, in contrast to aridity, which is a permanent feature of the climate and is restricted to areas of low rainfall. Drought is related to the timing of rains (e.g. season of occurrence, delays in the start of the rainy season, occurrence of rain in relation to crop growth stages) and the effectiveness of the rains (e.g. rainfall amounts, intensities, number of rainfall events). Thus, each drought has unique climatic characteristics and differs in its interpretation relative to their impacts (Wilhite *et al.,* 2000). Generally, four categories of drought may be identified (Wilhite and Glantz, 1985; FAO, 2004):

*Meteorological Drought:* This occurs with a reduction in rainfall supply over an extended period compared with the long term average conditions.

*Agricultural Drought:* This implies a reduction in water availability below the optimal level required by a crop during each of its different growth stages, resulting in impaired growth and reduced yields. Agricultural drought is highly influenced by other variables such as the crop water requirement, the water holding capacity of the soil and the magnitude of evaporation.

*Hydrological Drought:* This consists of a substantial reduction/deficit in natural and artificial surface and subsurface water resources in a specified area.

*Socio-Economic Drought*: This implies the direct or indirect impacts of drought on human activities. It relates to extreme climatic events outside of their normal range, which affect production and, thereby, the wider economy. A good example is the 1991/1992 drought in South Africa during which about 50 000 jobs were lost in the agricultural sector alone (FAO, 2004).

The impacts of droughts are largely non-structural and usually spread over a larger geographical area than are impacts from other natural hazards (Wilhite *et al.,* 2000). Like any other hazards, the impacts of drought can be reduced through assessment and mitigation measures (Wilhite *et al.,* 2000; Schulze, 2003).

### 2.4.1.2 Floods

Floods constitute the second frequently recurring natural hazard, causing significant loss of life and damage to property. The major flooding events in southern Africa, each of which lasted for several days during 1984, 1988, 1999 and 2000 are examples of devastating floods in the region. Factors that cause floods are either natural phenomena like tsunamis, earthquakes or extreme rainfall events, or structural failure such as dam or levee failure (Hallowes, 2002). According to Hoyt and Langbein (1995), there are four different types of floods:

*Flash Floods:* This type of localised flooding is hazardous and destructive and occurs within minutes or a few hours of heavy local rainfall or failure of dams, or of levees, or releases of ice jams.

*River Floods:* This is the most common type of flooding. The actual amount of runoff from the rainfall or snowmelt exceeds the capacity of channels or depressions and water overflows onto the adjacent low-lying floodplains.

*Urban Floods:* In most formal urban areas, paved roads and roofs drain into storm water systems which are connected directly the streams. With heavy rain, a large fraction of rain water falling onto impervious areas is not infiltrated into the ground and thus leads to urban floods.

*Coastal Floods:* These usually occur along coastal areas, when there are hurricanes and tropical or other local storms which produce heavy waves, or large waves are created by volcanoes or earthquakes, with ocean water then driven onto the coastal areas and causing coastal floods.

It is important to remember that floods can also play a positive role within our ecosystems and the environment at large, as floodwaters carry nutrients that create fertile floodplains which are important not only for agriculture, but also for various aquatic species (WWAP, 2003).

### 2.4.2   The Multi-Dimensionality of Vulnerability

Vulnerability is a multi-dimensional concept (Hossain, 2001, Bogardi *et al.,* 2005) and varies widely across communities, sectors and regions (Coburn *et al*., 1994). There are various concepts of vulnerability such as social vulnerability, economic vulnerability, environmental vulnerability, vulnerability to food security, vulnerability to natural hazards or vulnerability to climate variability and change (Hossain, 2001).

Although vulnerability is an intuitively simple notation, it is surprisingly difficult to define, quantify and operationalise (Hossain, 2001). Vulnerability has no universally accepted definitions (Downing *et al.,* 2003). Nevertheless, the literature on vulnerability has grown enormously over the past years. It has been defined in many studies according to the consequences of the hazards affecting the operations of concern. A sample of definitions on vulnerability is summarised below:

- Blaikie *et al.* (1994), cited by Hossain (2001), define vulnerability in terms of sociological characteristics, such as ethnicity, religion, gender and age which influence access to power and resources.

- Bohle *et al.* (1994) define vulnerability as "an aggregate measure of human welfare that integrates environmental, social, economic and political exposure to a range of harmful perturbations".

- Coburn *et al.* (1994) regard vulnerability as "the degree of loss to a given element at risk, or set of elements, resulting from a given hazard at a given severity level".

- In a particular context of Food Security Information and Early Warnings System, FSIEWS (2000), vulnerability is defined as "a measure of the susceptibility of some groups of persons or regions to food insecurity".

- Downing and Lüdeke (2002) used the term vulnerability in the sense of assessment methods. They defined vulnerability as "a set of relationships between exposure to an extent threat (e.g. extended drought) and its consequences (e.g. human mortality due to starvation)".

- The IPCC (2001) defines vulnerability as "the degree to which a system is susceptible, or unable to cope with, adverse effects of climate change, including climate variations and extremes". Under this framework, a highly vulnerable system would be one that is highly sensitive to modest changes in climate.

- In contrast to the definitions presented by the IPCC (2001), the International Strategy for Disaster Reduction, ISDR (2004) defines vulnerability as "the set of conditions and processes resulting from physical, social, economic and environmental factors, which increase the susceptibility of a community to the impact of hazards".

Whatever definition of vulnerability is adopted, within the context of climate variability it is a function of the characteristics, magnitude, and rate of climate variation to which a system is exposed, and its sensitivity and adaptive capacity (Schulze, 2003; Bogardi *et al.,* 2005). It deals mainly with two elements, *viz.* exposure to hazard and the coping ability of the people. People having more capability to cope with extreme events are naturally less vulnerable to a given hazard (Hossain, 2001; Schulze, 2003; Bogardi *et al.,* 2005). This was

substantiated by Hossain (2001) who reported that the people in the USA are three times more exposed to certain natural hazards than those of Bangladesh, but they are far less affected because of their better coping capacity. The degree of vulnerability can, therefore, be considered as the combination of existing conditions of such communities, or systems that make them prone to being affected when the external event manifests itself (Schulze, 2003; Bogardi *et al.,* 2005).

According to Hossain (2001) vulnerability relates to the consequences of perturbations, rather than the agent of the perturbation *per se*. People are more vulnerable to loss of life, livelihood, assets and income than to specific agents of hazards such as drought and floods. Thus vulnerability is explicitly a social phenomenon, a threat to human value system rather than the nature or magnitude of the hazard itself.

### 2.4.3  The Multi-Dimensionality of Risk

Risk is the product of hazard and vulnerability. It increases with increasing occurrence of damaging events and with the vulnerability of a population (Schulze, 2003; Tychon *et al.,* 2003; Baethgen *et al.,* 2004). It may be defined as:

Risks as a result of disasters arise from combinations of hazardousness and vulnerability that vary over seasonal to decadal timescales as well as geographically. Smith (1996) identified three possible scenarios that give rise to risk over time, and Schulze (2003) added one more scenario that can increase or decrease the risk over time. These four scenarios are illustrated in Figure 2.2.

- In Scenario 1, the band of tolerance to a hazard and its variability remain constant, but the mean value of the hazard changes to beyond the tolerance limits as result of variations in the physical system. The frequency of extreme events at one end of the tolerance scale also increases, as would be the case of decreases in runoff associated with changes in land use such as upstream afforestation.

- Scenario 2 represents a scenario where both the mean of the hazard and the band of tolerance remain constant, but the variability increases (e.g. enhanced variability due to climate change). In Scenario 2, the frequency of damage producing events increases at both ends of the tolerance scale.

- In Scenario 3 the physical variable (e.g. runoff) does not change, but the band of tolerance narrows due to an increase in vulnerability to extreme climate hazards (e.g. vulnerability increases as poor people migrate to floodplains and other flood-prone areas). In this particular scenario, the frequency of extreme events increases at both end of the tolerance scale.

- Scenario 4 shows a sudden change in both the variability and tolerance of a system, as would be the case with alterations in downstream flow characteristics following the construction of a dam across a river. Risk may increase or decrease in this particular scenario.
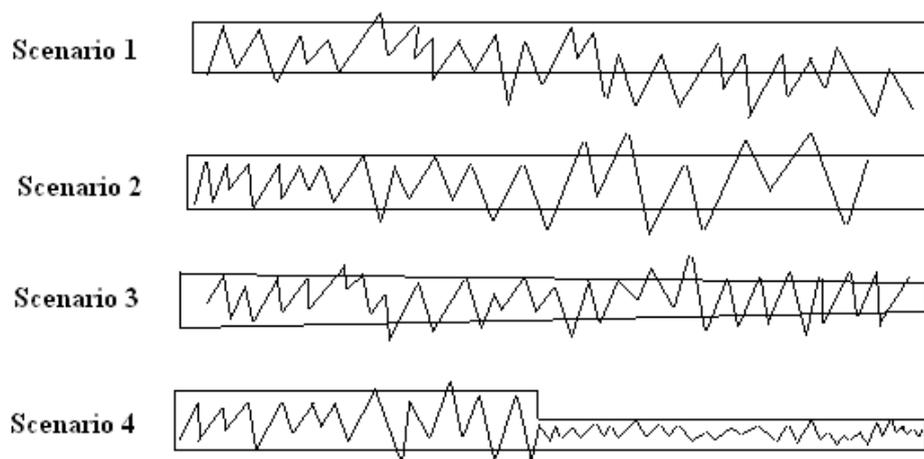


Figure 2.2    A schematic illustration in which risk varies as a result of changes in the physical and socio-economic systems (after, Smith, 1996; Schulze, 2003)

### 2.4.4   Risk Management from a Hydrological Perspective

The term risk refers to the expected losses from a given hazard to a given element of risk, over a specified future time period. According to the way in which the element at risk is defined, the risk may be measured in terms of expected economic losses, or in terms of numbers of lives lost, or the extent of physical damage to property (Coburn *et al.,* 1994). Natural risks may be unavoidable, but

with better understanding, planning and prevention (i.e. risk management) their impacts can be reduced (WWAP, 2003).

In agricultural activities, risk management requires the distinction between, and understanding of, different types of risks (e.g. production risks, market risks, institutional risks, financial risks, asset risks), as the actions required to reduce each of them are different (Baethgen *et al.,* 2004). Kundzewicz (2001) states that "it is impossible to design a system that never fails. What is needed is to design a system that fails in a safe way". This acknowledges that in water utilities, both the possible structural failure (e.g. of a dam) and non-structural failure (e.g. unreliable or unsafe water supply to water users) need to be considered in order to make good management decisions in situations where adversity and loss are possible (Plate, 2002; Schulze, 2003; WWAP, 2003).

There are many risk management paradigms (e.g. Grigg, 2000; Plate, 2002; UNDP, 2002; Baethgen *et al.,* 2004) to consider in hydrological hazards which threaten agricultural and water resource systems. However, these actions reduce to a few essential steps (Levitt, 1997), *viz.*

- identifying present and future hazards,
- measuring hazards,
- reducing hazards to their lowest practical levels, and
- devising means to deal with unavoidable and uncontrollable risks.

It is beyond the scope of this document to review each step of risk management. However, if hydro-climatic forecasting is going to be worth pursuing, a more complete understanding of the full procedures and practices to risk management from a more hydrological perspective is required. In attempts to the respond the negative impacts of hydrological hazards such as flooding and drought, Schulze (2001), using multiple source of information, developed a schematic overview of approaches to risk management from a more hydrological perspective (Figure 2.3), and classified the elements of risk management under two broad themes, *viz.* risk assessment and risk mitigation and control.
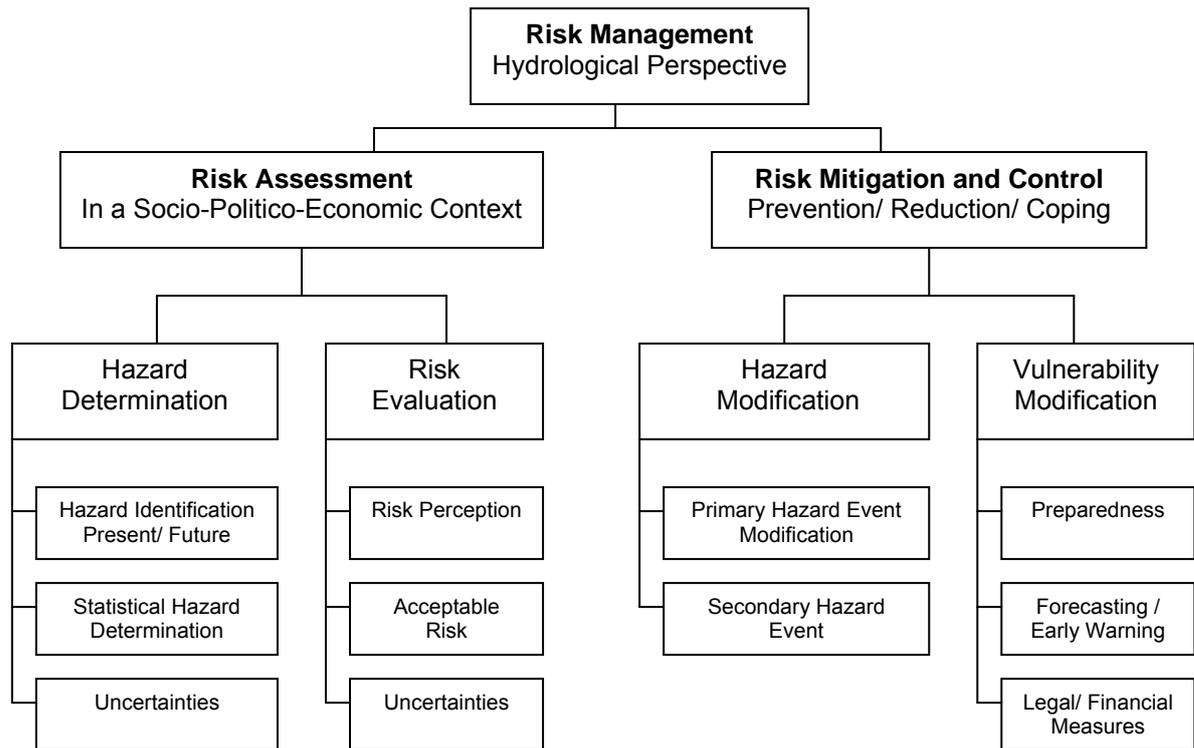
Figure 2.3    A schematic overview of approaches to risk management, developed from multiple sources (after Schulze, 2001)

### 2.4.4.1   Risk assessment within a socio-politico-economic context

The term risk assessment has been defined by many researchers (e.g. Smith, 1996; Hossain, 2001; Plate, 2002; UNDP, 2002; WWAP, 2003) in many various ways. Schulze (2003) recently defined risk assessment from a more generic perspective as "the process of assigning magnitudes and probabilities to the adverse effects of natural catastrophes or human activities using rigorous, formal and consistent forms of measurement and testing, to quantify the relationship between the hazard event and the responding effects, while acknowledging the inherent uncertainties involved". Risk assessment therefore includes objective determinations of both present hazards and future one (e.g. through increased variability in the future climates), and risk evaluation, which comprises of risk perception, acceptable risks and uncertainties (Schulze, 2001; Schulze, 2003). Risk assessment forms the necessary basis for the development of risk mitigation and control measures (WWAP, 2003).

### 2.4.4.2 Risk mitigation and control: Modifying hazards and vulnerability

Risk mitigation and control make up the second major component of risk management (Plate, 2002; Schulze, 2003), in which alternatives for possible risk measures are considered. Strategies for the mitigation and control of risk can be broadly classified into two categories (Smith, 1996; Schulze; 2001; Schulze; 2003), *viz.* hazard modification and vulnerability modification. According to Smith (1996) and Schulze (2003), hazard modification concentrates on modifying the physical processes that create the hazard either by modifying the primary processes (e.g. cloud seeding in situations of trying to modify rainfall) or the secondary processes (e.g. diverting the floodwater, land use planning).

Vulnerability modification is primarily concerned with reducing the impact of the event by rendering the human environment less vulnerable to, more prepared for, the event. Vulnerability modification has three different components (cf. Figure 2.3), *viz.* preparedness, forecasting/early warning systems and legal/financial measures. Preparedness includes the pre-arranged emergency measures which should be taken to minimise the loss of life and property damage following the onset of a hazard (Schulze, 2003). Forecasting/early warning, which has the potential to modify vulnerability in shorter and longer term planning, is briefly reviewed in the next chapter. Legal/financial instruments are designed to prevent certain activities in areas of high risks (e.g. building shacks in floodplain areas), or to provide aid that is able to accelerate the recovery of affected communities (e.g. insurance; Smith, 1996; Schulze, 2003).

## 2.5    Concluding Thoughts

This chapter commenced by defining climate variability and climate changes in order to obtain a better understanding of the mechanisms of climate variability and their relationships with hydrological extremes such as floods and droughts. This was followed by a discussion on climatic variations over southern Africa and their consequences on two of the most climate sensitive sectors, *viz.* the water and agriculture sectors. Even though it is difficult to quantify how extreme events have impacted these climate sensitive sectors, this chapter contained brief discussion on the impacts that climate variability can have on humans, property and the wider

economy by affecting the hydrological cycle. In many places in southern Africa, both agriculture and water resources face high risks, owing to the growing competition for water across economic sectors, and the natural variability of the hydrological system. The impacts may vary from one geographical region to the next, depending on the degree of vulnerability associated with different communities and societies. In order to manage such risks, current approaches to risk management were presented, first by defining and describing basic concepts such as hazard, vulnerability and risk. Finally, a conceptual framework for risk management, as presented by Schulze (2001), was reviewed in order to highlight the importance of the management strategies, procedures and practices that seek to minimise risks from a hydrological perspective.

Hydro-climatic forecasting is a crucial part of risk management and is discussed briefly in the following Chapter, in which an attempt is made to distinguish between key points on forecasting such as types of forecasts (e.g. short, medium, long), certainty of the forecasts (e.g. skill, accuracy, reliability) and the lead times of the forecast (e.g. days, months), as well as applications and limitations of the different types of forecasts.

# 3. FORECASTING AS A STRATEGY FOR VULNERABILITY MODIFICATION IN THE MANAGEMENT OF AGRICULTURAL AND WATER RESOURCE SYSTEMS

Vulnerability is not a consequence only of the year-to-year variability of climate *per se*, but also of its less predictability. Many critical agricultural and water resource decisions that interact with climatic conditions must be made in advance, based on available climate information and assumptions (Hansen, 2002). The emerging ability to provide timely and skilful short, medium and longer term hydro-climatic forecasts has the potential to reduce risk in the long term and to provide valuable support to meet the competing demands for increasingly scarce fresh water and agricultural resources. The incorporation of forecasting within the framework of risk management has, therefore, been acknowledged to play a vital role in modifying decisions, to either prepare for expected adverse conditions or to take advantage of expected favourable conditions (Hammer *et al.,* 2001; Hansen, 2002). Connecting climate forecasts with applications such as the management of agricultural and water resources decisions is, however, not straightforward, and remains an area in which much needs yet to be learned. While there has been a growing literature on potential applications of climate forecasts to mitigate risks in agricultural and water resources systems, there has been relatively little research done on the issue of applying climate forecasts in decisions to modify the vulnerability of humans and properties to the adverse impacts of climate variability. In this chapter, current forecast types and techniques are briefly reviewed.

## 3.1 Types of Forecasting

Although a considerable literature distinguishes between the terms forecasts, outlooks and predictions (e.g. Maidment, 1993; Schulze, 2003), in practice there are no naming conventions (Hartmann *et al.,* 1999). Forecasting is generally considered as being the estimation of conditions at a specific future time, or during a specific time interval, while prediction is the estimation of future conditions, without reference to a specific time. For very long lead times, however, the distinction between forecasting and predictions is blurred, as forecasting accuracy decreases with increases in lead times (Lettenmaier and Wood, 1993). There should, however, be a clear distinction between so-called official, operational,

experimental and research forecasts, as users can have direct access to all of these forecast types (Hartmann *et al.,* 1999). Forecasting techniques exist along a continuum of sophistication, ranging from simple implicit subjective processes (e.g. "feeling" that tomorrow's condition will be much like today's condition), to complex objective techniques which require many types of data, representations of the physical processes, and teams of scientific experts (Hartmann *et al.,* 1999). The broad range of forecasting encompasses various products (e.g. weather forecasts, crop forecasts, fire hazard forecasts, hydrological forecasts, marine forecasts, domestic aviation forecasts), with time scales ranging from minutes to seasons, and hence lead times from minutes to over a year. Therefore, the wide variety of forecast products and techniques can be categorised, according to several different perspectives (Hartmann *et al.,* 1999) which are discussed below.

### 3.1.1   Weather vs. Climate Forecasts

Although this chapter concentrates on agrohydrological forecasts, the distinction between weather, climate and agrohydrological forecasts is first highlighted in order to obtain a better understanding on issues of applying the various forecasting techniques and products.

According to commonly used definitions, weather forecasts track the movement and evolution of specific air masses and cover periods approaching less than one month, while climate forecasts are usually considered as extended weather outlooks and cover periods of one month and longer, i.e. climate forecasts describe the predictability of weather statistics, and not day-to-day variations in weather (Hartmann *et al.,* 1999; Kabat and Bates, 2002).  Because the climate system is so complex, it is almost impossible to take all the factors that determine the future seasonal climate into account. Therefore, climate forecasts are generally provided in terms of the probability that the rainfall or temperature will be either below normal, near normal or above normal (Kabat and Bates, 2002).

### 3.1.2   Agrohydrological Forecasts: Types and Potential Applications

Weather and climate forecasts are critical inputs to agrohydrological forecasts. Agrohydrological forecasts are predictable on scales equivalent to both weather

and climate forecasts owing to integrative behaviour of hydrological processes (Hartmann *et al.,* 1999). Although any classification of forecasts is subject to some overlap, there are four types of agrohydrological forecasting, depending on the lead times. These four types of forecasts and the potential usefulness of such forecasts are briefly presented on the sub-sections which follow.

### 3.1.2.1 *Near real time agrohydrological forecasts*

The temporal coverage of near real time agrohydrological forecasts varies from hourly to daily, with lead times from minutes up to several days in advance (Lettenmaier and Wood, 1993). These forecasts are most often used for flood warning purposes and for real time water resources and agricultural operations. Real time agrohydrological information is very important in areas of fast response because, as people place more pressure on the vulnerable areas (e.g. floodplains) for habitation or agriculture or other businesses, so there is a greater potential for loss of life and damage to property by catastrophic events such as flash floods (Pegram and Sinclair, 2002). The information received from automatic weather stations, together with information obtained from satellite and radar images, is integrated into hydrological models to produce near real time agrohydrological forecasts. Pegram and Sinclair (2002), for example, developed a linear catchment model for real time flood forecasting which accepts satellite, radar and raingauge data as input. The model has been installed in the Umgeni Water headquarter offices in Pietermaritzburg to provide flood warnings for disaster management operations (Umgeni Water is the water utility serving the Durban-Pietermaritzburg region with bulk water supplies).

Since structural measures are often insufficient to reduce risks associated with extreme events at the required local level, an important role is played by non-structural measures (Toth *et al.,* 1999). Some of the potential applications of near real time agrohydrological forecasts in agriculture and water resources include:

- dissemination of warning messages regarding the extent of extreme events, such as floods (Hossain, 2003),
- evacuation of people and mobile assets (e.g. pump equipment and machinery) from threatened high risk areas (Hossain, 2003; Schulze, 2005),

31

- providing information on the status of inflows into dams, such as timing of peak flows (Schulze, 2005),

- reservoir safety releases (Schulze, 2005),

- mobilisation of resources and planning relief and rehabilitation measures (Hossain, 2003),

- hydro-power scheduling (Lettenmaier and Wood, 1993), and

- precautionary measures (e.g. establishing embankments) to divert floods either into, or away from, agricultural areas, depending on the soil moisture status of the area.

### 3.1.2.2 *Short and medium term agrohydrological forecasts*

Short term agrohydrological forecasts are taken to be those with a temporal coverage from one day up to about three days, while medium agrohydrological forecasts cover time scales usually up to two weeks. The lead times of short and medium term agrohydrological forecasts vary from a day up to several days or a few weeks. Such forecasts are useful to making adjustments to agricultural planning and water management, for example, by allowing a farming community to react on time, especially at planting and harvesting times (Webster and Grossman, 2003). In addition, it provides a sufficient forewarning for catchment and disaster managers by allowing early decisions for flood mitigation and disaster management.

Some of the potential actions that can be taken by responding to the short and medium term agrohydrological forecasts in water resource operations include the following (Schulze, 2005):

- reservoir regulation decisions (e.g. formulation of reservoir release strategies),

- reservoir safety releases,

- environmental flow releases (e.g. freshettes),

- irrigation scheduling, and

- planning for water poverty relief, rehabilitation and reconstruction.

Some of the potential applications of short and medium term agrohydrological forecasts in agricultural operations are (Schulze, 2005):

- tillage, planting, transplanting and harvesting decisions,
- fertilizer and pest control application decisions,
- taking precautionary measures to protect assets, livestock and agricultural infrastructures, such as forage silos, embankments, roads etc (Hossain, 2003),
- firebreak burning operations,
- labour and equipment planning, and
- crop yield estimates.

### 3.1.2.3 Long term (seasonal) agrohydrological forecasts

Long term agrohydrological forecasts are those with longer lead times and time scales, usually up to several months ahead. At present, little forecast skill is possible for agrohydrological variables when forecast lead times extend beyond three months (Lettenmaier and Wood, 1993; SAWS, 2005). Reliable long term agrohydrological forecasts can improve the decisions in the management of water resource and agricultural systems by reducing risks associated with inter-seasonal and inter-annual climate variability. Some of the potential uses and benefits of long term agrohydrological forecasts include the following:

- reservoir management, such as status reviews and/or curtailment planning (Chiew *et al.,* 2003; Schulze, 2005),
- allocation of irrigation water and planning of irrigation timing, depending on forecasts such as soil moisture and streamflow (Lettenmaier and Wood, 1993; Schulze, 2005),
- greater efficiency in power generation and negotiation of hydro-power sales contracts (Lettenmaier and Wood, 1993; Collischonn *et al.,* 2005),
- planning for water poverty relief, rehabilitation and reconstruction (Hossain, 2003; Schulze, 2005),
- evaluation and implementation of mitigation measures, such as water conservation during droughts (Lettenmaier and Wood, 1993), and

- environmental flow releases, which depend on forecasted reservoir status or streamflow levels (Schulze, 2005).

Long term agrohydrological forecasts can assist farmers, agribusiness managers and governments in many ways to best manage their properties, strategies and short and long term policies. The potential applications in agricultural activities include (Klopper, 1999; Hossain, 2003; Schulze, 2005):

- crop variety selection (e.g. introduction of fast growing varieties),
- planting and harvesting decisions (delayed or earlier),
- conservative use of fertilizers, insecticides and pesticides,
- maintenance of conservation structures,
- fertilizer, planting and harvesting equipment orders,
- reducing stock (e.g. selling cattle before the drought season started),
- labour and equipment planning,
- transport and storage scheduling,
- crop yield estimates,
- planning national food import, storage and distribution programmes,
- adjustment of risk profiles, and
- development of drought and flood response policies.

### 3.1.3 Conclusions on Agrohydrological Forecasts

This section has outlined the potential measures that may be taken in agricultural and water resource management to decrease unwanted impacts and take advantage of expected favourable conditions, provided that skilful and timely agrohydrological forecasts, ranging from days to seasons were available in advance at the location of concern. The ability to forecast climate variability has progressed significantly in recent years. Despite the progress made in climate predictions, however, there appears little explicit application as yet of climate forecasts in the management of water resources and agricultural operations. The challenge to hydrologists and agriculturalists is on how to link the weather and climate forecasts with agrohydrological models in order for decision makers to benefit more fully from those climate forecasts. As a step towards addressing this

problem, the current techniques used to obtain agrohydrological forecasts from climate forecasts are briefly reviewed in order to obtain an idea of the complexity, uncertainty and applicability of these techniques that would help in the selection of an appropriate technique (or techniques) for use in the operation of agricultural and water resources decisions.

## 3.2    Agrohydrological Forecasting Techniques

Agrohydrological forecasting techniques may be categorised into two types, *viz.* short term forecasting and long term forecasting techniques.

### 3.2.1    Techniques for Short Term Agrohydrological Forecasting

Short term agrohydrological forecasts (e.g. streamflows, peak flows) may be made either by channel routing methods or by rainfall-runoff modelling. Channel routing methods simulate the routing of water down a river channel by using various hydrodynamic equations. Of the channel routing models the Dynamic Wave, Diffusion Equation, Muskingum, Kinematic Wave and Impulse Response Function are commonly used techniques around the world (Lettenmaier and Wood, 1993). Rainfall-runoff techniques represent the various water storage terms (e.g. interception, soil moisture and surface storage) and flux terms (e.g. infiltration, evapotranspiration, surface runoff, interflow, baseflow) with varying levels of complexity (Lettenmaier and Wood, 1993). There are numerous rainfall-runoff methods available. It is beyond the scope of this document to review each of these methods. However, it is helpful to discuss the criteria required to select an appropriate method for any particular application.

In many applications forecasts based on channel routing are preferred because of their simplicity (Collischonn *et al.,* 2005). Nevertheless, forecasting by rainfall-runoff models has been found to be more accurate, especially whenever the forecast lead time is significantly longer than the time taken to route flow along a river channel (Lettenmaier and Wood, 1993; Collischonn *et al.,* 2005). According to Lettenmaier and Wood (1993) the selection of short term hydrological forecasting methods depends on two criteria. The first relates the required forecast lead time, $T_f$, to the total time of concentration of the catchment at the forecast point. The total time of concentration includes both the hydrological response time,

$T_c$, which is the time of travel from the farthest point in the catchment to the forecast point, and the flood propagation time through the channel or river system, $T_r$. The second criterion is the ratio of the spatial scale of the meteorological event (e.g. rainfall) to the spatial scale of the catchment, $R_s$. Lettenmaier and Wood (1993) clarified by considering four cases.

*Case 1:* If the forecast lead time is larger than the time of concentration ($T_f > T_c + T_r$), then rainfall observations alone are not sufficient for forecasting purposes, since some of the water which is included in the flow forecast has yet to fall as precipitation on the catchment at the time the forecast is made. Some estimate of future rainfall is, therefore, required if flow forecasts are to be compiled up to the end of the lead time. Quantitative Precipitation Forecasts (QPFs), which can be supplied by Numerical Weather Prediction (NWP) models, are suitable for this purpose, provided that these are sufficiently accurate (Collischonn *et al.,* 2005).

*Case 2:* If the forecast lead time is shorter than the total time of concentration and the total time of concentration is dominated by the routing time of the flood wave through the channel system ($T_f < T_c + T_r$ and $T_c < T_r$), then streamflow forecasts can be based on observed flows at upstream gauged locations. For such systems, streamflow forecasts can be based on channel routing methods. This is often the situation for large river systems.

*Case 3:* If the forecast lead time is shorter than the total time of concentration and the total time of concentration is dominated by the hydrological response time of the catchment ($T_f < T_c + T_r$ and $T_r < T_c$), then hydrological forecasts can be simulated with rainfall-runoff models based on observed rainfall from a network of raingauges. This is a typical situation for small catchments and urban areas.

*Case 4:* If the ratio of the spatial scale of the meteorological event (e.g. rainfall) to the spatial scale of the catchment $R_s$ is < 0.7, then rainfall-runoff models that assume spatially uniform rainfall will not be able to produce accurate hydrological forecasts. This is a problem in large catchments. In such cases the catchment should be divided into subcatchments. The upstream catchment can be then forecasted by rainfall-runoff models, while the downstream channel flows can be forecasted using the channel routing methods.

### 3.2.2  Techniques for Long Term Agrohydrological Forecasting

Long term, i.e. seasonal, agrohydrological forecasting techniques in use nowadays are of two types (Landman *et al.,* 2001, Hallowes, 2002), *viz.*

- the direct method, which consists of statistical downscaling of atmospheric variability to agrohydrological variables (e.g. streamflow, soil moisture), and
- the indirect method, in which agrohydrological variables are generated from climate forecasts using a hydrological model.

Before reviewing these two methods, it is useful to first highlight the sources and current approaches behind the seasonal climate forecasts.

### 3.2.2.1  Approaches to seasonal climate forecasting

Approaches to seasonal forecasting rely on the fact that lower boundary forcing, which gives rise to atmospheric perturbations, evolves more slowly than the atmospheric perturbations and that the response of the atmosphere to this forcing is detectable (Murphy *et al.,* 2001). For long term atmospheric predictability, oceanic boundary forcing is the most important type of boundary forcing and it is normally monitored in the form of sea surface temperature (SST) and Southern Oscillation Index (SOI) anomalies (Landman *et al.,* 2001; Murphy *et al.,* 2001). A recurrent, quasi-periodic appearance of warm SSTs in the central and eastern equatorial Pacific Ocean is the El Niño phenomenon. An opposite phase, when the water in the central and eastern equatorial Pacific Ocean is cooler, it is termed the La Liña. The underlying atmospheric activity that drives El Niño events is the Southern Oscillation (Piechota *et al.,* 1998). El Niño and the Southern Oscillation combine to form the El Niño - Southern Oscillation, or ENSO phenomenon (UNEP, 1992; Sarachik, 1996; Piechota *et al.,* 1998).

The ENSO phenomenon remains the foundation of long term (seasonal) climate forecasting methods (Hammer *et al.*, 1996). Although the ENSO is the main source of atmospheric predictability at global scale, seasonal climate variability around the world has also been shown to be related to the North Atlantic Oscillation, NAO, the Pacific Decadal Oscillation, PDO, and the Indian Ocean

Dipole, IOD (Landman *et al.,* 2001; Kabat and Bates, 2002). The NAO, usually defined by an oscillation in sea-level pressure between stations in Iceland and the Azores, is responsible for a large portion of climate anomalies in Europe, North Africa, the Middle East and eastern North America (Hurrell, 1995; Hammer *et al.,* 2001). Southern African seasonal rainfall variability is also well identified as having connections to sea surface temperature anomalies in the central south Atlantic and western equatorial Indian Oceans (Mason, 1990; Landman and Klopper, 1998; Landman *et al.,* 2001).

### 3.2.2.2  Climate forecasting techniques

Climate forecasting techniques can be broadly divided into two categories, *viz.* statistical and dynamical (Hammer *et al.,* 1996; Landman *et al.,* 2001; Murphy *et al.,* 2001). A wide range of statistical techniques has been developed which relate predictand (e.g. seasonal rainfall) to predictor variables (e.g. SST, SOI). The commonly used statistical models include a linear statistical model, canonical correlation analysis (e.g. Landman and Mason, 1999), quadratic discriminant analysis, which is a non-linear statistical model (e.g. Mason, 1998) and Artifical Neural Networks, which are capable of modelling extremely complex functions (e.g. Hastenrath *et al.,* 1995). All statistical models are based on the basic principle of minimising the least-squares between predicted and observed variables (Murphy *et al.,* 2001).

Dynamical methods employ a fundamentally different approach to seasonal prediction (Hammer *et al.,* 1996; Landman *et al.,* 2001; Murphy *et al.,* 2001). In the case of the statistical models, SST persistence is assumed and observed SSTs are used to forecast the predictand. Dynamical methods, on the other hand, use a fully coupled ocean-atmosphere to mathematically model both the ocean and the atmosphere, thereby allowing SSTs to evolve (Hammer *et al.,* 1996). These evolved fields are then incorporated in a physically based model of the atmosphere system, known as General Circulation Models (GCMs), which are then used to simulate atmospheric variables.

The output from the statistical or dynamic models can be then used to generate hydrological variables (e.g. streamflow, soil moisture) either directly by

downscaling model outputs statistically, or indirectly by employing agrohydrological models (Figure 3.1). The two methods are described in more detail in the sub-sections which follow.

### 3.2.2.3  The direct method

The direct method employs stochastic methods to generate agrohydrological forecasts (e.g. streamflow), either based only on past history of measured agrohydrological variables (e.g. streamflow, reservoir storage) or from predictor variables (e.g. SST), thereby preserving the statistical properties that exist between the input and output variables. In the latter case, the predictors may be obtained either from the statistical climate models or the dynamical models or both, the output of which is then downscaled statistically to agrohydrological variables such as streamflow (cf. Figure 3.1).  The most widely used stochastic models that are employed to generate agrohydrological forecasts based only on historical observed agrohydrological variables are time series models.
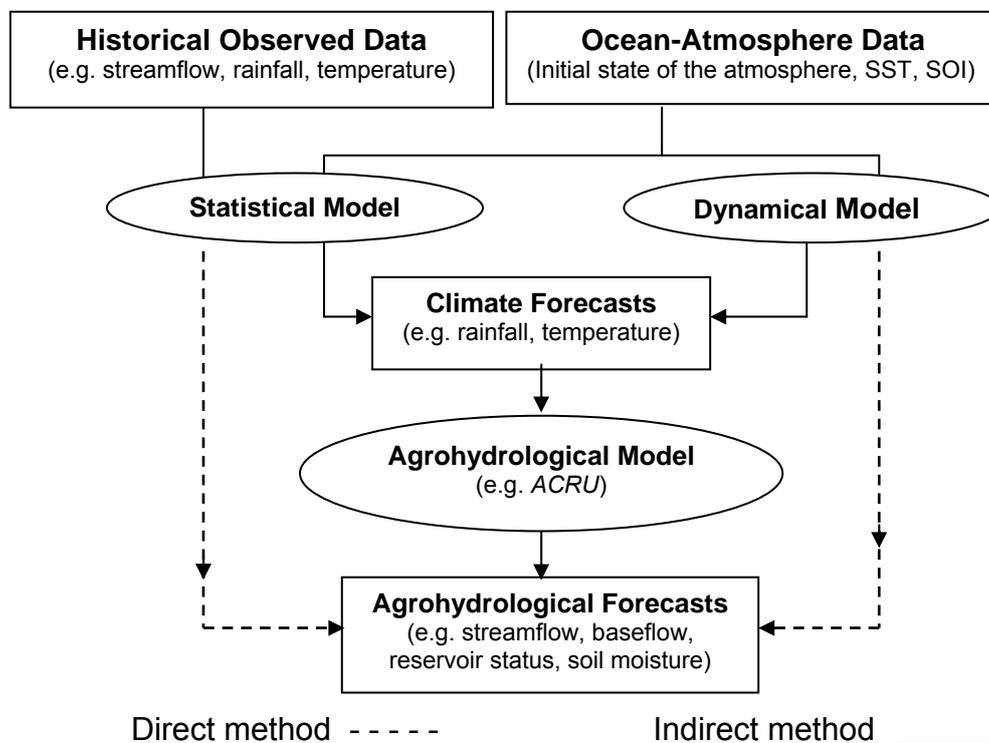


Figure 3.1    Methodologies used to generate long term agrohydrological forecasts (after Hallowes, 2002; with modification by the author)

The use of Markov chains is a good example of time series models that have been applied for the generations of short, medium and long synthetic agrohydrological variables (Yapo *et al.,* 1993). In the recent past, considerable attention has been focused on the time series models because of their simplicity. Yapo *et al.* (1993), for example, developed a streamflow forecasting model based on a Markov chain approach. The model relies on past information of streamflows to make forecasts. The concept here is that the forecasts consist of the probabilities that the next streamflow will be within specified flow ranges, where the probabilities are conditioned on the present state of the river.

The second option employs statistical models that have the ability to seek relationship between the predictand and predictors. As has been stated above, the predictors may be obtained from either, or both, the statistical climate models and the dynamical models. The basic premise is that if rainfall is strongly related to the predictors such as SST or SOI, then agrohydrological variables such as streamflow may also be related to these predictors in the same manner. Piechota *et al.* (1998) developed a seasonal forecast model based partially on the relationship between ENSO and streamflow. The model uses an optimal linear combination of four empirical statistical models, *viz.* climatology, persistence Linear Discriminant Analysis (LDA), SOI LDA and winter Sea Surface Temperature, $SST_w$, to obtain a consensus forecast. Similarly, Landman *et al.* (2001) developed a multi-tiered procedure to simulate real time operational seasonal forecasts of categorised (below-normal, near-normal, above-normal) streamflows at the inlets of 12 dams of the Vaal and upper Thukela river catchments in South Africa. A GCM model called COLA T30 was used to simulate atmospheric variability over southern Africa, the output of which was then statistically downscaled to streamflow using a perfect prognosis approach.

The major limitations of the statistical models are that they depend only on climatic factors such as precipitation and evapotranspiration. The non-climatic factors such as vegetation cover, land management, or soil characteristics which play an important role in the hydrological processes (Schulze, 1997) are not considered in the statistical relationship between the atmospheric variability and hydrological variables. The accuracy of such models also depends heavily on the quality of

observed data. Most of these models require reasonable periods of good quality of climatic and flow data (Kienzle *et al.,* 1997; Hallowes, 2002). Most statistical models are based on linear statistics and do not take into account the non-linear physical processes (Landman *et al.,* 2001; Hallowes, 2002). However, many important climate processes demonstrate strong non-linearities, and the forecast skill of statistical models is restricted owing to the exclusion of these important processes (Barnston *et al.,* 1994; Landman *et al.,* 2001). Algorithms are usually identified for a specific location, making algorithm transfers to other areas unsuitable (Kienzle *et al.,* 1997).

### 3.2.2.4  The indirect method

This method employs agrohydrological models to generate agrohydrological forecasts from climate forecasts (Hallowes, 2002). The models run up to the time of forecast with observed climatological inputs (e.g. rainfall, temperature). During the forecast period the climatological inputs can be of several forms (Lettenmaier and Wood, 1993). They are obtained either from:

- a forecast of future conditions predicted by climate models;
- observed historical data set which is similar to conditions up to the time of forecast; or
- models that generate synthetic sequences of future weather conditions.

A major advantage of the indirect method is the representation of physical hydrological processes in the model which can simulate the conversion of rainfall to runoff by using well established mathematical relationships. This method therefore tends to be applicable under diverse catchment conditions (Hallowes, 2002). Accurate information about the climate forecasts and representations of rainfall-runoff processes by the hydrological models are critical to the success of the indirect method of forecasting agrohydrological variables. Conceptual models, for example, represent a partial understanding of rainfall-runoff processes, in which the various catchment characteristics (e.g. soils, vegetation, terrain) are usually spatially averaged (Schulze, 1998). Deterministic models, on the other hand, represent the various spatial components and temporal variations in catchment hydrological processes (Kienzle *et al.,* 1997). Consequently,

deterministic models are expected to display relatively more accurate agrohydrological forecasts than simpler conceptual models. Although most physical-conceptual based models require reasonable periods of good quality flow data for verification purposes, they are considerably less reliant on good flow data sets when compared to statistical models (Hallowes, 2002).

In hydrologically heterogeneous regions such as southern Africa, forecasting by a suitable agrohydrological model would lead to improved estimates of agrohydrological outputs such as streamflows, reservoir status, irrigation demands and soil moisture status when compared with forecasting using the simple conventional statistical methods. Despite the many advantages of the indirect method, the use of rainfall forecasts as input to an agrohydrological model, thereby enabling the forecasting of agrohydrological variables across a range of time scales and lead times, is relatively new.  As has been mentioned in the introduction, the main reasons for this are due to the gaps that exist between the products of weather and climate forecasts, and their practical use in hydrological/crop models.

## 3.3    Conclusions

In this chapter the review was on types of agrohydrological forecasts and their potential applications to reduce vulnerability in the management of water resource and agricultural systems. Many other water related sectors in southern Africa may also benefit from the provision of an integrated time-varying agrohydrological forecast system. With the aim to understand the broad range of forecasting systems, a brief review was also conducted of the current state of climate and agrohydrological forecasting techniques. Key elements of the review were the techniques used to generate short and long term agrohydrological forecasts, and their qualities and limitations when used for operational decisions.

In Chapter 4 prerequisites, approaches and challenges related to forecasting are reviewed briefly for effective use of agrohydrological forecasts in agricultural and water resources decisions.

# 4. CHALLENGES AND APPROACHES TO MAXIMISE BENEFITS FROM HYDRO-CLIMATIC FORECASTS

In Chapter 3, a brief description was provided on the role that forecasting could play in modifying decisions, to either reduce expected adverse conditions or to take advantage of favourable conditions. However, the availability of hydro-climatic forecasts *per se* is not sufficient to ensure that decision makers will mitigate the potential negative consequences of climate variability or, alternatively, capitalise on potentially beneficial events (Podestá *et al.,* 2002). Benefits only arise when the use of hydro-climatic forecasts results in decisions that improve management of climate related risks in water resources and agricultural operations. According to many researchers (e.g. Pielke, 2000; Hansen, 2002; Podestá *et al.,* 2002), sustained and effective application of hydro-climatic forecasts requires three components to occur simultaneously, *viz.*

- the generation of skilful and timely hydro-climatic forecasts (i.e. forecast quality),
- the effective communication of that information, and
- the application of that climate information to modify decisions or policies (i.e. forecast value).

In practice however, the application of these components is not straightforward, let alone applying them simultaneously. Thus, it is important to explore the prerequisites, approaches and impediments associated with each of these components as a means of maximising the benefits from hydro-climatic forecasts.

## 4.1    Forecast Quality

Hydro-climatic forecasts should, in the first instance, be statistically valid (Ritchie *et al.,* 2004) and the information should be both

- prognostic (what is likely to happen?) and
- diagnostic (what has happened in the recent past, or what is happening now?).

The reason for latter is that diagnostic information can provide a relevant context within which to interpret a climate forecast (Podestá *et al.,* 2002). Sources of uncertainty and methods to evaluate forecast quality are therefore described briefly, as forecast quality is a central issue for anyone wishing to use hydro-climatic forecasts.

### 4.1.1 Sources of Uncertainty in Forecasting

Improved understanding of ocean-atmosphere interactions, more powerful remote sensing tools and the advances in simulation of complex non-linear systems with powerful computers has facilitated the generation of hydro-climate forecasts with increasingly more accuracy. However, there are some unavoidable errors in the generation of weather, climate and agrohydrological forecasts. These errors arise from three sources (Lettenmaier and Wood, 1993), *viz.*

- model errors,
- data errors, and
- forecast errors.

### *4.1.1.1 Model errors*

Errors in agrohydrological models often arise from an incorrect conceptualisation of the rainfall-runoff processes by the agrohydrological model (Lettenmaier and Wood, 1993). As has been highlighted in Chapter 3, agrohydrological models are limited by their representation of the local spatial heterogeneities and non-stationarities of rainfall-runoff processes. Errors arising out of inadequate model conceptualisation are ideally improved by research on relevant processes and incorporating the findings in improved algorithms (UKCIP, 2003; Schulze, 2007). Alternatively, as an interim solution when adequate agrohydrological observations are available good simulations of agrohydrological outputs may be obtained, by changing values of some internal variables or parameters (UKCIP, 2003; Collischonn *et al.,* 2005).

### *4.1.1.2 Data errors*

With hydro-climatic data it is often difficult to asses the "truth" of observed data because of several factors. Sources of errors in the observations include random

and biased errors as well as sampling errors. Errors in model inputs such as precipitation as a result of sparseness of the raingauge network, observer errors, raingauge splash errors and extrapolation errors will be amplified, for example, through the agrohydrological forecasts (Lettenmaier and Wood, 1993; Schulze, 1995a, UKCIP, 2003).

### 4.1.1.3  Forecast errors

Advances in computing and improved understanding of the atmosphere-ocean system, have enabled NWP and GCM models to respectively predict the weather in the near and more distant future. These models use equations of fluid motion, which are initialised with present conditions to predict the movement and evolution of disturbances such as frontal systems and tropical cyclones that cause rainfall (Ganguly and Bras, 2003). Despite the progress made in these models, weather forecasts have, as yet, obtained only limited success, i.e. their skill drops off with lead time and varies from one location to another. The reason for this is their limited representation of meso-scale atmospheric processes, terrain, land and sea distribution (Mecklenburg *et al.,* 2000; Schmidli *et al.,* 2006).

Moreover, no matter how good atmospheric models may become, the forecasts will always fail up to a point because the atmosphere is a chaotic dynamical system, and any error in the initial condition will lead to increasing errors in the forecast, eventually leading to a greater or smaller loss of predictability after a certain period of time (Toth *et al.,* 1997). The rate of the error growth depends on factors such as the circulation regime, season and geographic domain (Toth *et al.,* 1997). Thus, rainfall forecasts are still limited by the resolution of the simulated atmospheric dynamics and the sensitivity of sub-grid scale parameterisations of the rainfall forming processes (Lettenmaier and Wood, 1993; Toth *et al.,* 1997; Pappenberger *et al.,* 2005).

### 4.1.2  Improving the Quality of Forecasts

As was mentioned above, uncertainty is inherent in the forecasting process. However, minimising these uncertainties to acceptable levels promotes the value of the forecast.  A technique termed "ensemble forecasting" has been developed

45

by many weather forecasting centres around the world in order to assess the forecast uncertainty due to errors in the initial conditions of the atmosphere. In order to address the problems related with spatial resolution, several statistical and dynamical models have also been developed. These techniques are described in more detail in the sub-sections which follow.

### 4.1.2.1 Ensemble forecasting systems

Ensemble forecasting is a technique developed to assess the flow-dependent predictability of the atmosphere by running a NWP model several times, with slightly perturbed initial conditions which lie within the estimated cloud of uncertainty that surrounds the control analysis (Toth *et al.,* 1998). In non-linear dynamical systems this approach offers the best possible forecast with the maximum information content. In a statistical sense, averaging the ensemble members provides a more reliable forecast than simply using any one of the single forecasts, including that started from the control analysis (Toth *et al.,* 1997). Ensemble forecasting has become a common practice to assess the flow-dependent predictability of the atmosphere, and to create quantitative probabilistic forecasts at many NWP centres around the world, e.g. at the National Center for Environmental Prediction (NCEP) in the USA, the European Centre for Medium-Range Weather Forecasts (ECMWF), the Canadian Meteorological Centre (CMC), the Japan Meteorological Agency (JMA) and the SAWS (Toth *et al.,* 1997; Toth *et al.,* 2005). However, questions relating to the generation of adequate sampling of initial perturbations, and to estimating the analysis error in a probabilistic sense, remain major research issues for an ensemble forecasting system (Wei *et al.*, 2005).

According to Toth *et al.* (1997) and Wei *et al.* (2005) initial perturbation techniques are broadly classified into either as first or second generation techniques. The first generation initial perturbation techniques are commonly used at different centres for initial perturbations. These methods include the following:

*Singular vectors*: These have been developed at ECMWF, and identify the direction of fastest forecast error growth for a 2 day period at the beginning of the forecast (Toth *et al.,* 1997; Wei *et al.*, 2005). The estimation is neither explicit nor

flow-dependent. This technique would be the best sampling strategy if all possible analysis errors had occurred with the same probability. In this ensemble strategy initial perturbations are not consistent with the data assimilation systems that generate the analysis fields (Wei *et al.*, 2005).

*Breeding*: The so-called breeding technique has been developed at NCEP and has been adopted at many other NWP centres (e.g. SAWS, JMA, CMC). This technique captures the fastest growing errors that are most likely to be responsible for the error in the control forecast. The error in the short range forecast is the perturbation which is periodically rescaled at each analysis time by blending observations with the estimate (Toth *et al.,* 1997; Toth and Kalnay, 1997; Wei *et al.*, 2005). Breeding also cannot accurately assess the analysis errors as there are generally not enough observations to eliminate all errors from the short range forecast that is subsequently generated as the first estimate for the next analysis. Consistencies between ensembles and data assimilation are poor, owing to the use of fixed estimates of the analysis error variance and the lack of an orthogonalisation in the breeding procedure (Wei *et al.*, 2005).

*Perturbed Observations*: These have been developed at the CMC and generate initial conditions by assimilating randomly perturbed observations using different models in a number of separate analysis cycles. All observations are perturbed with random noise representing the error in observations. Similar to breeding, perturbed observations capture patterns that can occur in the analysis as errors. Breeding does not, however, take into consideration patterns that initially are not growing. The initial perturbations generated by the *perturbed observation* technique are, therefore, more representative of analysis uncertainties in comparison with the *singular vectors* and *breeding* techniques (Wei *et al.*, 2005).

The *Ensemble Transform Kalman Filter* (ETKF), *Ensemble Transform* (ET), *ET with breeding* and *singular vectors with Hessian norm* can be classified as the second generation initial perturbation techniques (Wei *et al.*, 2005). A common feature of the second generation techniques is that the initial perturbations are more consistent with the data assimilation systems when compared with the first generation initial perturbation techniques. Even though the *ET with rescaling* method looks the most promising one, with better consistency and efficient

computation the breeding method established at NCEP is widely used at many centres (Toth *et al.,* 2005; Wei *et al*., 2005).

### 4.1.2.2  Techniques for spatial downscaling

Knowledge of precipitation fields at fine resolution is a vital ingredient for agrohydrological forecasting. In the absence of full deterministic modelling of small-scale rainfall, it is common practice to use a spatial downscaling procedure (Rebora *et al.,* 2005). Many techniques have been developed for the spatial downscaling of rainfall. According to Schmidli *et al.* (2006) and Wood *et al.* (2004), the spatial downscaling methods that have been most widely used are categorised broadly into either

- statistical (e.g. Canonical Correlation Analysis, CCA; Multiple Linear Regression, MLR; Multivariate Autoregressive Model, MAR; Conditional Weather Generator, CWG; or Climate analogue), or
- dynamical (e.g. CHRM, HadRM3, HIRHAM).

*Statistical downscaling methods* use the observed relationships between large-scale circulation and the local climates to set up statistical models that attempt to translate anomalies of the large-scale flow into anomalies of some local climate variable (Zorita and von Storch, 1999; Schmidli *et al.,* 2006). Statistical downscaling methods are commonly used because of their relative simplicity and lower costs when compared with dynamical methods (Zorita and von Storch, 1999; Wood *et al.,* 2004). The climate analogue method is considered to be the simplest of the downscaling schemes and it compares the large-scale atmospheric circulation simulated by a GCM to historical observations. The most similar analogue is selected and simultaneously observed local weather data are then associated to the simulated large-scale pattern. A major problem associated with the climate analogue method is the need for accurate and long observations (Zorita and von Storch, 1999).

*Dynamical models* use the so-called Limited Area Models (LAMs) to account the regional and local characteristics such as topography, which influence rainfall patterns. These LAMs are atmospheric or oceanic models of limited geographical

area with finer horizontal resolutions than GCMs, but which use the large-scale fields simulated by the GCMs as boundary conditions and the local variables to provide weather forecasts at a regional scale (Zorita and von Storch, 1999; Wood *et al.,* 2004; Rebora *et al.,* 2005; Schmidli *et al.,* 2006). The LAMs are capable of simulating the regional climate conditions such as orographically induced precipitation. However, some systematic errors still exist in these models due to errors in sub-grid parameterisations, which are taken over from the parent GCMs (Zorita and von Storch, 1999).

Several researchers (e.g. Zorita and von Storch, 1999; Wood *et al.,* 2004; Rebora *et al.,* 2005; Schmidli *et al.,* 2006) have evaluated the differences between various statistical and dynamical downscaling methods, based on their implications for hydrological forecasts. For example, Wood *et al.* (2004) compared three statistical downscaling methods, *viz.* Linear Interpolation (LI), Spatial Disaggregation (SD) and Bias Corrected Spatial Disaggregation (BCSD), by using climate simulations produced by the Parallel Climate Model (PCM). Each method was applied to both PCM output directly and to dynamically downscaled PCM output with a Regional Climate Model (RCM). They concluded that dynamical downscaling does not lead to large improvements in hydrological simulations relative to the direct use of PCM output when BCSD was used. With LI of PCM and RCM outputs, the hydrological simulations were found to be poor, while applying SD improved sub-grid spatial variability and displayed better hydrological simulations (Wood *et al.,* 2004).

It should be noted that a rainfall field generated by any spatial downscaling method is one possible realization of the small scale field and should not be considered as providing the "true" rainfall distribution (Rebora *et al.,* 2005).

### 4.1.3  Verification of Forecasts

The quality of hydro-climatic forecasts is highly dependent on geographic location, season and lead times. Routine forecast quality control is usually performed by model developers and/or the forecast providers themselves. However, the quality of a forecast does not necessarily address its practical usefulness for a decision maker. The quality of hydro-climatic forecasts produced by various models needs to be assessed from users' perspectives before the products would have any

relevance to them. Hence, forecast performance assessments should include measures that express relevant properties of forecasts that help users to judge the usefulness of forecasts for their specific purposes (Hartmann *et al.,* 2002; Mailier *et al.,* 2006). Although research in forecast verification is continually growing, the nature of forecast products, the wide range of customer requirements and the different nature of delivery systems have complicated the development of standard measurements that would be useful to all the people making decisions (Mailier *et al.,* 2006). According to Jolliffe and Stephenson (2003), the three important reasons to verify the quality of forecasts are to:

- improve forecast quality by identifying the problems associated with the forecasts,
- compare the quality of different forecast systems in order to know to what extent one forecast system gives better results than another, and to
- monitor forecast quality in order to find out how accurate the forecasts are when compared to actual observations and to asses the degree of improvement over time.

Mailier *et al.* (2006) proposes the following points as being good practice in quality assessments:

- the assessment procedures should be clearly and full described, including descriptions/definitions of all technical terms used,
- forecast formats should be suitable to objective quality assessment, with qualitative terms avoided wherever feasible,
- the assessment methodology should, in principle, be repeatable by a user,
- the assessment methodology should be carefully chosen to produce information that is meaningful to the user,
- uncertainty about the forecasts should be presented in a simple format that the user can easily understand,
- users should be aware of the statistical properties and possible deficiencies of the methods used in the assessment,
- assessments should include the different facets of forecast performance, and

- the choice of the sample used for the assessment should be justified, in order to provide stable and representative estimates.

An assessment of forecast quality depends on the type of forecast, i.e. whether it is deterministic (i.e. non probabilistic), qualitative (e.g. scattered showers) or probabilistic (e.g. categorical, continuous). Qualitative forecasts are difficult to verify as different users will likely interpret them differently. Hence there is always a subjective interpretation, whether or not a forecast is a good one. Qualitative forecasts can only be verified in circumstances where a technical definition underlies a descriptive forecast (Mailier *et al.,* 2006). Most forecast techniques have some strengths, but all have some weaknesses (Jolliffe and Stephenson, 2003). This implies that more than one score (measure) is often needed for better decision making. The evaluation should consider all aspects of correspondence between forecasts and observations. In this regard, Murphy (1993) describes the following relevant terms

- *bias:* the correspondence between the average forecast and the average observation,
- *association:* the strength of the linear relationship between the forecasts and the observations,
- *accuracy:* the degree of correspondence between forecasts and observations,
- *skill:* the accuracy of forecasts compared to other forecasts produced using a standard strategy,
- *consistency*: the degree of correspondence between the forecaster's judgement and the forecast
- *reliability:* the correctness of forecast uncertainty,
- *resolution:* the extent to which outcomes differ from given forecasts,
- *discrimination:* the extent to which forecast depart from given observations,
- *sharpness:* the extent to which forecast depart from climatology, and
- *uncertainty*: variability of observations regardless of the forecast.

These are major attributes that contribute to the evaluation of forecast quality. A short definition of commonly used verification scores that can be used to assess the skill of continuous and categorical forecasts is given in this section.

In regard to continuous verification scores, *bias*, *relative bias*, *correlation coefficient, Root Mean Square Error* (RMSE) and *Mean Absolute Error* (MAE) are commonly used and they provide statistics on how much the forecast values differ from the observations. Most continuous verification scores are sensitive to large errors (Lettenmaier and Wood, 1993; Nurmi, 2003).

*Bias* measures systematic error in the forecast. It measures the degree to which the forecast is consistently above or below the observed value (Lettenmaier and Wood, 1993; Nurmi, 2003). It is expressed as

$$\text{bias} = \frac{1}{N}\sum_{i=1}^{N}O_i - \frac{1}{N}\sum_{i=1}^{N}F_i \qquad \text{4.1}$$

where

     $F_i$    =     forecast value of day i or pixel i,

     $O_i$    =     observed value of day i or pixel i, and

     N    =     total number of days or pixels.

The *correlation coefficient* measures the degree of linear association between the forecast and the observed values. However, it is important to bear in mind that the *correlation coefficient* evaluates forecast accuracy in terms of random error only (Lettenmaier and Wood, 1993). Thus, forecast errors could be large, even with a near-perfect correlation, if appreciable bias is present (Lettenmaier and Wood, 1993; Mason, 2000). The *correlation coefficient* is expressed as

$$r = \frac{\sum_{i=1}^{N}(F_i - \bar{F})(O_i - \bar{O})}{\sqrt{\sum_{i=1}^{N}(F_i - \bar{F})^2}\sqrt{\sum_{i=1}^{N}(O_i - \bar{O})^2}} \qquad \text{4.2}$$

where

     r    =     correlation coefficient

     $F_i$    =     forecast value of day i or pixel i,

$\overline{F}$ = average forecast value of all days or pixels,

$O_i$ = observed value of day i or pixel i,

$\overline{O}$ = average observed value of all days or pixels, and

N = total number of days or pixels.

RMSE measures the average error magnitude while MAE measures the average squared error magnitude and both methods measure systematic and random errors (Lettenmaier and Wood, 1993; Mason, 2000; Nurmi, 2003). They are expressed as

$$RMSE = \frac{1}{N}\sum_{i=1}^{N}(F_i - O_i)^2 \qquad\qquad 4.3$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|F_i - O_i| \qquad\qquad 4.4$$

where

$F_i$ = forecast value of day i or pixel i,

$O_i$ = observed value of day i or pixel i, and

N = total number of days or pixels.

The pixel-by-pixel scoring criteria, *viz.* the Critical Success Index, CSI, the Probability of Detection, POD and the False Alarm Ratio, FAR are also commonly used to assess the overall degree of positional accuracy over a selected area (Wilks, 1995). These statistics are calculated as follows:

$$CSI = \frac{H}{H + M + FA} \qquad\qquad 4.5$$

$$POD = \frac{M}{H + M} \qquad\qquad 4.6$$

$$FAR = \frac{FA}{H + FA} \qquad\qquad 4.7$$

where

CSI = the Critical Success Index,

POD = the Probability of Detection, i.e. the Hit Rate,

FAR = the False Alarm Ratio,

H = number of pixels for which both the estimated and observed values exceed a specified threshold,

M = number of pixels for which only the observed values exceed a specified threshold, and

FA = number of pixels where only the estimated values exceed a specified threshold.

A variety of categorical verification scores are used operationally to verify hydro-climatic forecasts. There are many textbooks, research papers and technical papers providing detailed information of these scores (e.g. Wilks, 1995; Potts *et al.,* 1996; Zhang and Casey, 1999; Joliffe and Stephenson, 2003; Livezey, 2003; Nurmi, 2003; Mailier *et al.,* 2006). What follows below, however, focuses on the discussion of the five more commonly used scoring methods, *viz.* the

- Heidke Score (HS),
- Revised True Skill Statistics (RTSS),
- Linear Errors in Probability Space (LEPS),
- Brier Score (BS), and
- Ranked Probability Skill (RPS).

Categorical forecasts are usually assessed by reducing them to a series of binary (i.e. yes and no) forecasts (Livezey, 2003). Often a 2 x 2 contingency table is constructed to transform categorical probabilistic forecasts into binary events based on decision probability thresholds (Table 4.1).

Table 4. 1 Schematic contingency table for categorical forecasts of a binary event, with the number of observations in each category being represented by A, B, C, D and N (Source: Livezey, 2003)

| Forecast | Observed | | Total |
|----------|-----|-----|-------|
| | Yes | No | |
| Yes | A | B | A+B |
| No | C | D | C+D |
| Total | A + C | B + D | A +B +C + D = N |

Given a set of forecasts, it is possible to calculate the number of times that the forecast was correct. The HS (Equation 4.8) is a simple measure of forecast accuracy for binary (i.e. yes or no) forecasts. It is simply the sum of points scored, divided by the total number of forecasts (Mason, 2000).

$$HS = \frac{A+D}{N} \times 100 \qquad\qquad 4.8$$

where

|     |     |                                |
|-----|-----|--------------------------------|
| HS  | =   | the Heidke Score,              |
| A   | =   | number of hits,                |
| D   | =   | number of correct rejections, and |
| N   | =   | total number of observations.  |

The problem with the HS is that a high score is achievable both if the forecasted event is rare or extremely common (Mason, 2000). The HS is often compared with some reference forecasts such as climatology, persistence or random chance to form a single index called Heidke Skill Score, HSS (Mason, 2000; Banitz, 2001), which is expressed as

$$HSS = \frac{A-E}{N-E} \times 100 \qquad\qquad 4.9$$

where

|     |     |                                                                 |
|-----|-----|-----------------------------------------------------------------|
| HSS | =   | the Heidke Skill Score,                                         |
| A   | =   | number of hits,                                                 |
| E   | =   | number of forecasts expected to be correct, based on a reference such as climatology, persistence or random chance, and |
| N   | =   | total number of observations.                                   |

The RTSS is another technique similar to the HSS. However, the RTSS score (Equation 4.10) measures the fraction of correct forecasts after eliminating those forecasts which would be correct due purely to random chance. It gives the best estimates on an "unequal" trial basis as it gives equal emphasis to the ability to forecast events and non-events (Zhang and Casey, 1999). The RTSS is given as

$$TSS = \frac{N_{CM} - N_{CCM}}{N - N_{CCO}} \qquad\qquad 4.10$$

where A, B, C, D, and N are the components in the Table 4.1 and

RTSS = the True Skill Statistics,

N = total number of observations,

$N_{CM}$ = number of correct forecasts from the forecast model, i.e. (A+D),

$N_{CCM}$ = number of correct forecasts that could be achieved by chance, i.e. $(A+C) * P_{yes} + (B+D) * P_{no}$,

$N_{CCO}$ = number of observed events that can be correctly forecasted by chance, i.e. $(A+B) * P_{yes} + (C+D) * P_{no}$.

$P_{yes}$ = climatological probabilities, i.e. (A+B)/N, and

$P_{no}$ = climatological probabilities, i.e. (C+D)/N,

It is important to bear in mind that the HS and RTSS scores do not penalise the errors in terms of their severity between each categories.

The LEPS scoring matrices are calculated from the distance between the forecasts and observations in continuous cumulative probability space (Figure 4.1). It rewards good forecasts, and penalises two-category misses much more than one-category misses (Zhang and Casey, 1999; Klopper and Landman, 2003; Livezey, 2003).

LEPS is then computed by the following equation (Equation 4.11):

$$LEPS = \frac{1}{N} \sum_{i=1}^{N} | CDF_o(F_i) - CDF_o(O_i) | \qquad\qquad 4.11$$

where

LEPS = the Linear Error in Probability Space,

$CDF_o$ = cumulative probability density function of observations, obtained from an appropriate climatology,

$F_i$ = forecast value of category i,

$O_i$ = observed value of category i.
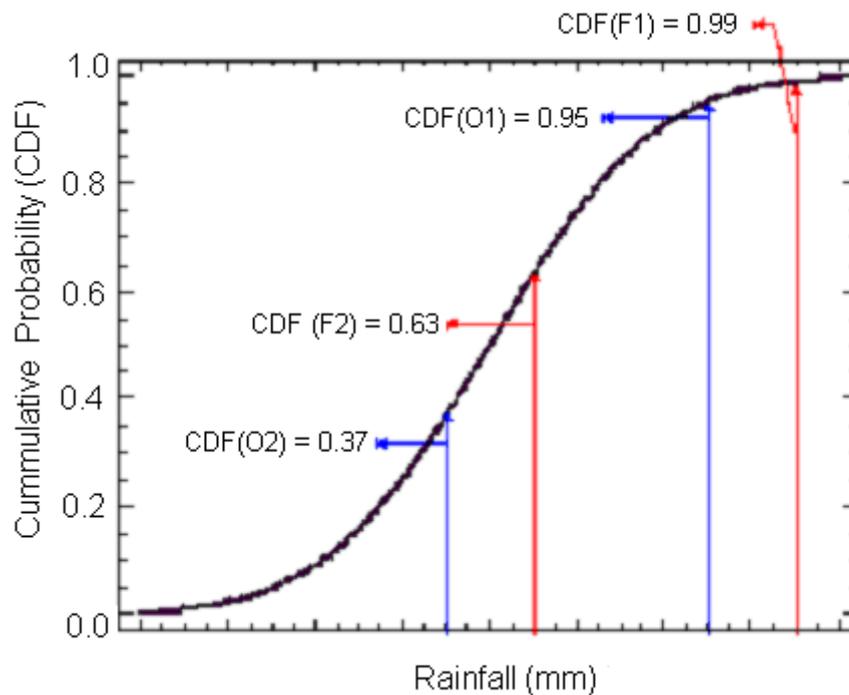
Figure 4.1    Schematic diagram for calculating the Linear Errors in Probability Space (Source: http://www.bom.gov.au)

Potts *et al.* (1996) derived an improved version of the LEPS score that does not discourage forecasting extreme values if they are warranted. It is given by:

$$LEPS = 3(1 - |CDF_o(F_i) - CDF_o(O_i)| + CDF_o^2(F_i) - CDF_o(F_i) + CDF_o^2(O_i) - CDF_o(O_i)) - 1 \qquad 4.12$$

In the LEPS matrix, a score of +100% will indicate perfect hits and a 0 score indicates a result as good as the climatology, while a score of -100% shows no hits. LEPS has been developed for continuous variables as well (Livezey, 2003).

The BS and RPS provide combined measures of reliability and sharpness. The RPS is similar to the BS, but is used for more than two categories (Mason, 2000). The BS and RPS measure the sum of squared differences in cumulative probability space for two categories and multi-category probabilistic forecasts respectively. They penalise forecasts more severely if the weight of the forecasts are not closer to the actual observed distribution (Zhang and Casey, 1999; Nurmi, 2003). RPS is given by

$$RPS = \frac{1}{M-1}\sum_{m=1}^{M}[(\sum_{i=1}^{m}F_i) - (\sum_{i=1}^{m}O_i)]^2 \qquad\qquad 4.13$$

where

| | | |
|---|---|---|
| M | = | number of categories, |
| $F_i$ | = | the forecasted probability in forecast category i, and |
| $O_i$ | = | an indicator (0 = no, 1 = yes) for observation in category i. |

The BS and RPS can also be expressed as skill scores indicating the fractional improvement relative to a reference forecast (Mason, 2000). Hence,

$$RPSS = \frac{RPS_{forecast} - RPS_{reference}}{0 - RPS_{reference}} = 1 - \frac{RPS_{forecast}}{RPS_{reference}} \qquad\qquad 4.14$$

where

| | | |
|---|---|---|
| RPSS | = | the Ranked Probability Score Skill (fraction), |
| $RPS_{forecast}$ | = | the probabilistic forecasted RPS (fraction), and |
| $RPS_{reference}$ | = | the RPS expected from the reference forecast (fraction). |

The RPSS ranges from -∞ to 1, with a score less than or equal to 0 indicating no skill when compared to the reference forecast, and a score of 1 indicating a perfect forecast. The RPSS is, however, highly unstable when applied to small data sets (Mason, 2000).

The above categorical verification techniques measure the skill, sharpness and reliability of forecasts relative to the quality of some other forecasts produced by standard procedures. Reliability addresses the questions as to whether repeated application of forecast procedures will produce similar results. It measures the forecaster's level of confidence to produce reliable forecasts (Scott and Collopy, 1992; Mason, 2000; Schneider and Garbrecht, 2003). According to Mason (2000), perfect reliability occurs if:

- forecasts are statistically consistent with the observations, but it does not necessarily mean that the forecasts are accurate, and

- the forecaster's confidence is appropriate.

Climatology, random, persistence and median values are simple forecast strategies used for a reference strategy (Mason, 2000; Hallowes, 2002). The forecast skill is usually defined as the percentage improvement in accuracy over the reference forecast (Zhang and Casey, 1999; Mason, 2000). Care should be taken to select appropriate reference forecasts so that the computed skill reflects the true usefulness of the forecast (Mailier *et al.,* 2006).

Not all categorical verification techniques account for possible near-misses across category boundaries, and they do not account for the accuracy of the forecasts within a category (Mason, 2000). In addition, part of the information from categorical forecasts will be lost during the transformation to binary forms (Zhang and Casey, 1999).

### 4.1.4  Procedures for Updating Forecasts

Most hydrological models use mathematical equations to describe the various components of spatially and temporally varying catchment hydrological processes. In most conceptual and parameter optimising hydrological models, the forecast errors may result from inadquecies in the model structure, incorrect conceptualisation of the model parameter and errors in the data, as well as errors induced by the lack of knowledge of the future rainfall (Toth *et al.,* 1999; Xiong *et al.,* 2004). When any hydrological model is intended for use in a real time forecasting system, it will be associated with explicit or implicit updating procedures whereby, at the time of making the forecast, errors already observed in recent forecasts will be used to modify the forecast (Xiong *et al.,* 2004).

Univariate linear statistical models such as the *AutoRegressive* (AR), the *AutoRegressive Integrated Moving Average* (ARIMA), the *Linear Transfer Function* (LTF) or, alternatively, *Artifical Neural Networks* (ANN), which is a non-linear statistical model, are  commonly used in the updating mode to post-process the forecasts made by the conceptual or physically based hydrological model. Descriptions of these models are widely available in research papers (e.g. Toth *et al.,* 1999; Madsen and Jacobsen, 2001; Xiong *et al.,* 2004; Goswami *et al.,* 2005). These statistical models are not alternatives to deterministic or conceptual models,

rather they are used to predict simulation errors induced by unsatisfactory model paramaterisation, or errors cascaded from rainfall forecasts. Various types of updating schemes may be implemented that may compensate for the deficiencies of the hydrological models. According to Anctil *et al.* (2003) and Goswami *et al.* (2005), four types of updating procedures exist. They are described below.

*Updating of Input Variables*: Additional input information from the most recently measured variables can be used in the updating procedure. Thus, the forecasting system can be corrected as and when daily observed rainfall, temperature and runoff data become available, in order to account for any spatio-temporal errors that may have occurred in previous forecasts.

*Updating of State Variables:* Day-to-day catchment state variables deviate from the the so-called average conditions simulated with a conceptual hydrological model (Anctil *et al.,* 2003). The catchment state variables need to be calibrated or updated continuously to render the potential of agrohydrological forecasting more useful to decisions in water resource and agricultural operations. Schulze *et al.* (1998) identified the following state variables that need day-to-day updates in the *ACRU* model (Schulze, 1995a and updates):

- baseflow store and baseflow releases,
- stormflow store and stormflow releases,
- soil moisture in the topsoil and subsoil,
- dam levels, abstractions, water transfers and return flows, and
- irrigation abstractions and return flows.

*Updating of Model Parameters:* This is the least favoured updating scheme because it is not sound practice to modify model parameters at each time step. Moreover, this is an iterative process, quite time consuming and computational demading, especially when the model includes a large number of parameters (Toth *et al.,* 1999; Anctil *et al.,* 2003; Goswami *et al.*, 2005).

*Updating of Output Variables:* This updating scheme is commonly used. Toth *et al.* (1999), for example, applied six different stochastic models, aimed at updating the discharge forecasts produced by a conceptual rainfall-runoff modelling called ADM

(Franchini, 1996). They found that all the six updating models were more efficient than the ADM model. Similar results have also been reported by Xiong *et al.* (2004) after three updating schemes using ANN discharge forecasting had been applied on ten catchments in various countries. The statistical models attempt to predict the simulation series error produced by the conceptual or deterministic hydrological models. The updated forecast is then the sum of the simulated plus the predicted error values (Toth *et al.,* 1999; Anctil *et al.,* 2003; Xiong *et al.,* 2004).

The selection of the updating scheme depends on what is considered by the modeller to be the main cause of any discrepancy between observed and forecasted values (Anctil *et al.,* 2003). In this study, updating with daily observed rainfall values was used for simulating one day streamflow forecasts with the *ACRU* model (Schulze, 1995a), in order to correct for any errors that may have occurred by the lack of knowledge in the forecast of the previous day. A description and evaluation of the updating procedure is presented in Chapter 8.

## 4.2    Communication of Hydro-Climatic Information

The communication process is the second component of an effective hydro-climatic system and it includes preparation of weather forecasts for public and private interests, as well as educating end users about forecast issues (e.g. contents, formats, limitations and dissemination). Communication using participatory approaches and collaborative learning is an important step in promoting use of hydro-climatic forecasts (Podestá *et al.,* 2002). Communication should flow in both directions, i.e. from scientists to practitioners or decision makers and vice versa, in order to create opportunities for mutual learning. Information received at one step may produce a demand for other information. Feedback is important as an indicator of users' reactions that allow scientists to improve forecasts for specific purposes, and stakeholders to learn about capabilities and limitations of hydro-climatic forecasts (Hobbs, 1980; Klopper, 1999; Podestá *et al.,* 2002).

The wide range of users and increasing demand for hydro-climatic forecasts implies a similarly broad range of requirements and expectations of the forecasts.

Requirements may vary in terms of the desired weather format and spatio-temporal scales (Hobbs, 1980).

The nature and speed of forecast dissemination are major issues that may influence the usefulness of forecasts. Advances in technology have facilitated the transmission of forecasts in a real time mode. Newspapers, radio, television, cellphone and the internet are important devices for the forecast dissemination to users. However, misinterpretation of the forecast by users and the media is a major problem (Hobbs, 1980). A survey conducted by Klopper (1999) indicates that some users do not fully understand the definition of the hydro-climatic forecast terms. Moreover, some believed the newspapers to be more desirable while others preferred to listen to radio or television broadcasts. The media may also be more interested in the style and attractiveness of the forecasts than the accuracy. These types of confusion indicate that the news media and end users should be educated on how the forecasts should be interpreted. Technical advices on how to respond to hydro-climatic forecasts should ideally come from trusted sources such as agricultural extension agents or technical consultants, and not directly from forecast provider institutions. The reason for this is that end users (e.g. farmers) may evaluate the credibility of forecasts based on its source. Usually they act positively when the information comes from sources that they already know and trust (Hobbs, 1980; Hansen, 2002).

The communication process is a challenging issue and is often impeded by financial, technical and cultural barriers (Glantz, 1996; Podestá *et al.,* 2002). Many societies have had long traditions of using a variety of different indicators to predict the weather conditions. However, more efforts must be made to ensure closer articulation with end users. Such interaction will provide better insights of their needs and expectations. It would also promote trust building communication between forecasters and end users (Podestá *et al.,* 2002)

## 4.3    Application of Hydro-climatic Information

Hydro-climatic forecasts must ideally contribute to a change in decisions, which leads to desirable outcomes, regardless of how accurate and well communicated the forecast is (Hammer *et al.,* 2001; Hansen, 2002; Podestá *et al.,* 2002; Ritchie

*et al.,* 2004). If a forecast system is validated, but fails to generate changed decisions, the information will have only academic value. However, if the forecast system has a positive value of information, but has not been statistically validated, then the system is not useful, as the value may be the result of chance (Ritchie *et al.,* 2004).

Decision makers should be able to examine the value of forecasts for a specific purpose, and evaluate its economic return in terms of cost-loss ratio analysis. In fact, it is not easy for a decision maker to make a rational decision that minimises the expected losses and maximises the expected benefits under uncertain forecasts. According to Podestá *et al.* (2002), changes in decision making processes depend on the following conditions:

- the quality of hydro-climatic forecasts, with appropriate lead time and geographic and temporal resolution,
- the feasibility of alternative actions that can be taken in response to a hydro-climatic forecast,
- the ability of decision makers to evaluate the outcomes of those alternative actions, and
- the willingness of decision makers to change their decisions in an already complicated decision-making environment.

A decision support system is another key element that can facilitate the use of hydro-climatic forecasts. Decision support tools allow the exploration of multi-dimensional decision space that would help decision makers to evaluate the consequences of alternative management in respond to forecasts (Podestá *et al.,* 2002). Recognising the importance of the three components, Hansen (2002) proposed a framework that represents the opportunity to benefit from hydro-climatic forecasts. The opportunity to benefit falls within the intersection of human vulnerability, hydro-climatic forecasting and decision capacity, as shown in Figure 4.2.
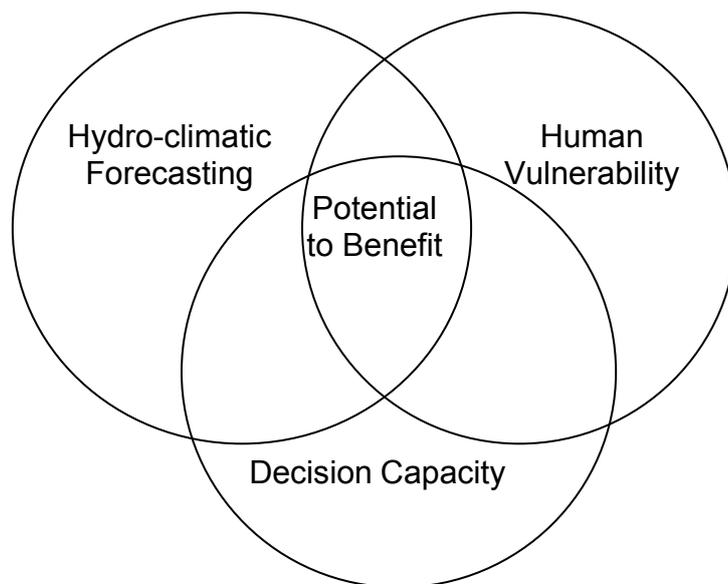
Figure 4.2    Determinants of the potential for human populations to benefit from hydro-climatic forecasts (after Hansen, 2002)

## 4.4    Concluding Remarks

This chapter commenced by outlining the approaches required to maximise benefits from the use of hydro-climatic forecasts.  It was found that the benefits which might accrue do not only depend on the scientific advances of hydro-climatic forecasts, but also on an effective way of dissemination as well as on appropriate education of forecast presenters and decision makers. Apart from forecast quality considerations, the format and speed of dissemination of forecasts, as well as the willingness and ability of decision makers to make a change, are critical elements in the usefulness of forecasts. Nonetheless, the production of skilful and timely forecasts continues to be one of the major issues challenging to hydrometeorologists. Owing to the inherent uncertainties in the weather and model limitations to account for the local rain-bearing features, weather and climate forecasts are not as accurate as desired. The accuracy of such forecasts will further be degraded during the rainfall-runoff transformation by hydrological models. The reason for this is that the complex and non-linear rainfall-runoff processes are not explicitly represented by most hydrological models.

A brief review was presented of some of the elements that contribute towards forecast uncertainties and techniques developed to minimise forecast errors,

followed by the description of some commonly used verification techniques for assessing forecast quality. The chapter further described the potential application of forecast updating by the combined use of conceptual physically based models in simulation mode plus stochastic models in the updating mode, in order to eliminate, or minimise, errors resulting from inadequecies in the hydrological model or the incorrect estimation of rainfall forecast by weather prediction models. Finally, the challenges and approaches in cummunication process and use of hydro-climatic forecasts to modify decisions were described briefly.

A brief outline of the GIS based framework for agrohydrological forecasting system is presented in Chapter 5.

# 5. A GIS BASED FRAMEWORK FOR AN AGROHYDROLOGICAL FORECASTING SYSTEM

An effective, operational agrohydrological forecasting system should provide the right information, at the right time, to address the needs of decision makers and operational users in agricultural and water resources management. Thus, the main aim of this study is to develop a framework that facilitates the application of near real, plus daily, multi-day to seasonal rainfall forecasts as a nested set of inputs to agrohydrological and/or crop yield models, thereby enabling the forecasting of agrohydrological variables across a range of time scales and lead times in southern Africa, defined here as the RSA plus Lesotho and Swaziland. This aim is to be achieved by integrating different sources of forecast information from radar, satellite, and weather/climate models. Generic methodologies are also developed for temporal downscaling of probabilistic categorical seasonal forecasts to a daily time series of values suitable for agrohydrological models. This chapter highlights the motivation behind the development of a GIS based framework in the agrohydrological forecasting system, followed by a brief description on how to use the outputs of the weather/climate models imbedded within the framework.

## 5.1 The Need for a GIS Based Framework

The effective and efficient management of water resource and agricultural operations relies on skilful and timely forecasts of agrohydrological variables such as streamflows, soil moisture, crop yields or reservoir levels. In turn, a key factor for accurate agrohydrological forecasts are accurate and prompt weather/climate forecasts on, for example, rainfall and temperature, as input to the agrohydrological model. Weather and climate forecasts (e.g. SAWS forecasts) for southern Africa have been shown to posses certain levels of skill when they are compared against observations (Klopper and Landman, 2003). The challenge, however, still lies in the improvement of the spatial and temporal resolution of the weather and climate forecasts, and the "translation" of these forecasts into suitable scales and forms that are required by agrohydrological models. These challenges must be addressed if hydrological and/or crop yield models are to contribute to the task of transformation of weather/climate forecasts into more tangible attributes

such as soil water content, streamflows, irrigation requirements, reservoir levels and crop yields.

This calls for the development of generic methodologies to link the outputs of weather and climate models with agrohydrological models. Owing to the complexity and iterative calculations of the translation process from climate to agrohydrological forecasts, manual calculations and data extractions are out of question. A Geographic Information System (GIS) based framework, therefore, becomes a very important platform for gathering, filtering, translating and generating information that can be used directly with agrohydrological models for an effective agrohydrological forecasting system. Within this framework, GIS organises spatial information, provides techniques for pre-processing data (including spatial disaggregation), provides data structure and format conversion and displays post-processed information through reformatting, tabulation, mapping and report generation. A schematic flow chart demonstrating the structure of the GIS based framework for the agrohydrological forecasting system developed in this project is provided by Figure 5.1.
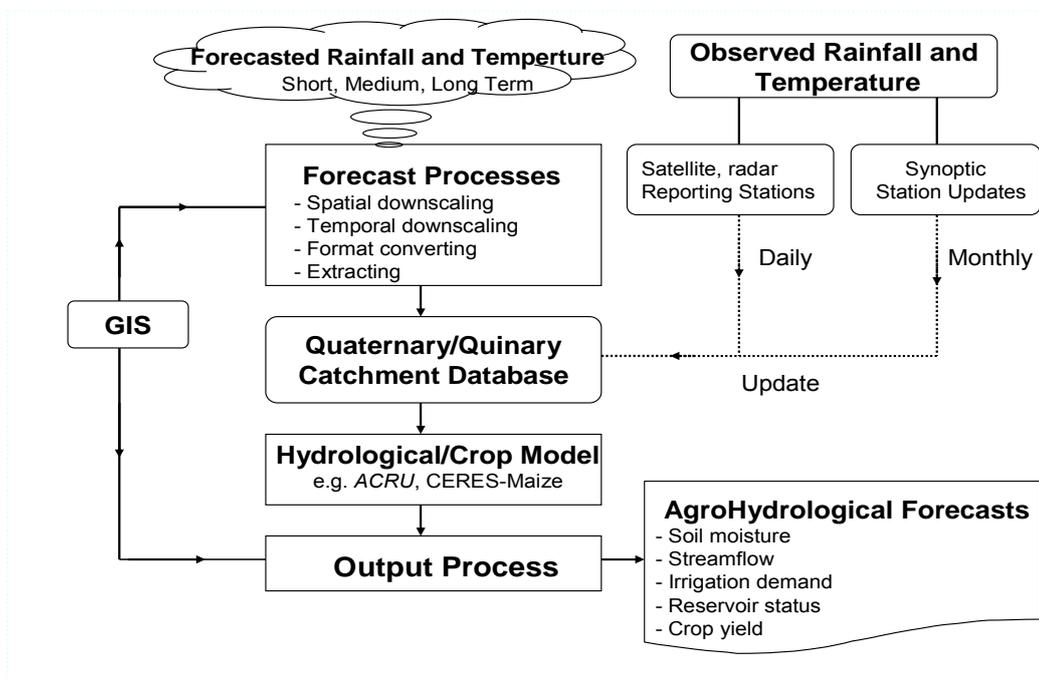


Figure 5.1    A schematic flow chart demonstrating the structure of the agrohydrological forecasting framework

Currently a number of different institutions are involved in producing weather and climate forecasts that could potentially benefit end users. At the present time (2007), forecasts issued by the South African Weather Service (SAWS) and the University of Pretoria (UP) are adopted for the generation of agrohydrological forecasts within this framework. In the future, it seems reasonable to also consider other institutions' forecast products, as different forecasters use different weather and climate models that may perform better than others under particular conditions and/or for specific locations. However, with so many providers and different formats, there is real potential for confusion among users as to which forecasts to use, especially when the forecasts are not similar. Weather and climate forecasts encompass a broad range of variables (e.g. rainfall, temperature, solar radiation, frost), but at this stage rainfall is the key variable of interest in this framework as it is the main determinant of both hydrological and agricultural responses in southern Africa. In the future, the development of the framework is therefore expected to continue beyond the current stage in order to incorporate other weather variables and forecast products issued by other institutions.

Based on the above framework, a GIS based computer program has been developed using the Visual Basic programming language that links to GIS and processes all the calculations required to translate the multi-day, monthly and/or to seasonal climate forecasts into daily quantitative values suitable for application with daily time step hydrological or crop yield models. The program runs on the Windows operating system. Once the program is initiated, the user has options to select the forecast types in the main window (Figure 5.2). In its present state the program is designed to operate at the spatial scale of 1 946 Quaternary Catchments (QCs) into which South Africa has been delineated by the Department of Water Affairs and Forestry (DWAF) for operational decision making. The program has three major components, *viz.*

- near real time observations derived from radar, satellite and daily reporting weather stations,
- short (up to 4 days) and medium term (up to 14 days) forecasts from various Numerical Weather Prediction (NWP) models, and
- monthly and long term (up to 3 months) forecasts from climate models.

These components are described in more detail in the sections which follow. A brief explanation has been made on how to use the GIS based program. However, it has not been written in the conventional style of a software user manual. In a later section, the *ACRU* agrohydrological modelling system (Schulze, 1995a and updates), which was selected in this research to generate agrohydrological forecasts, is also described briefly.
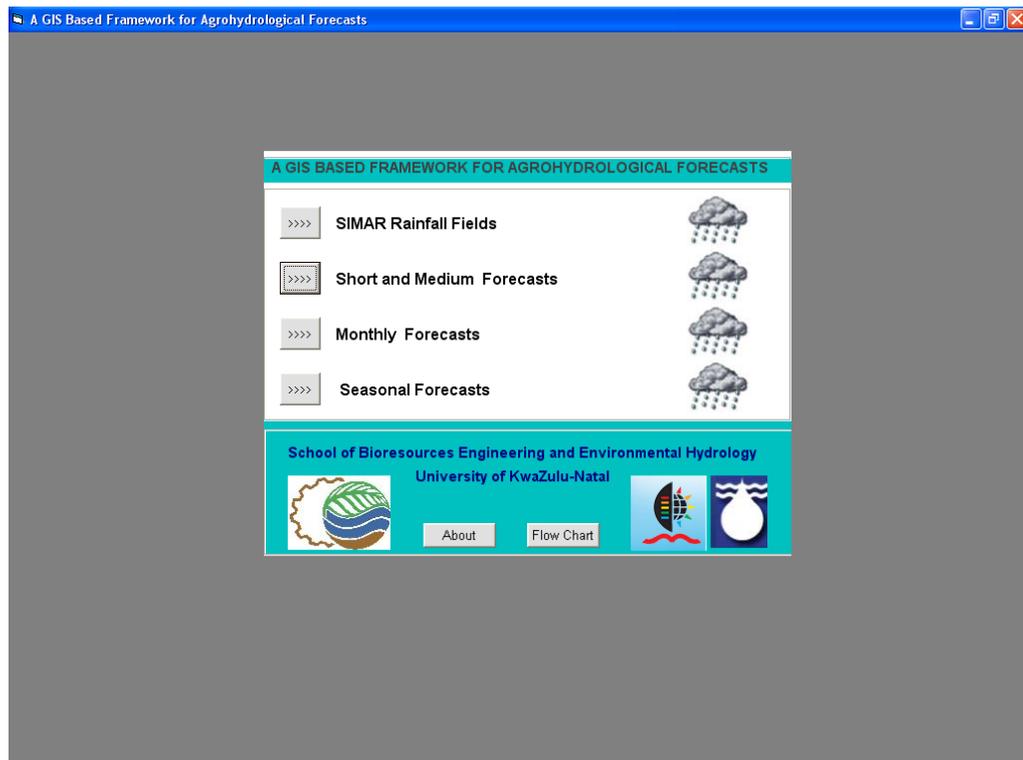


Figure 5.2    The main window showing options for near real time, short and medium as well as long range forecasting in the GIS based framework for the agrohydrological forecasting system

## 5.2    Near Real Time Estimates of Precipitation Derived from Satellite, Radar and Raingauge Data

Near real time weather information is, of necessity, required by hydrologists for nowcasting, especially in areas of fast hydrological response, and also to simulate the "now state" of various hydrological state variables such as soil moisture contents, streamflows, reservoir levels. Approaches for nowcasting are based mainly on rainfall estimated by conventional ground stations, radars, satellites and NWP models. These data and information sources have their respective strengths

and weaknesses. The use of conventional ground stations has become less efficient to meet the existing and anticipated management requirements in agricultural and water resources management because their distribution is sparse and data are frequently not available in mountainous where runoff is often generated, nor in other remote areas (Budhakooncharoen, 2003; Deyzel *et al.,* 2004; Kroese, 2004). For the above reasons, the use of near real time remotely sensed observations  from radar reflectivity measurements and satellite images has, therefore, been acknowledged to play a key role in agrohydrological applications, assisting in more timely decision making operations, especially for flash flood related disaster management. Presently, however, the outputs from satellite and radar images, although providing useful information on precipitation patterns, do not seem able to provide accurate rainfall values at the temporal and spatial resolution required by many hydrological models (Toth *et al.,* 2000). This is mainly so because of the problems related to ground clutter and false accumulation of rain fields when totals of rainfall are required. The raingauge networks then play a vital role in investigations regarding the elimination of ground clutter and also in verifications of radar and satellite derived rainfall on the ground (Deyzel *et al.,* 2004; Kroese, 2004).

By taking into consideration the merits and limitations of these data sources, the METSYS group of the SAWS and the School of Civil Engineering of University of KwaZulu-Natal, in collaboration with the Department of Water Affairs and Forestry (DWAF) and the national electricity utility ESKOM, have developed a new rainfall monitoring system termed SIMAR, i.e. **S**patial **I**nterpolation and **MA**pping of **R**ainfall. The system integrates raingauge, radar and satellite derived data in the production of daily rainfall maps of 24 hour accumulated rainfall at a resolution of one arc minute, i.e. approximately 1.7 km x 1.7 km over the southern Africa region. These maps are accessible on the Internet. A detailed description of these data sources is presented in Chapter 7.

The SIMAR project aims at producing one rainfall field that is acceptable by all water users (Deyzel *et al.,* 2004; Pegram, 2004). The generation of the merged radar/satellite/gauge rainfall field is a three step process, starting with the merging of the radar and raingauge fields followed by the merging of satellite and

raingauge fields. Thereafter the two resultant merged fields are combined (Pegram, 2004). In order to convert these maps into a suitable format and to downscale them to a particular location of interest (e.g. Quaternary Catchment) and use them as input into agrohydrological models, the following steps are required:

1. *Downloading rainfall maps*

   The accumulated rainfall for 24 hours, derived from daily reporting stations, radar and satellite across southern Africa arrives at the METSYS office in Bethlehem in the Free State by 09:00 daily. Daily rainfall maps from the radar, gauge and satellite information, together with the merged fields, are then completed by 11:30 and the results are posted on the METSYS website (Pegram, 2004). At this stage, however, these maps are considered to be demonstration versions which cannot be accessed in GIS. The accessible rainfall maps, which are given in ASCII format, can be downloaded from the SAWS ftp server on a daily basis.

2. *Converting formats*

   From the main window (Figure 5.2), clicking on the *SIMAR Rainfall Fields* option button initiates the *ASCII to Grid* window (Figure 5.3) to allow an ASCII format conversion to a grid format. This can be done by browsing the location into which the ASCII file is saved, and by specifying the output name and output directory, as shown in Figure 5.3. The grid layer will then be saved on the specified directory.
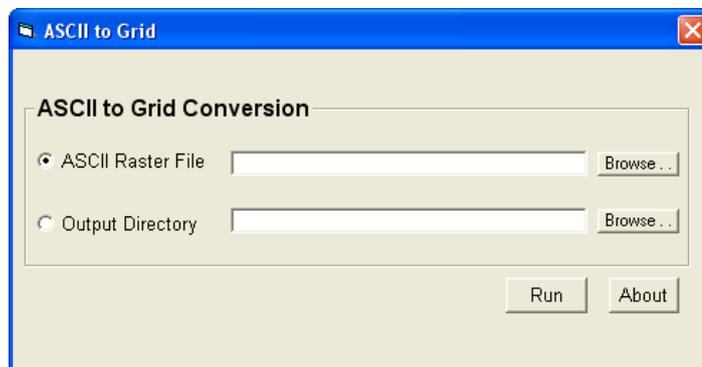


Figure 5.3    The *ASCII to Grid* window for format conversion

3. *Running ArcMap*

Once the format conversion has been completed, the *Forecasting tool* developed in the *ArcMap* shell automatically pops up. Clicking on the *SIMAR* button initiates the SIMAR window (Figure 5.4). The converted grid layer and shape file are added by browsing its path.
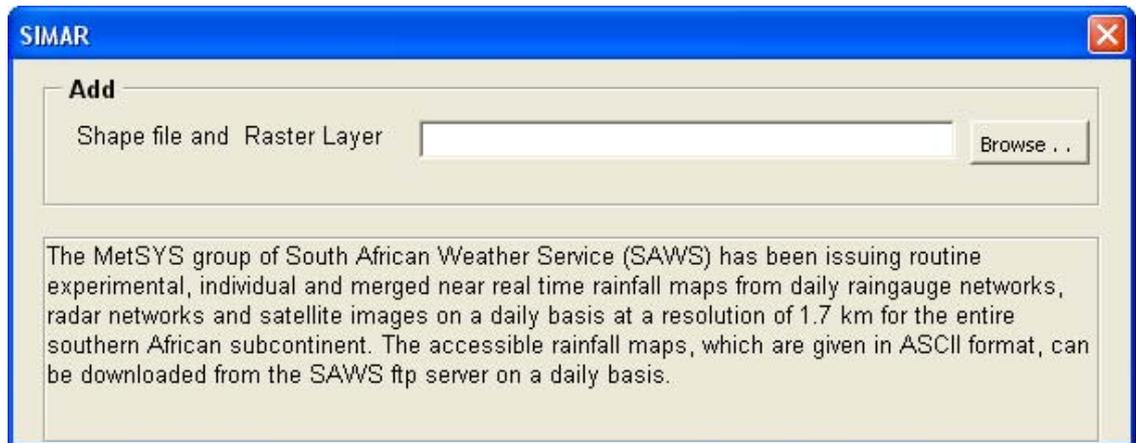


Figure 5.4     The screen for adding a grid layer for extracting daily rainfall values over a selected catchment

4. *Calculating catchment mean value and Joining of the Data*

The joining of the data can be done by averaging the points falling within each subcatchment of the chosen shape layer. Then the output is displayed automatically in "excel" format in the working directory.

5. *Converting  to ACRU format*

By inputting the forecast date as "yyyy/mm/dd" format in Figure 5.5, rainfall values representing each location will be extracted from the layer to respective *ACRU* model formatted input text files.
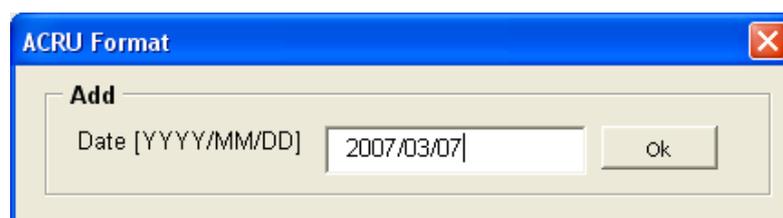


Figure 5.5     The screen for extracting daily rainfall values to *ACRU* formatted rainfall files

## 5.3    Short and Medium Forecasts from Weather Prediction Models

The SAWS is currently employing the Unified Model (UM) for short range weather forecasts (up to 2 days) and the National Center for Environmental Prediction for Medium Range Forecasting (NCEP-MRF) model for medium and extended range forecasts (up to 14 days) across the southern Africa. The rainfall forecasts from these two models and the forecasts issued by the University of Pretoria (UP) using the Conformal-Cubic Atmospheric Model (C-CAM) have, to date, been incorporated in the framework for short and medium range agrohydrological forecasting systems (Figure 5.6). The resolution, uncertainty and challenges associated with these models, as well as the procedures constructed to convert these forecasts into a suitable form, are described in detail in the sub-sections which follow.
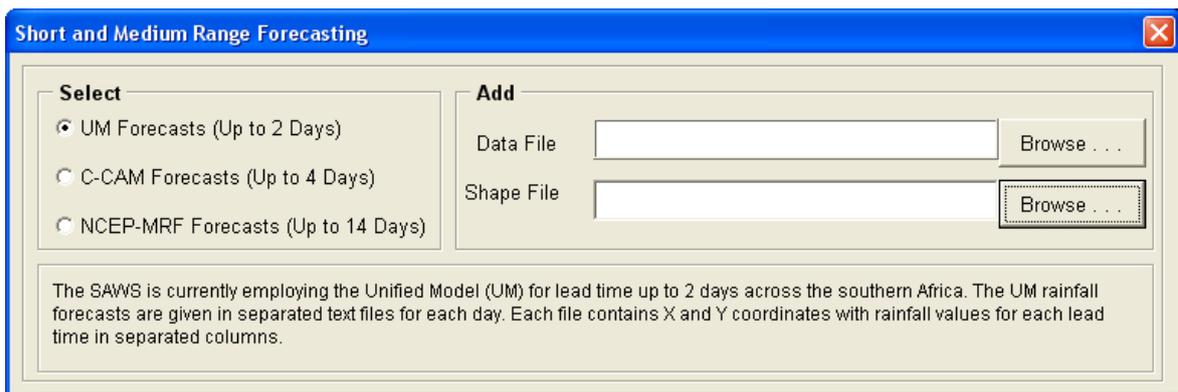


Figure 5.6    A screen showing the short and medium range forecasting model options

### 5.3.1  The C-CAM Rainfall Forecasts

Significant progress has been made at the University of Pretoria (UP) in simulating rainfall over southern Africa using the Conformal-Cubic Atmospheric Model (C-CAM). This progress had been made in collaboration with scientists in Australia. C-CAM is a relatively recent global model developed by the CSIRO Marine and Atmospheric Research in Melbourne, Australia (Reason *et al.,* 2006).  A key feature of C-CAM is its ability to stretch the model grid in order to focus the model resolution over any particular area of interest (Katzfey and McGregor, 2003; Engelbrecht, 2005; Reason *et al.,* 2006).  Initially, the model produces a coarse

resolution forecast with a grid spacing of 60 km over tropical and southern Africa. Far-field nudging is then used in order to force the higher resolution runs towards the low resolution portions and to dynamically downscale the 9-day global forecast with 60 km resolution into 4-day regional forecasts with a 15 km resolution. Further stretching of the model grid to an 8 km by 8 km resolution over the southern African region is currently under way (Engelbrecht, 2007). Although simulations are performed at a time step of five minutes, results are aggregated and issued on a daily basis. The 15 km resolution rainfall forecasts of four days' lead time are incorporated in this framework for application in the short term agrohydrological forecast system.

### 5.3.2 The UM Rainfall Forecasts

The UM is a non-hydrostatic weather forecasting model which had been developed in the UK Meteorological Office by the end of 1980s, but was introduced into operational service in 1992 (UK Met Office, 2007). The formulation of the model supports global and regional domains and is applicable to a wide range of temporal and spatial scales that allow it to be used for both numerical weather prediction and climate modelling as well as a variety of related research activities (Kershaw, 2006). The UM is designed to run either in atmosphere or ocean mode separately, or in a coupled mode. In each mode a run consists of an optional period of data assimilation followed by a prediction phase. Forecasts of a few days ahead are required for numerical weather prediction, while for climate modelling the prediction phase may be for tens, hundreds or even thousands of years (UK Office, 2007). The SAWS adopted the model for the southern Africa region in 2006. The UM model is run four times per day, providing model forecast guidance at a 12 km resolution for up to 2 days ahead (van Hemert, 2007).

### 5.3.3 The NCEP-MRF Rainfall Forecasts

At the NCEP the ensemble forecasting approach has been applied operationally for the short range forecasts by applying the ETA and Regional Spectral Models, and for the medium and extended range by using the Medium Range Forecast Model (MRF). Different ensemble based products have been generated and these

are distributed via File Transfer Protocol (ftp), to a wide range of users both nationally and internationally (Toth *et al.,* 1997).

Since 2003 the NCEP-MRF forecasts at grid spacing of 2.5° resolution with 22 ensemble members has been used operationally in South Africa for medium range forecasts up to 14 days ahead (Tennant *et al.,* 2006). At the present time, however, the SAWS has also been downloading a 1° x 1° grid spaced NCEP-MRF forecasts with 60 ensemble members every day, in addition to these 2.5° scaled forecasts (Tennant, 2007). One of the most challenging aspects of incorporating the NCEP-MRF rainfall forecasts into the framework for the agrohydrological forecasting system is that of condensing the vast amounts of model output and information into an operationally relevant and useful form. Currently, the SAWS are using the 2.5° grid spaced forecasts to produce one or two week lead time probabilistic rainfall forecasts by calculating the forecast probability that 24 hour precipitation amounts exceeding certain threshold values (usually 5 mm and 20 mm) over 2.5° by 2.5° grid boxes (Figure 5.7). For each day, 345 sets (i.e. 15 days and 23 ensembles) of unique forecasts are generated at each of the 2.5° by 2.5° grid boxes. Each ensemble represents an average probabilistic quantitative precipitation forecast (PQPF) for that 2.5° grid box.
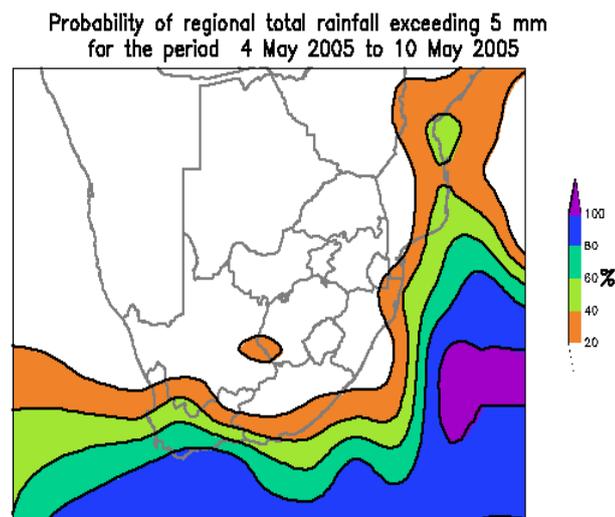


Figure 5.7    An example of a one week lead time probabilistic forecast from NCEP-MRF ensemble rainfall forecasts at 2.5° resolution over southern Africa (Source: SAWS, 2005)

At each grid box the number of ensemble members having a 24-hour precipitation amounts greater than the threshold limit are counted and the forecast probability is expressed as follows (Tennant, 2005):

$$F_P = \frac{M}{23} \times 100 \qquad\qquad 5.1$$

where

$F_P$     =     forecast probability (%), and

M     =     number of ensembles greater than a given threshold limit.

In order to take full advantage of the ensemble system, similar procedures should also be followed in the simulation of agrohydrological forecasts. However, most agrohydrological models run at a much finer spatial resolution than $2.5^{\circ}$ or $1^{\circ}$ and each ensemble member should be represented as geospatial (i.e. raster) data to be downscaled to the relevant catchment scale. By considering the computation time and file space required in the downscaling process, use of 23 or 60 ensemble forecasts becomes extremely cumbersome and difficult to comprehend. Hence, a single mean value of the ensemble members for each $2.5^{\circ}$ or $1^{\circ}$ grid box for a forecast lead time of 1 day up to 14 days will be used in the simulation of agrohydrological forecasts. Many studies (e.g. Toth *et al.,* 1997; Ebert, 2001) have shown that averaging the ensemble members allows not only the reduction of data sets and computational requirement, but also provides a more accurate forecast than any of the single ensemble forecasts. The evaluation of the ensemble forecast system to explain the uncertainties associated with selected forecasts is presented in Chapter 8, Sub-section 8.3.3.

## 5.3.4 Transferring and Reformatting of Forecasts

In May 2007 members of the WRC forecasting project (K5/1646) decided that all weather and climate forecasts would be fed to the University of Cape Town (UCT) in ASCII text format and the coarse spaced forecasts would be interpolated via cubic spline to a Quaternary (i.e. ~$0.25^{\circ}$) or Quinary (i.e. ~$0.1^{\circ}$) catchment scale. Forecasts would then be imported in ASCII format from the UCT for the application of agrohydrological forecasts.

The C-CAM, UM and NCEP-MRF rainfall forecasts are given in ASCII format in separated text files for each day. Each file contains X and Y coordinates with rainfall values for each lead time in separated columns. A program has been developed within this framework to convert these ASCII text file to a Data Base File (dbf) format in order to access it in GIS. The following steps are required to convert the C-CAM, UM and NCEP-MRF rainfall forecasts into suitable formats and to downscale them over a particular location of interest (a QC in this study) and use them as input in agrohydrological models:

1.  *Downloading rainfall forecasts*
    Rainfall forecasts should be imported or downloaded from the UCT ftp server on a daily basis.

2.  *Selecting the model option*
    From the main window of observation and forecast options in Figure 5.2, clicking on the *Short and Medium Forecasts* option initiates the *Short and Medium Range Forecasting* window which is nested in the *ArcMap* shell (Figure 5.6).

3.  *Browsing the location of a file and converting to a DBF file*
    After selecting one of the models in Figures 5.6, the location into which the data and shape files can be browsed by clicking on the *Browse* buttons. The text file will then be converted to a DBF file and the coarse spaced forecasts would be interpolated via Inverse Distance Weight (IDW) to a Quaternary (i.e. ~0.25$^\text{o}$) or Quinary (i.e. ~0.1$^\text{o}$) catchment scale.

From Section 5.2, steps 4 and 5 should be then followed to calculate catchment mean rainfall values, to join data to selected layer and, finally, to extract the rainfall forecasts to *ACRU* model formatted text input files.

Owing to their coarse spatial resolution, the NCEP-MRF rainfall forecasts are generally recommended for large scale agrohydrological forecasts. For small scale applications, dynamically downscaled rainfall forecasts from the C-CAM and UM models should be used. The high spatial and temporal resolution of the C-CAM

and UM forecasts allows the identification of features such as topography, land-sea distribution and land uses that influences the development of rainfall patterns over a particular region. Hence, the rainfall forecasts obtained from those models are expected to be more skilful than the forecasts made by NCEP-MRF model. However, since different GCM models exhibit different skill levels, the confidence that may be placed in downscaled rainfall forecasts is dependent foremost on the validity of the parent GCM model used to generate the large-scale fields. It must be noted that there is ongoing research (2007) at the University of the Free State (UFS) and University of Pretoria (UP) to combine dynamical downscaling with one or more statistical downscaling models. In the near future (2008), these models can then be incorporated in this framework.

## 5.4    Categorical Seasonal Forecasts from Climate Models

In southern Africa, seasonal (3-6 months) hydro-climatic forecasts are frequently required by different sectors of society as the region is severely affected by droughts and floods. Among the various sectors, water resources and agriculture obviously can benefit considerably from such long term forecasts.

As was described briefly in Chapter 3, Sub-section 3.2.2.2, a wide range of statistical and dynamical models have been developed by a number of institutions to issue seasonal forecasts for southern Africa. Until the recent past, the statistical models were most dominantly used in seasonal forecasts for southern Africa. Major improvements have been made in recent years in understanding southern Africa's seasonal climate by shifting from using only the empirical-statistical methods to more sophisticated forecast schemes involving the use of dynamical models (Landman and Goddard, 2005). In addition, the feasibility of producing probabilistic seasonal rainfall forecast skill for five equi-probable categories is in progress (Landman *et al.,* 2005). General Climate Models (GCMs) as forecast tools over southern Africa are currently available. GCMs are, however, unable to represent local sub-grid processes and tend to over-estimate rainfall over southern Africa. Moreover, the sub-gird representation of rainfall at mid-latitudes is highly complicated and may not be explicitly estimated by a GCM (Reason *et al.,* 2006). Recalibrated GCM output to regional levels was developed to overcome such systematic biases and this has the potential to outscore simple statistical models

(Landman *et al.,* 2001; Bartman *et al.,* 2003). Currently the SAWS compiles seasonal rainfall outlooks by combining output from Canonical Correlation Analysis (CCA), Quadratic Discriminant Analysis (QDA) and Atmospheric General Circulation Models (AGCMs). Results have shown that a combination of these different models consistently deliver a more skilful forecast than any individual model on its own (Klopper and Landman, 2003). Regional Climate Models (RCMs) have been used operationally in southern Africa since 2006 and they have the potential to simulate the seasonal rainfall variability and can subsequently be used to provide operational seasonal rainfall forecasts in the future (Reason *et al.,* 2006).

Seasonal forecasts of climate variables such as rainfall and temperature are often expressed as probabilities of occurrence within the above, near and below normal categories (Zhang and Casey, 1999). This approach has been adopted because of the inherent variability of the atmosphere and a lack of understanding of all the various components of the climate system (SAWS, 2005). A probability is assigned to each category, indicating the chance of a particular category to occur during the target season. The subsequent forecast probabilities indicate the direction of the forecast as well as the degree of confidence in the forecast. The higher the confidence in the forecast, the higher the assigned probability will be for that specific category. When there is no confidence in the forecast, climatological probabilities (33.3%) are assigned to each of the three categories (SAWS, 2005).

The SAWS has been producing seasonal forecasts in three equi-probable categories of below normal, near normal and above normal rainfalls for monthly and three consecutive months. These forecasts are available routinely on the SAWS website. However, production of seasonal climate forecasts in itself will not be enough for operational hydrological and agricultural decision making. Often in operational agrohydrological services, there is a need to estimate the consequences of seasonal climate forecasts with respect to agrohydrological variables that are closer to the actual problems faced by society such as streamflow amounts, reservoir levels, soil moisture contents and crop yield estimates. Hence, generic methodologies were developed in this framework for

temporal downscaling of categorical seasonal forecasts into a daily time series of values suitable for agrohydrological models.

## 5.4.1 Methods of Temporal Downscaling

Basically, weather generators and analogue methods are the most widely used methods for generating time series data that can be used as input to agrohydrological models. A stochastic *weather generator* employs stochastic methods to generate synthetic sequences of weather (Clark *et al.,* 2004). Stochastic weather generators have been used widely for simulating climate variables (e.g. precipitation, temperature, solar radiation) in climate change studies, but relatively little research had been done in relation to seasonal prediction (Feddersen and Andersen, 2005). The Markov Chain model is a widely used statistical technique to generate the sequence of rainy and dry (no rain) days.  It is based on the assumption that the state of any particular day is conditioned by the states of the previous day, or sequence of days. A distribution (e.g. Gamma) is fitted to the observed rainfall amounts for the target site. For the rainy days, rainfall values are sampled from the fitted distribution. Another set of weather generator methods generates weather by resampling data from historical records several times (e.g. Clark *et al*., 2004).

The second, relatively simple, approach is the *analogue* method which considers the assumption that a current synoptic situation will likely develop in the similar way as similar past synoptic situations have (WMO, 1992). Indices of climatic information, such as the ENSO status, SST or SOI and daily mean sea level pressure can be used to select analogue years from past records which had a similar status to that of the current situation, provided that these indices are well established for the target region. For example, indices of ENSO and SOI have been used in Ethiopia (e.g. Bekele, 1992) and Australia (e.g. Piechota *et al.,* 1998; Chiew *et al.,* 2003; Ritchie *et al.,* 2004) for similar purposes. The analogue approach has been used previously by several researchers in South Africa (e.g. Schulze *et al.*, 1998; Lumsden, 2000; Hallowes, 2002; Bezuidenhout, 2005) by first ranking the historical rainfall records in ascending order. The ranked rainfall totals are then grouped into three categories of seasonal rainfalls, *viz.* below normal, near normal and above normal. One approach is to select the median year

in each category as the analogue year and daily rainfall values representing the selected forecast season are then extracted from the selected median years.

The temporal downscaling method developed in this framework uses both the analogue and weather generator approaches. The analogue method used in this framework is also based on ranking of historical rainfall records, but analogue years are selected randomly, conditioned by the probability assigned to each category. Each category is weighted, based on the level of the confidence in the forecast. The higher the assigned probability, the higher the number of analogue years that will be sampled from that particular category. To generate the daily rainfall values representing the selected forecast season from each of the selected analogue years, two methods, *viz.* the *Historical Sequence Method* and the *Ensemble Re-ordering Based Method* (also termed the "Schaake shuffle") have been adopted in this study.

The *Historical Sequence Method* is based on the assumption that "daily rainfall values within the forecast season develop in similar sequences developed in the selected analogue years representing each category" (Schulze *et al.*, 1998). This approach provides one possible realisation of the past climate which is likely to occur in the future and attempts to preserve the historical temporal persistence of the past weather conditions that occurred in the selected analogue years. The *Historical Sequence Method* is described and evaluated in detail in Chapter 9.

Synthetic sequences of rainfall that are statistically consistent (in terms of the mean, variance, skew, long term persistency) with the observed characteristics of the historical data can provide alternative realisations that are equally likely to occur in the future and which can then be used to quantify uncertainty associated with climate variability. The synthetic sequences method randomly generates unique replicates (sequences), i.e. sequences of rain that have not observed. However, the approach should preserve the statistical moments of the historical time series from which they are populated (Clark *et al.*, 2004; Chiew *et al.,* 2005).
The "Ensemble Re-ordering" approach was applied by Clark *et al.* (2004) and uses random chance as the determining factor for an observation to be included in the sample that represent the forecast day. In this respect, the ensembles used to

populate the sequences are randomly selected from a mix of different dates of all historical years, or from a subset of preferentially selected years. For each forecast day, the ensemble members are re-ordered so as to preserve the spatio-temporal variability in the historical records. A description and evaluation of the *Ensemble Re-ordering Based Method* is presented in Chapter 10.

An algorithm has been coded within the framework that enables the processing of all the steps required for conditioning the random selection of analogue years on the probability assigned to each category (Figure 5.8). Moreover, the program has been designed to automatically extract daily data sets that represent estimates of future conditions for the targeted forecast season based on the *Historical Sequence* and *Ensemble Re-ordering Based Methods* (Figure 5.9). The following steps are contained in the algorithm and are applicable to both the monthly and seasonal (3 months) categorical climate forecasts:
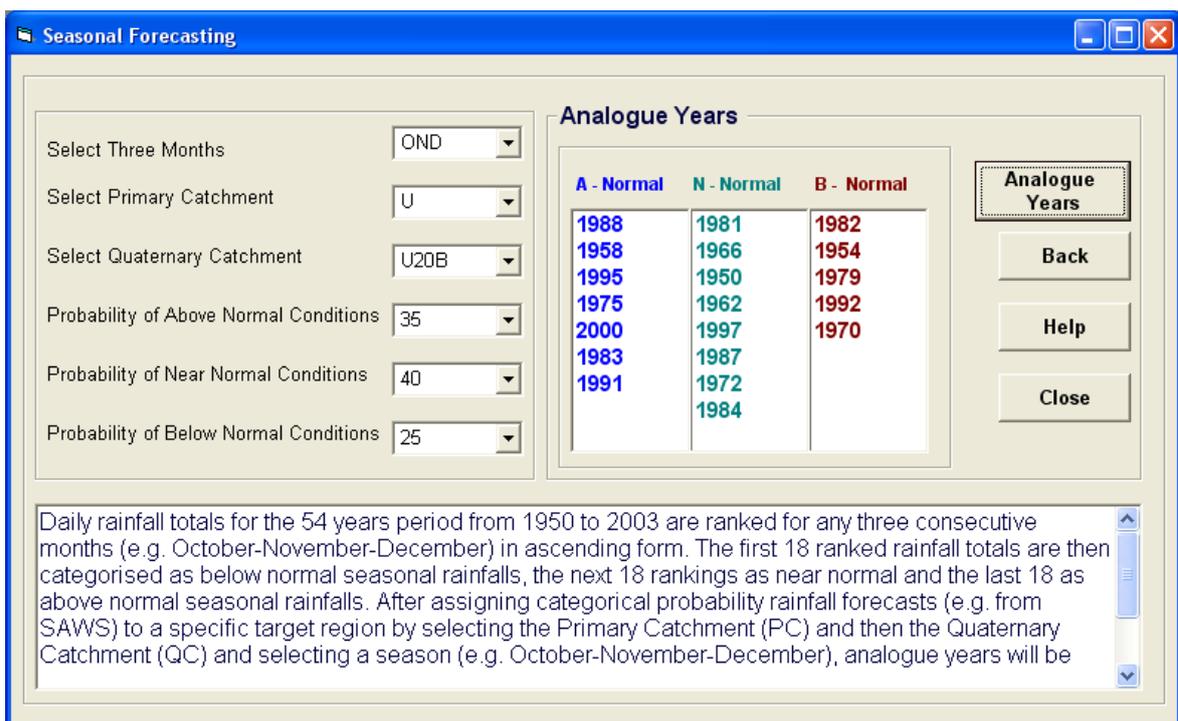


Figure 5.8    A window for translating seasonal categorical rainfall forecasts into daily time series values based on the analogue method

1.   *Ranking of daily rainfall totals*

Quality checked daily rainfall totals for the 54 years period from 1950 to 2003 are ranked for monthly and any three consecutive months (e.g. October-November-December) in ascending (lowest to highest) order. The first 18 ranked rainfall totals are then categorised as representing "below normal" seasonal rainfalls, the next 18 rankings as "near normal" and the highest 18 as "above normal" seasonal rainfalls.

2.   *Assigning inputs and selecting analogue years*

First, a season (e.g. October-November-December), the Primary Catchment (PC) and a Quaternary Catchment (QC) within the PC, as well as categorical probability rainfall forecasts obtained from various institutions (e.g. from SAWS) are selected from their respective drop down menus (Figure 5.8). Thereafter analogue years are randomly sampled, based on the probability assigned to each category. Since probabilities of categorical climate forecasts are usually given in multiples of 5 percentiles, the analogue years that represent each category are obtained by dividing the probability forecast by 5. In each run, therefore, 20 analogue years in total will be selected to represent the probability assigned to the three categories (Figure 5.8). For example, for each of the three categories, if the probabilities of above, near and below normal rainfall are 35%, 40% and 25% (as in Figure 5.8), the respective number of analogue years will be 7, 8 and 5.

3.   *Extracting daily rainfall values from selected analogue years*

Daily rainfall values representing the selected forecast season can then be extracted based on either the *Historical Sequence Method* or the *Ensemble Re-ordering Based Method* (Figure 5.9).

If the *Historical Sequence Method* is selected, 20 independent daily rainfall files from each of the analogue years will be generated. Each file has daily data sets extracted from the same dates in the historical records of the analogue years, and these files are then automatically used as the daily rainfall files for agrohydrological models.

If the *Ensemble Re-ordering Based Method* is chosen (Figure 5.9), the daily rainfall values from each of the selected analogue years for the target season are collected in a temporal array. The program then randomly resamples ten ensemble members for each forecast day of a given season from a mix of dates in the temporal array. Another random selection of dates from all historical years (1950-2003) of the same season is then used to re-order the temporal correlation structure of the ensembles selected from the preferentially selected analogue years. The random selection of dates from the historical records is only used for the first forecast day, and is persisted with for the subsequent forecast lead times. The re-ordered ensemble members can then be used as inputs into agrohydrological models. Unlike many other weather generator models, the temporal persistence is not preserved intrinsically, but is constructed as a post-processing step (Clark *et al.*, 2004). The concepts contained in the *Ensemble Re-ordering Based Method* are described more fully, and assessed quantitatively, in Chapter 10.
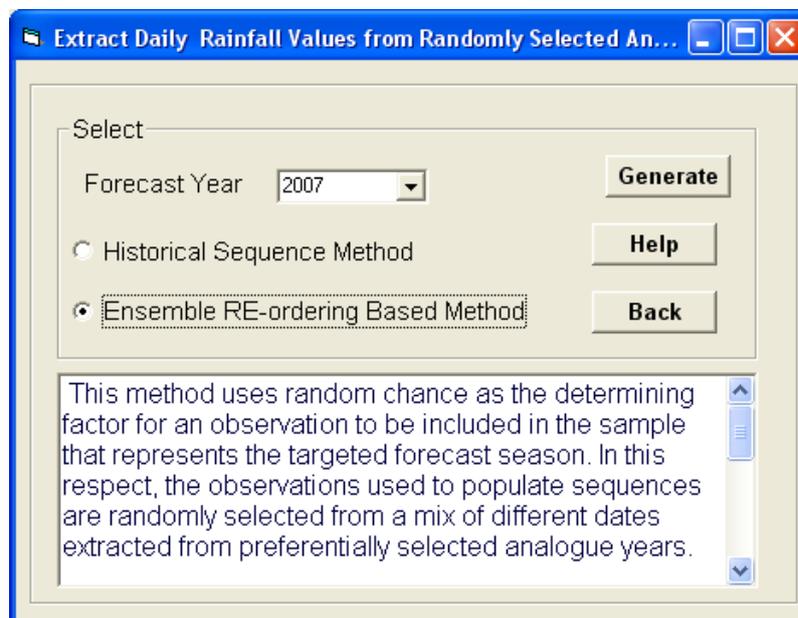


Figure 5.9    A window for extracting daily rainfall values from randomly selected analogue years

The main objective of developing the framework for southern Africa is to facilitate the translation of state-of the-art weather and climate forecasts into suitable

quantitative values which can be input into the daily time step hydrological and crop models. Once the translation process is completed, the subsequent step is the generation of agrohydrologically related forecasts (e.g. streamflows, reservoir levels, crop yields). For this purpose, the *ACRU* agrohydrological modelling system (Schulze, 1995a and updates) is employed in this study to generate agrohydrological forecasts. At a later stage it is envisaged that other daily models, such as CERES-Maize, will be imbedded within the framework. A brief overview of the *ACRU* model follows below.

## 5.5 The *ACRU* Agrohydrological Modelling System

### 5.5.1 Reasons for Selecting the *ACRU* Model

*ACRU* is a daily time step, multi-purpose and multi-level conceptual-physical agrohydrological simulation model. It was selected for this study because it has been widely verified under highly varying hydrological regimes on gauged catchments in southern Africa (cf. reviews by Schulze, 1995a; Schulze and Smithers, 2004) and elsewhere (e.g. Dunsmore *et al.,* 1986; Ghile, 2004). Furthermore, for southern Africa, *ACRU* is linked to extensive databases containing quality controlled daily rainfall, minimum and maximum temperatures for the period of 1950 to 2000 as well as to baseline land cover and soil information for each of the 1 946 hydrologically interlinked Quaternary Catchments (QCs) that make up southern Africa (Schulze, 2006).

The linking of the *ACRU* model to the databases is known as the Quaternary Catchments Database, QCD. A detailed description of the *ACRU* model in terms of inputs, simulation options and outputs is provided by Schulze (1995a) and Smithers and Schulze (1995; 2004). In the section which follows, only a brief overview of the concepts imbedded in the *ACRU* model is presented.

### 5.5.2 A Brief Description of the *ACRU* Model

As a conceptual-physical water budget model, *ACRU* (Schulze, 1995a and updates) integrates various water budgeting and runoff producing components of the terrestrial hydrological system, as well as operational aspects of water resource management, all with risk analysis (Schulze, 1995a; Schulze and

Smithers, 2004). The model was designed as a daily time-step, two layer soil water budgeting model which has been structured to be sensitive to land use changes on soil moisture, evaporative rates and runoff regimes. The model has been considerably updated from original versions to its present status (Schulze and Smithers, 2004) in order to simulate those components and processes of the hydrological cycle which are affected by the soil water budget, such as stormflow, baseflow, irrigation demand, sediment yield or crop yield, and to output any of those components on a daily basis (where relevant), or as monthly and annual totals of the daily values.

A summary of the concepts of the *ACRU* model with respect to inputs, operational modes, simulation options and objectives is given in Figure 5.10. Figure 5.11 represents a schematic of the multi-layer soil water budgeting by partitioning and redistribution of soil water, as conceptualised in the *ACRU* model.
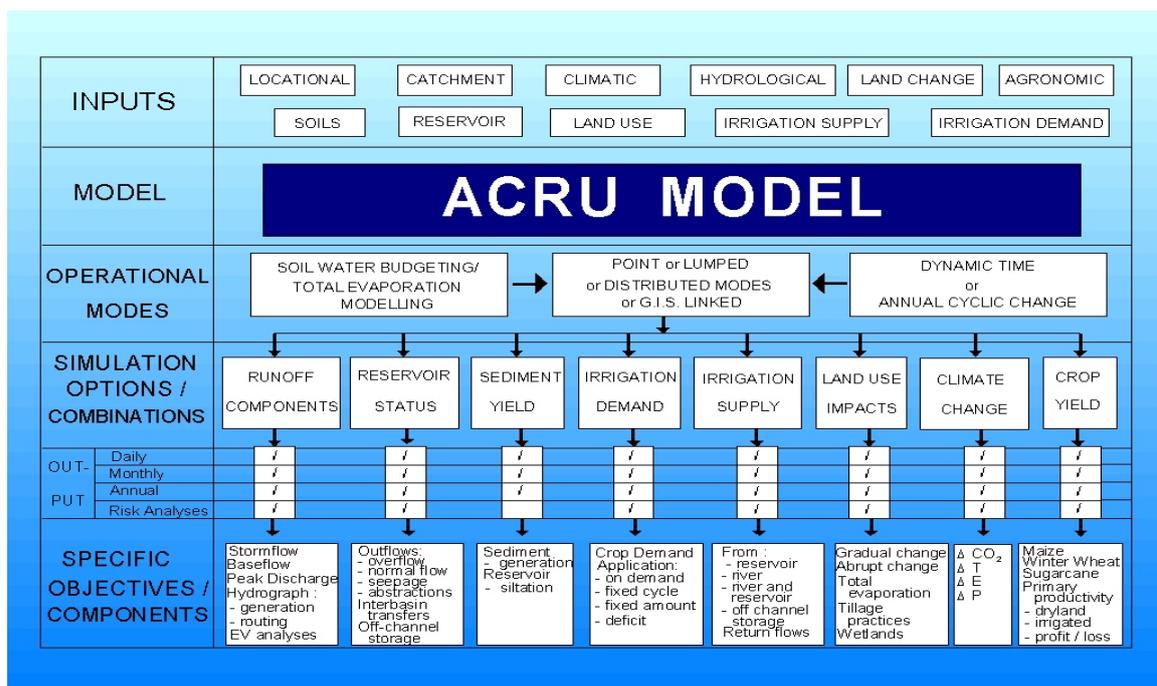


Figure 5.10   The *ACRU* agrohydrological model: Schematic of inputs, modes of operation, simulation options and objectives/ components (after Schulze, 1995a)
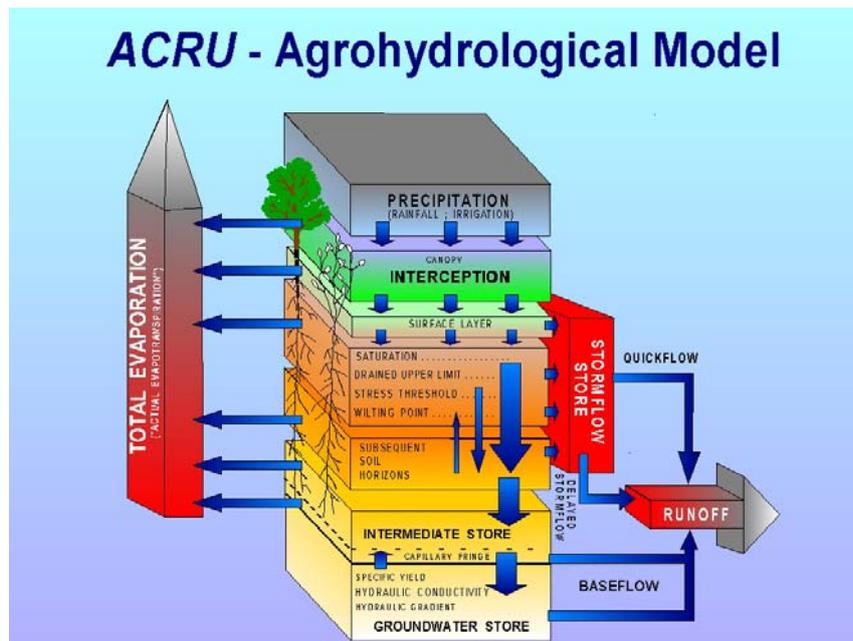
Figure 5.11   The *ACRU* agrohydrological model: Schematic of its multi-layer soil water budgeting and partitioning and redistribution of soil water (after Schulze, 1995a)

## 5.6   Summary

The development of effective procedures for the application of weather and climate forecasts into forecasts of various agrohydrological variables (e.g. streamflows, soil moisture, crop yields) plays a prominent role in operational decision-making in the agriculture and water sectors. For this purpose, a GIS based framework was developed to serve as an aid to process all the computations required in the translation of the daily to seasonal climate forecasts into daily quantitative values suitable as input in hydrological or crop models.  The framework was, and is being further, designed to include generic windows which allow users to process the near real time rainfall fields estimated by remotely sensed tools, as well as forecasts of weather/climate models into suitable scales and formats that are needed by many daily time step agrohydrological models. The key features of the framework are that it:

- facilitates the selection of near real time remotely sensed observations, as well as short term, medium term and longer term forecasts supplied by

various weather and climate models from different institutions across a range of time scales;

- links to comprehensive GIS functionality that provides tools for spatial disaggregation, data structure and reformatting, as well as for post-processing of data/information through tabulation, mapping and report generation;

- translates categorical seasonal forecasts into a daily time series of values suitable for agrohydrological models through generic algorithms developed within the framework;

- converts ensembles of rainfall forecasts into suitable formats which are understood by GIS;

- downscales grid layers to Quaternary Catchments; and

- finally, extracts rainfall data to *ACRU* formatted text input files.

The application of near real, plus daily to seasonal rainfall forecasts as a nested input to one or more agrohydrological models, thereby enabling the forecasting of agrohydrological variables across a range of time scales and lead times, is a new concept in southern African context. With further development and refinement, this framework has the potential to play an important role in bridging the gaps that exist between outputs of weather and climate models and their practical application in agrohydrological models. The development of the framework is an ongoing process and is expected to continue beyond the current stage, in order to incorporate other weather variables and forecast products issued by other institutions.

The outputs from the framework in which multiple forecasts are downscaled spatially to Quaternary Catchments and temporally to daily values needs to be evaluated either hydrologically or agriculturally. In this study a hydrological evaluation in a tested catchment has been opted for. A general background on the test catchment from a hydrological perspective, followed by a description of inputs required by the *ACRU* model for streamflow simulations is presented in Chapter 6.

# 6.  CASE STUDY: CONFIGURATION OF THE MGENI CATCHMENT FOR SIMULATION MODELLING WITH THE *ACRU* MODEL

It is widely accepted within the forecasting community that the skill of weather and climate models to realistically estimate the rainfall forecasts varies from storm to storm, from season to season, and from region to region (Seo *et al.,* 2000). The daily individual and merged rainfall fields derived from a network of daily reporting raingauges, radars and satellite images, as well as the estimated rainfall forecasts from the C-CAM, UM and NCEP-MRF models, are still experimental. Furthermore, the temporal downscaling methodologies developed in this study to translate the categorical rainfall forecasts into daily quantitative rainfall values are also new and untested in southern Africa. Verifications of the outputs of these weather and climate models and the two downscaling methodologies under a range of spatial and temporal scales are, therefore, of utmost importance if they are to be applied in decisions that improve management of climate related risks in water resources and agricultural operations.

For a comprehensive verification exercise, catchments representing a wide range of climates with complete hydrological and climatic data are, therefore, ideally needed for continuous assessment of the reliability of these models. In this study the Mgeni catchment was selected to serve as a point of departure in the verification phase, largely because of the availability and completeness of rainfall data from a relatively dense network of raingauges. Apart from its suitability for verification, the Mgeni is of major socio-economic importance in South Africa (cf. Section 6.1.1), making agrohydrological forecasting within the Mgeni catchment vital for many applications in climate sensitive sectors. The verification study will, at a later stage and beyond the timeframe of this thesis, be extended to other operational catchments of South Africa, in order to fully test the usefulness of the rainfall information/forecasts. As part of a wider WRC funded research programme, the configuration and collection of hydro-climatic data for the Modder, Berg, Mkomati and the Olifants catchments are currently underway by other researchers from other institutions, with the aim of conducting more detailed tests that will facilitate the operational use of these agrohydrological forecasts for the entire southern African region.

What follows in the next section is a general background on the Mgeni catchment from a hydrological perspective. A description of inputs required by the *ACRU* model for streamflow simulations and some concluding remarks are given in the subsequent sections.

## 6.1    General Background on the Mgeni Catchment

### 6.1.1  Location

The Mgeni catchment is home to over 5 million people in the Durban-Pietermaritzburg metropolitan area and produces approximately 20% of South Africa's gross national product (Schäfer and van Rooyen, 1993) from only 0.35% of the country's area. It is one of the South Africa's tertiary level catchments which have been delineated by the Department of Water Affairs and Forestry (DWAF). The catchment is located from 29º 13′ - 29º 46′ S and 29º 46′ - 30º 54′ E (Figure 6.1). The catchment, with an area of 4 469 km$^2$, ranges in altitude from zero to 2 103 m (Schulze *et al.,* 2004).
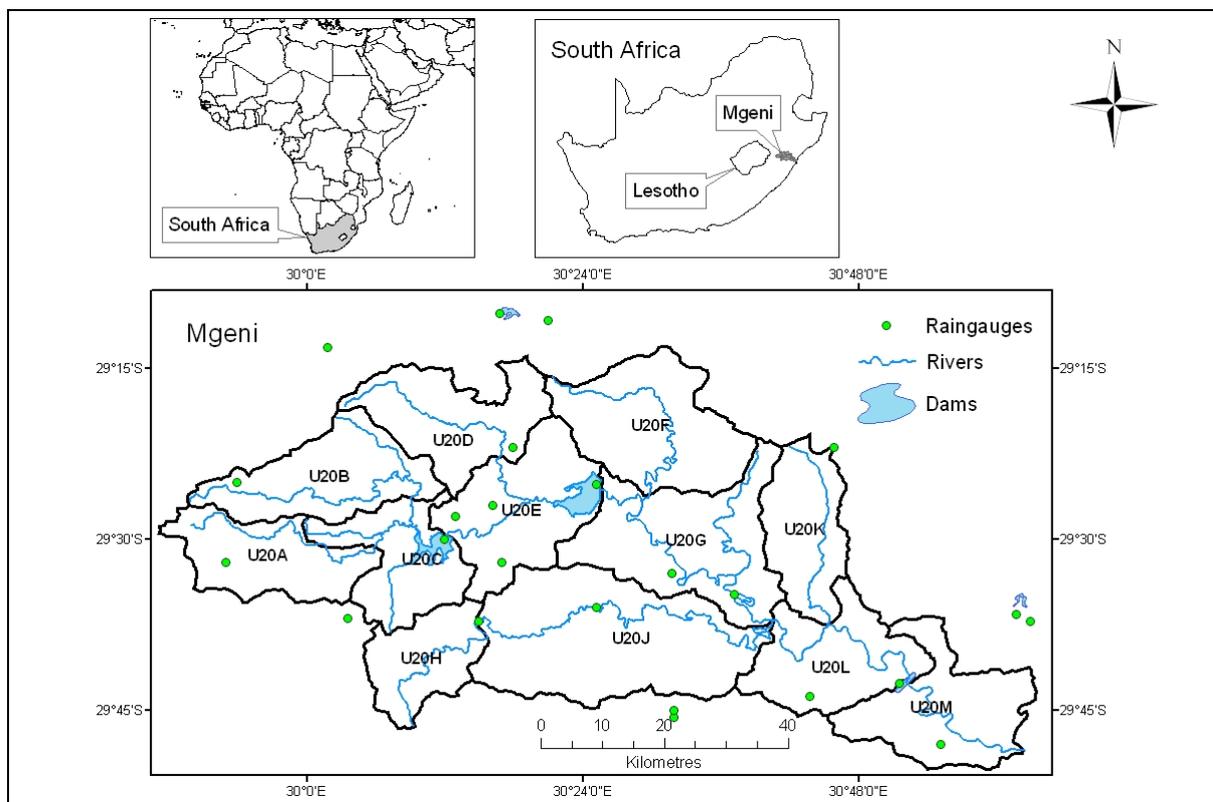


Figure 6.1     Overview of the Mgeni and its Quaternary Catchments

90

## 6.1.2 Climate and Hydrology

Rainfall is strongly seasonal and varies from 680 mm near the coast to 1 200 mm in the more rugged western parts of the Mgeni catchment, with 80% of the inland rainfall occurring largely as convective storms in the summer months (October-March), while along the coast lower intensity general rains in summer make up 65-70% of annual total (Schulze *et al.,* 2004). The catchment mean annual precipitation is 902 mm (Schulze *et al.,* 2004). Maximum daily temperatures are experienced in summer from December to February and minimum daily temperatures in winter in June and July (Schulze, 1997). Mean daily midwinter (July) maximum temperature increases from 12 $^{\circ}$C in the inland to 24 $^{\circ}$C on the coast on average, while means of daily maxima in midsummer (January) increase from 25 $^{\circ}$C in the inland to 28 $^{\circ}$C along the coast (Schulze, 1997). The catchment's mean annual temperature ranges form 16 to 18 $^{\circ}$C (DWAF, 2001). Snow occurs occasionally in winter at the higher altitudes of above 1 200 m near the Drakensberg, while the risk of hail also increases with proximity to the mountains (Rural Development Services, 2002).

The Mgeni catchment is characterised by high spatial and temporal variability of rainfalls and streamflows and is subjected to periodic droughts and heavy flooding (Kienzle *et al.,* 1997; Schulze, 1997; Schulze and Perks, 2000). Research conducted by Schulze (1997) has indicated that the inter- annual coefficient of variation (CoV%) over the Mgeni catchment ranges between 25 and 30%, while that of the annual runoff is between 50 and 100%. The conversion ratio of mean annual rainfall to mean catchment runoff is 18%. Climatically the Mgeni catchment is classified as a sub-humid zone (e.g. Van Zyl, 2003). However, considering the strong rainfall seasonality, low rainfall to runoff conversion and high ratio of annual evaporative demand, an extended area of the Mgeni catchment may be regarded as hydrologically semi-arid (Schulze, 1997).

## 6.1.3 Vegetation and Land Use

As shown in Figure 6.2, approximately 37% of the Mgeni catchment is under agriculture, consisting mainly of commercial production forestry, sugarcane plantations and subsistence farming, with some temporary commercial dryland

and irrigated agriculture. About 3% of the catchment consists of degraded bushland and shrubland, while 52% remains under natural vegetations and is comprised of grassland, bushland, and natural forest. Roughly 8% of the catchment land cover is urban, mostly residential, industrial, and commercial development associated with the cities of Durban at the coast and Pietermaritzburg inland (DEAT, 2001).
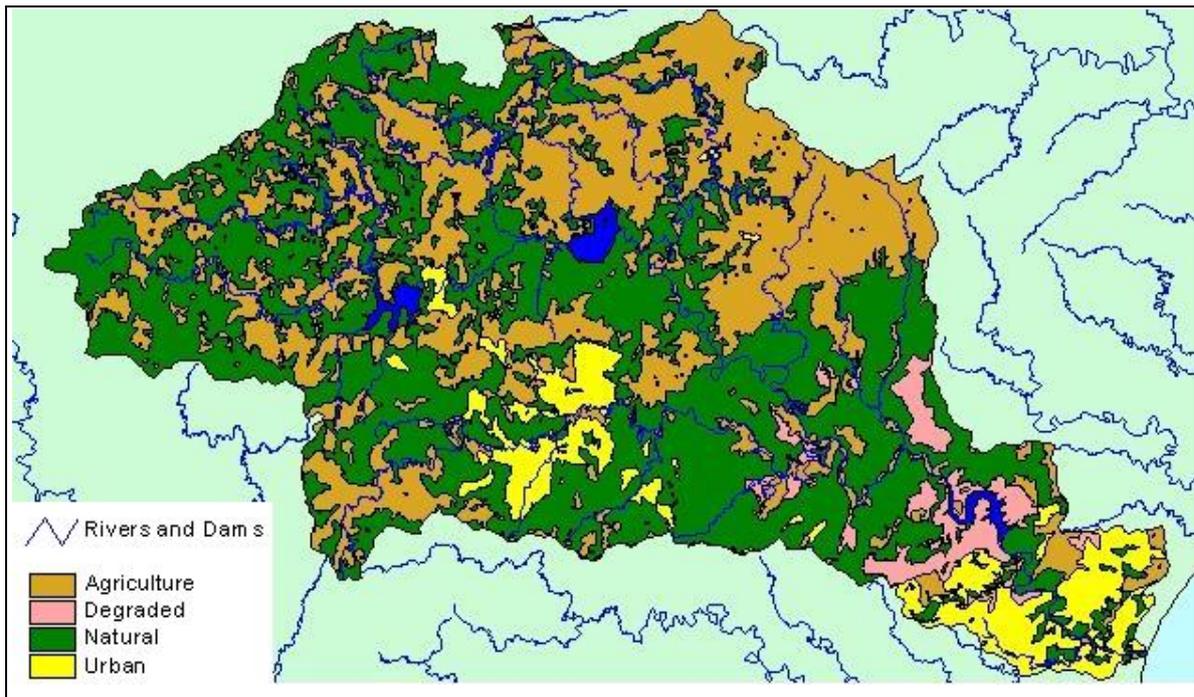


Figure 6.2    Catchment land cover and land use (Source: DEAT, 2001)

### 6.1.4  Water Use

A number of large storage dams have been constructed along the length of the Mgeni river from which water is abstracted and supplied to demand centres via various supply routes (Schäfer and van Rooyen, 1993; Kienzle *et al.,* 1997; Kjeldsen and Rosbjerg, 2001). The Mgeni river is approximately 232 km long (DEAT, 2001). The water resources of the Mgeni river basin are utilised for the supply of water to the Durban and Pietermaritzburg metropolitan complex, which is the third largest industrial and urban consumer base in southern Africa after Johannesburg and Cape Town (Schäfer and van Rooyen, 1993).

Since the mid 1980s, Umgeni Water, the authorised water board responsible for the management and bulk water supply of water, has supplied a mean volume of 20 million m$^3$ annually to consumers living within and adjacent to the Mgeni catchment area (Schulze *et al.,* 2004). Irrigation and afforestation are also the major water users in the Mgeni catchment. In addition to the surface water resources there are also groundwater resources which supply a considerable number of boreholes in the catchment.

Rapid rural, urban and industrial development within the Mgeni catchment, together with a predicted growth in the population to between 9 and 12 million by 2025, will increase water demand to be in excess of the available water resources (Tarboton and Schulze 1993; Kjeldsen and Rosbjerg, 2001). On the other hand, water quality of the streams, rivers and dams within the Mgeni river catchment has also been at risk. This is due mainly to the occurrence of irrigated and urban return flows, intensified agricultural practices and the unorganised growth of large informal settlements. Transport of suspended solids, pathogens and phosphorus during frequent convective thunderstorms is also common, leading to a severe deterioration of the water quality of the Mgeni river system (Kienzle *et al.,* 1997). Mean annual sediment yield within the Mgeni catchment ranges from 500-700 tonnes/km$^2$ (DWAF, 2001).

Considering the above water related problems, the DWAF and Umgeni Water have carried out a number of feasibility studies to assess how water could be transferred from other catchments to the Mgeni river system. Several alternatives have been attempted, including transfers of water from Mkomazi river. However, to date only the transfer of water from the Mearns diversion weir in the Mooi river to a tributary of the Mgeni river has proved to be economically viable (Fair, 1999; DWAF, 2004).

### 6.1.5 Geology

The geology of the Mgeni area is dominated by different types of sedimentary rocks with different strata of rock being exposed at varying altitudes. The majority of the western and central areas of the catchment are covered by *arenitic* and

*argillitic* rocks, with *tillites* in the centre of middle altitude range. The broad pattern of sedimentary rocks has, however, been altered by intrusions of undifferentiated assemblages of *dolerites*. The area near the mouth of the Mgeni river is exposed to a mixture of *tillites*, *shales* and *sandstones* (Rural Development Services, 2002).

## 6.2    Input Data to the *ACRU* Model for Streamflow Simulation

*ACRU* is a physical-conceptual (rather than a calibration model) and simulations are conducted based on measurable and/or derivable catchment characteristics (Schulze, 1995a). As already mentioned in Chapter 5 (Section 5.5), the *ACRU* model is linked to the Quaternary Catchment Database (QCD) which covers all of South Africa, Swaziland and Lesotho. The QCD has been automated through a Graphic User Interface (GUI) in the development of the *ACRU* Agrohydrological Modelling System (AAHMS), whereby information for a selected area can be automatically read into the *ACRU* model's input menu for the model to run (Smithers *et al.,* 2004). *ACRU*, being a multi-purpose model, can be used in the transformation rainfall forecasts into a range of agrohydrological forecasts. However, this case study focuses only on streamflow forecasts. Data, other than rainfall, which are required by the *ACRU* model for streamflow simulations at the exit of each of the QCs within the Mgeni catchment, were extracted from the QCD as described in the sub-sections which follow.

### 6.2.1  Subcatchment Information

As already stated in Chapter 5 (Section 5.5), *ACRU* is a multi-layer soil budgeting model in which the streamflow generation process is based on the premise that, after satisfying the initial abstractions (through interception, depression storage and infiltration before runoff commences), the streamflow produced is a function of the magnitude of rainfall and the soil water deficit from a critical response depth of the soil (Schulze, 1995a; Smithers and Schulze, 1995). Hence, detailed information on soils, land use and climate are required by *ACRU* to realistically simulate the soil water deficit antecedent to rainfall events on a daily basis.

In order to integrate the spatial variability of rainfall, soils and land cover in the Mgeni catchment, the *ACRU* model was configured in "distributed" mode with data from the QCD being extracted at the level of QCs. Although each QC is assumed to represent a relatively homogenous hydrological response unit, more than one soil type or land cover may still exist within it. In such cases, area-weighted values were assigned according to their respective areas within a QC (Smithers and Schulze, 1995).

Physiographic information for each of the 12 QC that make up the Mgeni catchment is shown in Table 6.1. Rainfall and temperatures in the Mgeni area tend to be closely related to altitude, with higher parts receiving higher amount of rainfalls and lower values of temperature. In addition to altitude, aspect has a major bearing on rainfall. The reason for this is that moist air enters the area from the southeast, and as a result the southeasterly slopes tend to be wetter than the northwesterly ones (Rural Development Services, 2002).

Table 6.1    Subcatchment physiographic information of the Mgeni catchment (after Schulze, 1997)

| Quaternary Catchment | Latitude (Degree. Minutes) | Longitude (Degree. Minutes) | Altitude (m) | Area (km$^2$) | MAP (mm) |
|---|---|---|---|---|---|
| U20A | 29° 32′ | 29° 57′ | 1595.1 | 295.0 | 1007 |
| U20B | 29° 24′ | 30° 03′ | 1420.2 | 355.0 | 989 |
| U20C | 29° 35′ | 30° 08′ | 1204.9 | 280.6 | 931 |
| U20D | 29° 21′ | 30° 13′ | 1318.7 | 340.4 | 1040 |
| U20E | 29° 29′ | 30° 19′ | 945.7 | 392.4 | 974 |
| U20F | 29° 19′ | 30° 28′ | 908.3 | 437.8 | 981 |
| U20G | 29° 31′ | 30° 34′ | 778.3 | 497.3 | 895 |
| U20H | 29° 41′ | 30° 08′ | 1270.0 | 221.0 | 942 |
| U20J | 29° 40′ | 30° 29′ | 761.1 | 683.0 | 840 |
| U20K | 29° 20′ | 30° 43′ | 778.7 | 272.9 | 952 |
| U20L | 29° 40′ | 30° 46′ | 437.4 | 331.0 | 808 |
| U20M | 29° 45′ | 30° 52′ | 262.6 | 362.7 | 923 |

The *ACRU* model was configured to simulate accumulated streamflows from subcatchments cascading downstream at the exit of each QC. Figure 6.3 shows the subcatchment configuration and flow cascading pattern of the Mgeni catchment. The shaded boxes are those QCs with major dams and with water flowing in from upstream.
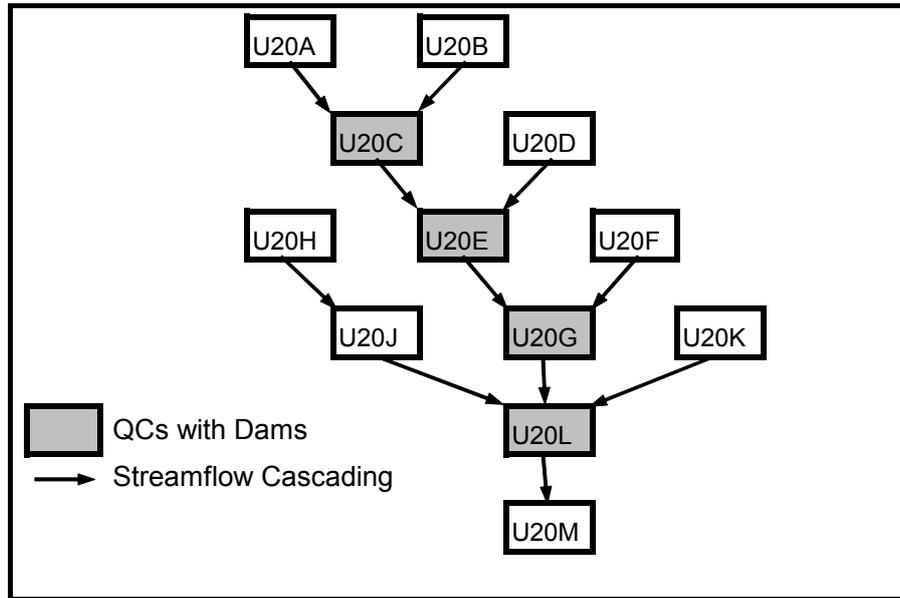
Figure 6.3    Subcatchment configuration and streamflow cascading pattern of the
              Mgeni catchment

## 6.2.2  Climate

The *ACRU* model operates at a daily time step and requires an input data file containing daily hydro-climatic values when simulating the hydrological processes. However, certain less sensitive variables (e.g. temperature) for which values can be input on a monthly basis if daily values are not available, are transformed internally in *ACRU* to daily values by Fourier analysis (Schulze, 1995a; Smithers and Schulze, 1995).

### 6.2.2.1    *Rainfall*

Rainfall is the fundamental driving force input behind most hydrological processes. This implies that the success of hydrological simulation studies depends to a large extent on the accuracy with which the rainfall data are observed temporally and spatially (Schulze *et al.,* 1995a). Two approaches, *viz.* the so-called "driver station method" and the "areal rainfall method" were used in this study, with the selection of the method based on the nature of rainfall information/forecasts.

The approach followed in the verification of the monthly and seasonal streamflow forecasts was the so-called "driver station method", in which one rainfall station per QC was selected to "drive" the hydrological response of a subcatchment. The reason for this is that the temporal downscaling of categorical rainfall forecasts into daily rainfall values for each QC is performed on the basis of the historical rainfall records measured by these driver stations. Rainfall files from these driver stations which represent each of the QC were updated to construct a complete daily record from January 1950 to March 2007. The issue of temporal downscaling, data extraction and format conversion of categorical rainfall forecasts are presented in detail in Chapters 9 and 10.

The SIMAR rainfall values as well as the short and medium range rainfall forecasts derived from the various Numerical Weather Prediction (NWP) models are given as daily averaged values over a grid box (pixel). One means of verifying this type of rainfall information is to extract pixel values of rainfall from each data source across the location of the driver stations. However, this limits the number of rainfall pixels used for evaluation purposes to only 12. The second alternative is to estimate areal rainfall from all the pixels falling within each of the QC. This option was chosen for this study in order to include all rainfall pixels within the evaluation statistics and to obtain a better idea of subcatchment areal rainfalls. Daily subcatchment average areal rainfalls were derived from the SIMAR data as well as from the outputs of NWP models for the selected evaluation period, and were successively converted into *ACRU* formatted rainfall input files.

For the purpose of their verification, observed areal rainfalls were estimated from point rainfall data measured by around 30 raingauges (including the 12 driver stations) distributed across the Mgeni catchment. The Natural Neighbours technique was employed for purposes of interpolation. This is discussed in more detail in Chapters 7 and 8.

### 6.2.2.2    *Temperature and potential evaporation*

Reference potential evaporation estimates are a second important climate input required by *ACRU* to simulate the soil water budget accurately, as most of the

rainfall is transformed into evaporation either through the soil or through the plant (Schulze and Kunz, 1995). Evaporation estimates have to be generated either directly from a pan or via surrogate means by physically based or empirical equations. However, since pan evaporation estimate is susceptible to measurement errors and are usually not available at the location of concern, it is a common practice in hydrology to use temperature information as a surrogate for estimating potential evaporation (Schulze and Kunz, 1995).

Like rainfall, daily to seasonal temperature forecasts can be obtained either from weather/climate models or from stochastic weather generator methods for use as input into the *ACRU* model for the simulation of streamflows. This case study, however, aims at evaluating the rainfall information/forecasts alone. Temperature forecasts were therefore not included in the streamflow simulations in order to avoid undesirable biases that could possibly affect the evaluation processes of the rainfall forecasts.

*ACRU*, being a multilevel model, contains a range of different options for deriving potential evaporation. One option in the absence of accurate pan evaporation or daily maximum and minimum temperature data is to use means of monthly totals of A-pan equivalent potential evaporation and monthly means of daily maximum and minimum temperatures, which can obtained from the "South African Atlas of Agrohydrology and –Climatology" (Schulze, 1997; Schulze, 2006). Each of the 1 946 QCs that make up southern Africa has been assigned areal-weighted values of monthly means of daily maximum and minimum temperatures, as well as means of monthly totals of A-pan equivalent potential evaporation. These parameters were clipped from the QCD for each of the QCs in the Mgeni catchment (Table 6.2).

Table 6.2    Monthly means of daily maximum and minimum temperatures ($^{o}$C) and means of monthly totals of A-pan equivalent potential evaporation (mm) for the 12 Quaternary Catchments of the Mgeni catchment, as input into the *ACRU* model (after Schulze, 1997)

| QC | Monthly Means of Daily Maximum Temperature, TMAX (ºC) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Jan | Feb | Mar | Apr | May | Jun | Jul | Aug | Sep | Oct | Nov | Dec |
| U20A | 23.4 | 23.4 | 22.4 | 20.6 | 18.7 | 16.4 | 16.7 | 18.4 | 20.1 | 20.6 | 21.4 | 23.1 |
| U20B | 24.4 | 24.4 | 23.4 | 21.6 | 19.6 | 17.3 | 17.6 | 19.3 | 21.0 | 21.6 | 22.4 | 24.1 |
| U20C | 25.4 | 25.4 | 24.6 | 22.7 | 20.7 | 18.3 | 18.6 | 20.2 | 21.9 | 22.5 | 23.3 | 25.1 |
| U20D | 24.7 | 24.7 | 23.8 | 22.0 | 20.1 | 17.8 | 18.1 | 19.7 | 21.4 | 21.9 | 22.6 | 24.4 |
| U20E | 26.6 | 26.6 | 25.9 | 24.1 | 22.1 | 19.8 | 20.0 | 21.5 | 23.1 | 23.7 | 24.5 | 26.3 |
| U20F | 26.5 | 26.5 | 25.9 | 24.1 | 22.2 | 20.0 | 20.2 | 21.6 | 23.1 | 23.7 | 24.3 | 26.2 |
| U20G | 26.7 | 26.8 | 26.2 | 24.5 | 22.6 | 20.5 | 20.6 | 21.9 | 23.2 | 23.8 | 24.5 | 26.4 |
| U20H | 24.8 | 24.9 | 24.1 | 22.3 | 20.3 | 18.0 | 18.3 | 19.8 | 21.4 | 22.0 | 22.8 | 24.5 |
| U20J | 26.9 | 27.0 | 26.4 | 24.7 | 22.7 | 20.6 | 20.7 | 22.0 | 23.4 | 24.0 | 24.8 | 26.6 |
| U20K | 26.2 | 26.3 | 25.8 | 24.1 | 22.4 | 20.3 | 20.4 | 21.6 | 22.8 | 23.3 | 24.0 | 25.8 |
| U20L | 27.7 | 27.9 | 27.4 | 25.7 | 24.1 | 22.1 | 22.0 | 23.0 | 24.0 | 24.7 | 25.5 | 27.3 |
| U20M | 27.4 | 27.6 | 27.2 | 25.6 | 24.2 | 22.3 | 22.2 | 22.9 | 23.6 | 24.1 | 25.0 | 26.9 |
| QC | Monthly Means of Daily Minimum Temperature, TMIN (ºC) |
| | Jan | Feb | Mar | Apr | May | Jun | Jul | Aug | Sep | Oct | Nov | Dec |
| U20A | 12.5 | 12.4 | 11.3 | 8.4 | 5.8 | 3.2 | 3.0 | 4.8 | 7.0 | 8.5 | 10.0 | 11.6 |
| U20B | 13.5 | 13.4 | 12.2 | 9.3 | 6.4 | 3.7 | 3.5 | 5.3 | 7.8 | 9.4 | 10.9 | 12.5 |
| U20C | 14.5 | 14.5 | 13.3 | 10.3 | 7.3 | 4.5 | 4.3 | 6.2 | 8.7 | 10.3 | 11.9 | 13.6 |
| U20D | 13.9 | 13.8 | 12.7 | 9.8 | 7.1 | 4.3 | 4.2 | 6.0 | 8.3 | 9.8 | 11.4 | 13.0 |
| U20E | 15.7 | 15.8 | 14.6 | 11.7 | 8.5 | 5.5 | 5.4 | 7.5 | 9.9 | 11.6 | 13.1 | 14.8 |
| U20F | 15.8 | 15.8 | 14.8 | 11.9 | 8.7 | 5.6 | 5.6 | 7.7 | 10.1 | 11.7 | 13.2 | 14.9 |
| U20G | 16.5 | 16.6 | 15.6 | 12.8 | 9.7 | 6.8 | 6.7 | 8.5 | 10.8 | 12.4 | 13.9 | 15.6 |
| U20H | 14.1 | 14.1 | 13.0 | 10.1 | 7.4 | 4.6 | 4.4 | 6.2 | 8.5 | 10.0 | 11.6 | 13.2 |
| U20J | 16.5 | 16.7 | 15.6 | 12.7 | 9.6 | 6.6 | 6.5 | 8.4 | 10.8 | 12.4 | 14.0 | 15.6 |
| U20K | 16.5 | 16.6 | 15.7 | 13.2 | 10.4 | 7.7 | 7.6 | 8.9 | 11.1 | 12.5 | 14.0 | 15.6 |
| U20L | 18.3 | 18.5 | 17.6 | 15.0 | 12.1 | 9.2 | 9.0 | 10.5 | 12.7 | 14.2 | 15.7 | 17.4 |
| U20M | 19.0 | 19.3 | 18.4 | 16.0 | 13.1 | 10.5 | 10.2 | 11.2 | 13.5 | 15.0 | 16.5 | 18.1 |
| QC | Means of Monthly Totals of A-Pan Equivalent Evaporation, E (mm) |
| | Jan | Feb | Mar | Apr | May | Jun | Jul | Aug | Sep | Oct | Nov | Dec |
| U20A | 164.7 | 140.3 | 135.3 | 117.5 | 100.3 | 97.8 | 107.6 | 127.7 | 143.9 | 148.8 | 147.1 | 171.5 |
| U20B | 172.3 | 146.4 | 141.0 | 119.6 | 102.0 | 96.5 | 106.8 | 129.2 | 146.3 | 154.6 | 153.2 | 177.3 |
| U20C | 179.0 | 153.5 | 145.5 | 121.3 | 103.3 | 95.8 | 105.6 | 129.4 | 146.5 | 158.4 | 157.9 | 182.3 |
| U20D | 172.2 | 146.4 | 141.4 | 120.0 | 102.8 | 97.0 | 107.0 | 129.5 | 145.9 | 154.3 | 152.5 | 176.6 |
| U20E | 181.4 | 155.8 | 149.3 | 124.8 | 105.6 | 94.9 | 105.6 | 132.0 | 150.0 | 160.5 | 161.1 | 185.3 |
| U20F | 177.8 | 153.3 | 147.9 | 125.4 | 107.1 | 95.7 | 106.6 | 131.8 | 149.9 | 160.2 | 160.7 | 183.9 |
| U20G | 175.4 | 154.9 | 147.6 | 123.3 | 106.8 | 95.1 | 104.7 | 127.5 | 145.5 | 158.7 | 160.7 | 183.2 |
| U20H | 174.7 | 149.6 | 143.0 | 120.4 | 102.0 | 96.2 | 106.2 | 128.2 | 144.2 | 154.0 | 153.4 | 178.9 |
| U20J | 178.5 | 157.7 | 148.7 | 123.2 | 106.0 | 94.6 | 104.1 | 128.1 | 146.0 | 158.4 | 161.8 | 185.1 |
| U20K | 169.2 | 152.4 | 145.5 | 121.4 | 106.5 | 95.1 | 104.5 | 123.8 | 141.4 | 158.5 | 158.5 | 179.9 |
| U20L | 177.7 | 162.0 | 151.3 | 122.7 | 107.5 | 93.5 | 101.4 | 121.3 | 140.0 | 159.7 | 163.5 | 185.8 |
| U20M | 168.8 | 156.8 | 149.5 | 120.0 | 106.1 | 92.4 | 98.5 | 112.5 | 130.4 | 157.8 | 160.1 | 181.3 |

## 6.2.3  Soils

Soil is a prime regulator through which many hydrological processes operate within a catchment (Schulze *et al.,* 1995b). The soils information contained in the QCD used in this study is derived from the Institute of Soil, Climate and Water's (ISCW) 84 Broad Natural Homogeneous Soil Zones. This soil information was converted into hydrological information which can be utilised by *ACRU* (Smithers and Schulze, 1995). The following soils variables are utilised in the *ACRU* soil

water budgeting routines for modelling the hydrological responses (Smithers and Schulze, 1995, 2004):

- Respective thicknesses of the topsoil (DEPAHO) and subsoil (DEPBHO);
- Soil water content at the permanent wilting point for the topsoil (WP1) and subsoil (WP2);
- Soil water content at the drained upper limit for the topsoil (FC1) and subsoil (FC2);
- Soil water content at saturation for the topsoil (PO1) and subsoil (PO2);
- Fraction of "saturated" soil water to be redistributed daily from the topsoil into the subsoil when the topsoil's water content is above its drained upper limit (ABRESP); and
- Fraction of "saturated" soil water to be redistributed daily from the subsoil into the intermediate/groundwater store when the subsoil's water content is above its drained upper limit (BFRESP).

Areal-weighted values of the above soil variables were assigned to each of the 1 946 QCs covering southern Africa (Schulze, 1997). The values for these variables were "clipped" from the QCD for each of the QC in the Mgeni catchment to be used as input into the *ACRU* model (Table 6.3).

Table 6.3    Horizon thicknesses (m),  critical soil  water retention constants $(m.m^{-1})$ and redistribution fractions for typical top- and subsoil horizons in the Mgeni catchment, as input into the *ACRU* model (Source: Schulze, 1997)

| QC | DEPAHO (m) | DEPBHO (m) | WP1 $(m.m^{-1})$ | WP2 $(m.m^{-1})$ | FC1 $(m.m^{-1})$ | FC2 $(m.m^{-1})$ | PO1 $(m.m^{-1})$ | PO2 $(m.m^{-1})$ | ABRESP (fraction) | BFRESP (fraction) |
|---|---|---|---|---|---|---|---|---|---|---|
| U20A | 0.28 | 0.49 | 0.15 | 0.20 | 0.25 | 0.29 | 0.42 | 0.41 | 0.39 | 0.39 |
| U20B | 0.30 | 0.55 | 0.16 | 0.21 | 0.25 | 0.31 | 0.42 | 0.42 | 0.42 | 0.42 |
| U20C | 0.27 | 0.48 | 0.15 | 0.20 | 0.24 | 0.29 | 0.39 | 0.40 | 0.35 | 0.35 |
| U20D | 0.31 | 0.55 | 0.17 | 0.22 | 0.27 | 0.33 | 0.40 | 0.42 | 0.38 | 0.38 |
| U20E | 0.28 | 0.50 | 0.16 | 0.20 | 0.25 | 0.30 | 0.39 | 0.40 | 0.34 | 0.34 |
| U20F | 0.30 | 0.57 | 0.15 | 0.20 | 0.25 | 0.30 | 0.42 | 0.42 | 0.39 | 0.39 |
| U20G | 0.29 | 0.51 | 0.15 | 0.18 | 0.24 | 0.27 | 0.44 | 0.42 | 0.39 | 0.39 |
| U20H | 0.30 | 0.53 | 0.16 | 0.20 | 0.25 | 0.30 | 0.42 | 0.42 | 0.40 | 0.40 |
| U20J | 0.28 | 0.46 | 0.16 | 0.20 | 0.26 | 0.29 | 0.42 | 0.42 | 0.35 | 0.35 |
| U20K | 0.29 | 0.53 | 0.14 | 0.17 | 0.23 | 0.26 | 0.44 | 0.42 | 0.42 | 0.42 |
| U20L | 0.25 | 0.38 | 0.13 | 0.15 | 0.22 | 0.23 | 0.41 | 0.39 | 0.37 | 0.37 |
| U20M | 0.24 | 0.40 | 0.15 | 0.15 | 0.24 | 0.24 | 0.44 | 0.43 | 0.41 | 0.41 |

### 6.2.4 Land Use Information

Land use and treatment measures play a significant role in altering the hydrological processes by influencing the interception, infiltration as well as the evaporation processes (Schulze, 1984). In order to account for these components of the hydrological cycle, *ACRU* requires the following land use related attributes for a variety of crops and/or land covers (Smithers and Schulze, 1995; 2004):

- A crop (i.e. water use) coefficient (CAY) that accounts for the vegetative water use relative to a reference potential evaporation in a given month;

- A land use/cover's canopy interception loss value (VEGINT). This monthly variable may be estimated by a number of methods in *ACRU* and accounts for the estimated interception loss by a plant canopy on a given day with rainfall;

- A fraction of the root system (ROOTA) that is actively extracting soil water from the topsoil in a given month; and

- A coefficient of initial abstraction (COIAM), which is used in the computation of the initial amounts of rainfall that do not contribute to the generation stormflow. This variable may change on a monthly or seasonal basis based on climatic and land use attributes.

In the case of the Mgeni catchment, in which 37% of the area is under intensive agriculture (cf. Section 6.1.3), the principal changes in agricultural crops and their level of management, as well as the rapid changes of land uses to residential and industrial lands as a result of population growth, migration and economic development, need to be considered explicitly, in order to realistically simulate the elements of streamflow. For this study, such updated changes in land use and managments were not readily available for the Mgeni catchment. Several options, however, exist for baseline land cover information to be derived for the above mentioned hydrological attributes through the *ACRU Menubuilder*. In this study the classification used to represent natural land cover conditions for the Mgeni catchment was that of Acocks' (1988) Veld Types (Figure 6.4), and area-weighted hydrological attributes for each of Acocks' Veld Types found in the Mgeni catchment were retrieved from the QCD (Table 6.4). No human land use

catchment or in-channel activities, such as irrigation abstractions, return flows or and dams were considered in this study.

Table 6.4    Land cover information used in the *ACRU* model for each of the Quaternary Catchments in the Mgeni catchment

| Monthly Means of Crop Coefficients, CAY (fraction) | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| QC | Jan | Feb | Mar | Apr | May | Jun | Jul | Aug | Sep | Oct | Nov | Dec |
| U20A | 0.70 | 0.70 | 0.70 | 0.50 | 0.31 | 0.20 | 0.20 | 0.20 | 0.50 | 0.65 | 0.70 | 0.70 |
| U20B | 0.70 | 0.70 | 0.70 | 0.50 | 0.31 | 0.21 | 0.20 | 0.20 | 0.51 | 0.66 | 0.70 | 0.70 |
| U20C | 0.73 | 0.73 | 0.73 | 0.50 | 0.36 | 0.20 | 0.20 | 0.20 | 0.53 | 0.68 | 0.73 | 0.73 |
| U20D | 0.70 | 0.70 | 0.70 | 0.50 | 0.32 | 0.22 | 0.20 | 0.20 | 0.52 | 0.67 | 0.70 | 0.70 |
| U20E | 0.71 | 0.71 | 0.71 | 0.50 | 0.38 | 0.23 | 0.20 | 0.21 | 0.55 | 0.70 | 0.71 | 0.71 |
| U20F | 0.70 | 0.70 | 0.70 | 0.50 | 0.35 | 0.24 | 0.20 | 0.20 | 0.55 | 0.70 | 0.70 | 0.70 |
| U20G | 0.71 | 0.71 | 0.71 | 0.61 | 0.50 | 0.33 | 0.32 | 0.41 | 0.58 | 0.69 | 0.69 | 0.71 |
| U20H | 0.70 | 0.70 | 0.70 | 0.51 | 0.36 | 0.26 | 0.23 | 0.24 | 0.54 | 0.67 | 0.69 | 0.70 |
| U20J | 0.73 | 0.73 | 0.73 | 0.59 | 0.49 | 0.30 | 0.29 | 0.37 | 0.58 | 0.69 | 0.71 | 0.73 |
| U20K | 0.71 | 0.71 | 0.71 | 0.61 | 0.49 | 0.36 | 0.35 | 0.42 | 0.58 | 0.68 | 0.68 | 0.71 |
| U20L | 0.76 | 0.76 | 0.76 | 0.69 | 0.59 | 0.33 | 0.33 | 0.49 | 0.65 | 0.75 | 0.75 | 0.76 |
| U20M | 0.82 | 0.82 | 0.82 | 0.79 | 0.69 | 0.53 | 0.53 | 0.66 | 0.78 | 0.82 | 0.82 | 0.82 |
| Canopy Interception Loss, VEGINT (mm.rainday$^{-1}$) | | | | | | | | | | | |
| QC | Jan | Feb | Mar | Apr | May | Jun | Jul | Aug | Sep | Oct | Nov | Dec |
| U20A | 1.60 | 1.60 | 1.60 | 1.42 | 1.24 | 1.05 | 1.05 | 1.05 | 1.32 | 1.60 | 1.60 | 1.60 |
| U20B | 1.56 | 1.56 | 1.56 | 1.37 | 1.17 | 1.04 | 1.04 | 1.04 | 1.33 | 1.56 | 1.56 | 1.56 |
| U20C | 1.58 | 1.58 | 1.58 | 1.51 | 1.37 | 1.26 | 1.26 | 1.26 | 1.43 | 1.58 | 1.58 | 1.58 |
| U20D | 1.54 | 1.54 | 1.54 | 1.34 | 1.14 | 1.05 | 1.05 | 1.05 | 1.35 | 1.54 | 1.54 | 1.54 |
| U20E | 1.56 | 1.56 | 1.56 | 1.44 | 1.27 | 1.24 | 1.23 | 1.23 | 1.46 | 1.56 | 1.56 | 1.56 |
| U20F | 1.53 | 1.53 | 1.53 | 1.33 | 1.13 | 1.13 | 1.13 | 1.13 | 1.43 | 1.53 | 1.53 | 1.53 |
| U20G | 1.81 | 1.81 | 1.81 | 1.66 | 1.50 | 1.46 | 1.42 | 1.46 | 1.68 | 1.81 | 1.81 | 1.81 |
| U20H | 1.52 | 1.52 | 1.52 | 1.35 | 1.16 | 1.09 | 1.09 | 1.10 | 1.37 | 1.52 | 1.52 | 1.52 |
| U20J | 1.79 | 1.79 | 1.79 | 1.69 | 1.55 | 1.49 | 1.46 | 1.49 | 1.66 | 1.79 | 1.79 | 1.79 |
| U20K | 1.67 | 1.67 | 1.67 | 1.55 | 1.40 | 1.35 | 1.33 | 1.38 | 1.58 | 1.67 | 1.67 | 1.67 |
| U20L | 2.49 | 2.49 | 2.49 | 2.29 | 2.01 | 1.89 | 1.83 | 1.95 | 2.29 | 2.49 | 2.49 | 2.49 |
| U20M | 2.94 | 2.94 | 2.94 | 2.87 | 2.37 | 1.99 | 1.97 | 2.34 | 2.87 | 2.94 | 2.94 | 2.94 |
| Fraction of Active Root System in Topsoil Horizon, ROOTA (fraction) | | | | | | | | | | | |
| QC | Jan | Feb | Mar | Apr | May | Jun | Jul | Aug | Sep | Oct | Nov | Dec |
| U20A | 0.90 | 0.90 | 0.90 | 0.95 | 0.99 | 1.00 | 1.00 | 1.00 | 0.95 | 0.90 | 0.90 | 0.90 |
| U20B | 0.90 | 0.90 | 0.90 | 0.94 | 0.98 | 1.00 | 1.00 | 1.00 | 0.95 | 0.90 | 0.90 | 0.90 |
| U20C | 0.90 | 0.90 | 0.90 | 0.94 | 0.96 | 0.99 | 0.99 | 0.99 | 0.94 | 0.90 | 0.90 | 0.90 |
| U20D | 0.90 | 0.90 | 0.90 | 0.94 | 0.97 | 1.00 | 1.00 | 1.00 | 0.95 | 0.90 | 0.90 | 0.90 |
| U20E | 0.89 | 0.89 | 0.89 | 0.93 | 0.95 | 0.99 | 0.99 | 0.99 | 0.94 | 0.89 | 0.89 | 0.89 |
| U20F | 0.89 | 0.89 | 0.89 | 0.93 | 0.95 | 0.99 | 0.99 | 0.99 | 0.94 | 0.89 | 0.89 | 0.89 |
| U20G | 0.86 | 0.86 | 0.86 | 0.89 | 0.93 | 0.97 | 0.97 | 0.96 | 0.91 | 0.86 | 0.86 | 0.86 |
| U20H | 0.90 | 0.90 | 0.90 | 0.93 | 0.97 | 0.99 | 0.99 | 0.99 | 0.94 | 0.90 | 0.90 | 0.90 |
| U20J | 0.87 | 0.87 | 0.87 | 0.90 | 0.93 | 0.98 | 0.98 | 0.96 | 0.91 | 0.87 | 0.87 | 0.87 |
| U20K | 0.87 | 0.87 | 0.87 | 0.89 | 0.94 | 0.97 | 0.97 | 0.96 | 0.91 | 0.87 | 0.87 | 0.87 |
| U20L | 0.80 | 0.80 | 0.80 | 0.83 | 0.87 | 0.93 | 0.93 | 0.90 | 0.86 | 0.80 | 0.80 | 0.80 |
| U20M | 0.76 | 0.76 | 0.76 | 0.77 | 0.78 | 0.81 | 0.81 | 0.80 | 0.78 | 0.76 | 0.76 | 0.76 |
| Coefficient of Initial Abstraction, COIAM (fraction) | | | | | | | | | | | |
| QC | Jan | Feb | Mar | Apr | May | Jun | Jul | Aug | Sep | Oct | Nov | Dec |
| U20A | 0.15 | 0.15 | 0.24 | 0.30 | 0.30 | 0.30 | 0.30 | 0.30 | 0.30 | 0.30 | 0.20 | 0.15 |
| U20B | 0.15 | 0.15 | 0.23 | 0.30 | 0.30 | 0.30 | 0.30 | 0.30 | 0.30 | 0.30 | 0.20 | 0.15 |
| U20C | 0.15 | 0.15 | 0.21 | 0.30 | 0.30 | 0.30 | 0.30 | 0.30 | 0.30 | 0.30 | 0.20 | 0.15 |
| U20D | 0.15 | 0.15 | 0.22 | 0.30 | 0.30 | 0.30 | 0.30 | 0.30 | 0.30 | 0.30 | 0.20 | 0.15 |
| U20E | 0.15 | 0.15 | 0.20 | 0.30 | 0.30 | 0.30 | 0.30 | 0.30 | 0.30 | 0.29 | 0.20 | 0.15 |
| U20F | 0.15 | 0.15 | 0.20 | 0.30 | 0.30 | 0.30 | 0.30 | 0.30 | 0.30 | 0.29 | 0.20 | 0.15 |
| U20G | 0.18 | 0.18 | 0.23 | 0.30 | 0.30 | 0.30 | 0.30 | 0.30 | 0.30 | 0.28 | 0.22 | 0.18 |
| U20H | 0.15 | 0.15 | 0.22 | 0.30 | 0.30 | 0.30 | 0.30 | 0.30 | 0.30 | 0.30 | 0.20 | 0.15 |
| U20J | 0.18 | 0.18 | 0.23 | 0.30 | 0.30 | 0.30 | 0.30 | 0.30 | 0.30 | 0.28 | 0.21 | 0.18 |
| U20K | 0.18 | 0.18 | 0.23 | 0.30 | 0.30 | 0.30 | 0.30 | 0.30 | 0.30 | 0.28 | 0.22 | 0.18 |
| U20L | 0.22 | 0.22 | 0.26 | 0.30 | 0.30 | 0.30 | 0.30 | 0.30 | 0.30 | 0.26 | 0.22 | 0.22 |
| U20M | 0.27 | 0.27 | 0.28 | 0.30 | 0.30 | 0.30 | 0.30 | 0.30 | 0.30 | 0.28 | 0.27 | 0.27 |
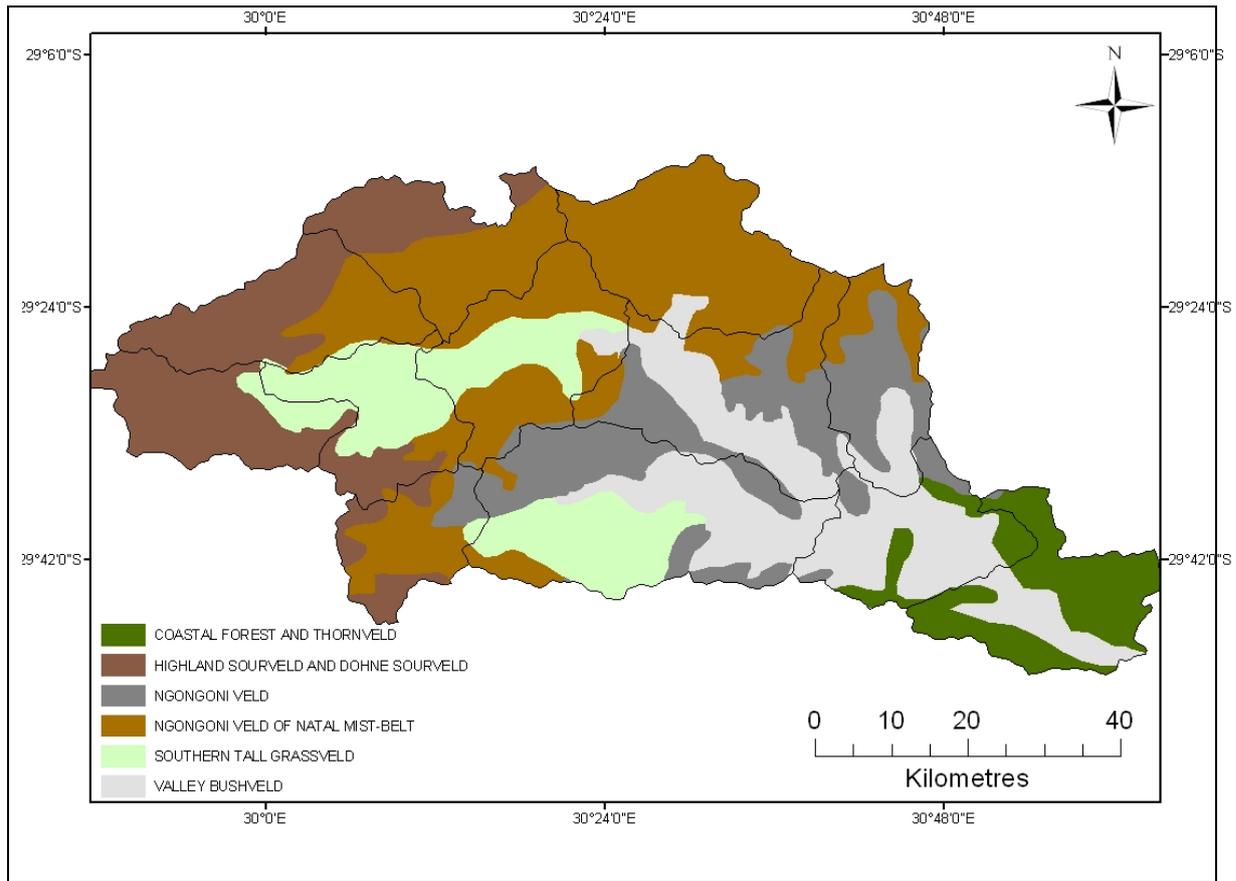
Figure 6.4    Baseline land cover in the Mgeni catchment, represented by Acocks'
              (1988) Veld Types

## 6.2.5  Streamflow Simulation Control Variables

In the *ACRU* model streamflow is derived from two components, *viz. baseflow* and
*stormflow*, with the *stormflow* component consisting of a *quickflow* response, i.e.
that fraction of *stormflow* generated which is released into a stream on the same
day as the rainfall event, and a *delayed stormflow* response, i.e. the remaining
*stormflow* entering a stream over several days due to interflow. *Baseflow* is
derived from slow releases from the intermediate and groundwater stores which
are recharged by the drainage out of the lower soil horizon when its water content
exceeds the upper drained limit (Schulze, 1995b).

In regard to streamflow simulation, the following control variables are required by
the *ACRU* model (Smithers and Schulze, 1995; 2004):

- The fraction of the total stormflow generated on a given day that will run off from a catchment or a subcatchment on the same day as the rainfall event that generated that stormflow (QFRESP);

- The fraction of water from the intermediate/groundwater store that becomes streamflow on a particular day (COFRU);

- The effective (critical) depth of the soil (m) from which stormflow is generated (SMDDEP);

- A variable to request the inclusion or exclusion of baseflow from the simulation of streamflow (IRUN);

- The fraction of the catchment which is occupied by adjunct impervious areas which are connected directly to a watercourse, from which precipitation contributes directly to quickflow (ADJIMP);

- The fraction of the catchment which is occupied by disjunct impervious areas which are not connected directly to a watercourse. Precipitation falling on these impervious areas does not contributes directly to streamflow, but is assumed to re-infiltrate on the adjunct pervious portion of the catchment (DISIMP); and

- The impervious surface's storage capacity (i.e. depression storage), which needs to be filled before stormflow commences (STOIMP).

Table 6.5 shows the streamflow simulation control variables for each of the QCs in the Mgeni catchment, which were extracted from the QCD.

Table 6.5    *ACRU* model streamflow simulation control variables for each of the Quaternary Catchments in the Mgeni catchment

| QC | QFRESP (fraction) | COFRU (fraction) | SMDDEP (m) | IRUN (0=No or 1=Yes) | ADJIMP (fraction) | DISIMP (fraction) | STOIMP (mm) |
|---|---|---|---|---|---|---|---|
| U20A | 0.3 | 0.009 | 0.28 | 1 | 0.03 | 0.12 | 1.00 |
| U20B | 0.3 | 0.009 | 0.30 | 1 | 0.02 | 0.07 | 1.00 |
| U20C | 0.3 | 0.009 | 0.27 | 1 | 0.01 | 0.05 | 1.00 |
| U20D | 0.3 | 0.009 | 0.31 | 1 | 0.02 | 0.07 | 1.00 |
| U20E | 0.3 | 0.009 | 0.28 | 1 | 0.01 | 0.04 | 1.00 |
| U20F | 0.3 | 0.009 | 0.30 | 1 | 0.01 | 0.04 | 1.00 |
| U20G | 0.3 | 0.009 | 0.29 | 1 | 0.01 | 0.02 | 1.00 |
| U20H | 0.3 | 0.009 | 0.30 | 1 | 0.01 | 0.04 | 1.00 |
| U20J | 0.3 | 0.009 | 0.28 | 1 | 0.01 | 0.04 | 1.00 |
| U20K | 0.3 | 0.009 | 0.29 | 1 | 0.01 | 0.03 | 1.00 |
| U20L | 0.3 | 0.009 | 0.25 | 1 | 0.02 | 0.05 | 1.00 |
| U20M | 0.3 | 0.009 | 0.24 | 1 | 0.02 | 0.03 | 1.00 |

### 6.2.6  Verification of Simulated Streamflows

In order to generate confidence in the output of hydrological simulations, verification against observations is required. In this study, however, for the analysis of the streamflow forecasts simulated by *ACRU* when using the rainfall forecasts provided by the different weather/climate models, the streamflow series used as the reference was not a series of observed streamflows, but the daily streamflows simulated by the *ACRU* model using as inputs the quality checked observed rainfalls (i.e. so-called "true" rainfalls). This approach was chosen for the following reasons:

- There are no complete flow data at the exits of *each* of the QCs within the Mgeni catchment for the selected evaluation periods;
- The Mgeni catchment was assumed for purpose of these simulations to be covered by natural vegetation, without any anthropogenic interference such as domestic or irrigation abstractions from dams, return flows and inter-catchment transfers; and
- The effects of any simulation errors induced by possible inadequacies in the process representation in the *ACRU* model were to play no role in verifications because the aim of this case study was to independently verify the various rainfall inputs/forecasts against a reference streamflow from quality checked observed rainfalls, and not to verify streamflow output *per se*.

The above points deserve future attention if streamflow forecasts are needed to be fairly treated in a comparison against observed flows. It is, therefore, important to bear in mind that wherever the term "observed" flows is used in this study, these are in fact made up of values simulated with the *ACRU* model using the observed quality controlled rainfalls, while "simulated" values are those derived from the various rainfall estimation or forecasting models.

### 6.3    Concluding Remarks

The routine availability of near real time to seasonal agrohydrological forecasts is not sufficient to ensure that they can be applied in decisions that benefit the water

resources and agricultural sectors. The usefulness of such forecasts will be enhanced by verifying them against observations for a wide range of environmental conditions. The Mgeni catchment was selected in this study as a starting point towards the operational use of these forecasts because, in addition to the availability of quality checked rainfall data from a relatively dense network of raingauges monitored by the South African Weather Service (SAWS), the Department of Water Affairs and Forestry (DWAF), the Agricultural Research Council (ARC) and Umgeni Water, this catchment is of major socio-economic importance in South Africa. The application of the various agrohydrological forecasts within the Mgeni catchment therefore plays a significant role in improved decision making in the agriculture and water sectors.

The distributed version of the *ACRU* model (Schulze, 1995a and updates) was employed for simulating the rainfall-streamflow transformation. The input data (other than rainfall), which were used for the streamflow simulations, were extracted from the Quaternary Catchments Database (QCD). The *ACRU* model was applied with a baseline land cover and it was decided to compare the simulated streamflows from the various rainfall forecasts against those streamflows simulated with the *ACRU* model using observed rainfalls as input.

In the chapters which follow a more detailed description is provided on how each of the rainfall forecasts was verified against the observed reference. The verification commences with the daily individual and merged rainfall maps from a network of radar and daily reporting raingauges as well as satellite images (Chapter 7) and is then followed in Chapter 8 by the verification of short and medium rainfall forecasts from Numerical Weather Models. Chapters 9 and 10 then cover, respectively, the verification of the two temporal downscaling methods, *viz.* the *Historical Sequence Method* and the *Ensemble Re-ordering Based Method*.

# 7. EVALUATING THE PERFORMANCE OF RAINGAUGE, RADAR, SATELLITE AND MERGED RAINFALL FIELDS

## 7.1    Introduction

Raingauge data are still widely used as the most abundant and precise sources of rainfall information for many applications in climatology, hydrology, agriculture and other environmental sciences. However, their uneven and often sparse networks, their limited sampling area and problems inherent in point measurements represent a substantial problem when dealing with effective spatial coverage of rainfall over a large area (Schulze, 1995a; Deyzel *et al.,* 2004; Pegram *et al.,* 2004; Schulze, 2006). Radar and satellite-derived rainfall estimates are widely accepted as promising strategies to address the above limitations, primarily because of much greater detail in their spatial and temporal resolutions over a catchment.

Weather radar has enormous potential to offer rainfall estimates in real time with high spatial resolution and temporal continuity (for example, Sun *et al.,* 2000). However, the accuracy of its quantitative estimation is highly sensitive to atmospheric conditions, sampling height of the radar beam, beam blocking, variations in the reflectivity-rainfall rate relationships, ground echoes and distance from the radar (Deyzel *et al.,* 2004; Jordan *et al.,* 2004; Pegram, 2004; Piccolo and Chirico, 2005).

On the other hand, the concept of real time rainfall estimates using satellite data provides a means of estimating rainfall at any point or area on the Earth's surface, regardless of country boundaries and/or unfavourable atmospheric conditions (Deyzel *et al.,* 2004; Pegram *et al.,* 2004). Satellites data are therefore essential in compensating for the lack of rainfall information over oceans as well as over remote and mountainous areas which are not covered by an adequate raingauge network and/or radars (Laurent *et al.,* 1998; Deyzel *et al.,* 2004; Pegram *et al.,* 2004). Unfortunately, satellites also suffer from some inherent shortcomings, although they display some useful information on rainfall patterns. The main reason for the errors is the assumption that precipitation at the surface is a function of cloud-top temperature. Identifying precipitating and non-precipitating

clouds is seen as a major constraint. Even if the precipitating clouds are correctly identified, the estimated rainfall amounts may still contain a substantial random error, largely because of the different dynamical processes occurring inside clouds (Deyzel *et al.,* 2004).

In 2002 the METSYS group of the SAWS and their collaborators launched a project called **S**patial **I**nterpolation and **MA**pping of **R**ainfall (SIMAR) that aimed at developing a near real time, spatially high resolution rainfall measuring and mapping system for southern Africa, based on both the surface raingauge networks and remote sensing techniques. The ultimate goal of the SIMAR project was to improve the outstanding issues of data quality and integrity that can be derived from daily reporting raingauge networks, radar networks and satellite images (Deyzel *et al.,* 2004; Kroese, 2004; Pegram, 2004). Daily individual and merged rainfall maps from these data sources are now available on a daily basis at a resolution of 1.7 km for the entire southern African subcontinent. The incorporation of these products into the framework of an agrohydrological forecasting system is thus of fundamental importance for many applications in agrohydrology. Their availability in near real time is a vital input in simulating the "now state" (i.e. of "this morning") of various hydrological fluxes such as effective rainfalls, soil moisture contents, streamflows, groundwater flows and reservoir levels on a daily basis. This, in turn, has the potential to improve the accuracy of near real time agrohydrological forecasts that provide guidance to decision makers in agriculture and water management, as well as to disaster managers issuing flood forecasts and warnings.

The incorporation of the SIMAR products into the framework for the agrohydrological forecasting system under development was one of the primary objectives of this study. For these products to be applied operationally in agrohydrology, the critical issue is to be able to estimate the expected magnitude of errors, not only in respect of accumulated rainfall depths, but also from an agrohydrological perspective. The aim of this chapter is therefore twofold. First, it makes a comparison of the rain fields estimated from reporting raingauges, radar, satellite as well as the merged rain fields against rainfall measured by independent synoptic rainfall stations, and assesses the accuracy and reliability of the SIMAR

products using various statistical measures. The second aim is to transform these rainfall estimates into streamflows using the *ACRU* model (Schulze, 1995a) and to investigate the significance of uncertainty that cascades from these rainfall estimates through the *ACRU* model to the streamflow simulations.

This chapter commences with a general description of the SIMAR products (Section 7.2). A comparison of the performances of the various rainfall fields is then presented, and a discussion ensues whether they can produce accurate simulations of streamflow for the Mgeni, which is used as a test catchment (Section 7.3). Finally, a conclusion in light of the results obtained is presented in Section 7.4.

## 7.2    General Description of the SIMAR Products

The SIMAR programme comprises of three major component projects. Detailed information pertaining to these components is given in three volumes, *viz.*

- Maintenance and upgrading of radar and raingauge infrastructure (Kroese, 2004);
- Radar and satellite products (Deyzel *et al.,* 2004); and
- Data merging for rainfall map production (Pegram, 2004).

Although an understanding of all the processes by which raingauge, radar, and satellite information is used to estimate rainfall is important to fully understand the SIMAR products, it is beyond the scope of this project to review all that detail. In this section, therefore, emphasis is given to the factors which have been shown to critically influence the quality of the daily rainfall fields generated from each of these data sources.

## 7.2.1   Kriged Raingauge Rainfall Field

Raingauges are simple, inexpensive and by far the most common method used to measure rainfall (Curtis and Humphery, 1995). They estimate rainfall at a point, but the measurements are often subjected to errors. Wind effects, calibration errors, the inclination of the gauge, splash into and out of the gauge funnel, evaporation of water inside the gauge and observer errors can all affect the

accuracy of raingauge rainfall measurement (Schulze, 1995a; Deyzel *et al.,* 2004). Areal rainfall is estimated indirectly by making assumptions regarding the amounts of rain falling between gauges (Curtis and Humphery, 1995). In order to accurately capture the estimates of areal rainfall, a very dense raingauge network is required. In southern Africa, the density of the current daily reporting raingauge network used in the process of producing daily rainfall maps is far less than the density needed to adequately cover the entire region (cf. Figure 7.1).
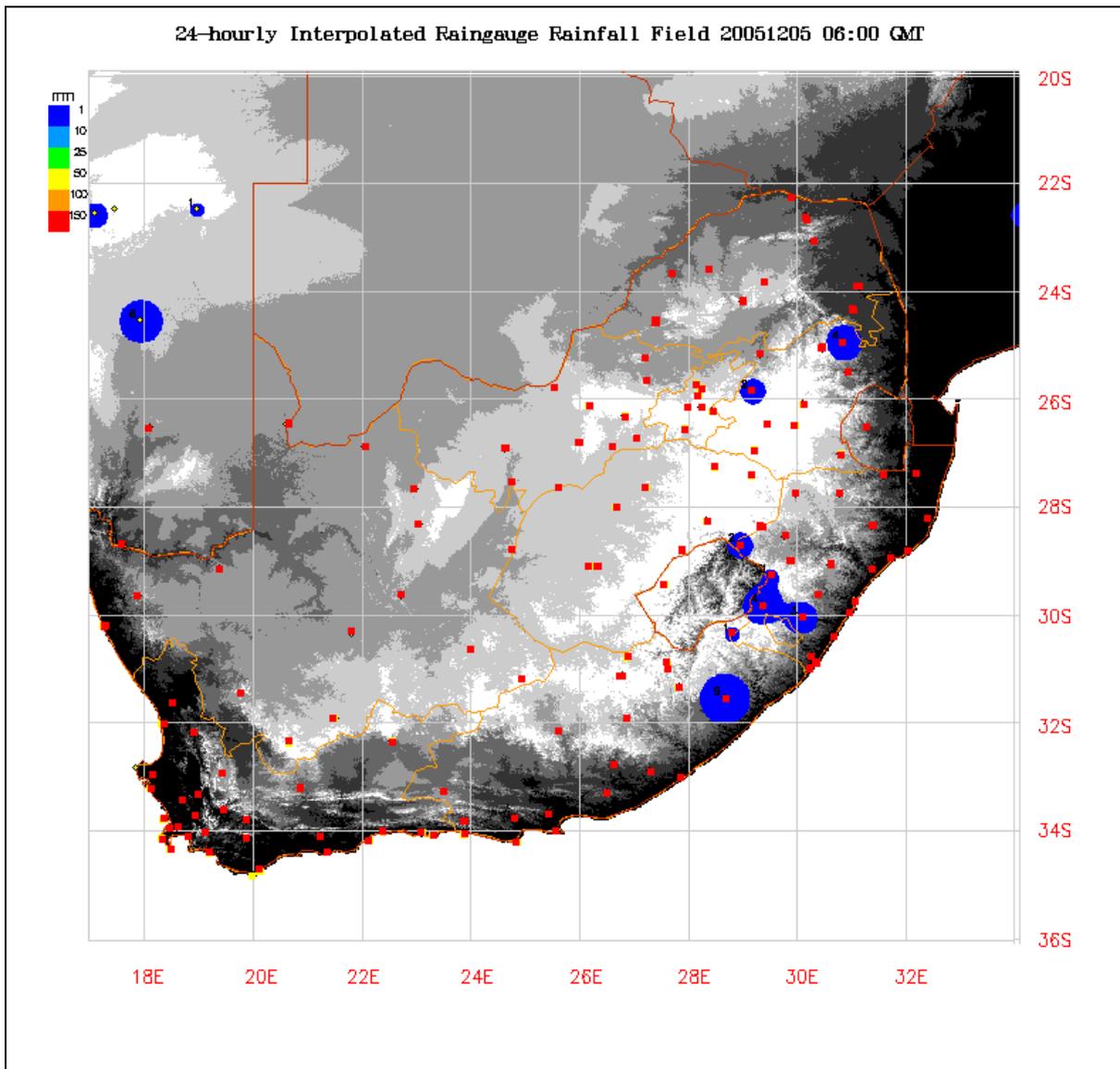


Figure 7.1    An example of the kriged raingauge rainfall field over southern Africa from the daily reporting raingauges of SAWS for 05 December 2005, with red squares corresponding to the daily reporting stations (Source: http://metsys.weathersa.co.za, 2007)

The raingauges number approximately 300 and are spaced unevenly. Of these daily reporting stations, 120 are automatic hourly reporting tipping bucket raingauges. Not all of the stations' data are available on an everyday basis, owing to occasional communication problems and/or absence of observers over weekends (Deyzel *et al.,* 2004).

In order to produce continuous rainfall fields, the point measurements of surface rainfall are converted to a spatial resolution of 1 minute longitude by 1 minute latitude (i.e. 1024 by 1024 grids) using the Nearest Neighbourhood method (Deyzel *et al.,* 2004). The irregularly gridded raingauge field is then interpolated with the kriging algorithm using radar covariance information to produce a kriged raingauge field (Deyzel *et al.,* 2004; Pegram, 2004). In areas with scattered reporting raingauges, the accuracy of the interpolated rainfall value suffers in that it does not adequately represent the spatial variability of rainfall (Figure 7.1). According to Kroese (2004), various initiatives have been undertaken by the SAWS to extend the daily reporting raingauges, principally by incorporating data from different sources such as the Department of Water Affairs and Forestry (DWAF), the Agricultural Research Council (ARC), Farmers' Co-Operatives and Agricultural Organisations (FCAO) and the Southern African Development Community (SADC). Unfortunately, to date, these data sources are not linked to the operation of SIMAR project due to various practical and political reasons. The SAWS initiative to replace the second and third order climate stations with a downscaled version of Automatic Weather Stations (AWS) is underway (Kroese, 2004). These new AWS data are expected to provide a sound foundation to secure a real time national database in the future for utilisation in the SIMAR and other projects.

### 7.2.2  Kriged Radar Rainfall Fields

Weather radar employs a fundamentally different approach to that of raingauges in estimating the amount of rain falling over a catchment. Radar is used to indirectly estimate rainfall amounts from direct areal measurements over a very wide area (Curtis and Humphery, 1995). Radar has many advantages when compared to raingauge networks, as single radar is able to scan an area of hundreds of square kilometres in a few seconds, with the ability to make informed "nowcasts" of future

rainfall (Jordan *et al.,* 2004; Piccolo and Chirico, 2005). Its ability to observe the three-dimensional structure of the system generating the rainfall provides improved information over raingauges in regard to the spatial and temporal variability of the rainfall (Deyzel *et al.,* 2004). However, the estimation of the rainfall from radar measurements is affected by technical problems and different sources of errors (Deyzel *et al.,* 2004; Jordan *et al.,* 2004; Piccolo and Chirico, 2005), as was mentioned in the introductory section of this chapter. According to Jordan *et al.* (2000), radar errors may be subdivided into three categories, *viz.*

- Error sources due to ground clutter, anomalous propagation and beam blocking, which can be removed by simply adjusting the radar reflectivity techniques;

- Systematic errors due, for example, to poor electronic calibration of the radar or contaminations in radar reflectivity measurements, where this type of error can be removed by making long term comparisons between radar derived rainfall and accumulated rainfall from a raingauge network over a large area; and

- Residual random errors caused mostly by temporal spacing of, for example, the radar scans, spatial sampling and height sampling.

The residual errors are the ones that have been least understood and are the most difficult to remove in their effects on rainfall measurement (Jordan *et al.,* 2000). During the course of the SIMAR project, a successful research application was conducted by Deyzel *et al.* (2004) to filter the negative impacts of ground clutter, to improve the conversion algorithm of radar reflectivity into the rain rate and to generate a merged rain field from rain fields generated at the individual radars. However, the existing 10 radars that constitute the National Weather Radar Network (NWRN) in South Africa are not sufficient to cover the entire study area (Figure 7.2). Vast areas of the North West Province are not under radar coverage, even though data are collected up to 200 km in range from each radar (Kroese, 2004). Daily rainfall maps are generated operationally from the mosaic of 10 radars in the NWRN. As with the raingauge data, the kriging technique is applied to the radar rain field for extrapolating information beyond the scope of the measurement and partly for enhancing image quality, especially for areal rainfall

(Deyzel *et al.,* 2004). The quality of these kriged radar rain fields remains in doubt as one moves to areas where currently no radar coverage exists (Deyzel *et al.,* 2004). Another important aspect is the inability of coast-located radars to sample rainfall adequately. Research conducted by Deyzel *et al.* (2004) confirmed that the radars located at the coast performed less well than the inland radars. This is due mainly to warm orographic forcing at the coast.
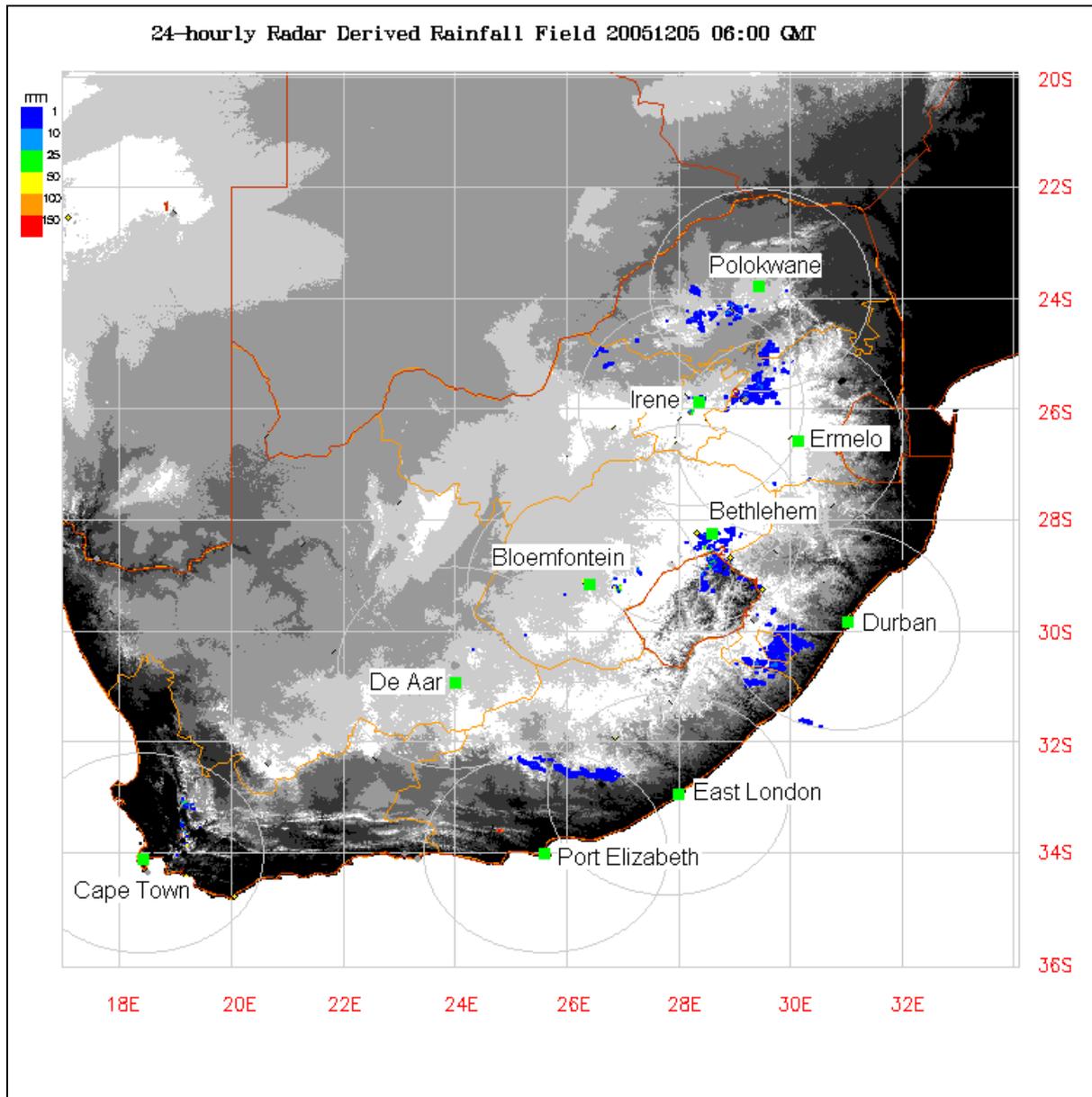


Figure 7.2     An example of the kriged radar rainfall field over southern Africa for 05 December 2005, with green squares representing the location of the radars (Source: http://metsys.weathersa.co.za, 2007)

### 7.2.3  Satellite Derived Rainfall Fields

Satellites, as distinct from radar, offer a complete spatial coverage over southern Africa and are essential in compensating for the low density, or lack of, raingauge and radar based rainfall measurements. Satellite data are cost-effective and available for any point on Earth, irrespective of country or inhospitable surface conditions (Deyzel *et al.,* 2004; Pegram *et al.,* 2004, Kamarianakis *et al.,* 2006). Visible (VIS) and Infrared (IR) imagery and passive microwave instruments are the satellite tools that have been extensively employed to estimate rainfall for a range of applications. Passive microwave radiometry from satellite platforms provide a direct information on rainfall (Kamarianakis *et al.,* 2006). However, the lack of sufficient temporal and spatial resolution of the current polar-orbiting sensors renders the passive microwave techniques not useful for studies of convective events (Deyzel *et al.,* 2004; Kamarianakis *et al.,* 2006). A number of techniques have been developed to estimate surface rainfall from VIS and IR satellite data indirectly. Most of these techniques are based on the premise that the cloud with a cold top produces more rain than that with a warmer top, and a statistical relationship exists between cloud top temperature and the rainfall falling at the surface (Yucel *et al.,* 2004; Deyzel *et al.,* 2004; Kamarianakis *et al.,* 2006). One disadvantage of the VIS/IR techniques is that they provide rainfall estimates that are technique-dependent. Because of varying rainfall characteristics under different climate regimes, the techniques are not transferable from location to location (Kamarianakis *et al.,* 2006).

The technique applied for the SIMAR is a multi-spectral rain rate (MSRR) method, in which active cloud identification is accomplished by the discrimination of high level cloud characteristics which are associated with rain and no-rain classes. Data from the three spectral bands, *viz.* visible (0.5 - 0.9 µm), infrared water vapour (5.7 - 7.1 µm) and thermal infrared (10.5 -12.5 µm), of the Meteosat -7 satellite are utilised in the rain area recognition rainfall estimation method. However, only the IR channel data are used for the estimation of surface rainfall values (Figure 7.3).
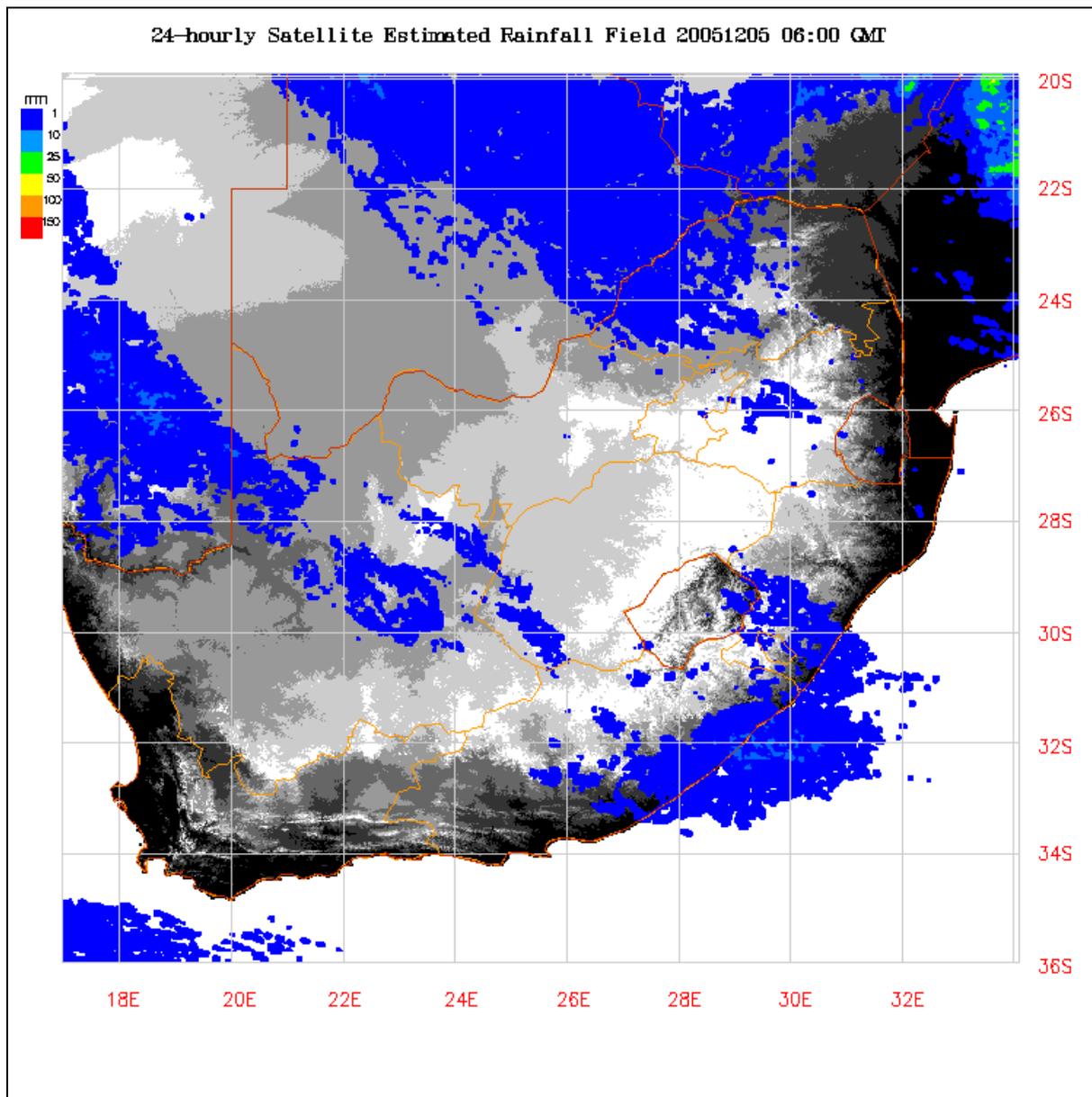
Figure 7.3    An example of the satellite rainfall field over southern Africa for 05 December 2005 (Source: http://metsys.weathersa.co.za, 2007)

Data from the Meteosat-7 satellite were used because the launch of the Meteosat Second Generation (MSG) was delayed due to some technical problems (Deyzel *et al.,* 2004; Pegram *et al.,* 2004). According to Deyzel *et al.* (2004), some of the inherent limitations in determining rainfall with the MSRR technique include the following:

- Clouds are opaque in both the VIS and IR spectral channels and precipitation is indirectly inferred from cloud top temperature. This indirect

115

method cannot intelligently identify the extent of convective rainfall. Furthermore, cloud structures may be incorrectly classified into rain or no-rain classes, especially for tropical rain systems. This is primarily due to the simplicity of the classification approach.

- The effects of warm coastal and orographic rainfall are not always correctly quantified.

- The variation of cloud top structure under different dynamical processes are handled at an average level and the estimated rain amounts may still contain a substantial random error, even if the clouds are correctly identified as precipitating.

In the process of producing the final satellite rainfall map for southern Africa two further steps are involved. The first is to mask the IR field in order to identify the precipitating clouds and the second is to then produce the rainfall estimates from the IR map (Deyzel *et al.,* 2004; Pegram *et al.,* 2004). These steps are detailed in a report of Deyzel *et al.* (2004). Unlike the raingauge and radar data sets, the satellite rainfall data set is sampled for every area on the gridded national rainfall map (Deyzel *et al.,* 2004).

### 7.2.4 Merged Rainfall Fields

The purpose of merging the gauge/radar/satellite estimations of the respective rainfall fields is to retain the fine sampling resolution of areal rainfall from radar, to compensate for the lack of observed rainfall information over the oceans and data sparse regions from satellite platforms, and to remove any quantitative biases by bringing the values closer to those of the "ground truth" raingauges (Deyzel *et al.,* 2004; Pegram, 2004; Pegram *et al.,* 2004). The merging operation is performed in three steps. First, the gauge and radar information is combined into a field, $R_G$. Then the gauge and satellite information is combined into a field, $S_G$, and finally the $R_G$ and $S_G$ fields are merged into the SIMAR field, which is published daily on the SAWS website along with the individual rainfall fields (Pegram *et al.,* 2004). Details on the process of merging and mapping of rainfall are presented more fully in the reports by Deyzel *et al.* (2004) and Pegram (2004). Based on those two reports, a brief overview of the merging process is outlined here, so as to highlight the principles behind it.

### 7.2.4.1 Merging radar and raingauge data

Merging the radar and raingauge data to produce the $R_G$ field is accomplished through the following steps (Pegram, 2004; Pegram *et al.*, 2004):

- *Extrapolate radar data*: The technique used was kriging by the Fast Fourier Transform (FFT) in conjunction with Iterative Constrained Deconvolution (ICD) to produce a kriged radar rainfall field, $R_K$.

- *Interpolate raingauge point measurements*: The covariance function derived from archived historical radar rainfall fields was used in kriging the raingauge data between the gauge locations to produce kriged raingauge rain field, $G_K$.

- *Compute the explained variance fields for radars and gauges*: The explained variance fields $R_V$ and $G_V$ for radars and gauges respectively, are computed using the FFT and ICD techniques in order to represent the accuracy of the data and the interpolated rainfall at non-measured locations of the fields. The explained variance is assumed to be 100% for the whole radar domain, corresponding to a 200 km range surrounding each radar, as well as for the pixels corresponding to the daily reporting raingauge positions. The explained variance diminishes with increasing distance from the locations of the radars or raingauges. The $R_V$ and $G_V$ fields remain unchanged from day to day, unless a radar or gauge is not operating.

- *Merge the radar and gauge fields*: Integrating the quantitative values of the raingauge with the superior spatial information of the distribution areal rainfall of the radar is performed using the following equation:

$$R_G = \frac{[R_K R_V + G_K G_V]}{[R_V + G_V]} \qquad 7.1$$

The weight of the raingauge field reduces the effects of bias in the radar rainfall field, while maintaining its spatial structure (Pegram *et al.*, 2004).

### 7.2.4.2　Conditioning satellite data on raingauge data

Conditioning the satellite data on the raingauge information to produce the $S_G$ field is accomplished through the following steps (Pegram, 2004; Pegram *et al.,* 2004):

- *Interpolate an average satellite rainfall field*: Satellite rainfall pixels are sampled at the positions of the raingauges to produce a set of satellite rainfall point estimates, $S_Z$. The average $S_Z$ data field is then interpolated to a regular grid using completely regularised splines (SRS).
- *Interpolate the raingauge data*: To interpolate the raingauge data, a kriged raingauge rain field, $G_K$ is produced, as was described in Section 7.2.4.1.
- *Condition the satellite on raingauge data*: The conditioned satellite field, $S_G$, is obtained simply by subtracting the mean satellite rainfall field, $S_Z$, from the satellite rainfall field, $S_R$, and the result is added to the interpolated gauge field, $G_K$, i.e.

$$S_G = S_R - S_Z + G_K \qquad\qquad 7.2$$

The $S_G$ field retains the spatial distribution of rainfall, as estimated from the satellite, but with reduced bias (Pegram *et al.,* 2004).

### 7.2.4.3　Producing the merged field

In producing the final combined rainfall field, the $R_G$ and $S_G$ fields are not treated equally; rather they are merged in proportion to their weighted fields. For the $R_G$ field the explained variance masks, $R_V$ and $G_V$, are integrated with the "Boolean" to produce the weighted field, $RG_V$. The bias skill scores, $S_{BSS}$, of the satellite rainfall in the neighbourhoods surrounding the raingauges are calculated and the results are interpolated using splines to a regular grid, $S_{SS}$, to be used as a weight field for the field $S_G$ (Deyzel *et al.,* 2004; Pegram, 2004; Pegram *et al.,* 2004). To calculate the bias skill score, a contingency table is formed for each of the pixels in a 9 by 9 neighbourhood surrounding the raingauge position (Table 7.1).

Table 7.1    A contingency table for bias skill score computation (Source: Pegram *et al.*, 2004)

| Gauge | Satellite | |
|---|---|---|
| | Rain | No-rain |
| Rain | Hits (H) | Misses (M) |
| No-rain | False Alarms (FA) | Correct Nulls (CN) |

The bias skill score is calculated in the vicinity of a raingauge. It is simply the sum of correct hits and correct nulls divided by the total number of pixels surrounding the raingauge position (Equation 7.3).

$$S_{BSS} = \sum_{i=1}^{81} \frac{[H_i + CN_i]}{81} \qquad\qquad 7.3$$

The bias skill score ranges from 0 to 1, and it measures the fraction of correct rain/no-rain classifications made by the satellite. A skill score of 1 indicates no bias, while a 0 score indicates a bias in the technique's ability to classify the rainfall classes correctly. The problem with the bias skill score is that a low value could be extended to areas where rainfall is identified correctly, if the satellite rainfall field $S_R$ has either a false alarm or a miss when compared to an isolated raingauge in areas with sparse data. To account for this limitation, the average skill score of the entire data domain, $S_{FSS}$, is used as a weight field in those areas far from a raingauge or a radar (Pegram *et al.*, 2004). The integration of the radar/raingauge and satellite/raingauge rainfall fields (Figure 7.4) is finally computed in proportion to their respective weight fields $RG_V$ and $S_{SS}/S_{FSS}$, as given by Equations 7.4 and 7.5 (Deyzel *et al.*, 2004; Pegram *et al.*, 2004).

$$R_{MERGED} = \frac{[S_G \times S_{SS} + R_G \times RG_V]}{[S_{SS} + RG_V]}, \quad \text{if } RG_V > 0 \qquad\qquad 7.4$$

$$R_{MERGED} = \frac{[S_G \times S_{FSS}]}{100}, \quad \text{if } RG_V = \text{null} \qquad\qquad 7.5$$
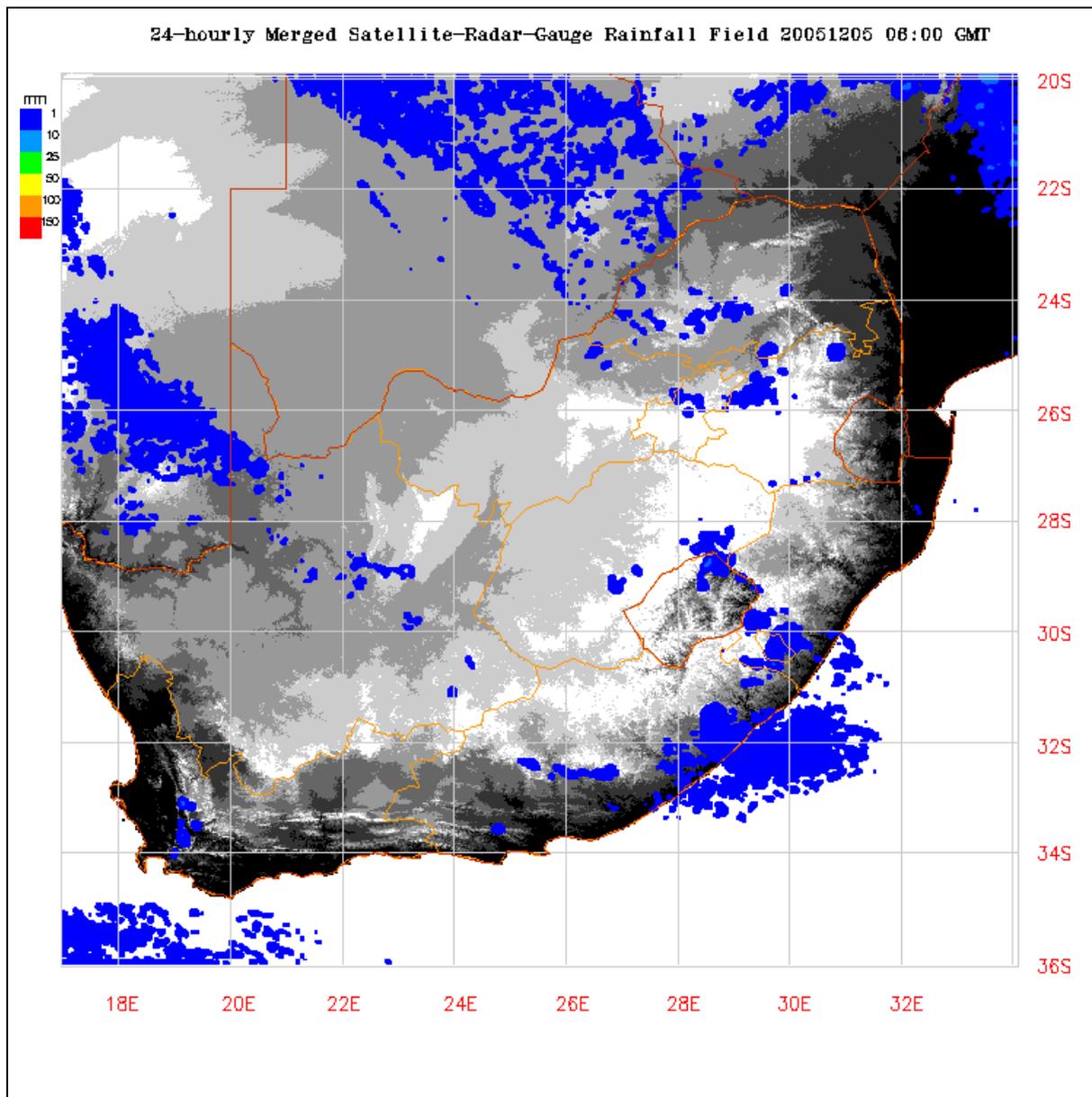
Figure 7.4    An example of the merged rainfall field over southern Africa for 05 December 2005 (Source: http://metsys.weathersa.co.za, 2007)

## 7.3    Verification of the SIMAR Products

Deyzel *et al.* (2004) attempted to verify the radar and satellite derived rainfalls as well as the merged rainfall fields, with the objective of minimising the specific errors and thereby improving the final SIMAR products. In a comparison made between the daily rainfall fields generated from the mosaic of 6 radars and rainfall measured by Automatic Weather Stations (AWS), they concluded that the Polokwane, Bloemfontein, Port Elizabeth and Irene radars were generally

performing reasonably well, albeit with the Bloemfontein and Port Elizabeth radars appearing to under-estimate the accumulated rainfalls in areas where the distance was more than 100 km from the radar. The Durban and Ermelo radars suffered from consistent under-estimation (cf. Figure 7.2 for the radar locations). The comparison was made in terms of accumulated rainfalls over the period 01 October to 30 November 2001.

The rainfall fields estimated from the satellite platforms were compared against the kriged raingauge rainfall field, assuming the interpolated values as a ground truth. Monthly data sets of satellite rainfall, corresponding to the raingauge neighbourhood area for the months December 2002 to February 2003 were used in the verification process. The correlation value for December 2002 was found to be the highest ($r^2 = 0.69$), but with a substantial bias. In contrast, the respective correlation values for January ($r^2 = 0.47$) and February 2003 ($r^2 = 0.28$), were quite low, even though the biases appeared to be low (Deyzel *et al.,* 2004).

In a similar way, the merged rainfall fields were verified against raingauge data. The conditioning technique for removing the bias of the MSRR satellite rainfall field was found very effective in improving the distribution of the rainfall values with respect to raingauge values. However, the explained variance of the kriged technique used for merging of radar with raingauge rainfall data was not as efficient as the conditioning process in minimising the bias, but produced reliable spatial rainfall fields. The final merged rain field may have lower errors than either of the individual rainfall fields. However, its reliability decreases when an extensive over-estimation occurs in the satellite estimates (Deyzel *et al.,* 2004).

It is clear that the SIMAR products are subjected to some spatial and temporal errors. They all suffer from some complexities, owing to the fact that they are affected by interpolation and scaling issues that may exacerbate the inaccuracy of these products. It is, therefore, appropriate to further extend the verification process, not only in terms of rainfall magnitudes, but also from a hydrological perspective. The inter-comparison of the different methods helps us to determine which one of them would provide better estimates when compared against the so-called ground truth. However, since their performances may depend on many

issues such as seasonality, spatial/temporal scales, distance from the radar/raingauge as well as distance from the ocean or aspect, it is useful to therefore test the SIMAR products under a wide range of environmental conditions and to utilize them in hydrological comparisons. In this study, the Mgeni catchment representing the summer rainfall areas and the Berg catchment representing the winter rainfall areas were initially selected to evaluate these data sources. However, only the Mgeni catchment was used, owing to the absence of merged rainfall fields for the winter rainfall period and the sparse raingauge networks in the Berg catchment. The results are presented in the sub-section which follows.

## 7.3.1 Methods of Comparison

There are now five source of spatial precipitation information, *viz.* the kriged raingauge rainfall field, the kriged radar rainfall field, the satellite rainfall field as well as the conditioned satellite and the merged rainfall fields to be evaluated in this study. Since 2002, the SAWS has archived some of these rainfall products. However, owing to the presence of significant missing records of the merged maps in the archived data, only data sets for the period from 01 January 2003 – 31 March 2003 were found to be suitable for this study. The synoptic raingauges distributed across the Mgeni catchment (Figure 7.5) were used as ground truth for the validation of the effectiveness of the SIMAR products.

The SIMAR data cannot be compared directly against raingauge data, because the radar and satellite data are averaged over an area while the raingauge data are averaged over time, but represent a point in space. An interpolation scheme has to be used to estimate areal rainfall values from the point measurements of rainfall observed by these daily synoptic stations. The chosen interpolation method is the Natural Neighbours technique, which is commonly used when the sample data points are unevenly distributed. Natural Neighbours interpolation creates a "Delauney" triangulation of the input points and selects the closest nodes that form a convex hull around the interpolation point, then weights their values by proportionate area (ESRI, 2005).
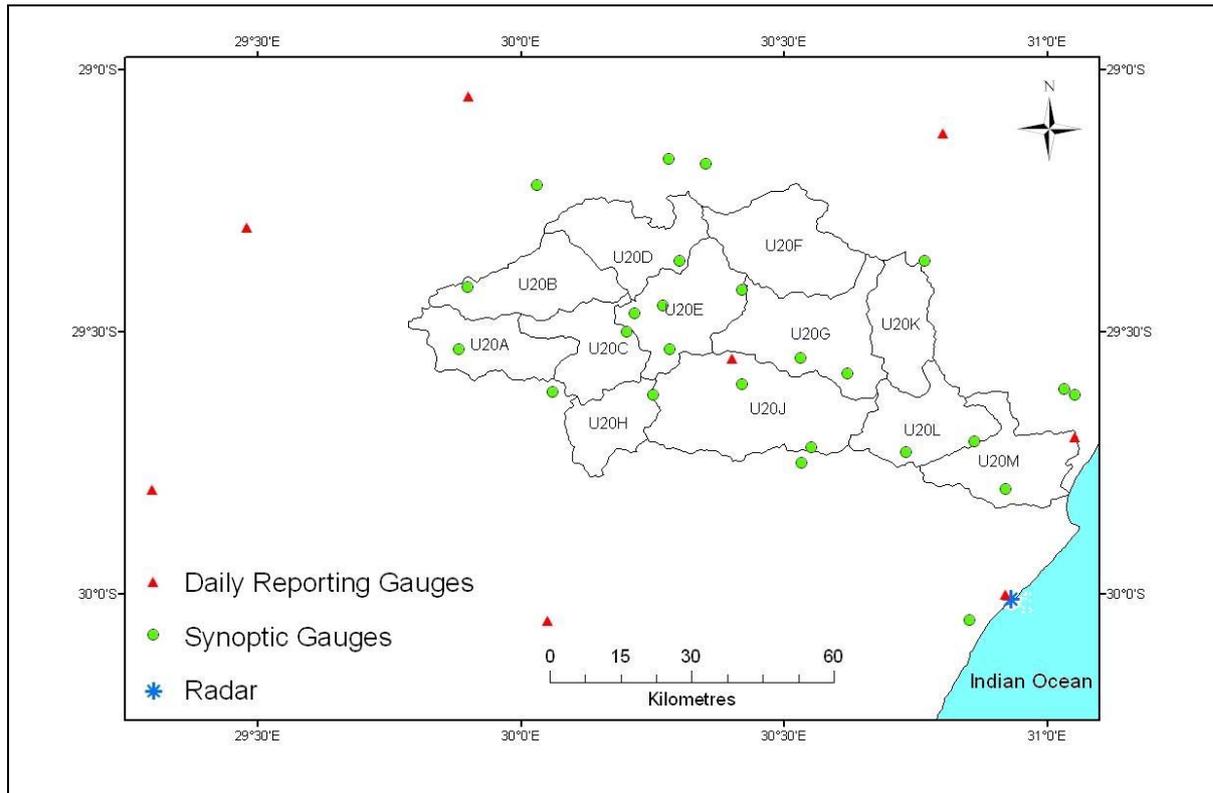
Figure 7.5 The distribution of radar, daily reporting gauges and synoptic stations over the Mgeni catchment

Various objective functions have been used to measure the strength of the statistical relationship between the estimated values and the reference values. A contingency table approach was first used as an indicator of skill to assess whether the rainfall distribution in the SIMAR images were correctly positioned in relation to the reference image within a selected area. For this specific case, the two highest daily rainfall events which occurred on 05 February and 19 March 2003 were chosen from the study period, and the Quaternary Catchment U20E (Figure 7.5) was selected, owing to its coverage by a relatively dense raingauge network (cf. Figures 7.6 and 7.7).

Each 1' x 1' pixel within U20E in the five sources of rainfall images was compared against the corresponding pixel in the reference rainfall image. For the purpose of comparison, threshold percentiles were defined based on the rainfall values recorded at 31 stations over the Mgeni catchment on these two days. The pixel-by-pixel scoring criteria, *viz.* the Critical Success Index, CSI, the Probability of

These statistics do not, however, indicate how severe the over-estimation or under-estimation in the SIMAR images is. Therefore, and additionally, the coefficient of determination ($r^2$), bias, Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) were used for the entire catchment to objectively evaluate the discrepancy between the SIMAR products and the observational reference data. All comparisons are based on a time series of daily rainfall totals, as the SIMAR values are intended for use with daily time-step hydrological/crop yield models.

To ensure a representative rainfall value, pixel rainfalls within each QC of the Mgeni catchment are averaged. The evaluation consists of two components. The first is a direct statistical comparison of the five SIMAR values with the gauge data. The second component employs the *ACRU* model (Schulze, 1995a and updates) in order to further evaluate the effectiveness of these data sources once transformed into streamflows. The semi-distributed catchment mode of the *ACRU* model was used for simulation of accumulated streamflows from subcatchments cascading downstream at the exit of each QC. The *ACRU* model was initiated with historical observed daily rainfall from year 2000 to the beginning of the simulation period in 2003 in order to create representative antecedent conditions and initial stores (e.g. soil moisture status of the top- and subsoil, baseflow store and releases). It is important to reiterate that wherever the term "observed streamflows" is used, these are in fact simulated streamflows with the *ACRU* model but when using the so-called ground truth rainfalls.

## 7.3.2 Results and Discussion

For the sake of comparison in this chapter, the interpolated reference data from the dense synoptic gauge network will be termed "observed" data, the kriged gauge rainfall field as "gauge", the kriged radar rainfall field as "radar", the raw satellite rainfall field as "satellite", the conditioned satellite field as "SG" and the gauge/radar/satellite rainfall field as "merged" in this chapter. Since the number and spatial distribution of the synoptic gauges and daily reporting gauges are not the same within the Mgeni catchment (cf. Figure 7.5), differences in accuracy are expected between the so-called "observed" and "gauge" data. Furthermore, the 05 February and 19 March events are named Event A and Event B respectively. A visual comparison of all six of the rainfall fields over QC U20E for Events A and B are presented in Figure 7.6. Plots of the various source data versus the reference are presented in Figure 7.7 to show the distributions of Hits (H), Misses (M), False Alarms (FA) and Correct Nulls (CN) within the boundary of 50th percentile. The gauge, SG and merged fields displayed similar patterns for both events with majority of their pixel rainfall values falling within the range of CN and FA for Event A, while for Event B the majority of their pixel rainfall values scattered within the ranges of CN and M. The performance of the radar and satellite is generally very poor, especially for Event B no skill was scored at all (Figure 7.7). The pixel-by-pixel analysis of the CSI, POD and FAR as a function of threshold rainfall percentiles also proves the superiority of the gauge, SG and merged data over the remaining data sources (Figure 7.8). Generally, decreases in the CSI and POD scores are expected with increasing threshold percentiles, while the reverse is expected for the FAR score. The reason for this is that most of the rainfall season is dominated by rainless days or by days with small events, with only few large events.

For Event A the radar displayed a high resemblance, with CSI and POD values of 100% up to the 30th percentile, followed by an abrupt collapse in correspondence at the remaining higher thresholds (Figure 7.8). A deterioration of the radar performance is much more noticeable for Event B, with zero scores of CSI, POD and FAR. The significant under-estimation of the rainfall pixels by the radar could be attributed to the fact that the Durban radar is influenced by coastal rain.

Another important explanation for the poor radar performance is the distance between the radar the QC U20E, which is approximately 100 km. Research conducted by Deyzel *et al.* (2004) also shows that the Durban radar significantly under-estimated rainfall, even within a 30 km distance from the radar.
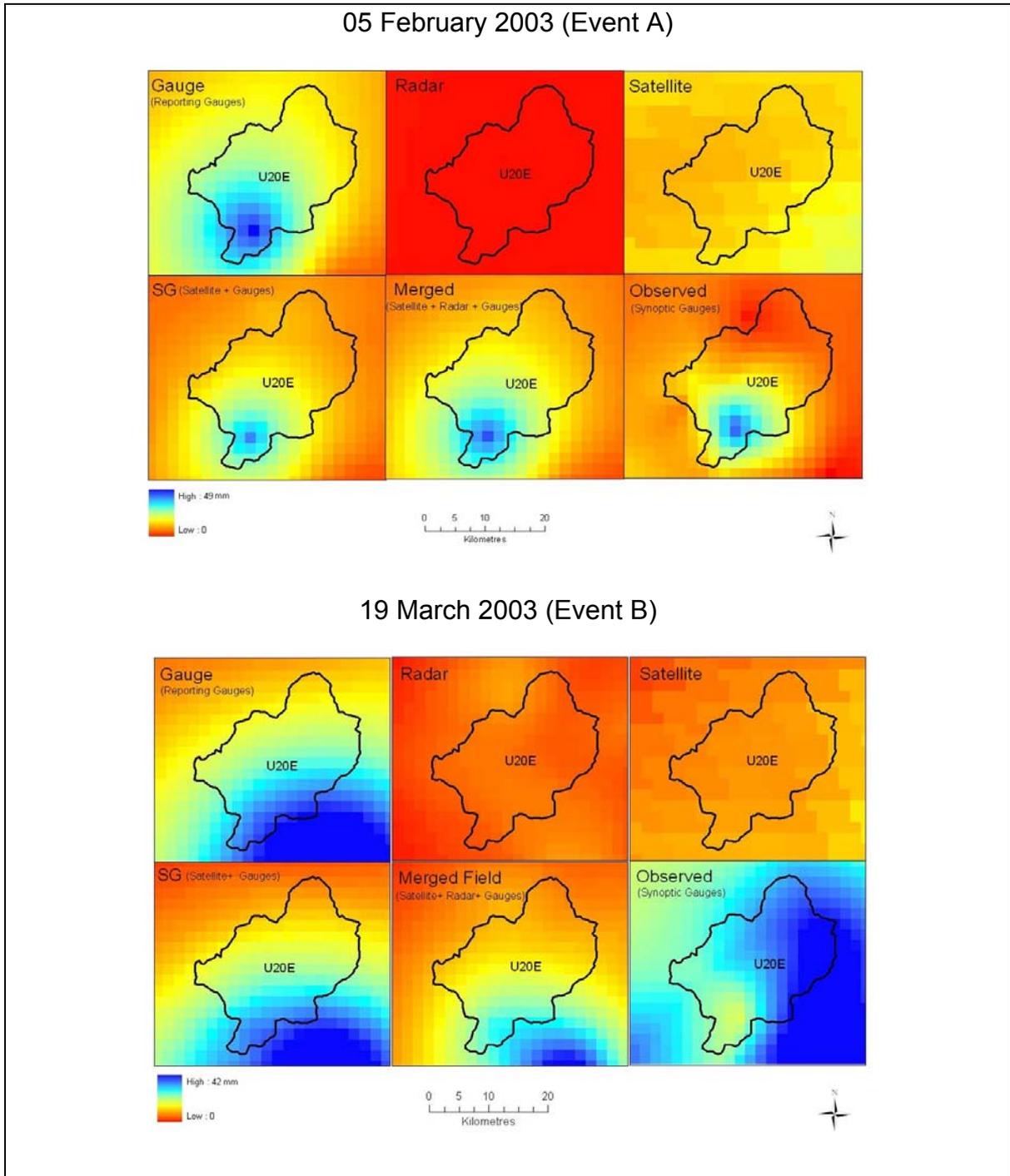


Figure 7.6    The estimated spatial distribution of rainfall over QC U20E for 05 February and 19 March 2003 using the various data sources
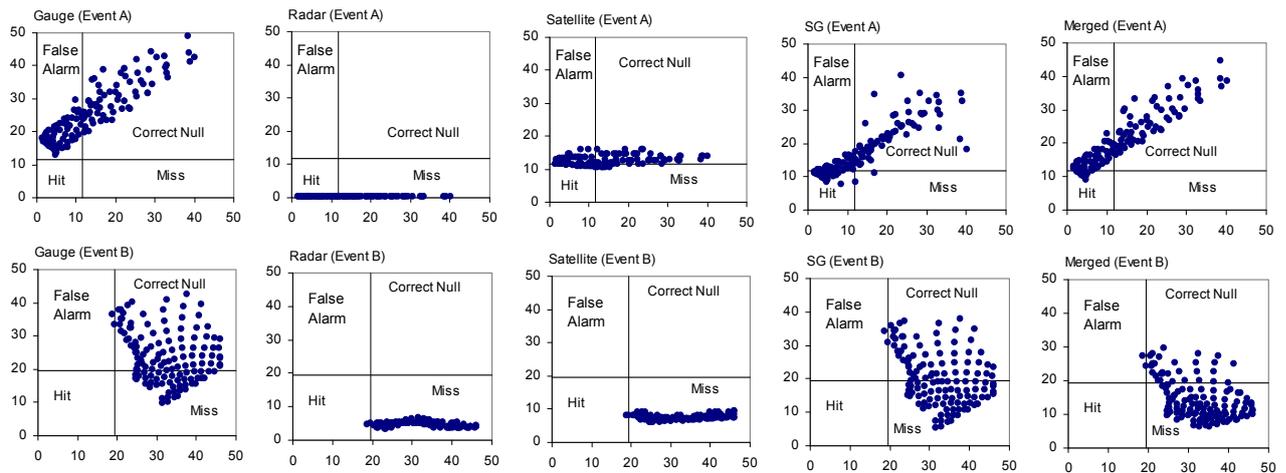
Figure 7.7    Scores of Hit, Miss, False Alarm and Correct Null defined by the threshold of 50th percentile for 05 February 2003 (Event A) and 19 March 2003 (Event B) using the gauge, radar, satellite, SG and merged rainfall fields

Similarly, the rainfall pixels from the raw satellite data significantly under-estimated those of the observed rainfall pixels for the majority of pixels in both events, even though it is positively biased over the majority of the study period (cf. Figures 7.10, 7.11 and 7.12). This may be inferred as being a reflection of the inherent limitations with the MSRR technique when classifying warm clouds into rain or no-rain classes (cf. Sub-section 7.2.3). The improvement in the gauge-based satellite estimates (SG) is highly noticeable in both events, suggesting that an adjustment based on the gauge observations has a profound influence in reducing the effects of bias in the satellite rainfall field.

Because of the limited number of daily reporting gauges in the Mgeni catchment, the gauge image failed to mimic the spatial structure of the observed rainfall distribution over QC U20E. Nevertheless, the performance of the gauge rainfall field is more or less similar to that of the merged and SG fields for both events. While the rainfall values from the final merged rain field have a better score for Event A than Event B, this is to be expected as the result of the influence of an extensive under-estimation of the spatial variability of the rainfall from the radar and satellite for Event B.

127
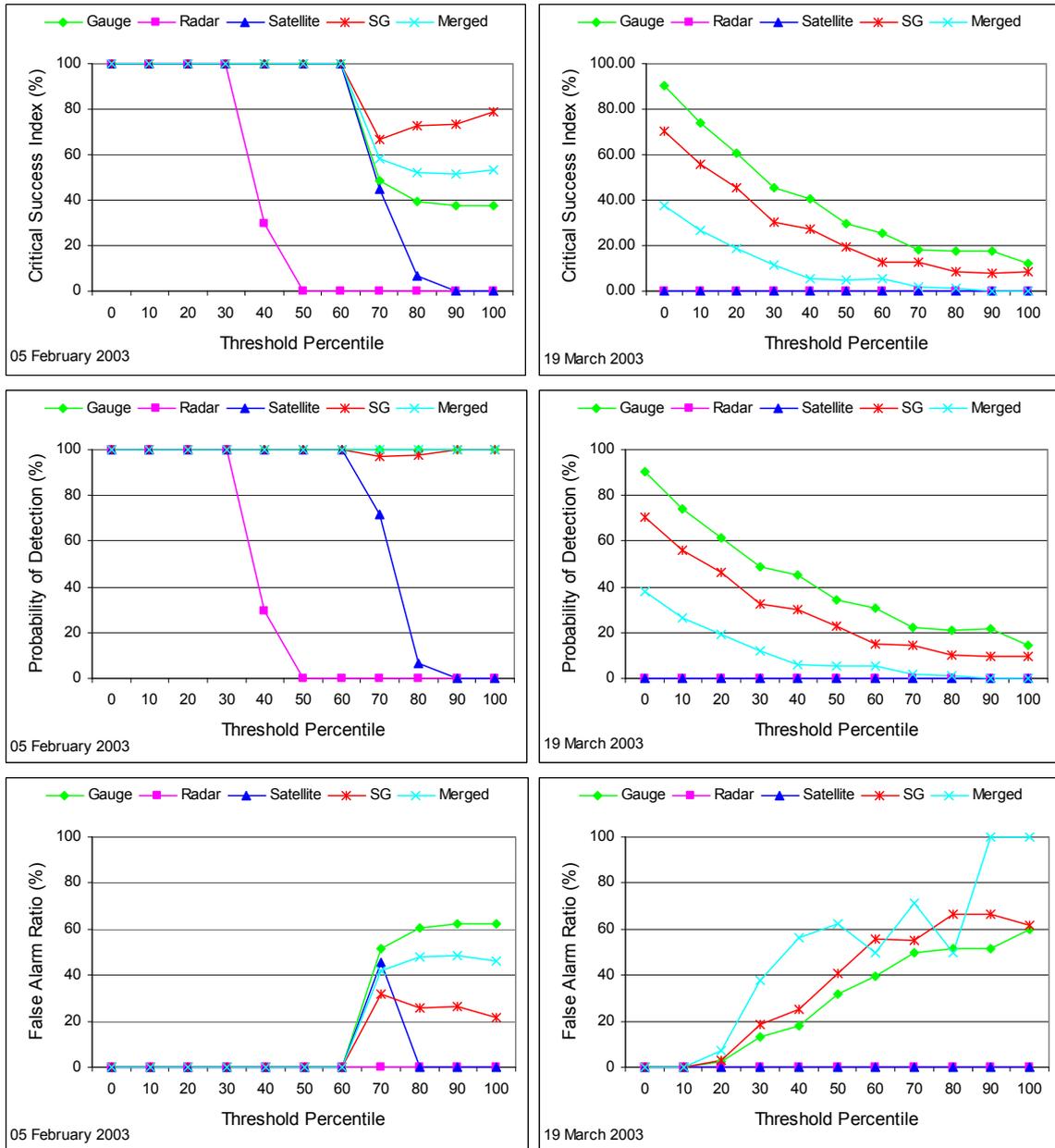
Figure 7.8    CSI, POD and FAR scores as a function of threshold percentiles for 05 February 2003 (Event A) and 19 March 2003 (Event B) using the gauge, radar, satellite, SG and merged rainfall fields

The pixel-by-pixel statistics analysed to this point provide only one insight into the quality of these data sources. Another insight is gained when the subcatchment (QC) rainfalls which were derived from each of the data sources were used as input into the *ACRU* model for streamflow simulation. The analysis is then further extended using $r^2$, bias, RMSE and MAE for both the rainfalls and streamflows for the evaluation period of 01 January to 31 March 2003. These statistics are plotted

in Figure 7.9 in order to investigate the uncertainty and systematic biases inherent in these data sources, and to provide a description of their overall performances.

In general, the association of all the data sources is low, with $r^2$ lying between 0.0 and 0.54 for the rainfalls, and between 0.0 and 0.96 for the simulated streamflows. The radar and satellite displayed the worst associations for each of the QCs that make up the Mgeni catchment, suggesting that direct use of radar and raw satellite data tends to produce large errors that can cause major errors in hydrological simulations. The differences in $r^2$ between the gauge, SG and merged fields are minimal for the majority of the QCs. The association is relatively stronger for U20C, U20E, U20H, U20J and U20M than for the remaining QCs, this being the result of the positive influence of the two daily reporting gauges located in U20J and U20M (cf. Figure 7.5). It is also suggested that the validity of information that a coefficient of determination can provide is highly sensitive to the sample size (i.e. evaluation period) and the size of values themselves (i.e. rainfalls or streamflows). With a small sample size, a single outlier is capable of considerably changing the slope of the regression line and, consequently, the value of the $r^2$. It is therefore important not to make important conclusions based solely on the values of the $r^2$.

In regard to the rainfall comparison, and using the bias as a measure, the SG slightly edged out the other data sources, but all are negatively biased, with the exception of satellite derived data. Again, in regard to rainfall, the raw satellite data had higher RMSE and MAE values than the other data sources, followed by radar (Figure 7.9).

The streamflow statistics, on the other hand, show different features to those of the rainfall statistics. The $r^2$ values of all the data sources were slightly degraded for all the QCs, with the exception of the $r^2$ in U20H and U20J (Figure 7.9). This reveals that the uncertain nature of the rainfall is amplified in its conversion to hydrological responses. However, the sensitivity of $r^2$ to small sample size and small values may again be a key factor. In regard to bias, the RMSE and MAE of the streamflow statistics show that the transformation of rainfalls to streamflows had a "smoothing" effect.
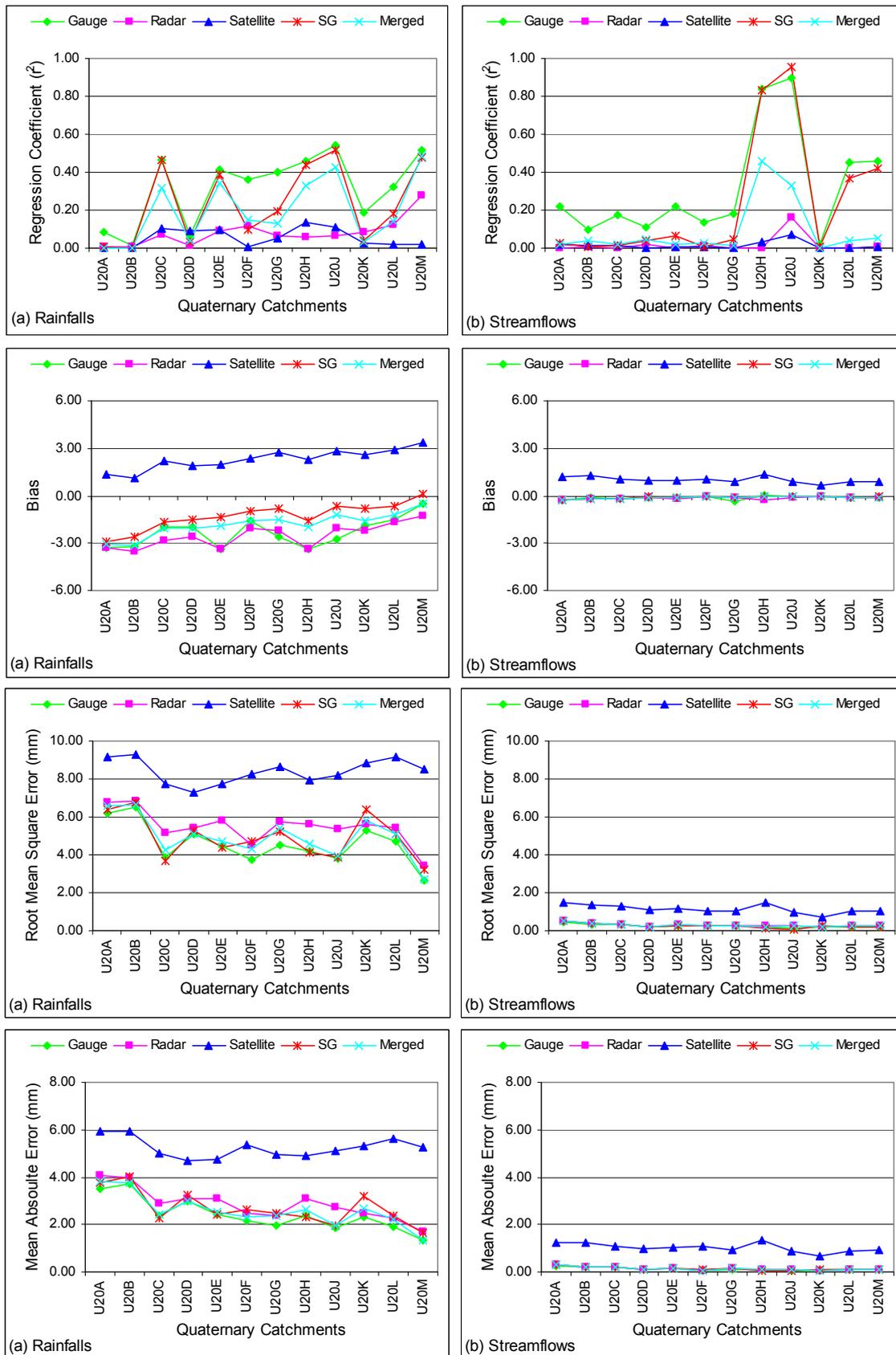
Figure 7.9    Coefficient of determination, bias, RMSE and MAE of subcatchment
               rainfalls (a) and accumulated streamflows (b) using the different data
               sources for the period 01 January to 31 March 2003

130

Except for the satellite derived values, all the data sources had similar errors for all the QCs (Figure 7.9). This is due mainly to the fact that the evaluation period is dominated by rainless days or small rainfall events. As a result, the response of the catchment to the minor rainfall events was very small as most of the rainfall was lost to satisfy the initial abstractions of the catchment before the commencement of streamflow. The generated daily streamflows from all the data sources range between 0 and 3 mm throughout the study period. Under these conditions, it is not possible to comprehensively investigate the influence of rainfall errors on streamflow errors since it is likely that for fast response catchments or for longer wet spells, the rainfall errors would be magnified in the streamflow simulations. Many researchers (e.g. Sun *et al.,* 2000; Pappenberger *et al.,* 2005) have confirmed that larger rainfall errors cause even larger errors in hydrological domains. It is important to bear in mind that the simulated streamflows at the exit of each QC includes the accumulated flows from upper QCs, while the rainfalls for each QC are independent of each other.

It is also important to note that, despite the overall failure of the radar to estimate the observed rainfall well over the Mgeni catchment, the estimates of rainfall have been shown to improve a little as one moves closer towards the location of the Durban radar. For example, the radar gave its worst estimates for U20A, U20B and U20D, which are far from the radar, while the estimates improved for U20L and U20M (cf. Figures 7.5 and 7.9), indicating that the distance from the Durban radar is partly responsible for its under-performance. Nevertheless, the radar estimates suffered from under-estimation even within the closest QC (U20M).

In order to visualise how well the trend of rainfalls and streamflows generated from each of these data sources correspond with the observations, a plot of accumulated daily rainfalls for the evaluation period is presented in Figure 7.10 for the QC U20M, followed by a plot of accumulated daily streamflows for U20M in Figure 7.11 and, finally, a plot of accumulated daily streamflows for the entire catchment, including cascaded streamflows from all QCs, in Figure 7.12. The three figures show the average ratio for the time series of rainfalls and streamflows. A perfect score is 1, while a score less than 1 indicates under-

estimation and a score greater than 1 indicates over-estimation above the reference.
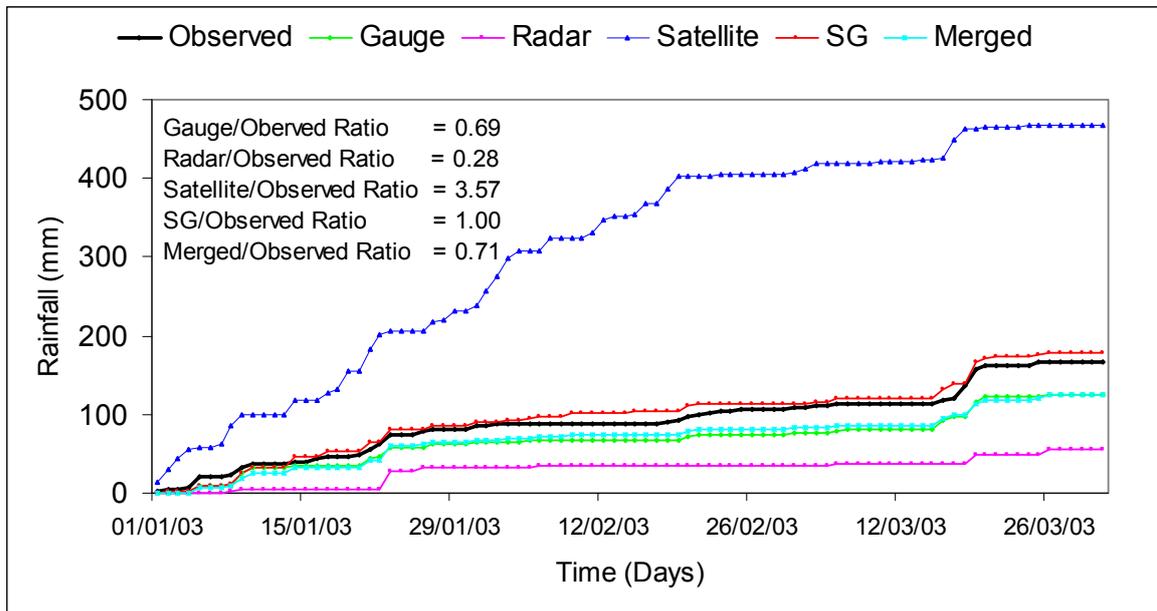


Figure 7.10   Time series comparisons of accumulated rainfall values derived from gauge, radar, satellite, SG and the merged rainfall field for QC U20M for the period 01 January to 31 March 2003

It is evident from the plots in Figures 7.10, 7.11 and 7.12 that the raw satellite data significantly over-estimated those of the observed values. The high rainfall ratio of 3.57 was amplified in the streamflow estimates to 5.01 and 5.18. The simplicity of the multi-spectral rain rate (MSRR) method to classify cloud structure into rain and no-rain classes is probably responsible for the huge over-estimations of rainfall during both rain days and rainless days of the study period. In contrast, the radar data failed to capture majority of the rainfalls.  The rainfall ratio of 0.28 was seen to improve in the streamflow ratios to 0.75 and 0.77. The relative dryness of the study period is the reason for the improvement. Nonetheless, the significant under-estimation may be caused by errors associated with ground clutter, distance from the radar and cloud height. The effect of distance is evident, as may be seen in Figure 7.9.
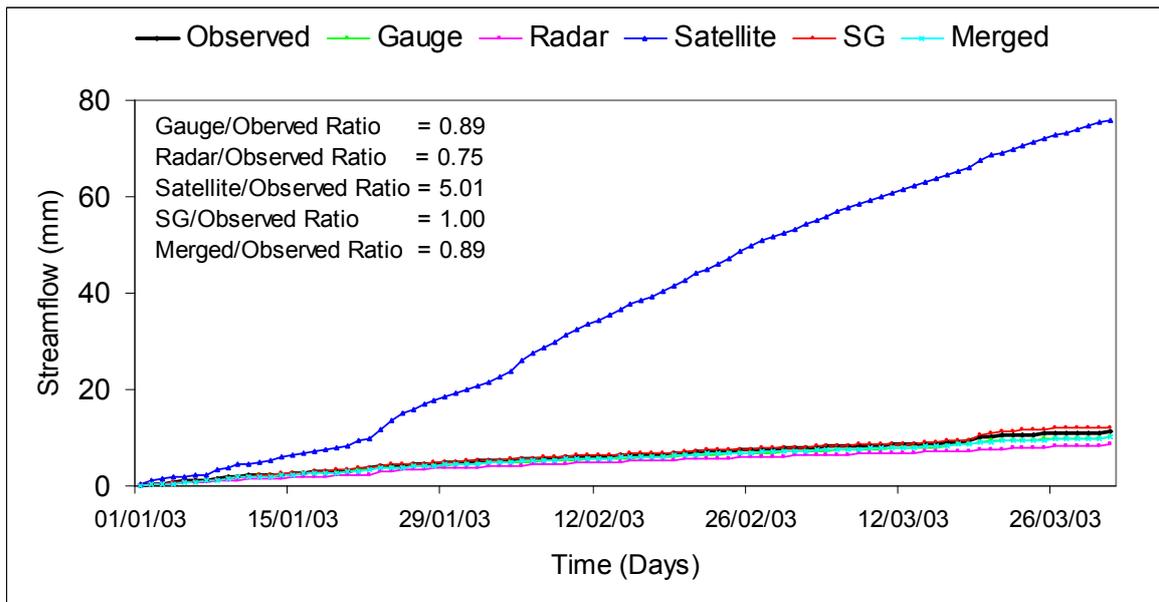
Figure 7.11 Time series comparisons of accumulated streamflows derived from gauge, radar, satellite, SG and the merged rainfall field for QC U20M for the period 01 January to 31 March 2003
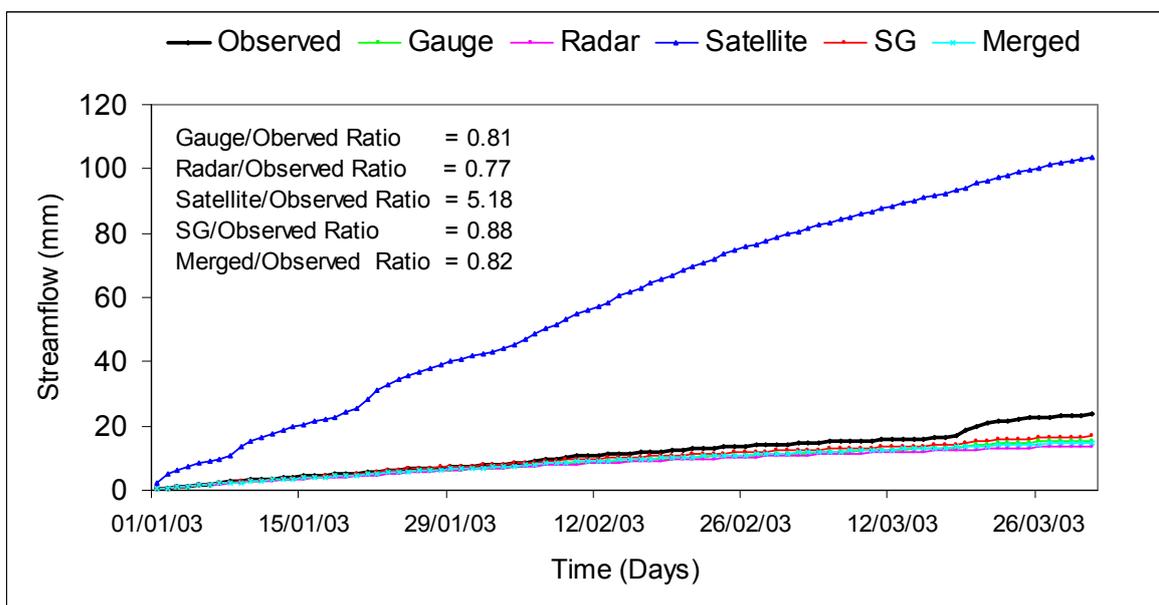


Figure 7.12 Time series comparisons of accumulated streamflows derived from gauge, radar, satellite, SG and the merged rainfall field for the entire Mgeni catchment for the period 01 January to 31 March 2003

The improvement that the SG offers is well illustrated in Figures 7.10 and 7.11. A perfect score was achieved in both the rainfall and transformed streamflow ratios for U20M only. The ratio was degraded to 0.88 when streamflows were considered for the entire catchment, suggesting that the SG slightly under-estimated the observed accumulated flows. The curves of gauge and merged data followed each other reasonably well in both the rainfall as well as the streamflows, but they suffered from under-estimation of the observed values (Figures 7.10, 7.11 and 7.12).

## 7.4    Summary and Conclusions

One of the most important steps in the development of a framework for agrohydrological forecasting is the near real time estimation of rainfall and, therefore, the choice of a reliable rainfall estimation method, because it is the near real time rainfall that will define the "now state" of a catchment at the beginning of a forecast period. Since 2002, the METSYS group of the SAWS and the School of Civil Engineering at the University of KwaZulu-Natal, in collaboration with the Department of Water Affairs and Forestry (DWAF) and ESKOM, have launched a project named **S**patial **I**nterpolation and **MA**pping of **R**ainfall (SIMAR) in order to produce near real time, spatially high resolution rainfall fields using data from around 300 daily reporting gauges and 10 radars as well as from satellite images. Individual and merged rainfall maps from these data sources are now available on a daily basis at a spatial resolution of one arc minute (i.e. ~1.7 km) for southern Africa. The benefits potentially obtainable from incorporating such products into agrohydrological forecast operations are well recognised. However, before such products can be used effectively in an operational agrohydrological forecasting system, a number of important questions have to be addressed. With rainfall being one of the most difficult components of the hydrological cycle to estimate, its uncertainty is the main one and is seen as a major constraint in the use of these products.

In this chapter the various SIMAR products have been reviewed and an attempt has been made to address the uncertainties associated with each of these products by applying a range of statistical tests. Rainfall fields derived from gauge, radar, satellite, conditioned satellite (SG) and the merged (gauge/radar/satellite)

field have been applied to the Mgeni catchment for the period of 01 January to 31 March 2003. The performances of each data source were evaluated by comparing results against observed rainfall data and also by comparing the streamflows simulated with the *ACRU* model using observed and estimated rainfalls as input. The observed rainfall fields in the evaluation were obtained from 31 daily rainfall stations in the Mgeni catchment, and they were assumed to be the "perfect reference", without consideration of interpolation errors.

First, the quality of the data from the various sources for two major rainfall events were quantified using the Critical Success Index (CSI), Probability of Detection (POD) and False Alarm Ratio (FAR) as a function of threshold percentiles, in order to evaluate their positional accuracy in relation to the observed rainfall pixels over QC U20E. Generally, the observed grid cell values for the two events were significantly under-estimated from the radar and satellite estimates. However, the gauge, SG and merged rainfall fields have shown their capability to reasonably depict the spatial patterns of observed rainfall for both events. Extending the verification, four commonly used statistical measures, *viz.* the coefficient of determination ($r^2$), bias, Root Mean Square Error (RMSE) and the Mean Absolute Error (MAE) were employed to distinguish between errors of the data sources in time series analyses. Overall, the SG seemed to be more promising than the other data sources used, and its superiority over the remaining four data sources was also confirmed by a score of perfect ratio (SG/observed ratio = 1). The performances of the gauge and merged rainfall fields in the time series analyses were found to be satisfactory. They had similar scores, but the merged data showed less bias and lower absolute errors than the gauge data. Again, the poor performances of the radar and satellite were shown in the time series analyses. The radar suffered from significant under-estimation, while the satellite estimates were over-estimating in the analyses of both rainfalls and streamflows.

Even though the results presented in this chapter are based on only a short data set, it highlights the fact that the radar and satellite alone cannot yet provide the accurate rainfall that is required for operational hydrological applications. The SG data seem to be preferable to the other data sources. However, the use of gauge and the merged rainfall fields as inputs into agrohydrological models could also

provide realistic results, without much difference in their hydrological outputs from that of the SG. Most importantly, however, their performances may not be similar to the ones presented for the Mgeni for different periods of record, or for other areas in southern Africa with a different rainfall season or different attributes of rainfall events. The evaluation of the SIMAR products should, therefore, be extended to other parts of southern Africa in order to account for a wider range of environmental conditions.

It is also important to point out that the rainfall-streamflow transformation by the *ACRU* model had a dampening effect on the accumulated flows, resulting in the levelling of the performances of the different data sources. This is hypothesised to be a consequence of the study period not being dominated by high runoff producing rainfall events. The influence of rainfall errors on hydrological estimates should be investigated in future studies by including a larger number of heavy rainfall events within the evaluation period. Moreover, observational uncertainty through interpolation schemes needs to be considered within the evaluation statistics.

# 8. EVALUATION OF SHORT AND MEDIUM RANGE RAINFALL PREDICTION MODELS FROM A HYDROLOGICAL PERSPECTIVE

## 8.1    Introduction

Most hydrological models employed for short and medium range forecasting depend on quantitative precipitation forecast (QPF) inputs, which are issued as either deterministic or probabilistic forecasts or as ensembles of probabilistic forecasts, over a pre-determined lead time (Goswami and O'Connor, 2006). The basis of current short (i.e. 1-3 day) and medium (i.e. 4-15 day) range forecasting practice is Numerical Weather Prediction (NWP), a science that has been developed rapidly over the past few decades (Anstee, 2004). NWP models can be categorised into global, regional or mesoscale, based on the extent of their spatial cover. Global models have global extent, while regional models cover only a fraction of the globe such as a continental land surface and surrounding oceans. Mesoscale models cover a relatively smaller area, ranging up a few hundreds of square kilometres. Since the spatial resolution of NWP models is constrained by the computational time and memory capacity of the computers used to run them, global models have the coarsest resolution of the three categories whereas mesoscale models have, relatively, the finest resolution (Anstee, 2004).

Generally improvements of the NWP models with respect to spatial and temporal resolution, as well as to more detailed representations of the atmospheric processes, have led to a significant improvement of weather forecasts (Golding, 2000; Habets *et al.,* 2004). In spite of these improvements, the skill of the NWP models has not yet reached an acceptable level of confidence, especially for longer lead time forecasts (Federico *et al.,* 2004; Habets *et al.,* 2004; Roads, 2004; Bocchiola and Rosso, 2006). The reasons for this are that rainfall is hugely variable both in space and time, and that great uncertainties affect the performances of the NWP models (Bocchiola and Rosso, 2006). In NWP models, the physical processes, which are at sub-grid scale, are represented in parameterised form. Thus, NWP models cannot account for local environmental attributes that influence the production of rainfall (Maini *et al.,* 2004). Another key problem in NWP modelling is the instability of the atmosphere, as well as the

sensitivity of the rainfall forecasts to small changes in initial conditions of the atmosphere (Ahrens and Juan, 2007).

According to Habets *et al.* (2004) the application of NWP precipitation forecasts into hydrological models to predict streamflows or peak flows is limited by three types of error:

- localisation of the events,
- timing of the events, and
- precipitation intensity.

However, NWP precipitation forecasts are often associated with other tools in order to correct some of the errors prior to their application with hydrological and/or crop yield models. Commonly used techniques that may improve upon these global-scale models are regional climate modelling and statistical post-processing methods (Hay *et al.,* 2003; Habets *et al.,* 2004; Maini *et al.,* 2004). Such methods can account for the local topographic and other environmental variables that control precipitation (Maini *et al.,* 2004). The introduction of Ensemble Forecasting Systems (EFS) to account for the probability distribution of atmospheric states arising from uncertainties in the initial state has also enabled some NWP models (e.g. NCEP) to display better results than using only a single deterministic forecast that is initiated by the best known, but nevertheless uncertain, atmospheric state (Golding, 2000; Hay *et al.,* 2003; Ahrens and Juan, 2007).

In South Africa, several institutions such as the SAWS, the University of Pretoria, and the University of Cape Town have been actively involved in research relating to quantitative precipitation forecasting in order to make short (1- 3 day) and medium (4 -15 day) rainfall forecasts operationally feasible for application into daily time-step hydrological and/or crop yield models. Incorporating such forecasts within the framework for an agrohydrological forecasting system has been a major task of this study. At the present time, experimental forecasts issued by the SAWS from the National Centre for Environmental Prediction for Medium Range Forecasting (NCEP-MRF) model and the Unified Model (UM), as well as forecasts given by the University of Pretoria from the Conformal-Cubic Atmospheric Model

(C-CAM), are incorporated within the framework for short and medium range agrohydrological applications (cf. Chapter 5). However, since these models have not been extensively tested in southern Africa, there is a strong need for objective assessments both in regard to rainfall characteristics and hydrological results in order to evaluate the skill and confidence of these models.

This chapter therefore aims at evaluating the archived rainfall forecasts from the C-CAM, UM and NCEP-MRF models on the Mgeni catchment. Methods of comparison and forecasting procedures are described in Section 8.2, while the results obtained from each model are briefly discussed in Section 8.3. Conclusions are presented in Section 8.4.

## 8.2    Methods of Comparison

The C-CAM and UM models have only recently (i.e. 2006) been adopted for southern Africa and the archived rainfall hindcasts from these two models are therefore only available for the period from May 2006 to date. As was the case in Chapter 7, the four highest daily observed rainfall events which occurred on 17 November 2006, 21 December 2006, 30 January 2007 and 04 March 2007 were selected for a pixel-by-pixel comparison over Quaternary Catchment U20E within the Mgeni catchment. Each 1' x 1' pixel value within U20E in the forecast lead time of the C-CAM and UM models was compared against the corresponding pixel value in the reference rainfall image and the Critical Success Index, CSI, the Probability of Detection, POD and the False Alarm Ratio, FAR (Wilks, 1995), the definitions of which are given in Chapter 4, were then used to assess the overall degree of their positional accuracy.

Observed and forecasted pixel rainfall values within each Quaternary Catchment (QC) of the Mgeni catchment were averaged to be used as input into the *ACRU* model for subsequent streamflow analysis. The semi-distributed catchment mode of the *ACRU* model was run with historical observed daily rainfall from year 2000 up to the time of the forecast start in order to create representative antecedent conditions and to initialise stores (e.g. soil moisture status in the top- and subsoil, the baseflow store and releases). Two scenarios were then used for the simulation of accumulated streamflows from subcatchments cascading downstream at the

exit of each QC for the period from 01 November 2006 to 31 January 2007. For the first scenario, the *ACRU* model was run with rainfall forecasts obtained from both the C-CAM and UM models from the time of the forecast start up to the end of the forecast period, while for the second scenario the *ACRU* model was initiated at each day of the forecast period with observed rainfall of the previous day (i.e. up to the "this morning" state) before a hydrological forecast was made for the next day with rainfall forecasts obtained from the NWP models. In other words, the bias due to the incorrect initial state of the catchment was corrected and only the error in the rainfall forecast generates some differences with the reference run.

As was the case in Chapter 7, daily rainfall values measured by raingauges distributed across the Mgeni catchment (Figure 7.5) for the selected evaluation period were interpolated using the Natural Neighbour method to serve as the "ground truth" for the verification. The "observed streamflows" were the simulated streamflows with the *ACRU* model using the so-called ground truth rainfalls. The coefficient of determination ($r^2$), bias, Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) were then computed in order to assess the skill of the C-CAM and UM models with and without updating schemes. Owing to the scale gap, the NCEP-MRF forecasts at grid box of 2.5º can not be applied directly with the *ACRU* model. These forecasts were therefore verified only against observed rainfalls.

## 8.3    Results and Discussion

The four selected rainfall events, *viz.* on 17 November 2006, 21 December 2006, 30 January 2007 and 04 March 2007 have been named Event A, Event B, Event C and Event D, respectively, for the sake of comparison in this chapter. Since each model runs for different lead times and at different spatial scales, individual comparisons against observations were first presented in Sub-sections 8.3.1, 8.3.2 and 8.3.3 in order to assess to the extent to which the lead time of each model is skilful, while in Sub-section 8.3.4 the comparison between the combined output of the C-CAM and UM models and the observational reference are discussed.

### 8.3.1  Evaluation of the C-CAM Rainfall Forecasts

C-CAM is a variable-resolution, hydrostatic model developed in Australia for regional and mesoscale weather prediction (Anstee, 2004). The model is formulated on a quasi-uniform grid, derived by projecting the panels of a cube towards the surface of the earth. The two dimensional projection of the squares onto the sphere forms the horizontal grid pattern used for the atmospheric model (Katzfey and McGregor, 2003; Anstee, 2004; Engelbrecht, 2005). An innovation that makes the C-CAM model more powerful is the ability to stretch the conformal-cubic grid over any selected region by a method termed the "Schmidt transformation" (Anstee, 2004; Engelbrecht, 2005). The C-CAM model has been adapted not only for short and seasonal term forecasting, but also for future climate change projections over southern Africa. Engelbrecht (2005) has, for example, applied the model for climate change simulations over southern and tropical Africa for the period 2070 -2100.

Work is underway (2007) at the University of Pretoria to extend the lead time of the model up to 40 days and to further stretch the model's grid from 15 km to an 8 km resolution over selected areas over South Africa. However, at the present time (November 2007) the archived hindcasts are available only for the 15 km spatial resolution rainfall forecasts of 4 days' lead time (Engelbrecht, 2007). The evaluation of these forecasts is presented here.

A visual comparison of the 4 day lead time rainfall forecast over QC U20E is presented in Figure 8.1 for the four selected events. The distributions of Hits (H), Misses (M), False Alarms (FA) and Correct Nulls (CN) within the threshold of 50th percentile are given in Figure 8.2. For Event A the 1 and 2 day lead time forecasts displayed similar distribution with 5.7% FA and 92% CN, while 92% of the pixel rainfall values were missed in the 3 and 4 day lead time forecasts.  For Event B the 1, 3 and 4 day lead time forecasts missed majority of the pixels with large rainfall values. By way of contrast, the 2 day lead time forecast for Event B tended to over-estimate, even though majority of the rainfall pixels are scattered in the CN range. Similar to Event B, the 1, 3 and 4 day lead time forecasts for Event C are clustered within the range of H and M ranges, while most of rainfall pixels for the 2

day lead time forecast fell within the ranges of FA and CN. For Event D the C-CAM model captured the pixels with low rainfall values, but most of the pixels with large rainfall values are missed (Figure 8.2). The pixel-by-pixel comparisons of the CSI, POD and FAR (Equations 4.5, 4.6, 4.7 in Chapter 4, Section 4.13) as a function of threshold rainfall percentiles for these four events are presented in Figures 8.3 and 8.4. For Event A the 1 and 2 day lead time forecasts displayed similar results with CSI and POD scores up to the 70th percentile, while the skill for the 3 and 4 day lead time forecasts was only up to the 20th percentile. No FAR was scored for Event A except for the 1 and 2 day lead time forecasts at the 60th and 70th threshold percentiles.
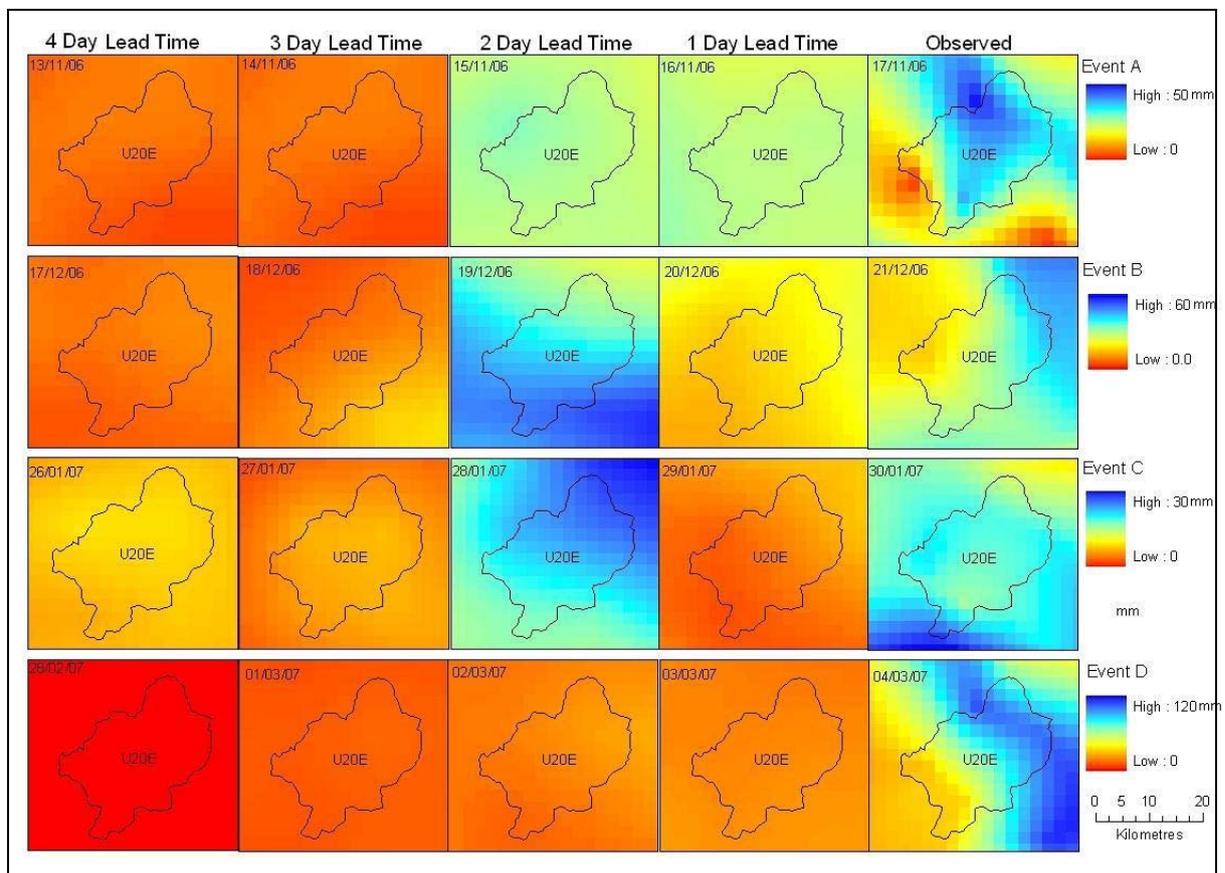


Figure 8.1    Four day lead time rainfall forecasts from the C-CAM model over QC U20E in the Mgeni catchment for Events A to D on 17 November 2006, 24 December 2006, 30 January 2007 and 04 March 2007
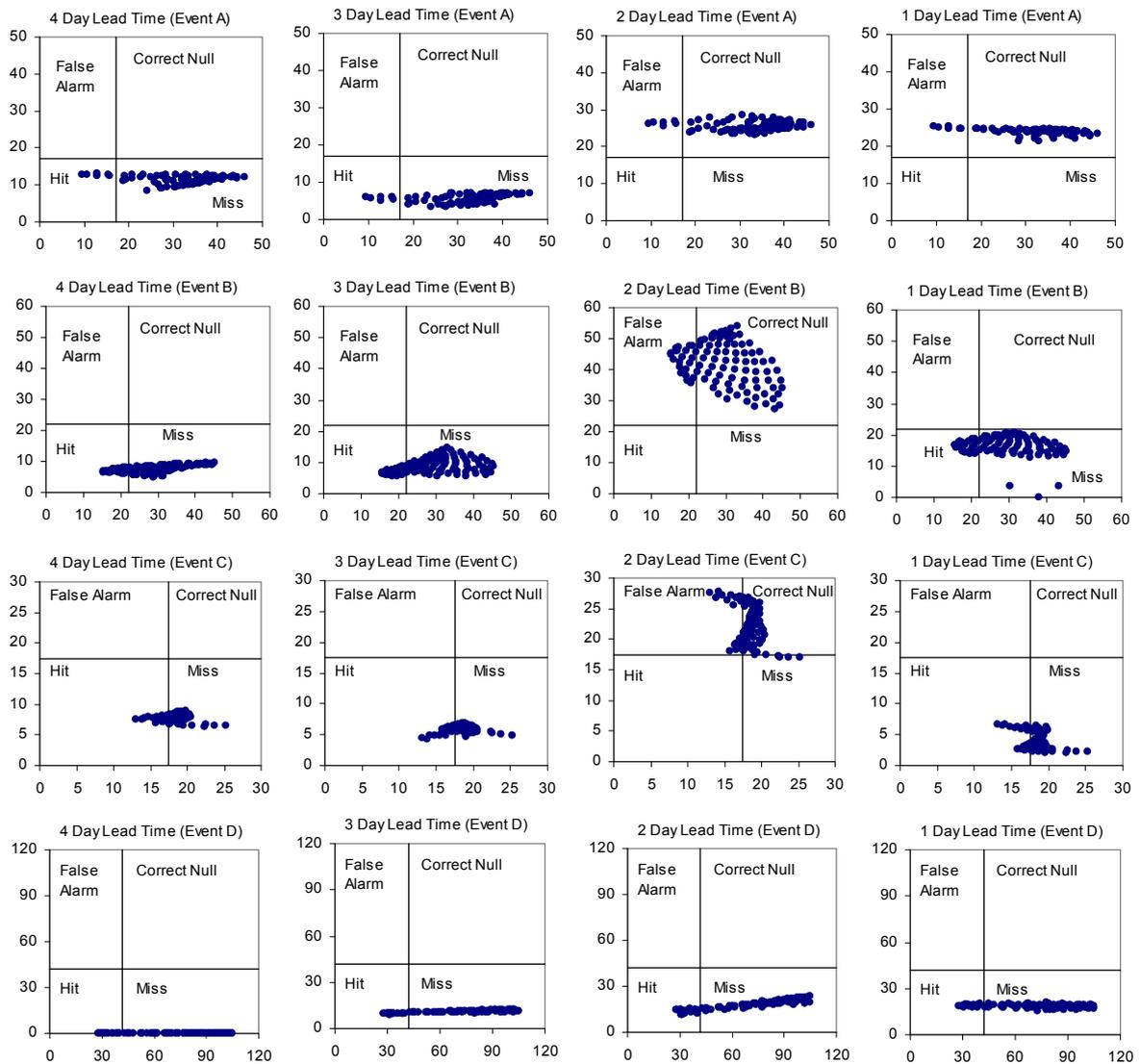
Figure 8.2    Scores of Hit, Miss, False Alarm and Correct Null defined by the threshold of 50th percentile for 17 November 2006, 24 December 2006, 30 January 2007 and 04 March 2007

Surprisingly, for Events B and C the 2 day lead time forecasts edged out the 1 day forecast, with CSI and POD scores up to the 70th and 80th percentiles, respectively. However, most pixel rainfalls in the 2 day lead time forecasts for these two events were significantly above their corresponding pixel rainfalls from observations, resulting in high FAR scores in most of the higher threshold percentiles (Figures 8.3 and 8.4). On the day of Event D more than 100 mm of rainfall was recorded by five raingauges located around QC U20E. However, the

C-CAM model has failed to capture the higher pixel rainfalls even on the 1 day lead time forecast. Most pixel rainfalls in the forecasts were significantly below their corresponding pixel rainfalls in the observations (Figure 8.4).



Figure 8.3    CSI, POD and FAR scores of 4 day lead time forecasts from the C-CAM model as a function of threshold percentiles for the rainfall events of 17 November and 21 December 2006

Figure 8.4    CSI, POD and FAR scores of 4 day lead time forecasts from the C-CAM model as a function of threshold percentiles for the rainfall events of 30 January and 04 March 2007

The 1 day lead time forecast displayed relatively more skilful forecasts than the longer range forecasts and was used as input into the *ACRU* model for streamflow simulations both with and without updating scenarios for the period of 01 November 2006 to 31 January 2007. Plots of daily and accumulated daily streamflows cascaded from all QCs to the mouth of the Mgeni catchment for the two scenarios are presented in Figures 8.5 and 8.6 respectively.

145

Figure 8.5    Time series comparisons of daily streamflows simulated with the *ACRU* model at the mouth of the Mgeni catchment, derived from the C-CAM rainfall forecasts both with and without updating procedures for the period 01 November 2006 to 31 January 2007



Figure 8.6    Comparisons of accumulated streamflows simulated with the *ACRU* model at the mouth of the Mgeni catchment, derived from the C-CAM rainfall forecasts both with and without updating procedures for the period 01 November 2006 to 31 January 2007

146

The streamflow forecasts when using the C-CAM rainfall forecasts without the updating scenario displayed significant under-estimation. The explanation for this under-estimation obviously lies in the fact that the *ACRU* model was initiated with uncorrected rainfalls on each successive day throughout the study period. The error cascade in the rainfall forecasts of each day had a significant influence on the *ACRU* streamflow simulation state variables such as the fraction of water that become a streamflow from the topsoil, subsoil and intermediate/groundwater stores on a given day, and consequently on the streamflow forecasts. However, the C-CAM based streamflow forecasts were seen to improve considerably when the *ACRU* model was initiated with observed rainfalls at the start of each day in the forecast period. As may be seen in Figures 8.5 and 8.6, the daily time series and accumulated daily streamflows simulated with the updating scenario appeared much closer to the reference streamflows, with the total streamflow ratio was improving from 0.56 to 0.90.

Statistical comparisons with respect to the coefficient of determination ($r^2$), bias, the Root Mean Square Error (RMSE) and the Mean Absolute Error (MAE) are presented in Figure 8.7, so as to highlight the significance of the improvements made by the updating procedure. As was expected, the $r^2$ values are relatively higher for the updated forecasts, ranging from 0.23 to 0.72 for the various QCs that make up the Mgeni catchment, thereby indicating a better agreement than the uncorrected scenario for which the $r^2$ range was only between 0.01 and 0.37. The updating procedure has also reduced the bias, RMSE and MAE values to minimum levels (Figure 8.7), suggesting that the daily correction of the *ACRU* streamflow state variables based on observed rainfall has a significant influence in reducing both the systematic and random errors in the accumulated streamflow forecasts.

In conclusion, the C-CAM model has suffered from both under-estimation and over-estimation in the analysis of the four individual daily rainfall events, indicating the variability of the model's performance from storm to storm. Overall, the observed rainfalls over the entire study period were under-estimated by the model. Consequently, the streamflow forecasts were consistently below their corresponding observed flows. However, the under-estimation was seen to

improve significantly when a daily correction with observed rainfalls was made to initiate the *ACRU* model with the correct "now-state" of the catchment. An error of 34% in the total streamflow forecasts of the first scenario was attributed to an incorrect initialisation of the *ACRU* model used for each forecast run.
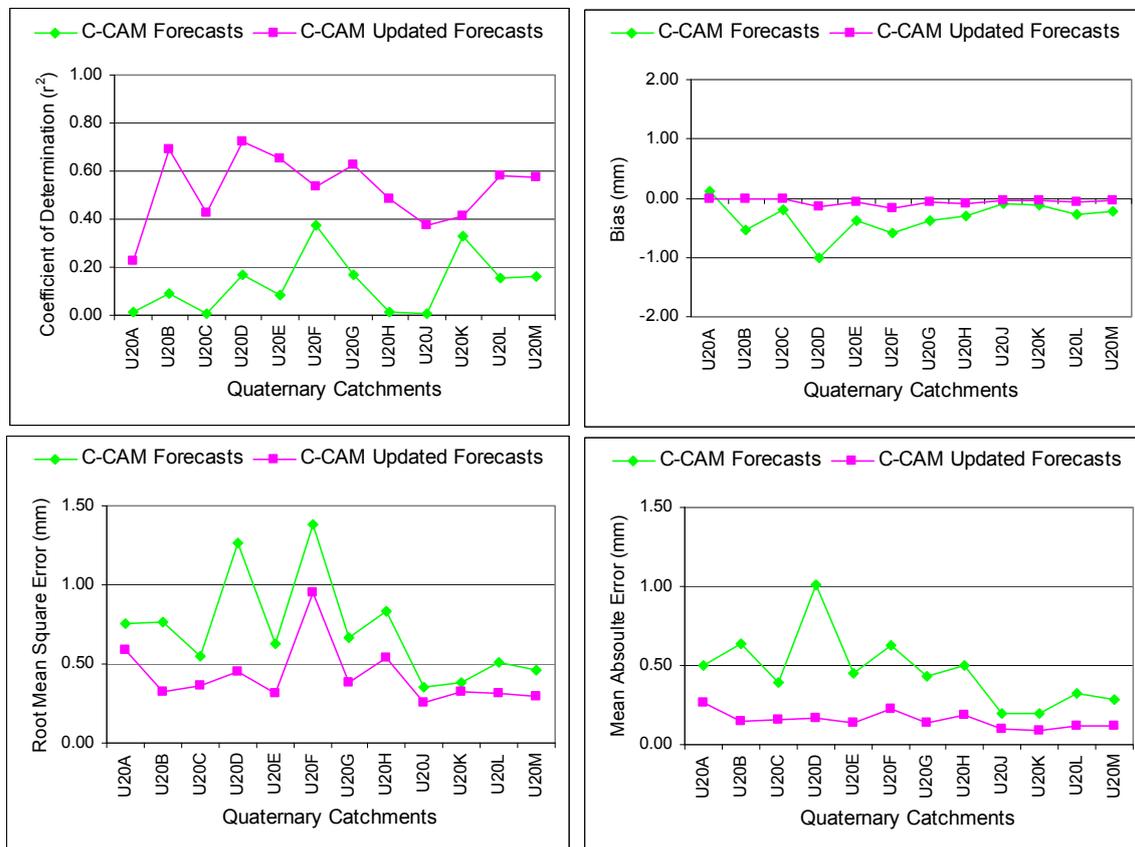


Figure 8.7    Coefficient of determination, bias, RMSE and MAE of accumulated Quaternary Catchment streamflows simulated with the *ACRU* model when using the C-CAM 1 day lead time rainfall forecasts both with and without updating, for the period of 01 November 2006 to 31 January 2007

## 8.3.2  Evaluation of the UM Rainfall Forecasts

The Unified Model (UM) is made up of atmospheric, oceanic, wave and sea-ice numerical submodels and can cover either all, or part, of the Earth's surface area with multiple atmospheric layers. The various submodel components have been designed to run individually or in a merged mode for a specific modelling application (UK Met Office, 2007). Operationally, the UK Meteorological Office

runs a number of configurations of its UM model, ranging from the global model, with a spatial resolution of 100 km, down to a high resolution of 4 km local model. The choice of horizontal and vertical resolution may be varied by a user (UK Met Office, 2007). Since 2006 the SAWS has been actively working on the implementation of this new Unified Model as a new NWP system for southern Africa. The 12 km resolution rainfall forecasts of 2 days' lead time is currently available and could be used for short term agrohydrological applications. The evaluation of these forecasts in the Mgeni catchment is demonstrated below. In Figure 8.8 the 1 and 2 day lead time rainfall forecasts over QC U20E are shown for events A, B, C and D, along with the reference observation, while in Figure 8.9 their Hits (H), Misses (M), False Alarms (FA) and Correct Nulls (CN) scores are plotted within the boundary of the 50th threshold percentile.
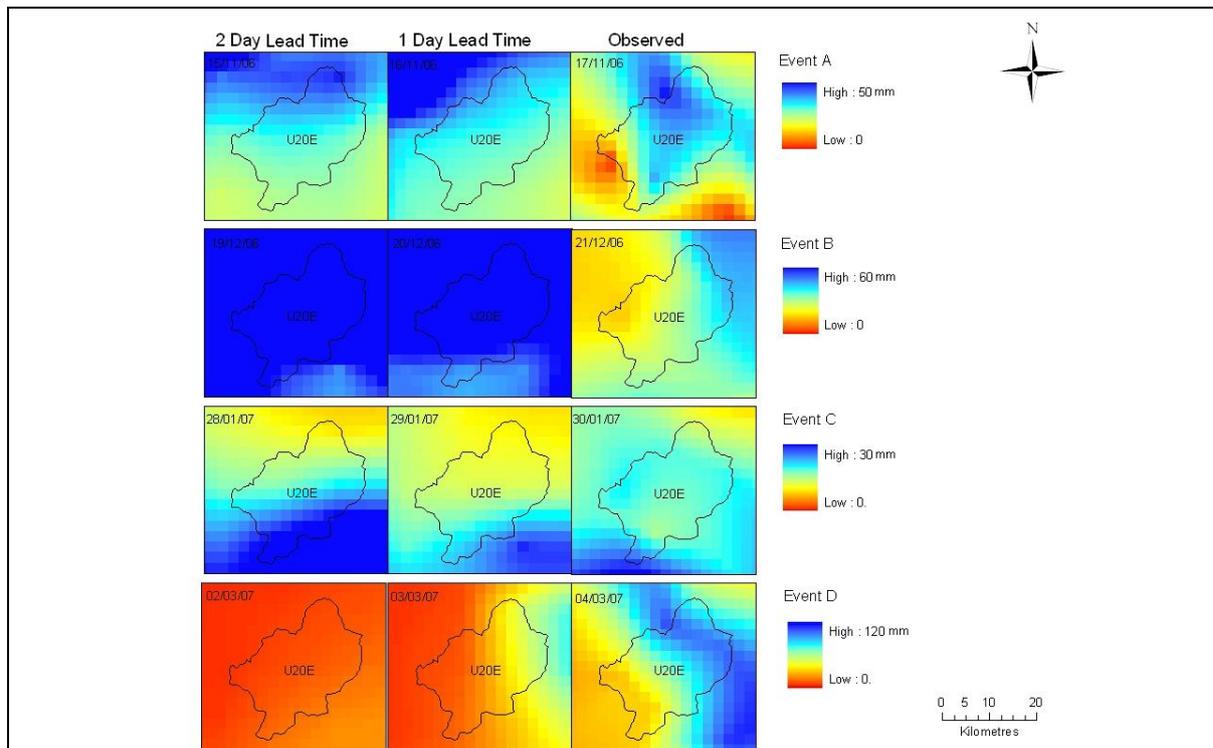


Figure 8.8    Two day lead time rainfall forecasts from the UM model over QC U20E in the Mgeni catchment for Events A to D on 17 November 2006, 24 December 2006, 30 January 2007 and 04 March 2007

For Events A and B the 1 and 2 day lead time forecasts are positively biased with high scores of FA. However, for Event C the 1 day lead time forecast is clustered along the boundaries of the four categories, while the 2 day lead time forecast is

more stretched to the FA side. For Event D the UM model failed to capture most of the pixels with high rainfall values, even though the model skill is better in the 1 day lead time forecast than the 2 day ahead forecast. The CSI, POD and FAR scores for the 1 and 2 day lead time UM forecasts as a function of threshold percentiles are illustrated in Figures 8.10 and 8.11.
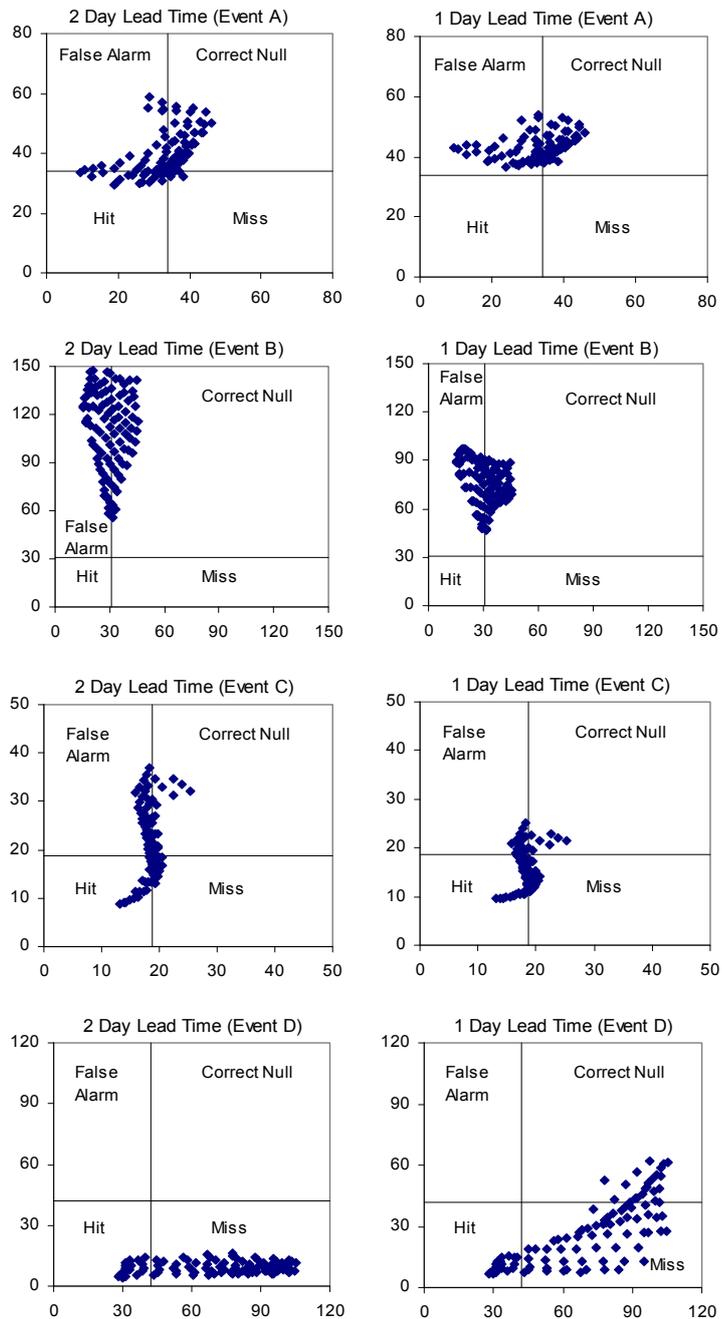


Figure 8.9    Scores of Hit, Miss, False Alarm and Correct Null defined by the threshold of 50th percentile for 17 November 2006, 24 December 2006, 30 January 2007 and 04 March 2007
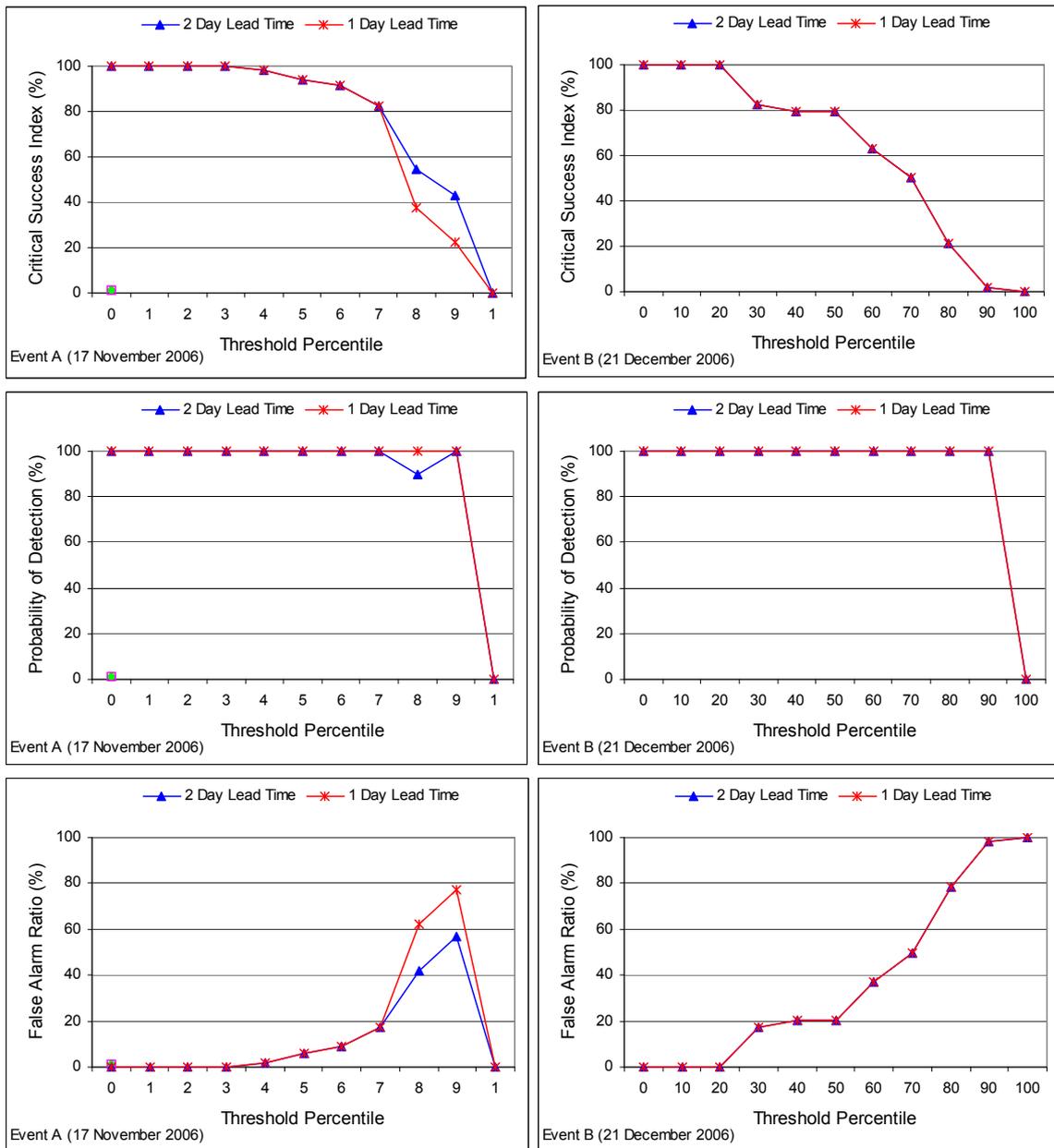
Figure 8.10   CSI, POD and FAR scores of 2 day lead time rainfall forecasts from the UM model as a function of threshold percentiles for the rainfall events on the 17 November and 21 December 2006

For Events A and B the 1 and 2 day lead time forecasts showed the same pattern over the entire range of threshold percentiles. The probability to detect a rainfall event up to the 90th percentile was 100% for both events. The CSI score dropped quickly with the increasing rainfall rate, ranging from 80% at the 70th percentile for Event A to 23% and 43 % at the 90th percentile for the 1 and 2 day lead time forecasts, respectively, while for Event B the CSI ranged from 80% at the 50th

151

percentile to 0% at the 90th percentile for both lead times. The probability of FAR for Event A for the 1 and 2 day lead time forecasts was seen to increase respectively from 5% at the 50th percentile to 77% and to 56% at the 90th percentile. For Event B, the FAR score for both lead times increased, starting from 18% at the 30th percentile to 100% at the highest threshold percentile (Figure 8.10).
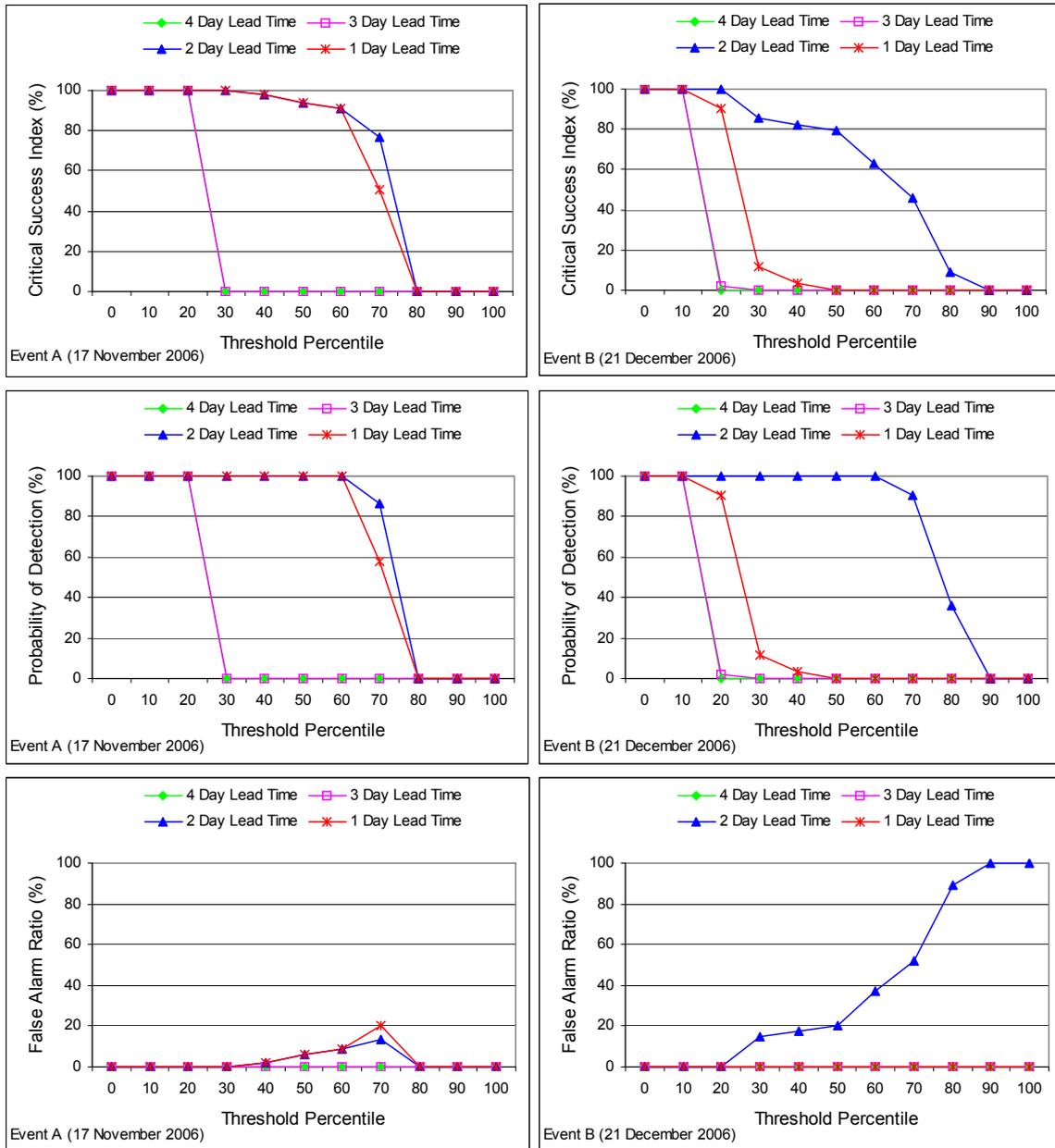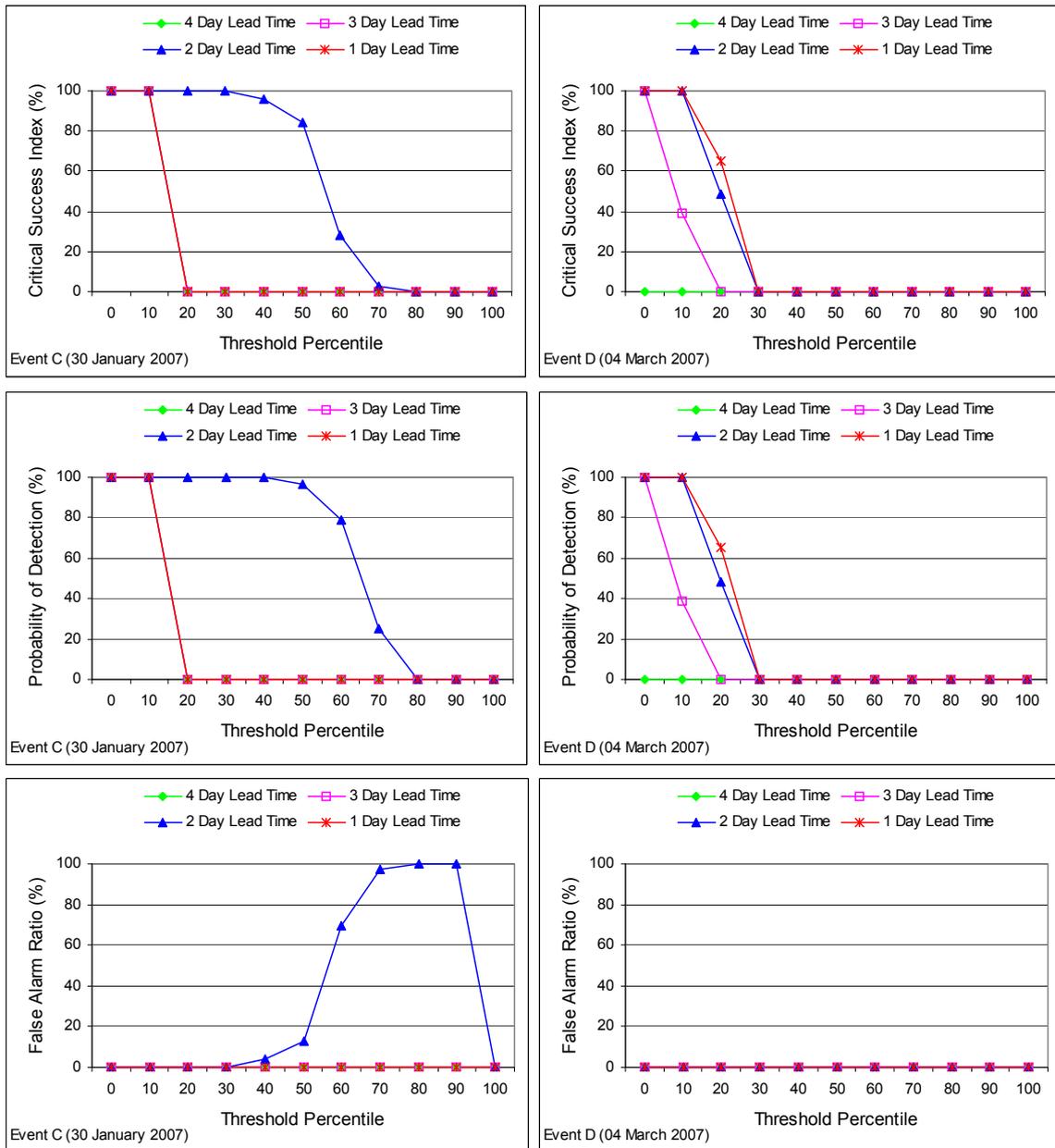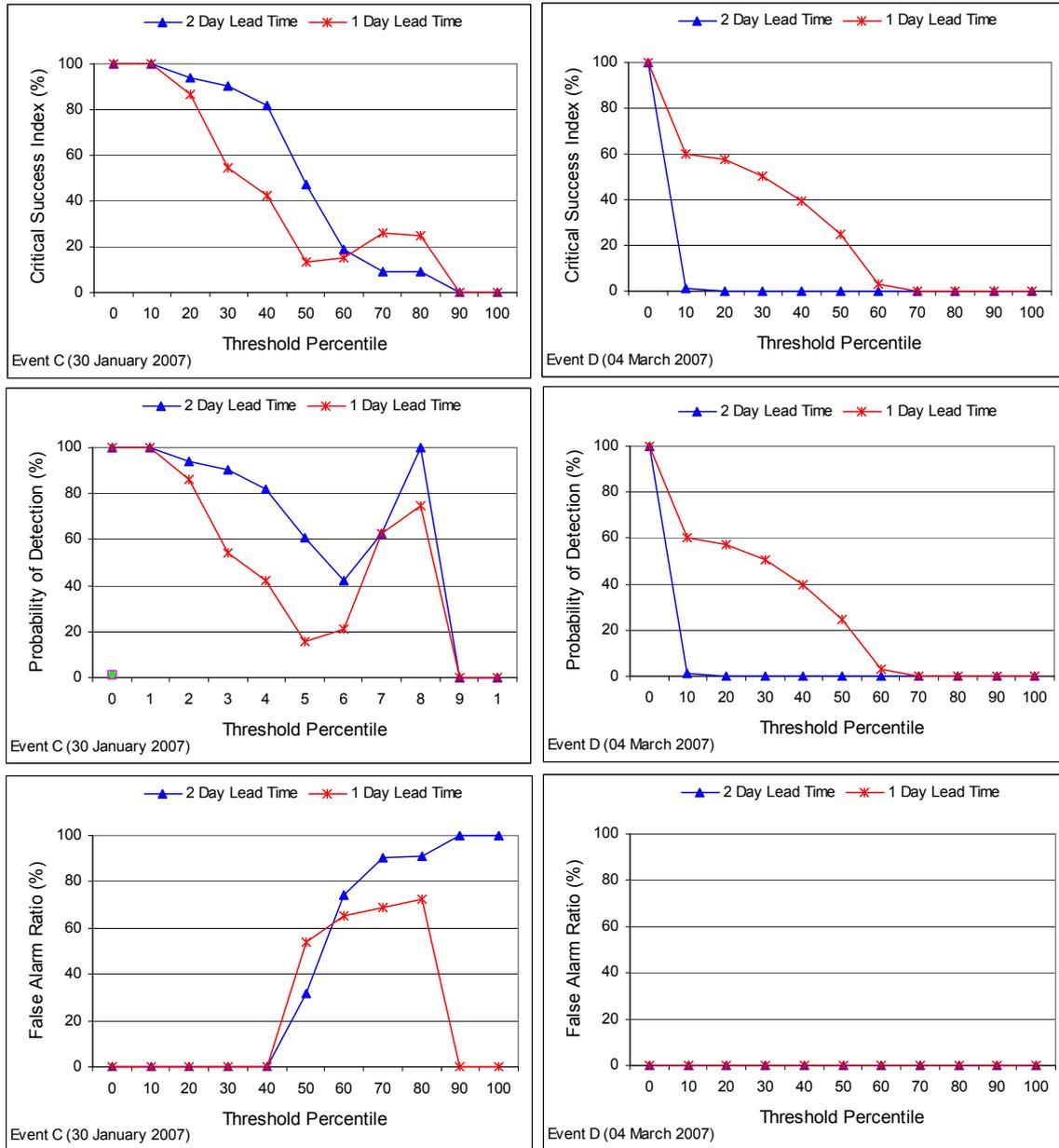


Figure 8.11   CSI, POD and FAR scores of 2 day lead time rainfall forecasts from the UM model as a function of threshold percentiles for the rainfall events on 30 January and 04 March 2007

For events C and D the CSI and POD scores dropped quickly with the threshold percentiles rate. The 2 day lead time forecast displayed a better statistical performance than the 1 day lead time forecast for Event C, but with higher FAR values for high rainfall thresholds. A significant under-estimation of the rainfall pixels was observed over the entire area (i.e. U20E) on the day of Event D. The 2 day lead time forecast displayed no skill scores over the entire domain except at the lowest threshold percentile (0%), while the skill of 1 day lead time forecast was extended up to the 50th percentile (Figure 8.11).

The UM model slightly over-estimated the rainfall for the majority of rainfall pixels in Events A, B and C and the results obtained from the 1 and 2 day lead time forecasts were generally very similar for these three events. The model's performance for Event D was poor, even though the skill increased slightly with decreasing lead time.

As was the case in Sub-section 8.3.1, the evaluation was extended by transforming the 1 day ahead UM rainfall forecasts into streamflow forecasts with the *ACRU* model, both with and without updating scenarios for the evaluation period of 01 November 2006 to 31 January 2007. Plots of daily time series and accumulated daily streamflows cascaded from all QCs to the mouth of the Mgeni catchment are presented in Figures 8.12 and 8.13, respectively. In Figure 8.14 plots of $r^2$, bias, RMSE and MAE are shown for each of the QCs that make up the Mgeni catchment. In general, the $r^2$ values with and without updating are very close to one another, ranging between 0.0 and 0.44. However, the improvement made by the updating scenario is highly noticeable in term of improvements to the bias, RMSE and MAE values.

It is evident from the plots in Figures 8.12, 8.13 and 8.14 that throughout the study period the UM model consistently over-estimated values compared to those of the observed. The streamflow ratio 3.91 was decreased to 1.59 when the updating scenario was used. Nevertheless, the updated forecasts are still positively biased by 59% according to the reference observed run, which is significant.
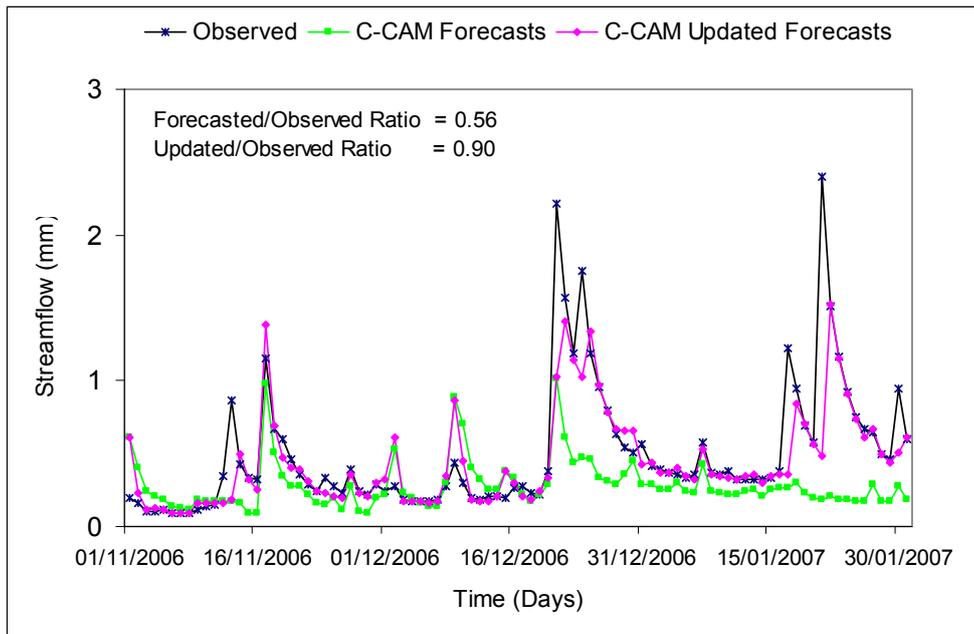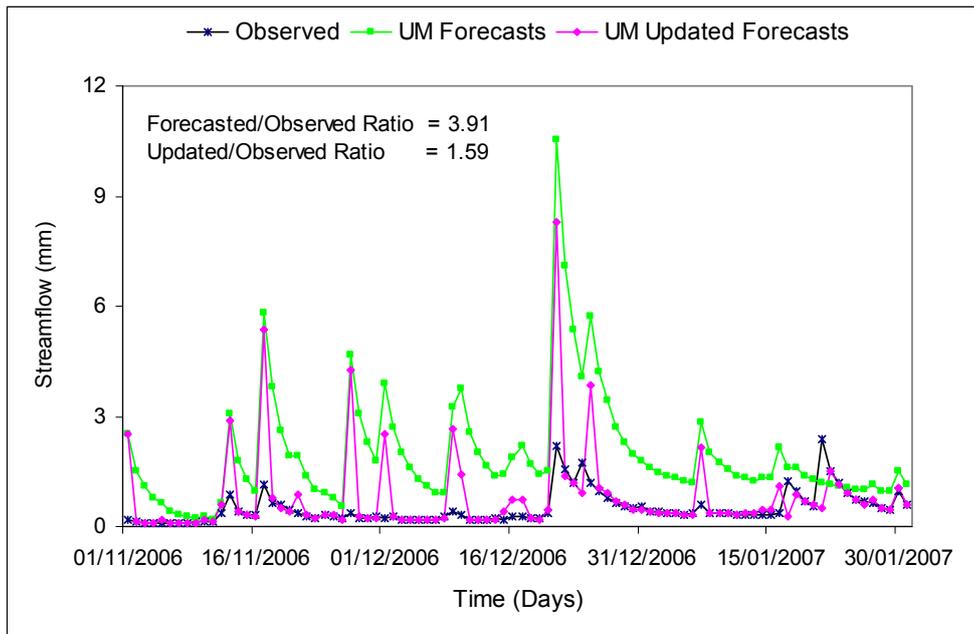
Figure 8.12   Time series comparisons of daily streamflows simulated with the *ACRU* model at the mouth of the Mgeni catchment, derived from the UM rainfall forecasts both with and without updating procedures for the period 01 November 2006 to 31 January 2007
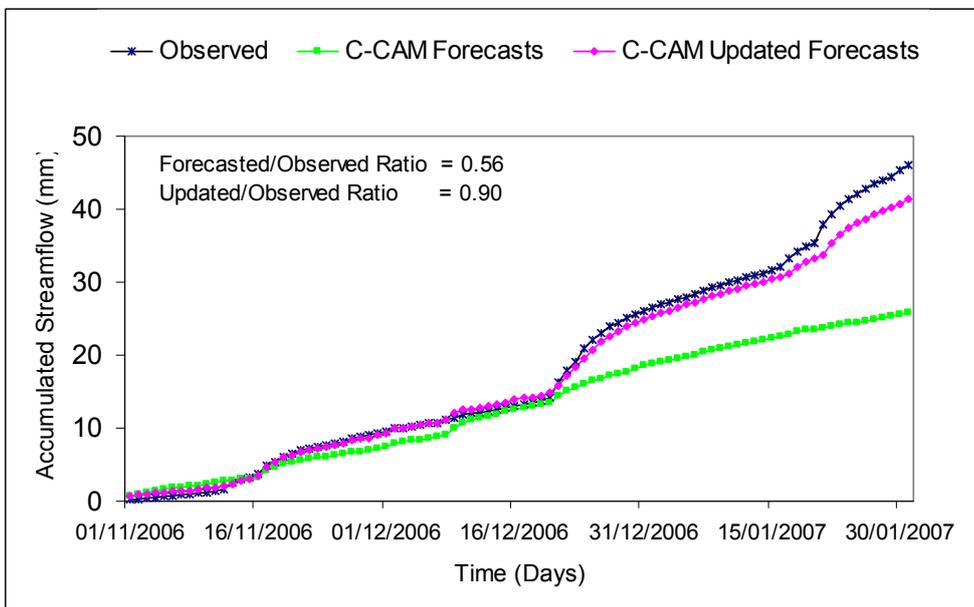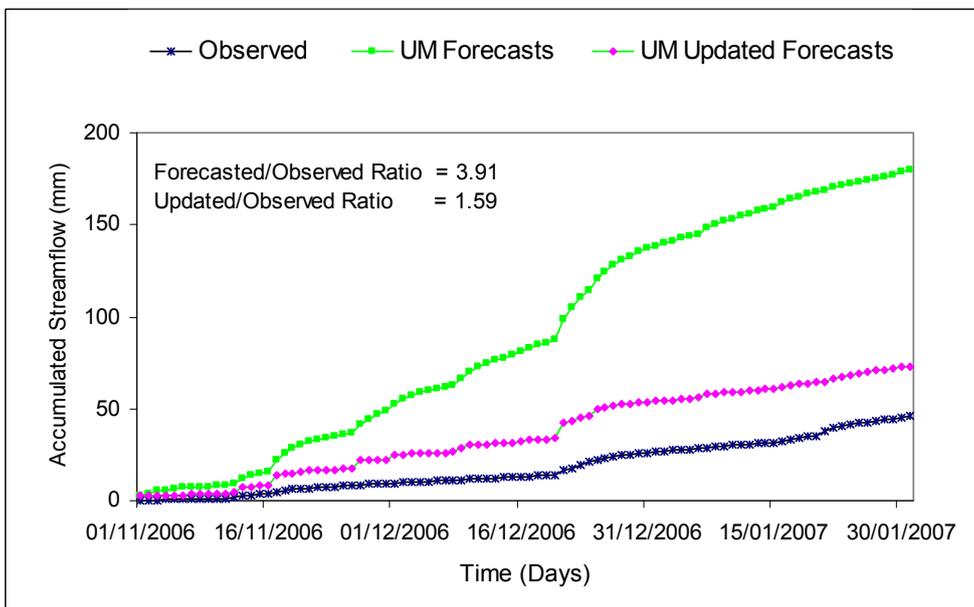


Figure 8.13   Comparisons of accumulated streamflows simulated with the *ACRU* model at the mouth of the Mgeni catchment, derived from the UM rainfall forecasts both with and without updating procedures for the period 01 November 2006 to 31 January 2007

154
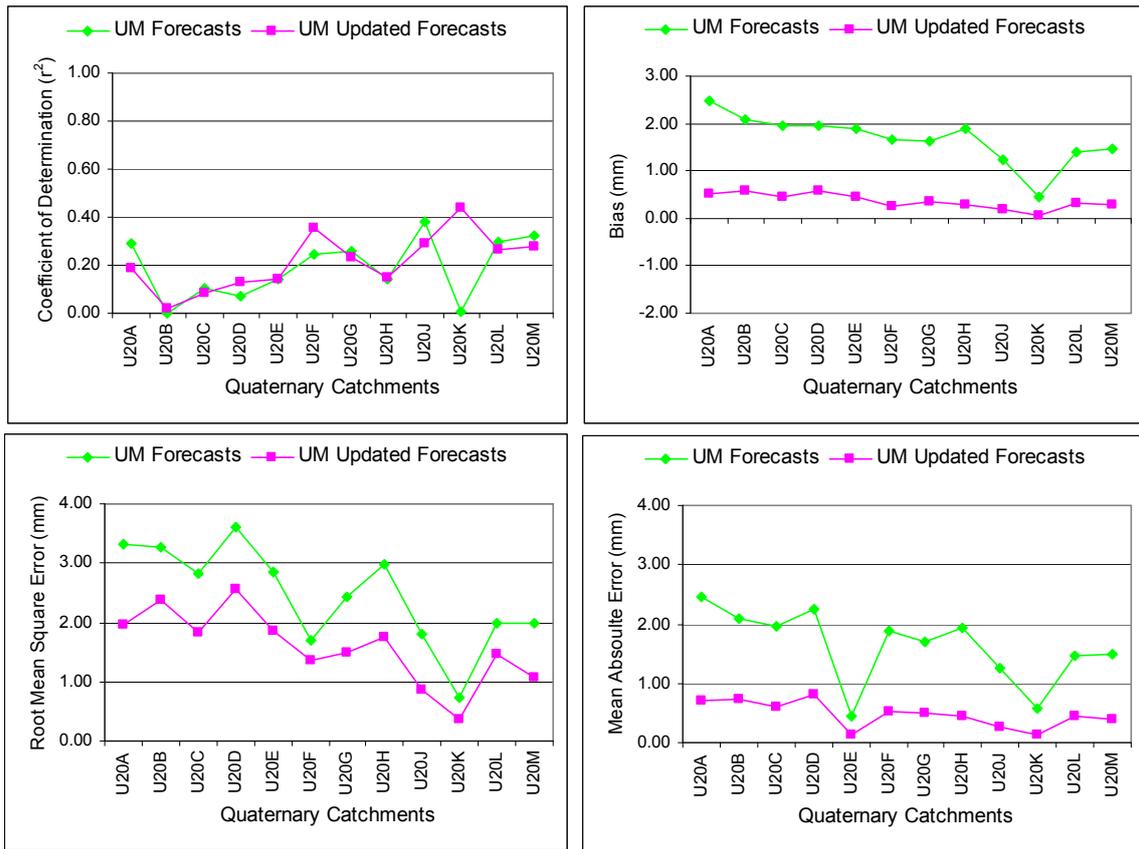
Figure 8.14   Coefficient of determination, bias, RMSE and MAE of accumulated Quaternary Catchment streamflows simulated with the *ACRU* model when using the UM 1 day lead time rainfall forecasts with and without updating for the period of 01 November 2006 to 31 January 2007

## 8.3.3   Evaluation of the NCEP-MRF Rainfall Forecasts

A number of ensemble based forecast products are being produced at the National Center for Environmental Prediction (NCEP) and distributed to a wide range of users both nationally in the USA and internationally (Toth *et al.,* 1997). The output of the Medium Range Forecast model (NCEP-MRF) has been used operationally for medium range forecasts in South Africa since 2003 (cf. Chapter 5, Sub-section 5.3.1). Unlike the C-CAM and UM models, the NCEP-MRF model uses the so-called Ensemble Forecasting System (EFS) to estimate the probability distribution of the "true state of the atmosphere" around the control analysis. The motivation for use of the EFS is that probabilistic forecasts initiated from slightly different initial states and model parameters provide better results than a single

deterministic forecast initiated by the best known state (Ahrens and Jaun, 2007). The current NCEP ensemble forecasts are generated every day, one with 22 members at a grid spacing of 2.5$^{\circ}$ x 2.5$^{\circ}$ and another with 60 members at a grid spacing of 1$^{\circ}$ x 1$^{\circ}$, both running for up to a14 day lead time. The latter product has recently been applied in South Africa by the SAWS and the historical archive of forecasts available for this study are the 2.5$^{\circ}$ grid spaced values. The evaluation of these coarse scaled forecasts at the Mgeni catchment is presented below.

Owing to the coarse resolution of the data (2.5$^{\circ}$), the Mgeni catchment is entirely contained within one grid box (Figure 8.15). The verification is undertaken against raingauge data by computing the average rainfall of all rainfall stations which fall inside the grid box. With a grid space of 2.5$^{\circ}$, only a crude representation of observed precipitation distribution could be achieved, especially in southern Africa where large scale rain bearing frontal systems are enhanced by local topography (Tennant *et al.,* 2006). Tennant *et al.* (2006) have attempted to verify the 2.5$^{\circ}$ grid spaced forecasts against SAWS station data by averaging the rainfall values of the stations within a grid box. Approximately 30 to 200 rainfall stations fall into each 2.5$^{\circ}$ grid box, with the lower station density found in the more arid western interior of South Africa. They found that the NCEP-MRF model over-estimates rainfall amounts by up to 300% over the summer rainfall areas of South Africa. This significant bias becomes greater for higher rainfall amounts. In contrast to the summer rainfall areas, rainfall is under-estimated in the winter rainfall areas of South Africa.

Nevertheless, the model performance has been continuously improving through upgraded model physics, resolution and data assimilation, and these effects are automatically manifest in the 2.5$^{\circ}$ grid spaced outputs (Tennant *et al.,* 2006). With further improvement and refinement, these forecasts have the potential to play an increasingly important role for large scale catchments in the short and medium range of an agrohydrological forecasting system. The present study is aimed at examining if the coarse resolution (2.5$^{\circ}$) is sufficient to resolve weather systems responsible for the summer rainfall over the Mgeni catchment. For operational use, however, these large scale forecasts should be downscaled to a finer resolution

based on the use of a statistical or dynamical rainfall downscaling model, which is beyond the scope of this study.
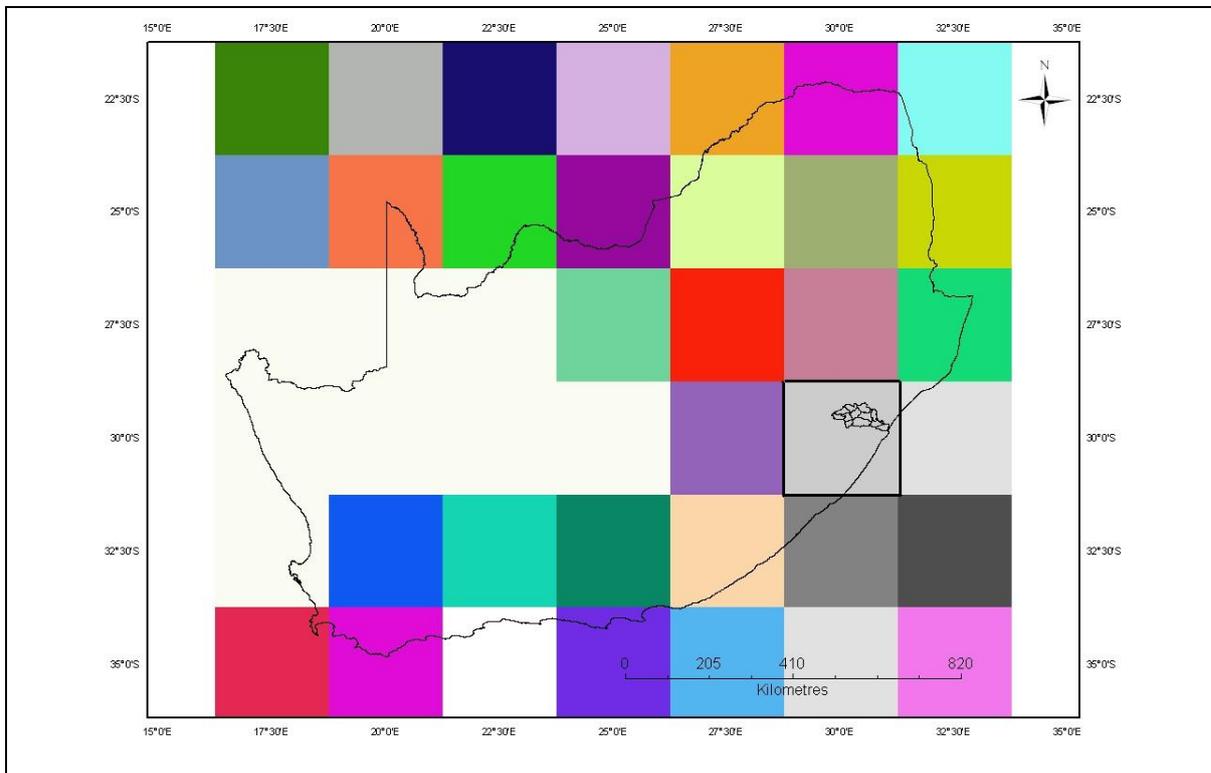


Figure 8.15    The 2.5º grid boxes over southern Africa for the NCEP-MRF forecasts, with the Mgeni catchment shown in its relevant grid box

The evaluation commenced with the investigation of ensemble members for the four selected rainfall events in order to assess the extent to which they could explain the uncertainty associated with a particular forecast. In Figure 8.16 the inter-quartile range of 24-hour accumulated precipitation amount is shown, plotted from 22 ensemble members for Events A, B and C, but from only 11 members for Event D. The spread of the ensemble describes the breadth of the range of forecasts made by the EFS. For a good ensemble forecast the "observed value lies somewhere within the range of the forecasts given by the ensemble members" (Ebert, 2001). In the case of Event A, the observation is significantly less than the driest ensemble member, even though the spread of the ensembles is large in ranging from 33.0 - 91.8 mm. This large spread suggests a lack of confidence in the forecast for that particular day. The spread for Event B is also large (6.4 - 67.0 mm), but the observation was captured within the lower inter-quartile range. In the

cases of Events C and D the spread was relatively small. The observed value for Event C lies within the lower inter-quartile range of the ensemble values, whereas in the case of Event D the observation is higher than the wettest ensemble member. The model's under-forecast for Event D is possibly due to the absence of 11 ensemble members of the 22.
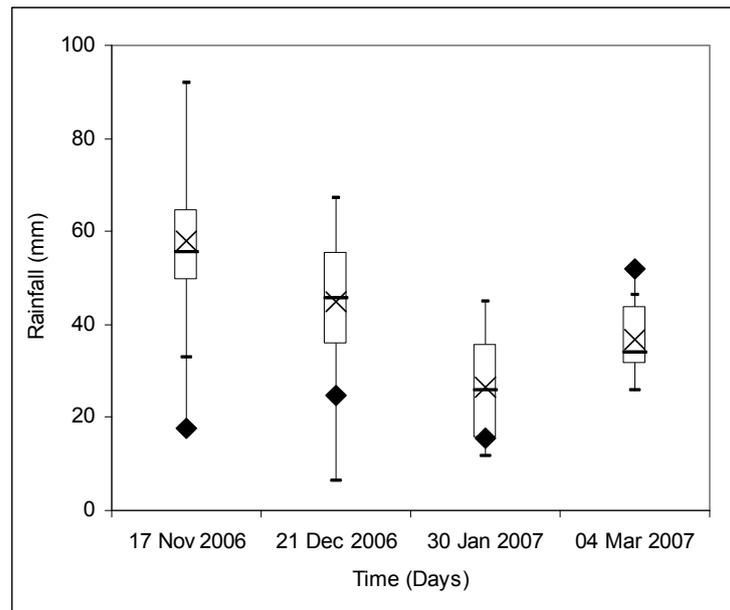


Figure 8.16    The spread of NCEP ensemble rainfall forecasts for 17 November 2006, 21 December 2006, 30 January 2007 and 04 March 2007. The box-and-whiskers represent the minimum, lower quartile, median, upper quartile and maximum values of the ensemble members. The x signs indicate the ensemble mean value. Diamonds represent the observed values

The most important benefit that the ensemble forecasts can offer is that they can be used to provide Probabilistic Quantitative Precipitation Forecasts (PQPFs), as is done in many centres (e.g. NCEP, SAWS). It has been shown by many researchers (e.g. Toth and Kalnay, 1997; Ebert, 2001; Zhu *et al.,* 2002)  that NCEP ensemble forecasts based on probabilistic values have the potential to provide a more meaningful indication not only for the temporal distribution, but also of possible spatial distributions of rainfall in the short and medium range forecasts. PQPFs are computed by counting how many of the ensemble members exceed a daily accumulated rainfall, or any given threshold, and then dividing that number

by the total number (in this case 22) of ensemble forecasts (Toth *et al.,* 1997). In future the same procedure can also be followed to generate probabilistic ensembles of streamflow forecasts by ingesting each of the ensemble rainfall forecasts into a hydrological model, provided that the spatial scale of these forecasts are comparable to those for which the hydrological model is applied.

In this study it is hypothesised that the most likely spatial representation of the rain field is given by the ensemble mean. In order to assess the extent to which the lead times of the NCEP forecasts are skilful, the ensemble means for each of 5 day lead time forecasts at the 30ºS 30ºE grid box for the period of 01 November 2006 to 31 January 2007 were compared against average rainfall values of all stations that fall into the box (Figure 8.17). It was found that the 1- 5 day forecasts show very similar patterns throughout the study period, although the quality of the forecasts increases with decreasing lead time, as expected. Nonetheless, the NCEP-MRF model showed a tendency to over-forecast throughout the study period.
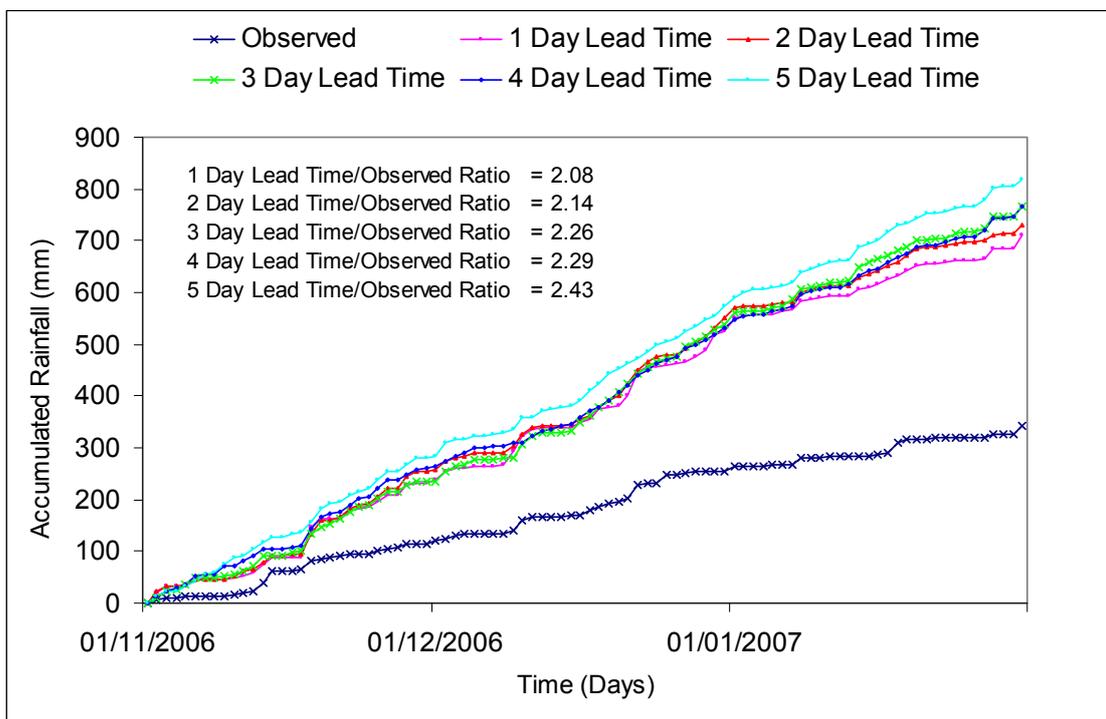


Figure 8.17    Time series comparisons in the 30ºS 30ºE grid box of accumulated rainfalls of 5 day forecasts derived from the NCEP-MRF rainfall model for the period of 01 November 2006 to 31 January 2007

A plot of 1 day forecasts, which are relatively more skilful than the 2 to 5 day ones, versus observed rainfalls reveals a positively biased performance of the NCEP model (Figure 8.18). The association is quite strong for less extreme events of < 20 mm per day, while there is more scatter with higher rainfalls. Taking the space scale limitation into account, however, the model's performance is considered satisfactory. These results reflect that the NCEP-MRF model is capable of identifying a rainfall event, but with a tendency to over-estimate the amount.
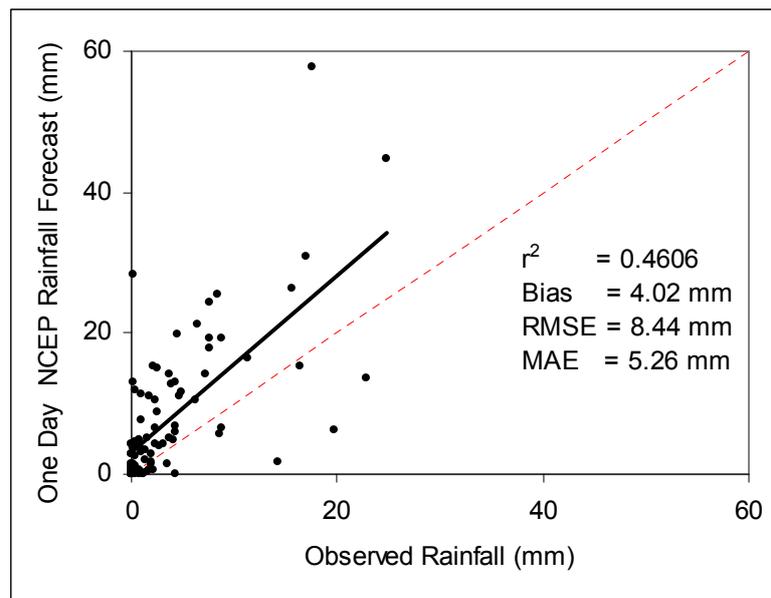


Figure 8.18    Scattergram of NCEP simulated 1 day forecasts from the means of 22 ensembles versus observed rainfalls for the period of 01 November 2006 to 31 January 2007

## 8.3.4   Combined Use of the C-CAM and UM Rainfall Forecasts

All NWP models predicting weather at shorter ranges, or its various statistics at a longer time ranges, are based upon the same laws of physics (Toth *et al.,* 2006). However, the quality of the forecasts is often constrained by the model formation through the variation of assumptions and approximations as to how the physical processes are parameterised in the models, as well as by their the levels of vertical and horizontal resolutions, forecast methodologies and data assimilation methods (Ebert, 2001; Anstee, 2004). As a result, no two models will display the same forecasts in exactly the same manner. Generally, different models will tend

to "cluster" around the perfect forecast, with some a little too wet while others are a little too dry (Ebert, 2001). In this instance, the C-CAM and UM models responded differently for the same season and the same area. Although they displayed similar patterns to the reference run, the C-CAM model showed a tendency to under-forecast whereas the UM model tended to over-forecast throughout the study period. This is particularly noticeable on the occasions of heavy rains (Figure 8.19).

The daily QPFs of the two models for the period of 01 November 2006 to 31 January 2007 were combined by "weighted averaging" in order to evaluate the extent to which their combined prediction could improve the accuracy of the forecasts. No particular model was favoured and the success of the "weighted averaging" to produce a better combined QPF is dependent on the performance of the models relative to each other on a given day. It was found that the combined forecast was influenced more by the outputs of the UM model than the outputs of the C-CAM model. As a result, the combined output was superior in relation to the UM than the C-CAM forecasts, both in terms of individual daily and accumulated flows (Figures 8.19 and 8.20). The under-estimation in the total streamflow forecasts of the C-CAM forecasts was reduced from 34% to 10%, while the over-estimation in the UM Model was decreased from 291% to 89%.

The advantage of using multiple models to determine rainfall is the ability to estimate the probability of receiving rain (Ebert, 2001). For example, if the C-CAM and UM models both predict that at least 10 mm would fall at a particular location, then the probability of receiving at least 10 mm will be 2/2, or 100%. If there is disagreement the chance will be 1/2, or 50%. Likewise, the probability of streamflow exceeding a given threshold can also be calculated and mapped at catchment or national scale. The greater the number of NWP forecasts the greater the skill will be of the probabilistic forecasts. Decision makers can then have more confidence in such probabilistic forecasts than any of the individual deterministic estimates.
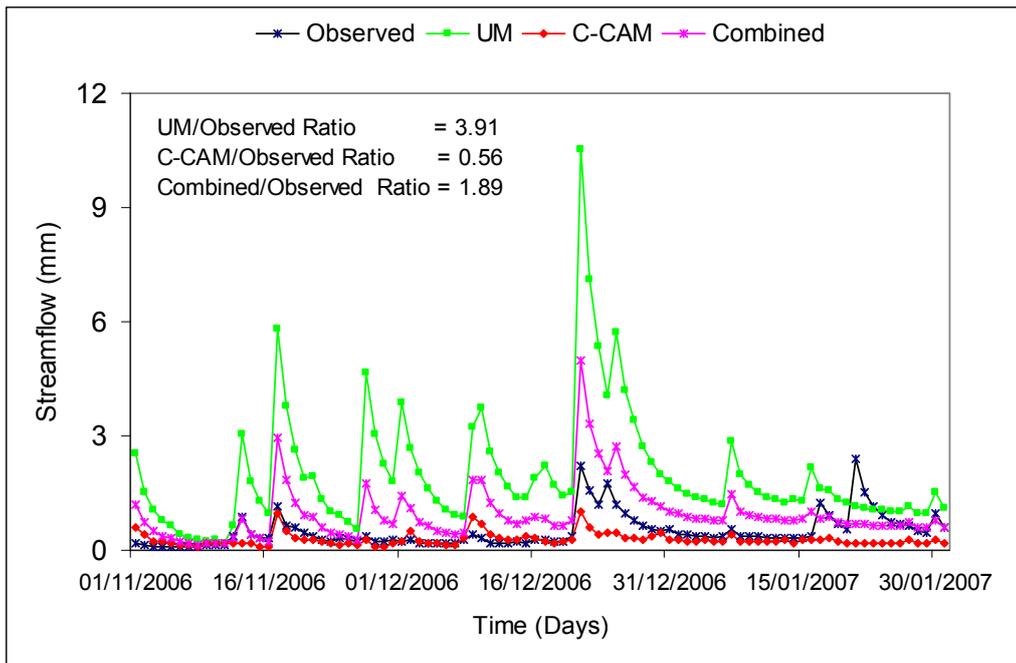
Figure 8.19    Time series comparisons of daily streamflows at the mouth of the Mgeni catchment, derived from the C-CAM, UM and combined rainfall forecasts for the period of 01 November 2006 to 31 January 2007
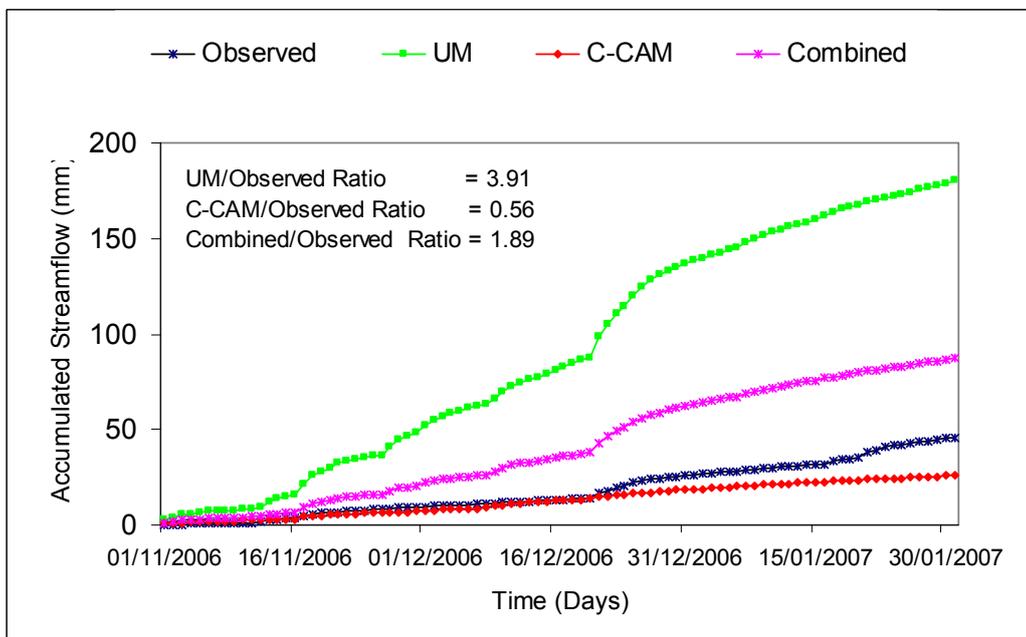


Figure 8.20    Comparisons of accumulated streamflows at the mouth of the Mgeni catchment, derived from the C-CAM, UM and combined rainfall forecasts for the period of 01 November 2006 to 31 January 2007

## 8.4    Conclusions

In this chapter three experimental NWP models were evaluated from a hydrological perspective, *viz.* the C-CAM, UM and NCEP-MRF. The results obtained from each model are encouraging. However, the quality of the results varied between the three models, between the modes of simulation (i.e. with and without updating) and between the selected events. Each model was evaluated separately for four selected days with relatively high amounts of rainfall, as well as for a continuous period of 92 days.

For the four selected events, the C-CAM model scored some skill with the 1 and 2 day lead time rainfall forecasts, whereas for the 3 and 4 day forecasts the skill was low and unreliable.  Except for Event D (04 March 2007), there was no significant difference between the 1 and 2 day lead time UM rainfall forecasts. The 2 day forecast was slightly superior to the 1 day forecast for Event C (30 January 2007), but for Event D the 2 day forecast showed no skill. Results obtained for the continuous period showed that streamflow forecasts based on the C-CAM model suffered from consistent under-estimation, while conversely the UM based streamflow forecasts suffered from consistent over-estimation. Since the degree of over-estimation by the UM model was more significant than the degree of under-estimation of the C-CAM model, their combined output was positively biased. However, considerable improvement was achieved in their individual streamflow forecasts when the state variables of the catchment were updated at the start of each day with observed rainfalls up to the previous day.

The NCEP-MRF rainfall forecasts were verified only against observed rainfalls owing to spatial scale differences. It was shown that these forecasts over-predicted those of the observed values for both the selected single events (except for Event D) and over the continuous period of time, although the quality of the forecasts increased slightly with decreasing lead time. The ensemble approach was effective for Events B (21 December 2006) and C (30 January 2007), but failed to capture Events A (17 November 2006) and D (04 March 2007). Despite the limitations of the coarse spatial scale, the correlation between the 1 day forecast and the reference was fair.

In conclusion, when taking into account the discrepancies between the forecast period (02:00 to 02:00) and observed period (08:00 to 08:00), scale issues and uncertainties in the reference run, the performances of the three models seem to be reasonable. The occurrences of the rainfall were signaled correctly over most of the study period, especially by the C-CAM and UM models, but with a tendency to respectively under- and over-estimate the correct amount.

The results obtained from this research reveal that there is still room for improvements in each of these models, especially in making the models' spatial scales more compatible with requirements of hydrological models for application in small and medium sized catchments and in improving the rainfall forecast skill, especially for longer lead times.

In Chapter 7 an attempt has made to evaluate the near real time ground based and remotely sensed rainfall fields from a hydrological perspective, while in this chapter the evaluation was extended for the rainfall forecasts generated from the C-CAM, UM and NCEP-MRF models. The next two chapters cover, respectively, the evaluations of the two temporal downscaling methodologies, *viz*. the *Historical Sequence Method* and the *Ensemble Re-ordering Based Method*.

# 9. TEMPORAL DOWNSCALING OF PROBABLISTIC CATEGORICAL FORECASTS USING THE HISTORICAL SEQUENCE METHOD

In South Africa, monthly and seasonal forecasts of rainfall are presented as a probability of occurrence within categories designated as above, near or below normal, i.e. as three-category tercile forecasts. These forecasts have been shown to possess certain level of skills when they are compared against observations (Klopper and Landman, 2003). However, such forecasts are not necessarily directly usable for many types of decisions required in the agricultural and water sectors. In order estimate the consequences of such seasonal climate forecasts, hydrological and crop yield models generally require quantitative *daily* time series climate information. This has necessitated the development of a methodology of translating categorical rainfall forecasts into quantitative daily values to be used as input into daily time step hydrological/crop yield models, thus allowing for the generation of agrohydrological forecasts (e.g. of streamflows or reservoir levels, or of irrigation demands or crop yield estimates) which are then suitable for operational use in water resources and agricultural decisions.

As was described briefly in Section 5.4 of Chapter 5, two approaches, *viz.* the *Historical Sequence Method* and the *Ensemble Re-ordering Based Method*, have been developed in this study in order to generate daily rainfall values from preferentially selected analogue years, for subsequent use as input into hydrological/crop yield models. If these approaches are to be applicable with a high degree of confidence, they should be evaluated in various parts of southern Africa. The Mgeni catchment in KwaZulu-Natal is used in this study to demonstrate the potential usefulness of the two approaches. The purpose of this chapter is to evaluate the performance of the streamflow simulations of the *ACRU* model derived using the *Historical Sequence Method*. Evaluation of the *Ensemble Re-ordering Based Method* is presented in Chapter 10.

## 9.1 Datasets Used in this Study

The categorical seasonal rainfall forecasts been used in this study were obtained from the South African Weather Service (SAWS). The SAWS has been producing seasonal rainfall forecasts in 3 equi-probable categories of below normal, near

normal and above normal rainfalls for one month and for any consecutive 3 months, i.e. seasonal. In this chapter the focus is on the seasonal forecasts and investigations are undertaken on the skill of retrospective forecasts for the three consecutive month blocks of October-November-December (OND), November-December-January (NDJ), December-January-February (DJF) and January-February-March (JFM) for the period October 2003 to March 2006. The October to March season was chosen as it makes up the main rainfall months in this southern hemisphere summer rainfall region. The forecasts were produced at the beginning of each season, referred to as a "zero" month lead-time. Categorical probabilities of seasonal forecasts issued by the SAWS over the Mgeni catchment for the selected seasons are shown in Table 9.1, and they cover the range from dry-average, to average-dry to average-wet.

Table 9.1   Categorical probabilities of seasonal forecasts issued by the SAWS over the Mgeni catchment for the main rainfall periods of 2003/04 – 2005/06, and qualitative descriptions of the forecasts

| Rainfall Year | 3 Months Period | Categorical Probabilities (%) | | | Qualitative Description |
|---|---|---|---|---|---|
| | | Above Normal | Near Normal | Below Normal | |
| 2003/04 | OND | 20 | 45 | 35 | Average to Dry |
| | NDJ | 25 | 45 | 30 | Average to Dry |
| | DJF | 25 | 35 | 40 | Dry to Average |
| | JFM | 25 | 35 | 40 | Dry to Average |
| 2004/05 | OND | 25 | 35 | 40 | Dry to Average |
| | NDJ | 25 | 45 | 30 | Average  to Dry |
| | DJF | 15 | 40 | 45 | Dry to Average |
| | JFM | 25 | 40 | 35 | Average to Dry |
| 2005/06 | OND | 35 | 40 | 25 | Average to Wet |
| | NDJ | 35 | 40 | 25 | Average to Wet |
| | DJF | 25 | 45 | 30 | Average to Dry |
| | JFM | 30 | 45 | 25 | Average to Wet |

## 9.2   Methodology

Wide ranging spatial variability within the Mgeni catchment with regard to its meso-climate, soils and land uses made it necessary to apply the *ACRU* simulations in distributed catchment mode in order to also simulate accumulated streamflows from subcatchments cascading downstream at the exit of each QC. The *ACRU*

model was run with historical observed daily rainfall from year 2000 to the beginning of the forecast period in order to create representative antecedent conditions and initial stores (e.g. soil moisture status in the top- and subsoil, the baseflow store and baseflow releases). The steps described in Section 5.4 of Chapter 5, were applied for the 12 QCs that make up the Mgeni catchment for the three month periods OND, NDJ, DJF and JFM from October 2003 to March 2006. For each QC and the selected seasons, 20 daily rainfall sets are re-sampled from the same dates of the preferentially selected 20 analogue years. The 20 daily rainfall sets extracted for each QC for the selected seasons were then applied, together with the antecedent conditions generated previously by the *ACRU* model, to generate 20 daily streamflow ensembles for each forecast season.

Before discussing the results, it is useful to know the forecast format types and the evaluation methods when seasonal forecasts are considered. Seasonal forecasts are often displayed either in deterministic or probabilistic (continuous or categorical) forms. A deterministic format contains no expression of uncertainty, while probabilistic forecasts convey the uncertainty associated with the forecasts in a quantitative way (Zhang and Casey, 1999; Klopper and Landman, 2003). Seasonal forecasts are often provided in probabilistic format. However, when users need to make a "yes" or "no" type decision under uncertain situations, conversion from probabilistic to deterministic format becomes vital. For this reason the output of the *ACRU* model, derived using the *Historical Sequence Method* from seasonal rainfall forecasts, was evaluated in both continuous and categorical formats against corresponding baseline land cover streamflows generated by the *ACRU* model from gauged rainfall.

## 9.3    Results and Discussion

The 20 generated daily ensembles of streamflow forecasts are first expressed as three equi-probable categories of below normal, near normal and above normal conditions for the selected three rainfall seasons from October 2003/March 2004 to October 2005/March 2006. In this study, three category tercile classes of streamflow were designated for each QC and each season by using daily streamflow values generated by the *ACRU* model for the period of 1950 to 2003 (Figures 9.1 and 9.2).
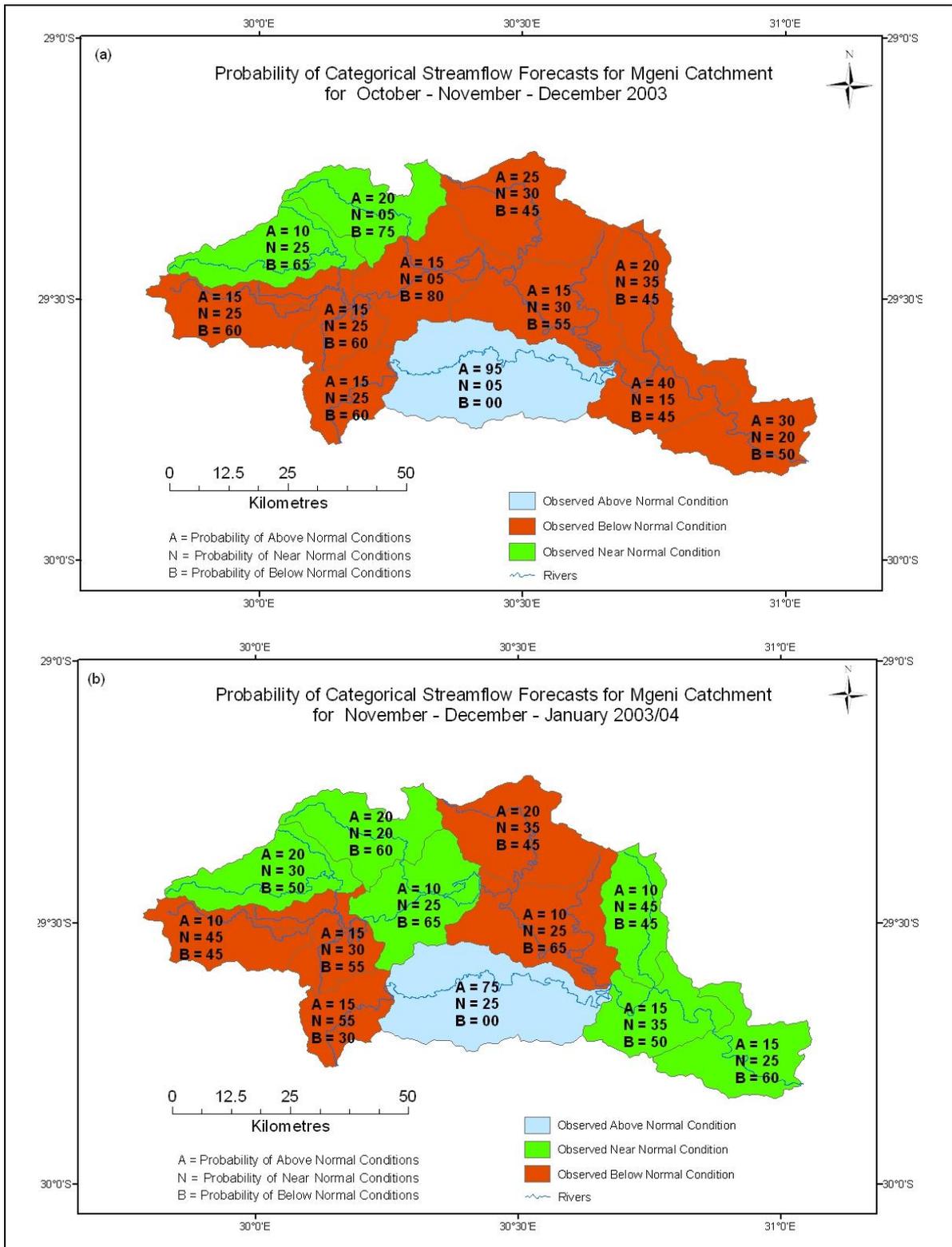
Figure 9.1    Probabilities of categorical streamflow forecasts for Quaternary Catchments making up the Mgeni catchment for OND of 2003 (a) and NDJ of 2003/04 (b), with the red shading representing observed below normal conditions, and the shades of green and blue representing near and above normal conditions, respectively
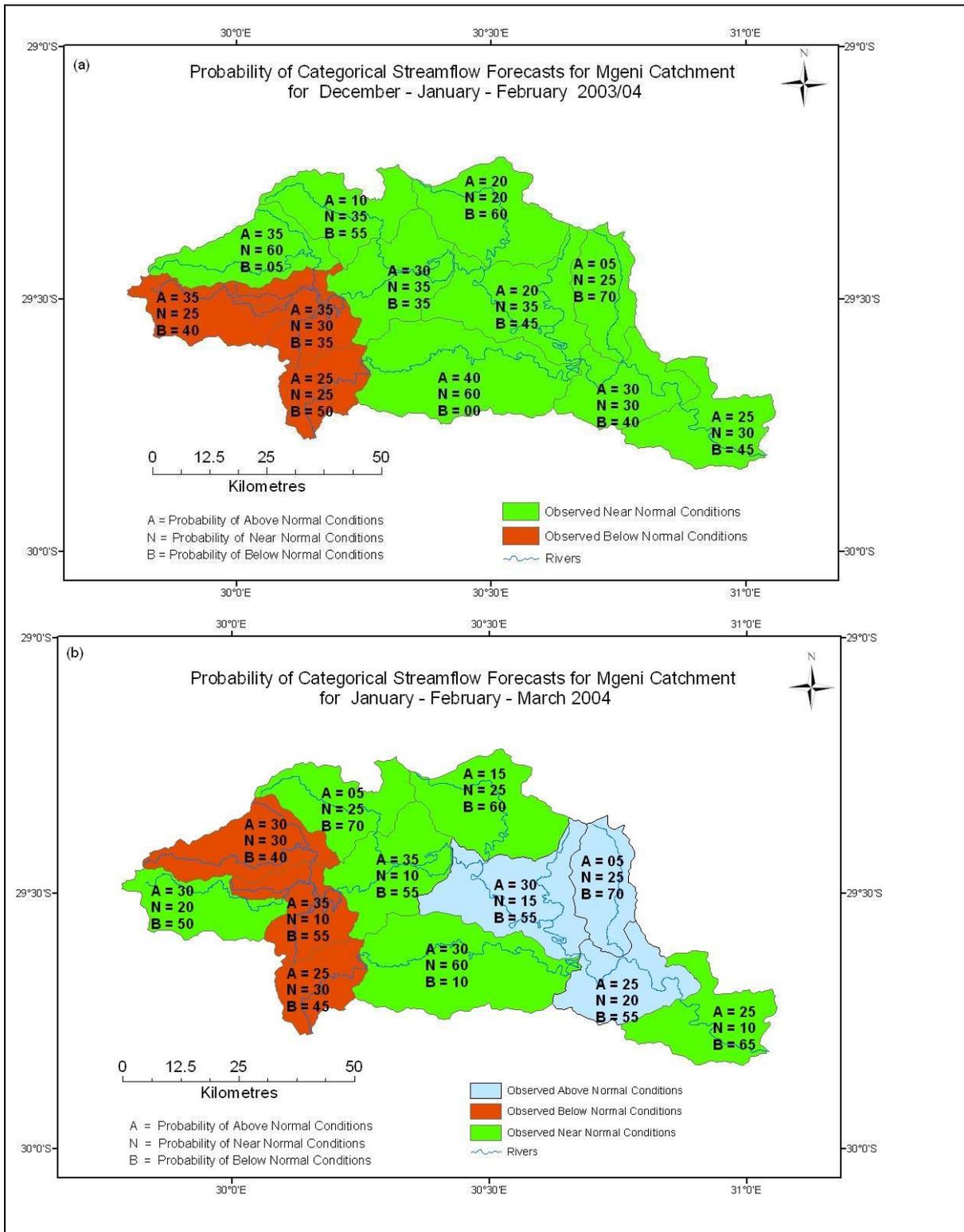
Figure 9.2    Probability of categorical streamflow forecasts for Quaternary Catchments making up the Mgeni catchment for DJF of 2003/04 (a) and JFM of 2004 (b), with the red shading representing observed below normal conditions, and the shades of green and blue representing near and above normal conditions, respectively

The 54 years of seasonal streamflow accumulations were then ranked from the lowest to highest value for the season under consideration. If the observed streamflow falls below the 33rd percentile it is considered below normal, while above the 66th percentile is considered to be above normal, and in between the two it is considered to be near normal. The future probability of the streamflow falling in each of the three categories is computed by dividing the number of ensembles falling within each category by the total number of ensembles, i.e. 20.

The Revised True Skill Statistics, RTSS (Equation 4.10), the Linear Errors in Probability Space, LEPS (Equation 4.12) and the ranked probability skill score, RPSS (Equation 4.14) were used to asses the skill of the categorical probabilistic forecasts. Often a 2 x 2 contingency table is constructed to calculate the ratio of hits and misses with respect to the possible totals. In this study a 3 x 2 contingency table, introduced by Zhang and Casey (1999), was used for transforming categorical probabilistic forecasts into binary events (i.e. yes, no and non-applicable) based on decision probability thresholds. Users may set different probability thresholds between forecasts and their random expectations for various purposes, and are expected to take protective action when the forecast probability of the event exceeds the specified threshold values. Zhang and Casey (1999) have defined *(1/m)/m* as a significant departure from the normal conditions for m-categorical forecasts. A "No" forecast is assigned if the forecasted probability is less than ($1/m - 1/m^2$), and a "Yes" forecast if it is equal or greater than ($1/m + 1/m^2$). If the forecasted probability is between ($1/m - 1/m^2$) and ($1/m + 1/m^2$) it is considered a "Non-applicable" forecast, as there is no significant shift in probability from the random expectation of each category. In this respect, for "Yes" forecasts the probability threshold is 44.4% and for "No" forecasts it is 22.2%. Figure 9.3 shows the RTSS score for each of the 12 Quaternary Catchments making up the Mgeni system for the selected three seasons for the period 2003/04 to 2005/06.

The RTSS score should be interpreted as a percentage improvement over the climatological probabilities (random chances). Overall, the RTSS scores demonstrate that the streamflow forecast on each of the Quaternary Catchments performed much better than would be expected by random chance. However, below zero scores were observed for DJF and JFM at QCs U20C and U20M

respectively (Figure 9.3), indicating worse probabilities than by random chance. It is important to bear in mind that the RTSS score does not penalise the errors in terms of their severity between each categories. In addition, part of the information from probability forecasts will be lost during the transformation to binary forms (Zhang and Casey, 1999).
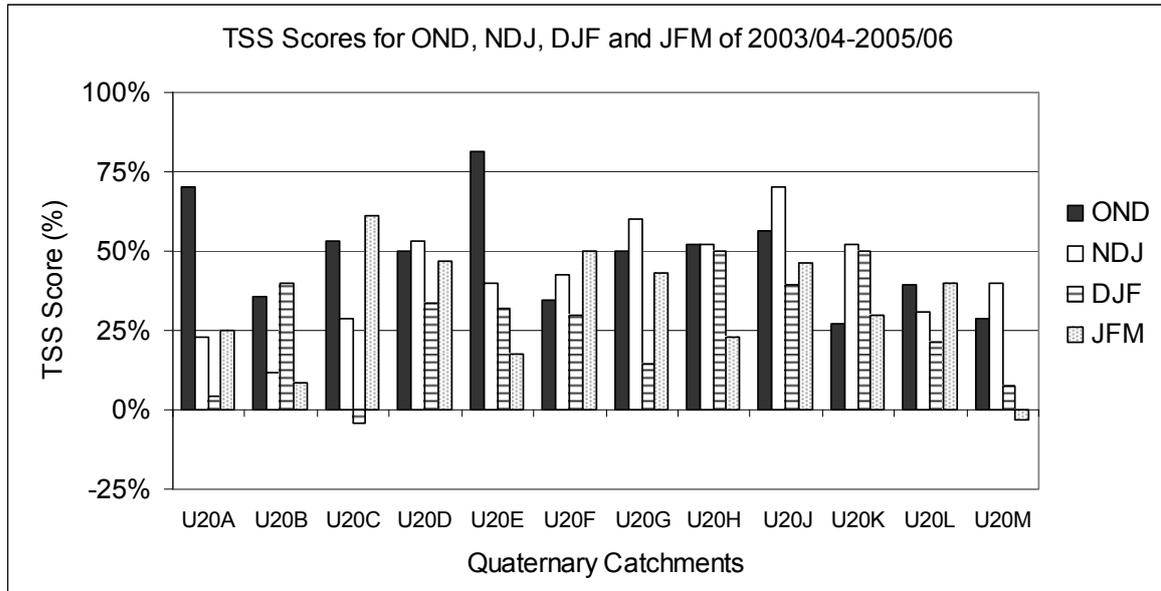


Figure 9.3    RTSS scores for retroactive Quaternary Catchment accumulated streamflow forecasts on the Mgeni catchment for OND, NDJ, DJF and JFM of 2003/04 - 2005/06

Considering the high risk that poor hydrological forecasting can have in water resource management, a skill test should be employed that heavily penalises a forecast that is two categories in error rather than only one category in error.  For this reason, the LEPS (Potts *et al.,* 1996) and RPSS techniques (Wilks, 1995) are used. In this contest, the reference forecast was climatology, with an expected hit rate of 33.3% for each of the below normal, near normal and above normal categories. The LEPS and RPS scores over the retroactive forecast period are depicted in Figures 9.4 and 9.5.

Assuming that a LEPS score of 10% is statistically significant (Letcher *et al.,* 2004), the OND season showed relatively high scores over the remaining three month periods, with nine out of 12 QCs scoring above 10%, but one below zero at

QC U20B (-6.4%). For the NDJ period, the scores are satisfactory, with five QCs scoring above 10%, two QCs scoring between 5-10%, and five QCs scoring below zero. Results for DJF and JFM periods show poor forecasting, with more than half of the QCs scoring below zero. Comparing LEPS scores on a basin scale, the highest score was achieved in U20J, with all seasons scoring above 17% and the lowest score being at the mouth of the Mgeni catchment (U20M), in which three periods scored below -25% (Figure 9.4).

The RPSS has similar magnitudes to the LEPS scores. Among the three seasons, the OND scored the highest RPSS, while the DJF and JFM had considerable negative skill scores (Figure 9.5). However, it is important to bear in mind that the RPSS is highly sensitive when applied to short data sets (Nurmi, 2003). In addition, the RPSS measures how well the probability forecast predicted the category that the observation fell into, and not how serious the errors was.
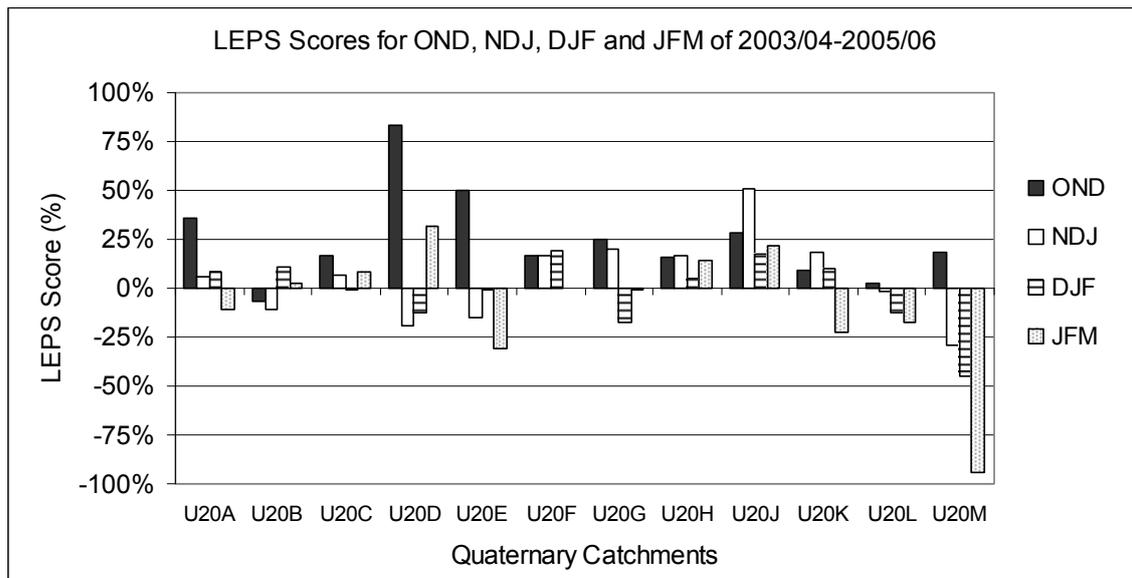


Figure 9.4    LEPS scores for retroactive Quaternary Catchment accumulated streamflow forecasts on the Mgeni catchment for OND, NDJ, DJF and JFM of 2003/04 - 2005/06
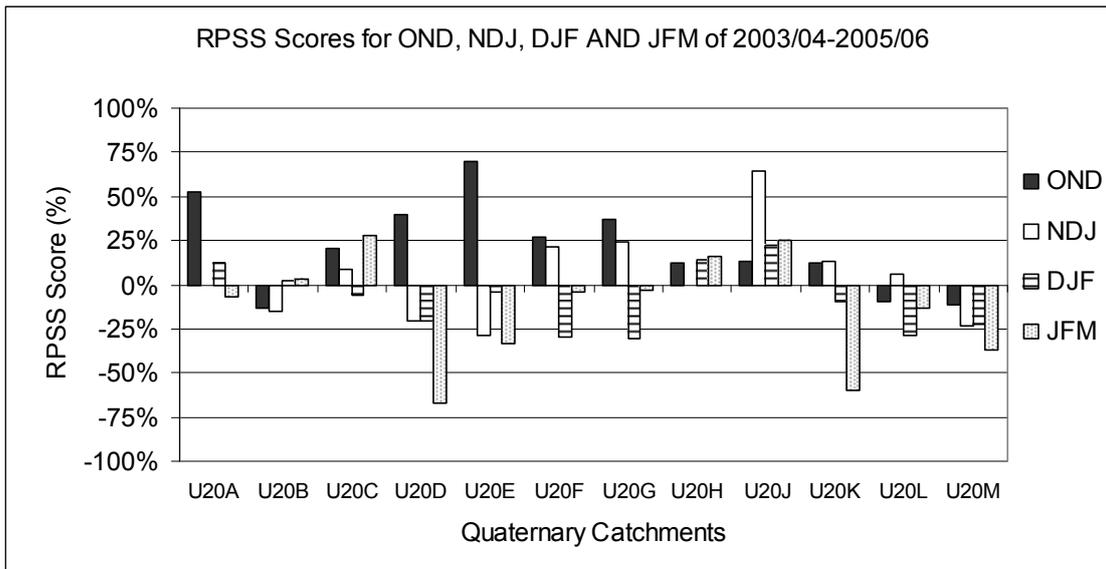
**Figure 9.5** RPSSs for retroactive Quaternary Catchment accumulated streamflow forecasts on the Mgeni catchment for OND, NDJ, DJF and JFM of 2003/04 - 2005/06

Time series streamflow forecasts are another important tool for water management decisions. One major advantage of producing a time sequence of streamflows during the forecast period is the ability to conduct risk-based scenario analyses. With a different probability threshold, it can provide important information to, for example, reservoir managers. For drought contingency analysis, for example, worst-case forecasts can be made by taking only the low flow scenarios. However, the set of daily ensembles of streamflows generated using the *Historical Sequence Method* should not be interpreted as an exact sequence of daily rainfall; rather, the forecast should interpreted in terms of summary statistics such as the mean, standard deviation, skewness and coefficient of variation in order to provide information on the probable total flows, be they high or low, during the target season. The reason for this is that the timing of individual rainfall events within a forecast season is unknown.

For reason of space, only ensembles of streamflows for the OND, NDJ, DJF and JFM of 2003/04 are given in this chapter. Figure 9.6 shows box-and-whisker plots of the generated mean (a), standard deviation (b), skewness (c) and coefficient of variation (d) from 20 ensembles of streamflows for the OND, NDJ, DJF and JFM of the 2003-2004 rainy season at the mouth of the Mgeni catchment. A box-and-

173

whisker in each box plot indicates the lower extreme, lower quartile (the 25th percentile), median (i.e. the line across the box), mean (x sign) upper quartile (the 75th percentile) and upper extreme of the forecasted streamflow sequences. The observed mean, standard deviation, skewness coefficient and coefficient of variation are depicted in Figure 9.6 as diamonds, with circles indicating the values outside of the simulated range.
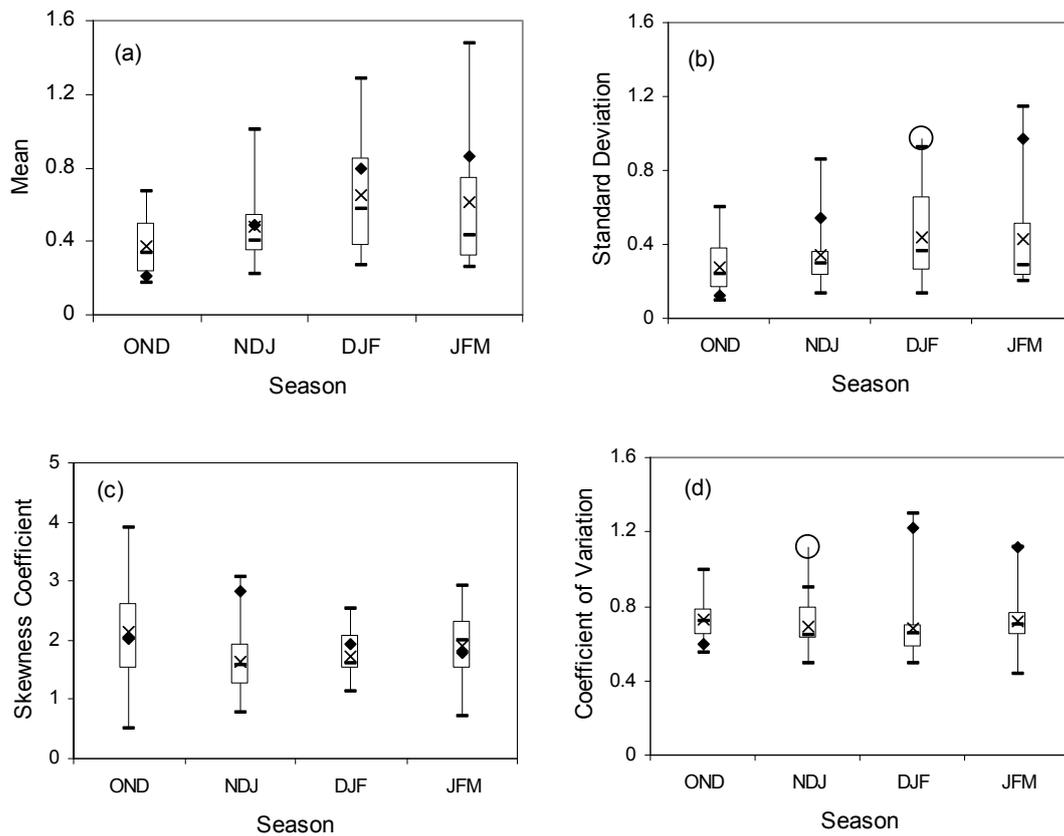


Figure 9.6    Box plots of statistics of generated streamflows along with the observed values at the mouth of the Mgeni catchment for OND, NDJ, DJF and JFM of 2003/04. The box-and-whiskers represent the minimum, lower quartile, median, upper quartile and maximum of the forecasted streamflow sequences, while the x signs represent the simulated mean values. Diamonds represent the observed values, with circles indicating the values outside the simulated range

The box plots show the centre, the spread and the overall range of distribution in the statistics from the 20 simulations and also show the capability of the *Historical Sequence Method* to simulate the observed statistics. The statistical moments shown in Figure 9.6 generally illustrate that the observed streamflow statistics are

well simulated within the range of the four quartiles, except for NDJ and DJF where, respectively, the coefficient of variation and the standard deviation exceeded the upper extremes. The observed means and skewness coefficients are reproduced well by the simulations. However, except for OND, the standard deviations and coefficients of variation are slightly under-estimated, suggesting that the observed dispersion (spread) is not as well simulated as desired.

To extend the analysis, ensembles of accumulated daily flows were computed at the Mgeni catchment outlet from the beginning of each season (Figure 9.7).
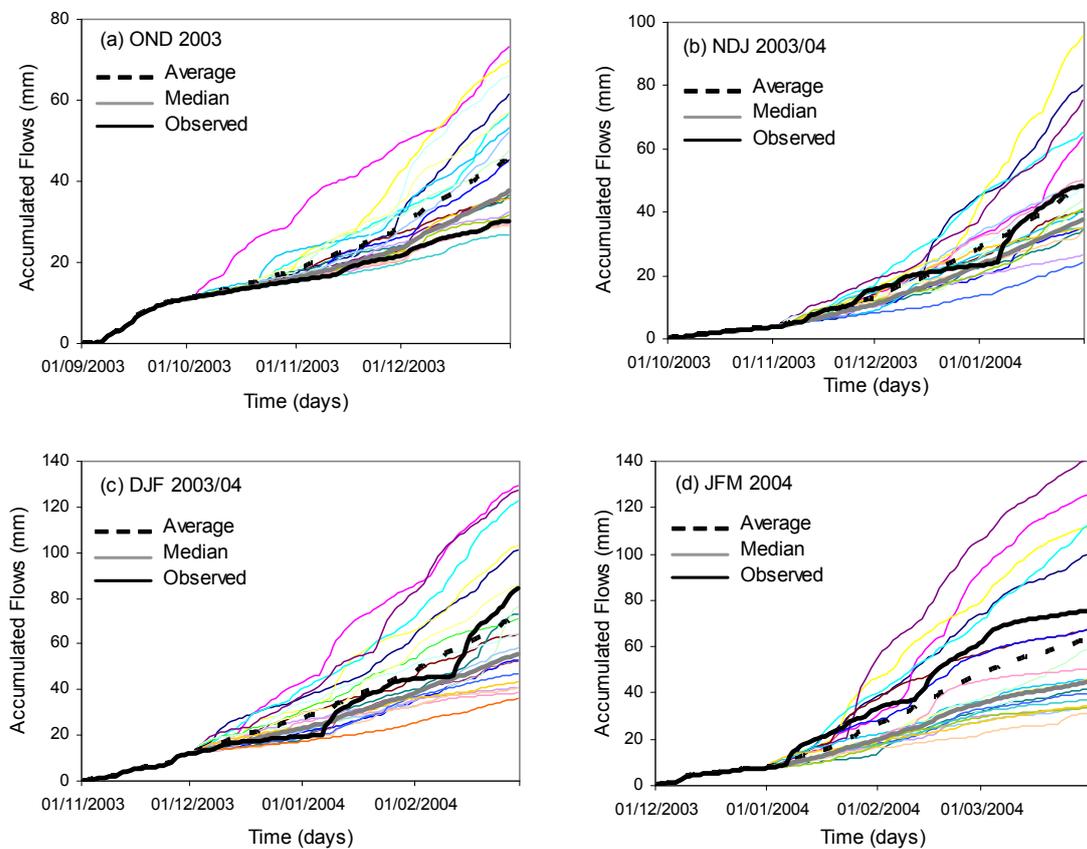


Figure 9.7    Forecasted versus observed seasonal accumulated flows at the mouth of the Mgeni catchment for OND (a), NDJ (b), DJF (c) and JFM (d) of 2003/04

From these ensembles of forecasts, the mean and median are computed as the "best" estimates that can be used to help agricultural and water managers in their decision making. Visually, the average accumulated streamflow values are much closer to the actual accumulated streamflow values than the median accumulated values for the selected seasons, except for the OND season.

In order to assess if the forecasted accumulated flows are statistically acceptable, a plot of average versus observed accumulated daily flows for the four selected seasons was performed, as shown in Figure 9.8. The coefficient of determination ($r^2$), bias, Relative Mean Square Error (RMSE) and Mean Absolute Error (MAE) were computed in order to assess the accuracy of averaged cumulative forecasted flows against their corresponding cumulative observed flows.
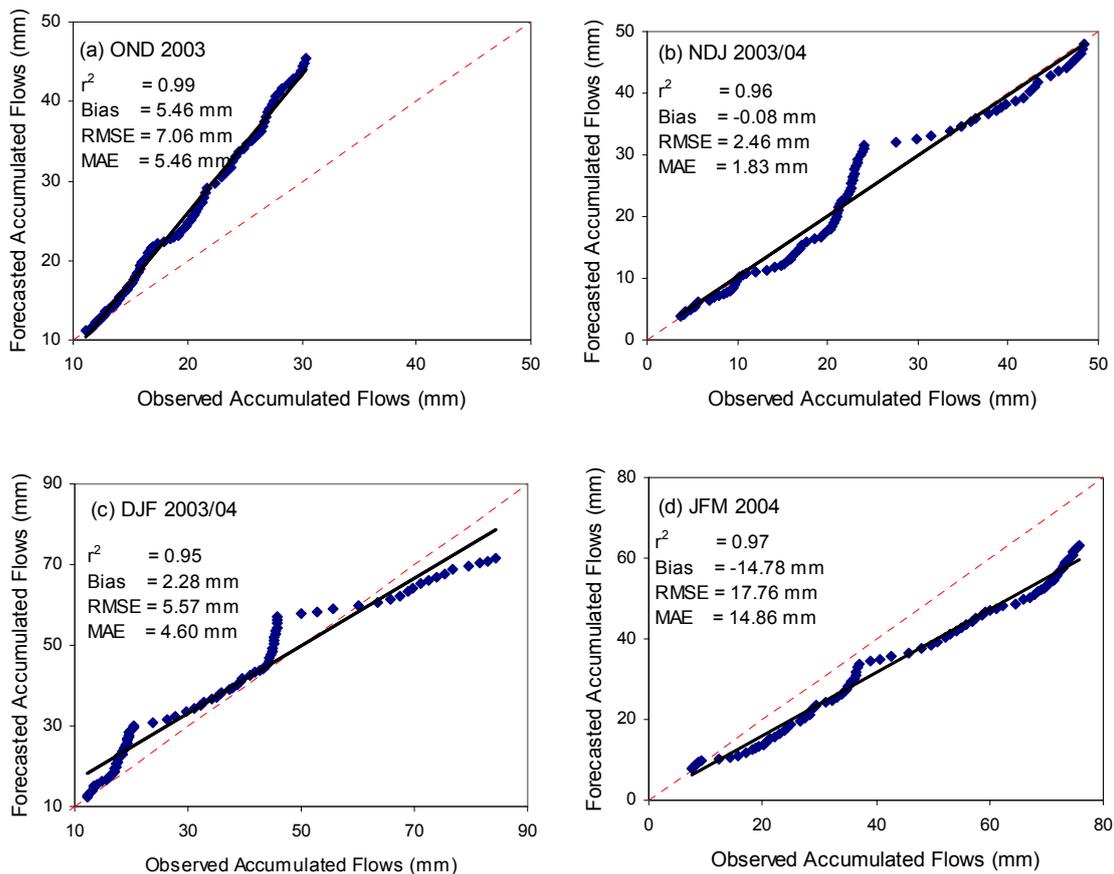


Figure 9.8    Plots of accumulated daily forecasted flows simulated with the *ACRU* model versus simulated flows from observed rainfall data at the mouth of the Mgeni catchment for OND (a), NDJ (b), DJF (c) and JFM (d) of 2003/04

Figure 9.8 reveals a very high performance of the *Historical Sequence Method*, with $r^2$ > 0.95 in all selected seasons. The coefficient of determination is especially strong for OND (0.99), but appears to be slightly biased (5.49), indicating a slight over-estimation over the accumulated observed flows. For NDJ, the MAE (1.83) and bias (-0.08) are small, revealing that the accumulated average streamflows have mimicked the corresponding observed accumulated flows excellently. Statistics for DJF also show good relationships between the forecasted and observed, albeit with a slight systematic bias. Although the coefficient of determination is high for the JFM period, substantial biases, RMSEs and MAEs were found, indicating significant systematic error in the forecast. The negative bias indicates that the averaged cumulative forecasts were consistently below their corresponding cumulative observed flows.

Another important technique is the cumulative density function (CDF) that helps to visualise the cumulative probability distribution (Figure 9.9). The cumulative probability is constructed from the generated accumulated average streamflows from forecasts and the simulated accumulated streamflow values from observed rainfall in Figure 9.8.

As is expected from the results in Figure 9.8, the forecasted cumulative probability for NDJ mirrored the corresponding the cumulative probability of observed accumulated streamflows well. For OND and DJF, the cumulative probability distribution is biased slightly on the wetter side, while the cumulative probability distribution for JFM is significantly drier than the corresponding observed cumulative distribution, especially for the higher streamflow values. As may be seen in Figure 9.9 (d), for accumulated streamflows of less than or equal to 50 mm, the cumulative probability is 84% with the rainfall forecasts, but with the same cumulative probability there is a chance of getting less than or equal to 65 mm of accumulated streamflows from the rainfall observation.
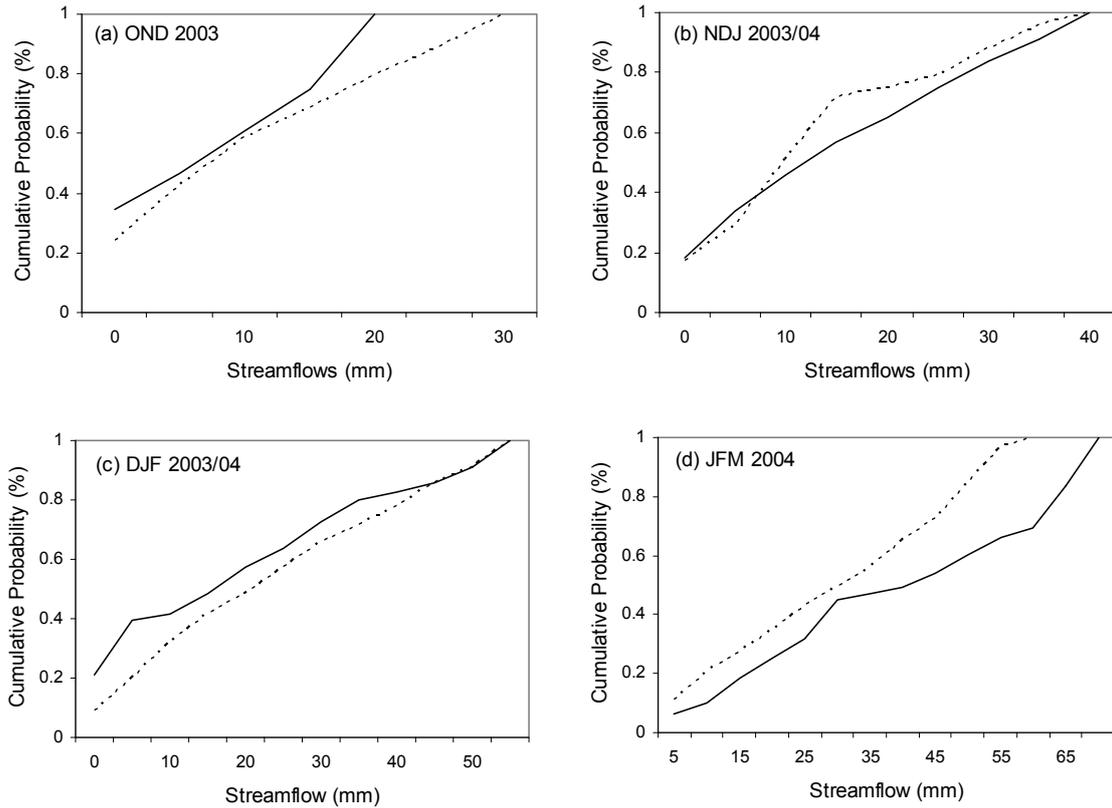
Figure 9.9    Cumulative probability of daily flows simulated with the *ACRU* model versus observed flows at the mouth of the Mgeni catchment for OND (a), NDJ (b), DJF (c) and JFM (d) of 2003/04, with the dashed line representing the cumulative probability of forecasted streamflows, and the solid line representing the cumulative probability of observed streamflows

All the forecasting skill assessment approaches used (in both the categorical and continuous probabilistic forecasts) reveal that the proposed analogue method had more success in the early summer season (OND) in the Mgeni catchment, and that forecast skills are seen to degrade with subsequent periods during the rainy season. This is most likely due to the much greater influence of high, and especially convective, rainfall events occurring during the late summer months. While there might have been occasional misreading of a raingauge, some extremely high rainfall events resulted in more than 50% of the season's total rainfall occurring in a one or two days of consecutive rain, resulting in a sudden shift from a below normal to a near or above normal streamflow season. For

178

example, at QC U20G (cf. Figure 6.1), three days of rainfall amount resulted in 51% of the DJF rainfall in 2003/04. Given the high spatial variability of convective events that characterise the summer seasons in the Mgeni catchment, the use of only one rainfall station per QC may also cause appreciable errors. Apart from data quality problems, a possible reason for this deterioration in forecast accuracy could also be the limitation of the climate models in predicting extreme rainfall seasons adequately, or high magnitude individual rainfall events within a season. Downscaling the coarse spatial resolution categorical forecasts to the much smaller scale QCs is another problem, as seasonal rainfall forecasts have shown to posses skills only for large areas. The poor association between January rainfall and sea surface temperature (Landman and Klopper, 1998) could be another possible explanation that may have affected the forecasts for the NDJ, DJF and JFM periods.

## 9.4    Concluding Remarks

The results from this study have demonstrated clearly the soundness of applying the *Historical Sequence Method* of translating 3 month categorical rainfall forecasts into ensembles of daily quantitative values suitable for application in hydrological/crop yield models. With an 11.1% probability of departure from its random expectation, high values of RTSS were scored over the selected seasons in the Mgeni catchment. Some skills are lost in the LEPS and RPSS skill assessments, especially in the late summer seasons. However, considering the uncertainties that cascade from the rainfall forecasts, through the translation process to the *ACRU* model, the results are considered to be generally acceptable. By averaging the 20 daily ensembles of streamflow forecasts, statistically realistic forecasts of accumulated streamflows were also obtained for the three selected seasons. The results reflect the assumption that the daily rainfalls during the forecast period could mirror those of the selected analogue years for the same calendar period. The *Historical Sequence Method* is conceptually simple, but computationally cumbersome. Nevertheless, it can be used with confidence to translate the skilful categorical rainfall forecasts into a daily quantitative values required by agrohydrological models.

# 10. GENERATING CONDITIONED RAINFALL BASED ON THE ENSEMBLE RE-ORDERING METHOD FOR APPLICATIONS IN AGROHYDROLOGY

## 10.1 Introduction

The use of the *Historical Sequence Method* as a means of generating daily rainfall values from the same dates of selected analogue years provides sequences that are based on only the realisation of the past climate. The *Historical Sequence Method* assumes that historical weather sequences can be surrogates for future weather sequences. It has many advantages, including its relative simplicity in concept and attempts to preserve the spatial and temporal distribution of the past weather conditions which occurred in the selected analogue years. However, the set of historical sequences of rainfall may not always repeat itself in exactly the same manner in the future. In such cases, modifying the analogue method by using a random generator method to generate ensembles of synthetic data which have statistical characteristics similar to those of historical data, can provide alternative realisations that are equally likely to occur in the future time. Such methods can generate long term daily weather sequences of values and the generated values can then be used as input into agrohydrological models in order to assess climate related impacts on agriculture and water resources which allow decision makers the option of a wide range of management alternatives.

The aim of this chapter is to disaggregate categorical probabilistic rainfall forecasts into synthetic daily rainfall ensembles that are statistically similar to the observed data. In order to do this, a conditional precipitation generator has been developed which uses an approach termed the "Ensemble Re-ordering" (or "Schaake shuffle"). Weather generators can produce daily values of precipitation, minimum and maximum temperature and solar radiation. This study, however, focuses only on the generation of conditioned sequences of daily precipitation. The remainder of this chapter is structured as follows: the *Ensemble Re-ordering Based Method* (ERBM) is described briefly in Section 10.2. Results, including an assessment of probabilistic forecast skills, biases and correlations are described in Section 10.3, followed by concluding remarks in Section 10.4.

## 10.2   The Ensemble Re-ordering Based Method

Because the dynamics of rainfall occurences are chaotic and dependent on the interactions between the atmosphere and the underlying ocean/land distributions, it is a difficult task to predict the evolution of indiviudal rainfall events with certainty, especially at a longer time scale.  However, by considering rainfall occurrences as random phenomena, there are currently several rainfall simulation models that have been developed to reproduce stochastic rainfall characterstics that reflect first and higher order statistics of observed rainfall. Reviews of commonly used weather generators can be found in Lall *et al.* (1996), Semenov *et al.* (1998), Rajagopalan and Lall (1999), Wilks (2002), Clark *et al.* (2004), Federsen and Andersen (2005) and Chiew *et al.* (2005). Stochastic weather generators based on Markov chains, kernel and the k-nearest neighbours are commonly used methods to yield realisations of daily weather that resemble the historical data with respect of certain statistical moments. The use of weather sequences generated by such methods can be applied to quantify the uncertainty and risks associated with climate variability over short and long time scales.

However, according to many researchers (e.g. Sharma and Lall, 1999; Yates *et al*., 2003; Clark *et al*., 2004), the existing stochastic weather generating methods have problems with under-prediction of precipitation when they are extended to longer time scales and/or when they are used at multiple sites.  Many researchers (e.g. Mason, 1990; Mason *et al.,* 1996; Kabat and Bates, 2002) also suggest that the summer rainfall over southern Africa is strongly related to sea surface temperature anomalies in the tropical Pacific, central south Atlantic and western equatorial Indian Oceans. Thus, it may not be appropriate to consider rainfall occurrences in this region as a purely random process. For these reasons, the rainfall generator that has been developed in this study is designed to account the local probabilistic categorical forecasts, and uses the "Ensemble Re-ordering" approach (Clark *et al.,* 2004) to preserve the statistical moments of the historical time series from which the rainfall sequences are populated. Clark *et al.* (2004) applied this technique of ensemble member construction for temperature and precipitation in four catchments in the USA. They found that the reconstruction methodology is capable of preserving the temporal persistence, inter-site and

inter-variable correlations. Most of the information contained in this section has been taken from Section 3 of Clark *et al.* (2004) to illustrate the theoretical background behind the ERBM method.

For a given forecast day a three-dimensional matrix, $X_{i,j,k}$, is constructed to resample data sequences from a subset of preferentially selected years, where *i* refers to the number of ensemble members, *j* refers to the individual station and *k* refers to each variable. Unlike other weather generating methods, resampling of data is not conditioned on previous simulated days to preserve the space-time variability of the station time series. Instead, another identical three-dimensional matrix, $Y_{i,j,k}$, derived from all historical years of the respective variables, is constructed so as to preserve the space-time variability in the historical time series, and the dates used to populate the Y matrix are persisted with for subsequent forecast days. As in the X matrix, *i* refers to an index of dates in the historical time series, while *j* refers to each station and *k* refers to each variable. The main assumption in this approach is that an ordered selection of dates from all days in the historical record enables the reconstruction of the spatio-temporal correlation structure for a subset of data (Clark *et al.*, 2004). Mathematically, the ERBM can be formulated as follows: For a given station (*j*) and variable (*k*), let X be a vector of "n" observations (*x*) resampled from preferentially selected analogue years and let $\vec{\chi}$ be the ordered vector of X. Therefore,

$$X = (x_1, x_2, \ldots\ldots, x_n)$$ 
$$\tag{10.1}$$

and

$$\vec{\chi} = (x_{(1)}, x_{(2)}, \ldots\ldots, x_{(n)}), \text{ where } x_{(1)} \leq x_{(2)} \ldots\ldots \leq x_{(n)}.$$
$$\tag{10.2}$$

Similarly, let Y be a vector of "n" observations (*y*) resampled from all historical years of the same season. Let $\vec{\gamma}$ be the sorted vector of Y and $\beta$ the vector indices describing the original observation number corresponding to the values in the ordered vector $\vec{\gamma}$. Therefore,

$$Y = (y_1, y_2, \ldots\ldots, y_n)$$
$$\tag{10.3}$$

and

$$\vec{\gamma} = (y_{(1)}, y_{(2)}, \ldots\ldots, y_{(n)}), \text{ where } y_{(1)} \leq y_{(2)}, \ldots\ldots, \leq y_{(n)}.$$
$$\tag{10.4}$$

The final re-ordered vector, $\vec{X}^{ss}$, is then constructed according to the order in vectors $\vec{\chi}$ and $\beta$, as follows:

$$\vec{X}^{ss} = (x_1^{ss}, x_2^{ss}, ..., x_n^{ss})$$   10.5

$$x_q^{ss} = x_{(r)}$$   10.6

$$q = \beta[r]$$   10.7

$$r = 1, .......n.$$   10.8

The implementation of the ERBM approach is further illustrated by the following example (Clark *et al.,* 2004). For the first day of January 2007, the randomly selected ensembles of 10 members (X) from a mix of dates of selected years and the corresponding selected ensembles of 10 members (Y) from all historical years of the same month are shown in Table 11.1(a) and (b) respectively.

Table 10.1   The ERBM for 01 January 2007, showing (a) the randomly selected ensemble members from selected years (X), (b) the randomly selected ensemble members from the same season of all days in the historical records (Y), and (c) the final re-ordered output, $\vec{X}^{ss}$ (Clark *et al.,* 2004)

|            (a)            |            (b)            |            (c)            |

| Ensemble # | X |
|---|---|
| 1 | 14.1 |
| 2 | 8.0 |
| 3 | 2.5 |
| 4 | 0.0 |
| 5 | 0.2 |
| 6 | 0.0 |
| 7 | 2.8 |
| 8 | 4.7 |
| 9 | 0.8 |
| 10 | 0.0 |

| Ens # | Date | Y |
|---|---|---|
| 1 | 03/01/1976 | 58.0 |
| 2 | 09/01/1952 | 7.4 |
| 3 | 27/01/1982 | 1.0 |
| 4 | 11/01/1998 | 0.8 |
| 5 | 19/01/1956 | 0.5 |
| 6 | 06/01/1996 | 2.8 |
| 7 | 09/01/1969 | 4.3 |
| 8 | 23/01/1995 | 10.5 |
| 9 | 17/01/2001 | 2.0 |
| 10 | 02/01/1984 | 0.0 |

|  | Reordered Ens # | $\vec{X}^{ss}$ |
|---|---|---|
| 1 | (1) | 14.1 |
| 2 | (8) | 4.7 |
| 3 | (5) | 0.2 |
| 4 | (4) | 0.0 |
| 5 | (6) | 0.0 |
| 6 | (3) | 2.5 |
| 7 | (7) | 2.8 |
| 8 | (2) | 8.0 |
| 9 | (9) | 0.8 |
| 10 | (10) | 0.0 |

The vector X may provide the sorted vector, $\vec{\chi}$, and the vector Y may provide the sorted vector, $\vec{y}$, with the vector indices $\beta$ as shown below:

| | | |
|---|---|---|
| X | = | (14.1, 8.0, 2.5, 0.0, 0.2, 0.0, 2.8, 4.7, 0.8, 0.0) |
| $\vec{\chi}$ | = | (0.0, 0.0, 0.0, 0.2, 0.8, 2.5, 2.8, 4.7, 8.0, 14.1) |
| Y | = | (58.0, 7.4, 1.0, 0.8, 0.5, 2.8, 4.3, 10.5, 2.0, 0.0) |
| $\vec{y}$ | = | (0.0, 0.5, 0.8, 1.0, 2.0, 2.8, 4.3, 7.4, 10.5, 58.0) |
| $\beta$ | = | (10, 5, 4, 3, 9, 6, 7, 2, 8, 1) |

According to Equations 10.6 and 10.7, the final re-ordered output for 01 January 2007, also shown in Table 10.1 (c), is:

| | | |
|---|---|---|
| $\vec{X}^{ss}$ | = | (14.1, 4.7, 0.2, 0.0, 0.0, 2.5, 2.8, 8.0, 0.8, 0.0) |

The process is repeated for all the forecast days within January 2007. However, the random selection of dates from all years that are used to populate the Y matrix are used only for the first day (i.e. 01 January 2007), and are persisted for subsequent forecast days. In this example, the dates that would be used to populate the Y matrix for the next day (i.e. 02 January 2007) are, 04 January 1976 for the first ensemble member, 10 January 1952 for ensemble two and 28 January 1982 for ensemble three and so on (cf. Table 11.1(b)). Since the ranks of the ensemble members match the ranks in the historical time series, the resulting ensembles of forecasts preserve the temporal patterns of the station time series (Clark *et al.,* 2004).

Clark *et al.* (2004) selected the analogue years based on ENSO indices. However, in this study an algorithm has been developed that generates ensembles of 10 members for each within the forecast period based on monthly (or seasonal) probabilistic categorical forecasts. The ERBM (Clark *et al.,* 2004) is then adapted as a post-processing step to reconstruct the temporal persistence of the generated ensembles. The approach is adapted in this study as follows:

For a given region (e.g. a catchment) with one representative climate station (j = 1) per QC and one variable (i.e. rainfall), let January 2007 be the forecast month of concern, with a triplet of probabilities $P_B$ = 30, $P_N$ = 40 and $P_A$ = 30 pertaining, respectively, to the forecasted percentiles of the three categories of "below", "near" and "above" normal rainfall conditions. In this study, quality checked daily rainfall totals for a 54 years period from 1950 to 2003 are then ranked for the month of January in an ascending (lowest to highest) order. The first 18 ranked rainfall totals out of the 54 are then categorised as representing "below normal" seasonal rainfalls, the next 18 rankings as "near normal" and the highest 18 as "above normal" seasonal rainfalls. A random selection of analogue years ($A_{(years)}$) is then made, based on the weighting criteria given in Equation 9. Since the sum of the three forecast probabilities is constrained to 100%, and the forecast probabilities are at interval of 5%, the total number of analogue years is always 20. Therefore,

$$A_{(years)} = (\frac{P_B}{5} + \frac{P_N}{5} + \frac{P_A}{5}) = 20 \qquad\qquad 10.9$$

In this example, the respective numbers of analogue years sampled from below, near and above normal rainfall years would therefore be 6, 8 and 6. Let Z be a vector of the selected 20 analogue years ($A_{(years)}$) as shown below, i.e.

$$Z = (A_{(year)}, year = 1, 2, ....., 20) \qquad\qquad 10.10$$

Each of the selected analogue years $A_{(years)}$ contains "n" observations (a), depending on the number of dates in the forecast period (e.g. for January n = 31). The subscript (i) refers to an index of dates in the historical time series of the selected month. Hence

$$A_{(year)} = (a_{(i)}, i = 1, 2, ..........., n). \qquad\qquad 10.11$$

$$Z = (a_{20(i)}, i = 1, 2, ..........., n). \qquad\qquad 10.12$$

To populate the X matrix, a random selection of 10 ensemble members from a mix of dates in the vector Z is made for each forecast day of the month, in this example for January 2007. Every member, $a_{(i)}$, has 1/20n chance to be included in the X matrix. Another random selection of 10 ensembles from all historical years (1950-2003) of the same month is made to populate the Y matrix only for the first day of the forecast season. For subsequent forecast days, the dates will persist to populate the Y matrix. The generation of rainfall ensembles using this approach can then be used to generate probabilistic agrohydrological forecasts. In turn, such risk based forecasts have the potential to assist users to better prepare for extreme events by giving early warnings.

The ERBM was applied to the 12 Quaternary Catchments (QCs) that make up the Mgeni catchment. The extent to which the ERBM is able to mimic the daily precipitation variability within a month and a season (i.e. three months), is then evaluated with respect to dry and wet spell frequencies, and simulated streamflows by the *ACRU* model. The results are discussed in the section which follows.

## 10.3   Results and Discussion

The various verification techniques presented in Chapter 9 have also been employed in this chapter to assess the performance of the ERBM derived rainfalls and the *ACRU* simulated streamflows. The statistical means, standard deviations, skewness coefficients and coefficients of variation are summarised with box plots. Although the simulations were made on a daily timescale, the statistics from the daily data have been aggregated into monthly and seasonal (three months) time scales to see how well they reproduce aggregated quantities such as monthly (or seasonal) totals, means and standard deviations, and number of dry (wet) spells within a month (or a season).

### 10.3.1  Simulating Dry and Wet Spells

Agriculture in southern Africa is at risk due, in part, to the erratic nature of rainfall, where individual events are often characterised by severe convective storms in many parts of the region. The within-season variability of rainfall in amount,

frequency and distribution may affect sustainable agriculture adversely as well as impacting on other social and economic functions of the region. This is particularly true for rainfed agriculture in the more semi-arid and arid areas where extended dry spells often result in poor crop establishment and in yield reduction (Schulze, 2006). At the start of the rainy season, few rainfall events may wet the soil sufficiently for planting, but these events are often followed by prolonged dry spells that can seriously reduce crop production. Hence, modelling the probability distribution of temporal rainfall characteristics is crucial to many water and agricultural management strategies. Prediction of the onset of the rainy season, the number of rainfall days yielding a measureable amount of rain, and the probability of long dry spells during critical periods of the growing season can provide a basis for decision makers to best manage their operations and strategies at the time when a decision has to be made. Such predictability is of great importance particulary in rainfed agriculture in assisting farmers in pre-planting farming operations and management strategies for different cultivars and crop productions. With the advancement of mathematically and physically based approaches, the analysis of within-season rainfall characteristics has also received much attention for several other applications such as irrigation scheduling, water harvesting, reservoir management and the design for the disposal of hazardous wastes (Adiku *et al*., 1997). Furthermore, it can also help disaster managers for defining drought and flood severity in order to mitigate the effects of these extreme hazards.

Recognising the benefits of accounting for such variations in agriculture and water resource management, the ERBM is used to investigate the within-season temporal variability of rainfall. Unlike parametric probability models, this approach does not require the estimation of parameters describing the sequence of the state of the previous day. The ERBM is designed to reproduce the following:

- The fraction of wet and dry days in a given month (season) of the year: In this study, a wet day (*w*) has been defined as a day with rainfall of 0.3 mm or more, and a dry day (*d*) is defined as a day with less than 0.3 mm rainfall. Thus, in a rainy month (season) of *n* days, a time series of rainfall records, $x_1, x_2, ..., x_n$ is truncated at the threshold rainfall value of 0.3 mm, where $x_i \geq$

0.3 mm = $w_i$ and $x_i$ < 0.3 mm = $d_i$. The probability of the fraction of wet days is then obtained by dividing the number of wet days by the number of days ($n$) of a given month (season). The same connotation also applies to the probability of the fraction of dry days.

- The probability that a dry day follows a wet day, i.e. *P(w/d)*, and the probability that a wet day follows a dry day, i.e. *P(d/w)*, for each forecast month (or season): These transitional probabilities can provide some insight into the problems related to the intermittency of rainfall. They are also widely used to simulate the probabilities of rainfall occurrences in parametric probability (e.g. Markov chain) models.

- The mean and maximum length of wet and dry spells in a given month (season) of the year: In a time series of daily rainfall during a rainy month (season) of $n$ days, there will occur runs of uninterrupted wet ($x_i \geq$ *0.3 mm*) and dry ($x_i <$ *0.3 mm*) days, as shown in Figure 10.1. Wet ($W_s$) and dry ($D_s$) spells are defined respectively as a series run of wet and dry days within a prescribed period of time. In this instance, the mean wet spell length in Figure 10.1 is 4.33 days and the mean dry spell length is 6 days. In a similar way, the longest wet and dry spell lengths are 6 and 9 days, respectively.
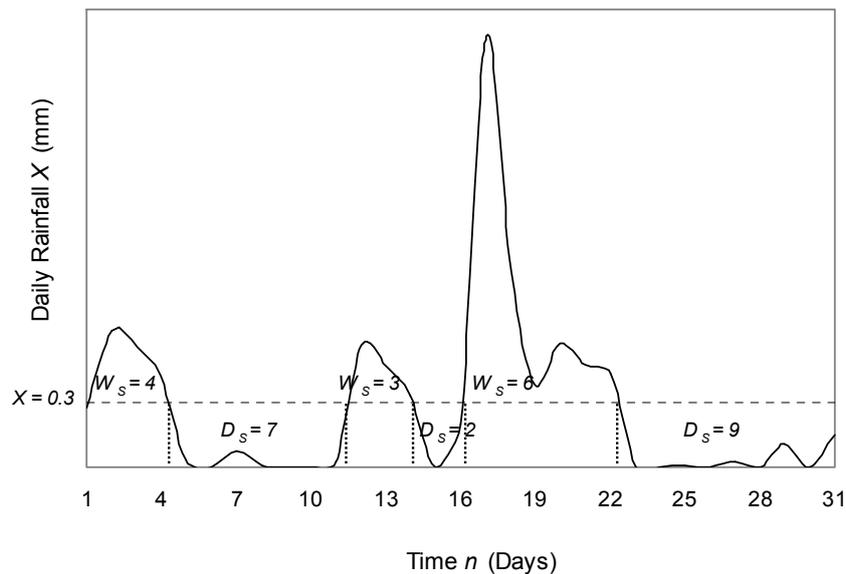


Figure 10.1   Schematic representation of wet and dry spells in a rainy month of 31 days

The ERBM was applied to assess its ability to adequately reproduce the aforesaid statistics of the observed rainfall characteristics at QC U20M in the Mgeni catchment (cf. Figure 6.1) for the year 2004. The categorical probabilistic forecasts for each month of the year 2004 were extracted from the SAWS website. The statistical moments of the simulated records are compared with those for the observed record using box plots. A box in the box-and-whisker plots (Figure 10.2) indicates the inter-quartile ranges of the statistics computed from ensembles of 10 members, the line in the middle of the box indicates the simulated median value, and the x indicates the simulated mean value. The solid continuous line corresponds to the observed record, with circles indicating the values outside the simulated range. In general, if the statistics of observed data lie within the box of simulated values, it suggests that the simulated values have reproduced the statistics of the observed data adequately.

Figure 10.2 illustrates the fraction of wet and dry days, the probability that a dry day follows a wet day *P(w/d)*, and the probability of that a wet day follows a dry day *P(d/w)* for each month. Average and longest of the wet and dry spell lengths are shown in Figure 10.3. The box plots illustrate the variability of rainfall in each statistic across the 10 ensemble members.

According to Clark *et al.* (2004), "the ERBM has difficulties dealing with the intermittency of rainfall when the ensembles of generated rainfall from analgoue years for a given day have fewer zero rainfall ensemble members than the ensemble from the persisted observed data. The reason for this is that while the generated ensemble members with zero rainfall days may match the observed precipitation values, their assignment to a given ensemble member will be entirely random". This discrepancy could potentially cause some biases in the model simulations. Nevertheless, except for March, the simulated fraction of wet and dry days as well as the *P(d/w)* and *P(w/d)* agreed well with the corresponding observed statistics.
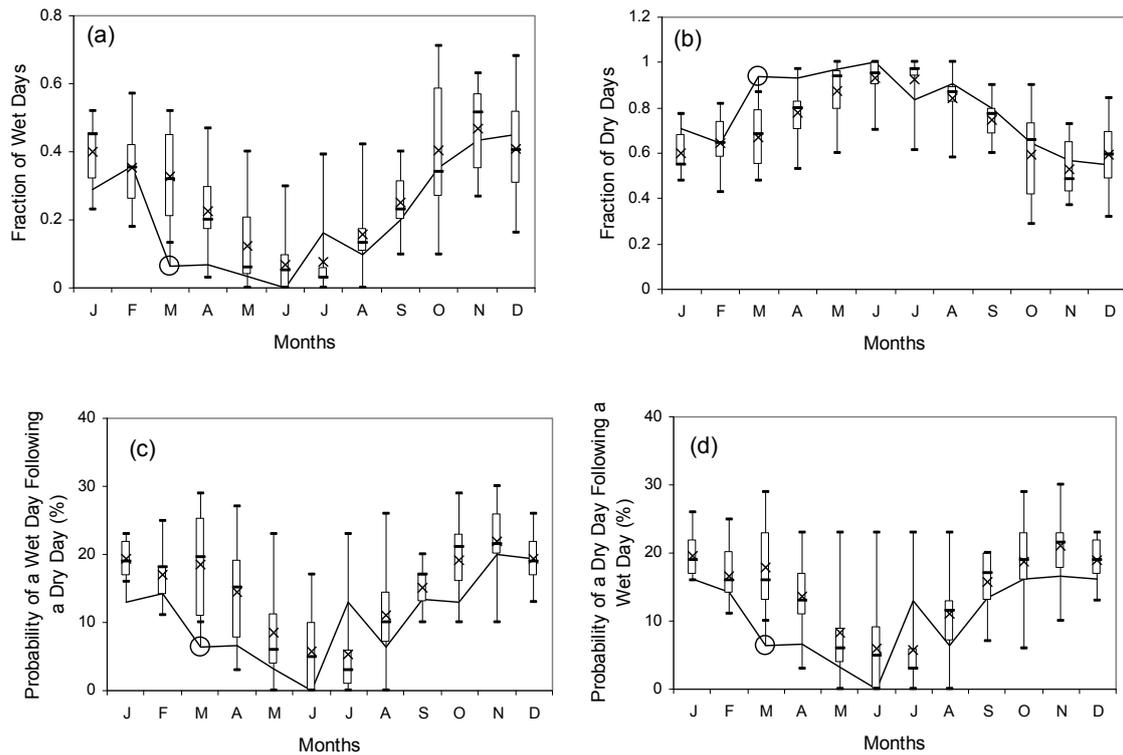
Figure 10.2   Box plots of statistics of generated fractions of wet (a) and dry (b) days, the probability that a wet day follows a dry day (c), and the probability of that a dry day follows a wet day (d) for each month, along with the observed values, at QC U20M for the year 2004, using the ERBM. The box-and-whiskers represent the minimum, lower quartile, median, upper quartile and maximum of the forecasted values. The x signs indicate the simulated mean values. The solid line corresponds to the observed record and the circles indicate the values outside the simulated range

It is important to note that the seasonal trend is reproduced well in the simulations from the ERBM. This indicates the potential capability of the method to capture the within-month statistics of the summer rainfall in southern Africa. It is well observed again that the trends of simulated means and longest spells of dry and wet conditions follow the pattern of observed values in all months of the year 2004, even though they were not explicitly mimicked (Figure 10.3). For ease of comparison, the within-month statistics of the simulated (mean) and observed rainfalls for the year 2004 are summarised in Table 10.2.

190
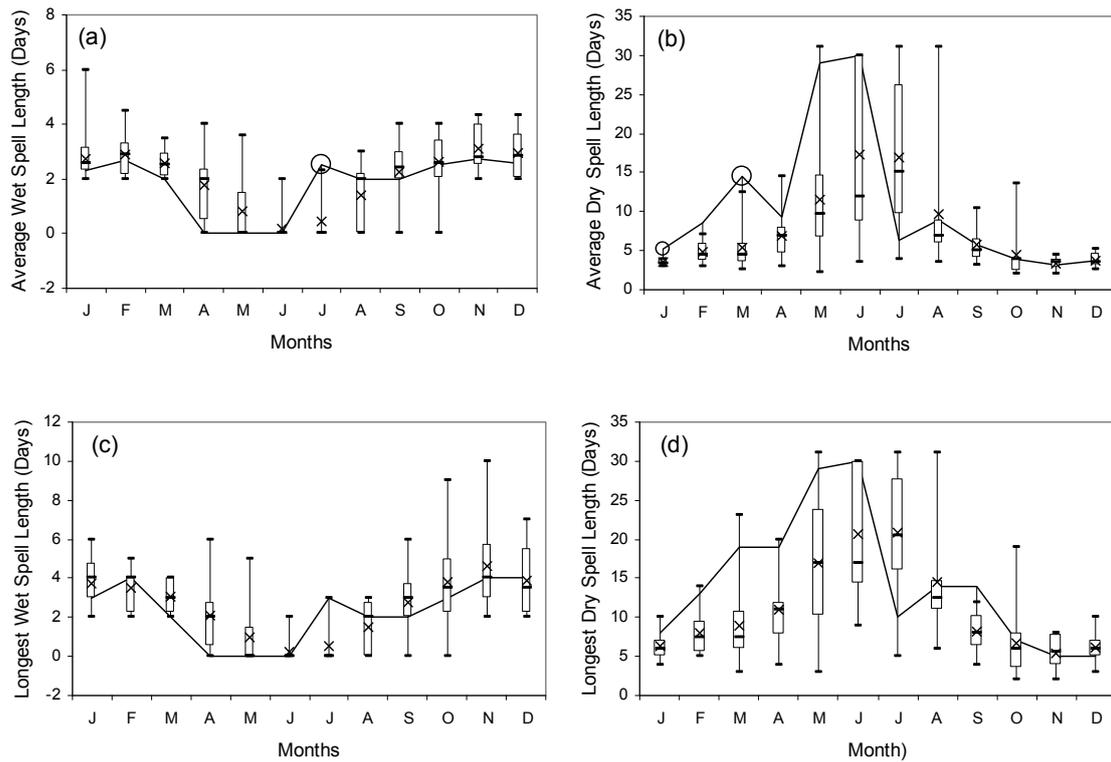
Figure 10.3   Box plots of statistics of generated mean ( a, b) and longest (c, d) lengths of wet and dry spells along with the observed values at QC U20M for the year 2004 using the ERBM. The box-and-whiskers represent the minimum, lower quartile, median, upper quartile and maximum of the forecasted values. The x signs indicate the simulated mean values. The solid line corresponds to the observed record and the circles indicate the values outside the simulated range

At the beginning of the rainy season (October), the agreement between the simulated mean and observed statistics was good, and the good agreement was seen to increase for November and December when the rainfall frequencies increased. For January, even though the simulated mean and longest spells of wet days are captured well, the fractions of wet/dry days as well as the mean and longest dry spells are biased towards the wetter side.  For February, the simulation model did well to capture the observed statistics, but the mean and longest dry spells are highly under-estimated. For the autumn (March and April) and early

winter (May-July) seasons the differences between the simulated and observed statistics are very high, suggesting poor forecasts. As these months exhibit a higher probability of no-rain days, the difference could be attributed to the weaknesses of the ERBM which were mentioned earlier. However, these differences are not necessarily significant agrohydrologically, as these months experience very little rainfall. For the early spring months (August-September), the simulated within-month statistics agreed reasonably well with the corresponding observed data.

Table 10.2    The within-month statistics of simulated and observed rainfalls at the mouth of the Mgeni catchment for year 2004, as estimated by the ERBM

| Statistc | | Jan | Feb | Mar | Apr | May | Jun | Jul | Aug | Sep | Oct | Nov | Dec |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Fraction of wet days | Observed | 0.29 | 0.36 | 0.06 | 0.07 | 0.03 | 0.00 | 0.16 | 0.10 | 0.20 | 0.35 | 0.43 | 0.45 |
| | Simulated (mean) | 0.40 | 0.35 | 0.33 | 0.22 | 0.13 | 0.07 | 0.08 | 0.16 | 0.25 | 0.40 | 0.47 | 0.41 |
| Fraction of dry days | Observed | 0.71 | 0.64 | 0.94 | 0.93 | 0.97 | 1.00 | 0.84 | 0.90 | 0.80 | 0.65 | 0.57 | 0.55 |
| | Simulated (mean) | 0.60 | 0.65 | 0.67 | 0.78 | 0.88 | 0.93 | 0.92 | 0.84 | 0.75 | 0.60 | 0.53 | 0.59 |
| P (%) that a wet day follows a dry day | Observed | 12.90 | 14.29 | 6.45 | 6.67 | 3.23 | 0.00 | 12.90 | 6.45 | 13.33 | 12.90 | 20.00 | 19.35 |
| | Simulated (mean) | 19.30 | 17.10 | 18.60 | 14.40 | 8.50 | 5.70 | 5.40 | 11.00 | 15.10 | 19.20 | 22.00 | 19.30 |
| P (%) that a dry day follows a wet day | Observed | 16.13 | 14.29 | 6.45 | 6.67 | 3.23 | 0.00 | 12.90 | 6.45 | 13.33 | 16.13 | 16.67 | 16.13 |
| | Simulated (mean) | 19.60 | 16.70 | 17.90 | 13.60 | 8.20 | 6.00 | 5.70 | 11.00 | 15.70 | 18.70 | 21.00 | 19.00 |
| Mean wet spell length (days) | Observed | 2.33 | 2.67 | 2.00 | 0.00 | 0.00 | 0.00 | 2.50 | 2.00 | 2.00 | 2.50 | 2.75 | 2.60 |
| | Simulated (mean) | 2.75 | 2.91 | 2.60 | 1.80 | 0.79 | 0.20 | 0.43 | 1.43 | 2.23 | 2.61 | 3.13 | 2.93 |
| Mean dry spell length (days) | Observed | 5.25 | 8.50 | 14.50 | 9.33 | 29.00 | 30.00 | 6.25 | 9.00 | 5.75 | 4.00 | 3.20 | 3.75 |
| | Simulated (mean) | 3.53 | 4.76 | 5.43 | 6.92 | 11.58 | 17.38 | 16.90 | 9.67 | 5.86 | 4.40 | 3.32 | 3.80 |
| Longest wet spell length (days) | Observed | 3.00 | 4.00 | 2.00 | 0.00 | 0.00 | 0.00 | 3.00 | 2.00 | 2.00 | 3.00 | 4.00 | 4.00 |
| | Simulated (mean) | 3.70 | 3.50 | 3.10 | 2.10 | 1.00 | 0.20 | 0.50 | 1.50 | 2.80 | 3.80 | 4.60 | 3.90 |
| Longest dry spell length (days) | Observed | 8.00 | 13.00 | 19.00 | 19.00 | 29.00 | 30.00 | 10.00 | 14.00 | 14.00 | 7.00 | 5.00 | 5.00 |
| | Simulated (mean) | 6.30 | 8.00 | 9.00 | 10.90 | 17.00 | 20.60 | 20.80 | 14.60 | 8.20 | 6.70 | 5.40 | 6.20 |

Another important feature of using the ERBM is to utilise the generated ensembles of rainfalls for agrohydrological modelling. A major of concern in many water dependent sectors is to predict streamflows, crop yields, soil moisture and reservoir storages at short and long time scales. To assess the usefulness of the ERBM in such applications, streamflows were generated by the *ACRU* model at the mouth of the Mgeni catchment. The results are discussed in the sub-section which follows.

**10.3.2  Forecasting Monthly and Seasonal Streamflows**

The monthly forecasts for year 2004 and the seasonal forecasts from October 2003 to March 2006 for the three month periods OND, NDJ, DJF and JFM were used for this evaluation. The *ACRU* model was run with historically observed daily

rainfalls from year 2000 to the beginning of the forecast period to create representative antecedent conditions and initial stores. The daily re-ordered ensemble of 10 members extracted for each QC and for the selected months (seasons) were then applied together with the antecedent conditions generated previously by the *ACRU* model in order to generate the corresponding 10 daily ensembles of streamflow for each forecast month (season). Simulated streamflows obtained from the *ACRU* model using observed rainfall serve as a baseline for evaluating streamflows derived from the ERBM based rainfall simulations. It is important to note that the simulated streamflows are accumulated at the mouth of the Mgeni catchment from all the upstream subcatchments (cf. Figure 6.1).

The monthly simulation of streamflows obtained from the ERBM derived rainfalls at the mouth of the Mgeni catchment are shown by way of box plots in Figure 10.4. Overall, the simulations adequately mimicked the means of daily streamflows for each month, even though there is a slight tendency of under-estimation for January and February, and a slight over-estimation for April (Figure10.4 (a)). Figure 10.4 (b) shows the standard deviations of daily streamflows for each month, where these measure the spread of the streamflows about the monthly mean value. As expected, the range of variability is higher for the summer rainfall months (November-March) than for the remaining months. Except for October, the observed standard deviations are adequately reproduced by the simulations (Figure 10.4 (b)).

The skewness coefficient is another important statistical moment in measuring the symmetry of the streamflow distribution (Figure 10.4 (c)). Typically, rainfall data are positively skewed, as rainfall typically occurs as many small events with a few large events that elevate the mean (Hobson, 1997). Consequently, the simulated streamflows are also positively skewed, placing the monthly mean value in the upper quartile range. The dry months May, June, July and August generally exhibit a wide range of values of the skewness coefficient. The reason for this is that they are characterised by almost no rainfalls with occasional small events which create zero streamflows. The historical skewness coefficients in these months were

consistently under-estimated by the model simulations. However, for the remaining months the simulations captured the observed skewness well (Figure 10.4 (c)).
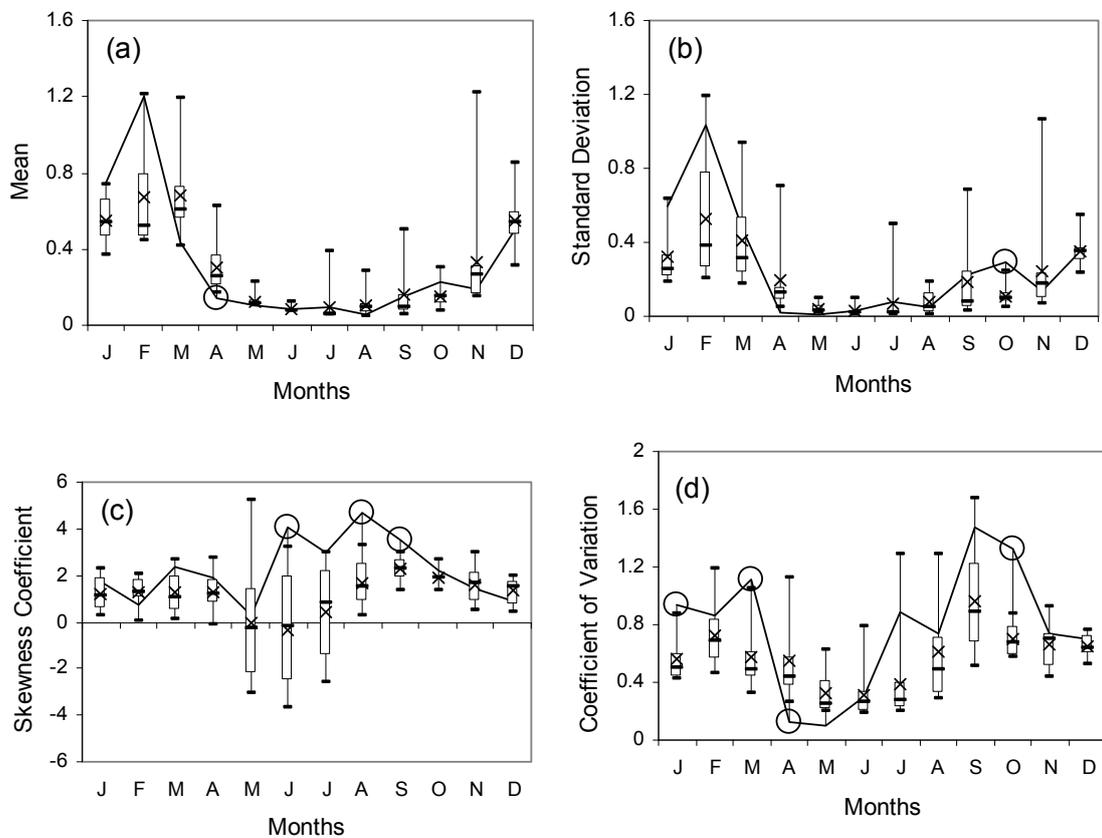


Figure 10.4   Box plots of statistics of generated streamflows along with the observed values at the mouth of the Mgeni catchment for the year 2004 using the ERBM. The box-and-whiskers represent the minimum, lower quartile, median, upper quartile and maximum of the forecasted streamflow sequences. The x signs indicate the simulated mean value. The solid line corresponds to the observed record and the circles indicate the values outside the simulated range

The coefficient of variation (CoV) provides a relative measure of dispersion in streamflows with respect to the mean value (Figure 10.4 (d)). The high range of rainfall variability in the summer months resulted a more spread-out and widely dispersed distribution of streamflows (Figure 10.4 (b)). The observed dispersion is reproduced fairly well by the simulations except for the months January, March,

April and October, where the CoV values were outside of the simulated inter-quartile ranges (cf. Figure 10.4 (d)).

Figure 10.5 (a) provides a visual plot of monthly totals of daily simulated vs observed streamflows on a month-by-month basis for the year 2004 at the mouth of the Mgeni catchment. Figure 10.5 (b) depicts the corresponding accumulated monthly flows for the same period. The monthly totals of daily simulated streamflows from the *ACRU* mimicked the corresponding observed flows adequately, with the exception of January and February, when the simulations under-estimated observations by around 26.3% and 44.4%, respectively (Figure 10.5(a)). The performance of the ERBM derived simulations by the *ACRU* model was generally better during the drier months (April - September) than the rainy months (October - March). The reason for this is the high rainfall variability that typically occurs during the rainy months. Nonetheless, the observed trend was captured fairly well throughout the year (cf. Figure 10.5(a)).
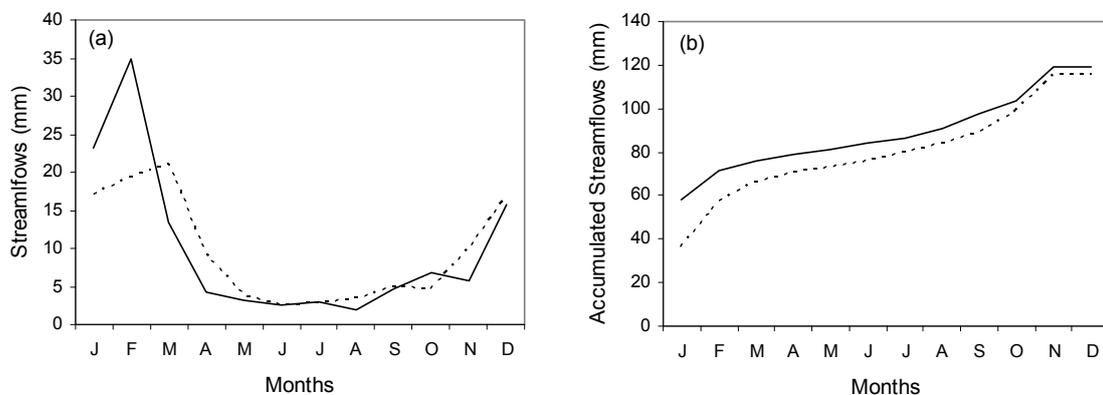


Figure 10.5   Comparison of *ACRU* simulated flows versus observed flows (mm) at the mouth of the Mgeni catchment for the year 2004, using the ERBM method, where (a) is a visual plot of monthly totals of daily simulated against observed flows and (b) depicts the corresponding accumulated monthly flows. The dashed line represents the simulated flows and the solid line represents the observed flows

The accumulated monthly totals of daily simulated streamflows from the *ACRU* model appeared to consistently under-estimate the corresponding accumulated monthly observed flows. This is due partially to the accumulated errors that cascade from one month to the subsequent months. Clearly, the under-predicted flows during January and February had a significant influence especially on baseflows in subsequent months, by consistently under-estimating the accumulated observed flows. Nevertheless, the overall trend was captured quite well (cf. Figure 10.5(b)).

Figure 10.6 presents the cumulative probability distribution for the year 2004 at the mouth of the Mgeni catchment. Since the flows during the dry season months April to September are very low, the cumulative probability distribution is constructed only for the remaining six rainy months. Previously it was demonstrated in Figure 10.5 (a) that a significant under-prediction of the January and February monthly totals of daily flows occurred. The cumulative probability distribution for January and February is therefore expected to be biased to the drier side (Figures 10.6 (a) and (b)). For March, the model simulation was good enough to estimate the observed cumulative probabilities of accumulated flows up to 10 mm, but changed to over-predict the probabilities of higher accumulated flows (Figure 10.6 (c)). In contrast to the above, the model simulation failed during October to simulate the cumulative probabilities of low accumulated flows (< 4 mm), but captured the probabilities of higher accumulated flows well (Figure 10.6 (d)). For November and December, the simulation appeared to consistently over-predict the observed cumulative probabilities. The over-prediction during November is more significant than for December, suggesting a higher probability of wet conditions than the observed cumulative probability (Figures 10.6 (e) and (f)).

The seasonal simulation of a streamflows by *ACRU* from the ERBM derived rainfalls for OND (a), NDJ (b), DJF (c) and JFM (d) of 2003/04 at the mouth of the Mgeni catchment are shown in Figure 10.7.

Figure 10.6   Cumulative probabilities of daily flows simulated with the *ACRU* model versus observed flows at the mouth of the Mgeni catchment for January (a), February (b), March (c), October (d), November (e) and December (f) of the year 2004, using the ERBM. The dashed line represents the cumulative probability of forecasted streamflows, while the solid line represents the cumulative probability of observed streamflows
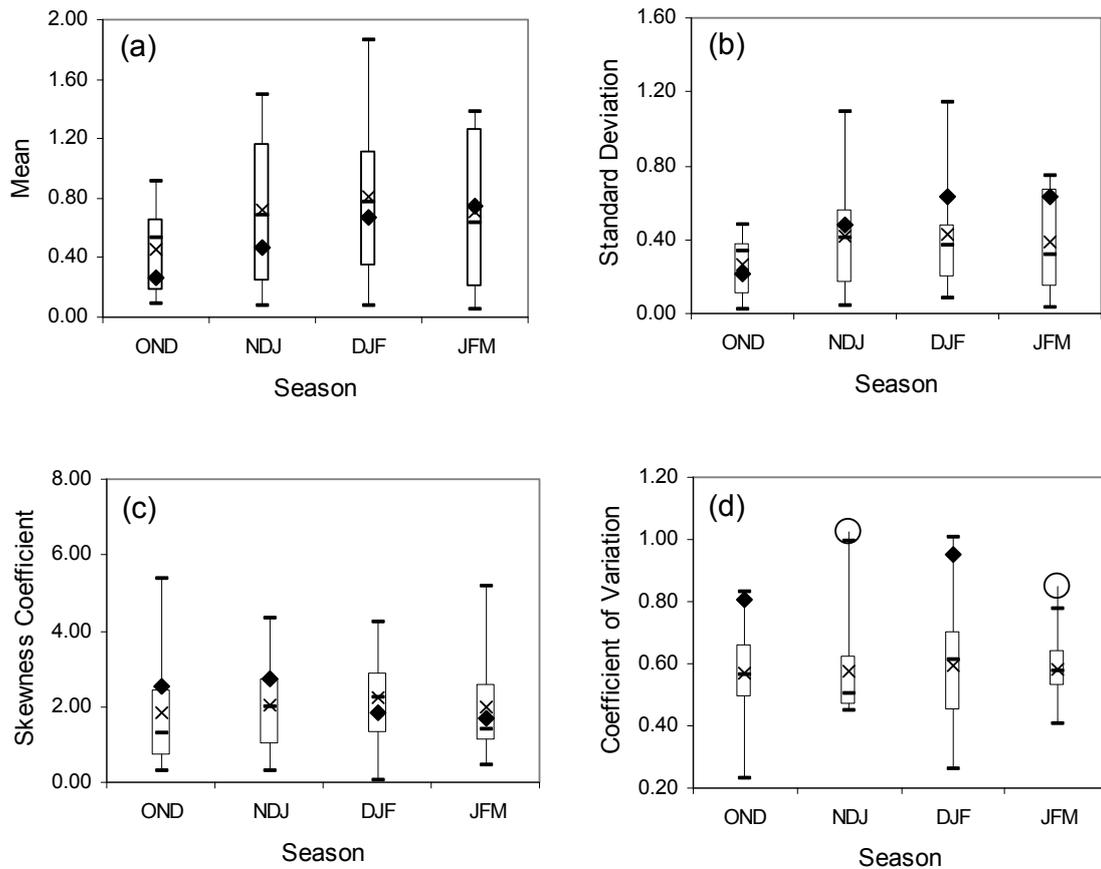
Figure 10.7 Box plots of statistics of generated streamflows along with the observed values at the mouth of the Mgeni catchment for OND, NDJ, DJF and JFM of 2003/04, using the ERBM. Diamonds represent the observed values with circles indicating the values outside the simulated range. The x signs represent the simulated mean values, and the box-and-whiskers represent the minimum, lower quartile, median, upper quartile and maximum of the forecasted streamflow sequences

The upper panels show the means and standard deviations of streamflows while the bottom panels show the skewness coefficients and coefficients of variation (CoV) for the same variable. Statistics of model performance for each season show that the ERBM derived streamflows reproduced the corresponding observed flows well with regard to means, standard deviations and skewness coefficients, but with a slight over-estimation of the observed means during the first three seasons and a slight under-estimation of observed standard deviations during the

last three seasons (cf. Figure 10.7 (a) and (b)). Consequently, the dispersion of the streamflows with respect to the mean values (COV) was consistently under-estimated, with COV values for the NDJ and JFM appearing just outside the simulated inter-quartile ranges (Figure 10.7 (d)).

To extend the analysis for the seasonal simulation of streamflows by *ACRU* from the ERBM derived rainfalls, ensembles of accumulated daily flows were constructed for each of the four forecast periods (Figure 10.8). From these ensembles of forecasts, the mean and median are computed as the "best" estimates that can be used to conduct alternative scenario analyses. Visually, the accumulated median streamflow values are much closer to the accumulated actual streamflow values than the accumulated mean values for the selected seasons, except for the JFM season.

In Figure 10.9 the $r^2$, bias, RMSE and MAE values between simulated (median) and actual accumulated streamflows are compared for the selected three month forecast periods. Even though forecast accuracy appeared to decline with the subsequent periods within the rainy season, the overall performance of the *ACRU* simulated flows is excellent. For OND, the accumulated simulated flows are in close agreement with the corresponding observed flows, especially for low flows, but with a slight bias (3.56). For NDJ and DJF the accumulated simulated flows mimicked the accumulated observed flows fairly well, although the simulated flows are slightly higher than the observed flows. For JFM, a significant correlation was achieved between the simulated and observed accumulated flows ($r^2$ = 0.96), but there is an appreciable bias in the simulation. The negative bias indicates a tendency for a consistent under-prediction of the observed flows.

The generated ensembles in Figure 10.8 were then applied to produce probabilistic forecasts of total streamflows for each of the selected three month forecast periods as shown in Figure 10.10.
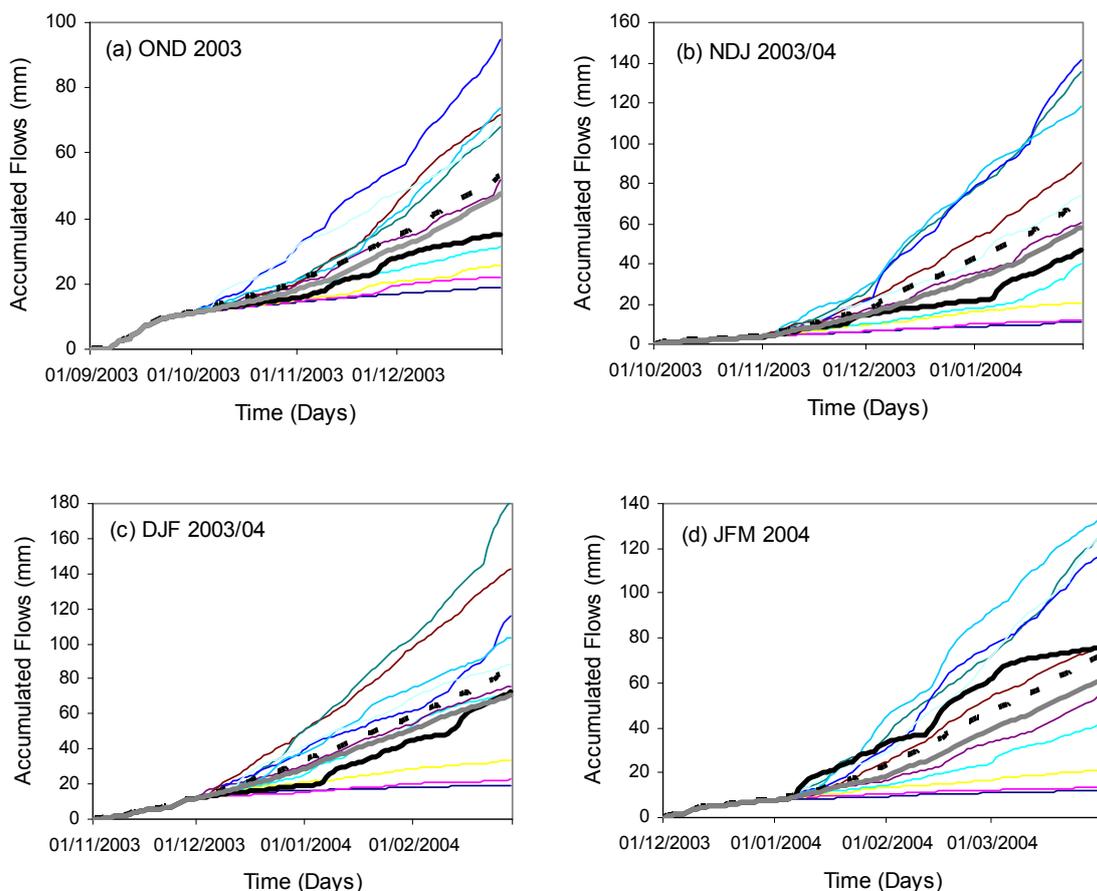
Figure 10.8    Forecasted versus observed accumulated flows at the mouth of the Mgeni catchment for OND (a), NDJ (b), DJF (c) and JFM (d) of 2003/04, using the ERBM, with the thick black solid line representing the observed accumulated streamflow, and the thick grey solid and dashed lines representing the forecasted accumulated median and average flows, respectively

The simulated cumulative probability of total streamflow for each of the selected three month forecast periods is represented by the dashed line, and the corresponding observed cumulative probability is represented by a solid line. As expected from Figure 10.9, the cumulative probability distributions for OND, NDJ, and DJF are positively biased, suggesting a wetter condition than the climatological cumulative probability. The cumulative probability distribution for JFM is, however, biased towards the drier side.
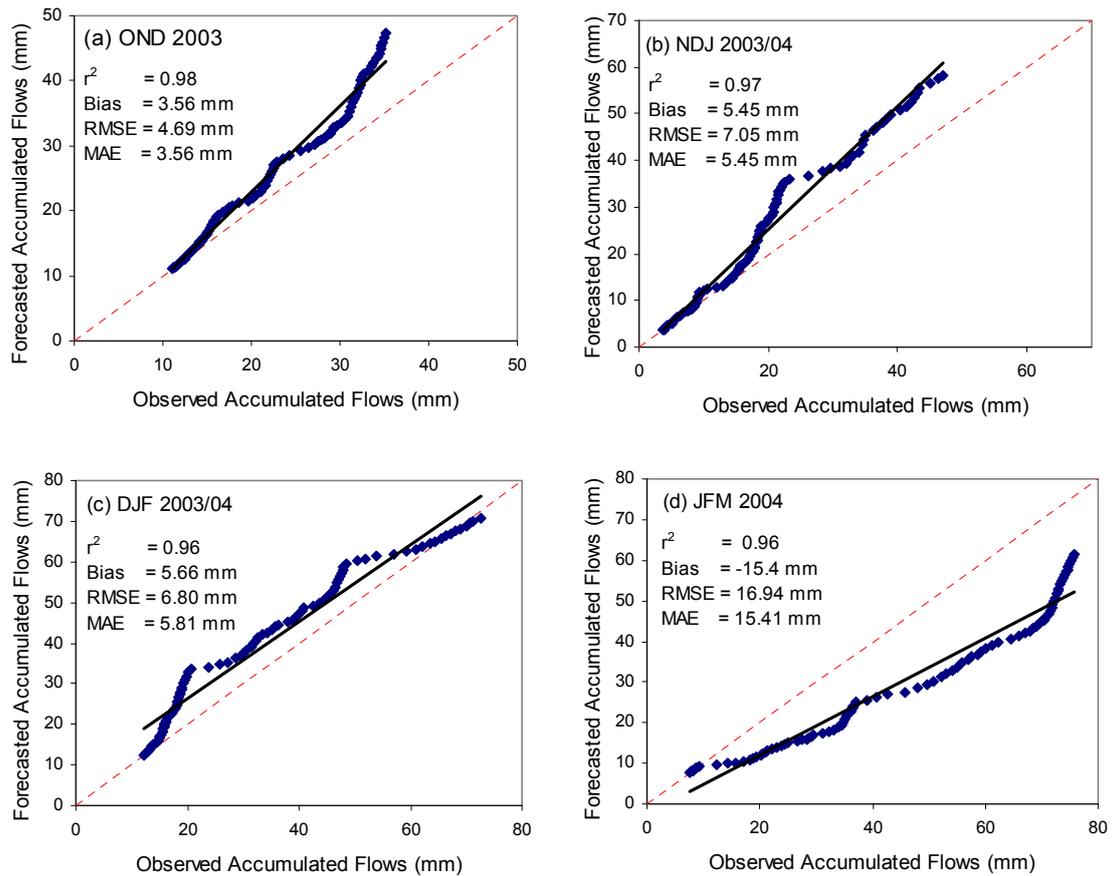
Figure 10.9    Plots of accumulated daily flows simulated with the *ACRU* model
versus observed flows at the mouth of the Mgeni catchment for OND
(a), NDJ (b), DJF (c) and JFM (d) of 2003/04

Extending the results, a comparison is made in Figure 10.11 of the three month
totals of daily *ACRU* simulated versus observed flows at the mouth of the Mgeni
catchment for the year 2004. The model simulations appear to over-estimate the
three month totals of daily observed streamflows for OND, NDJ and DJF, but
capture them well for JFM. The overlap of the months allowed the forecast errors
to persist through the subsequent three month periods.
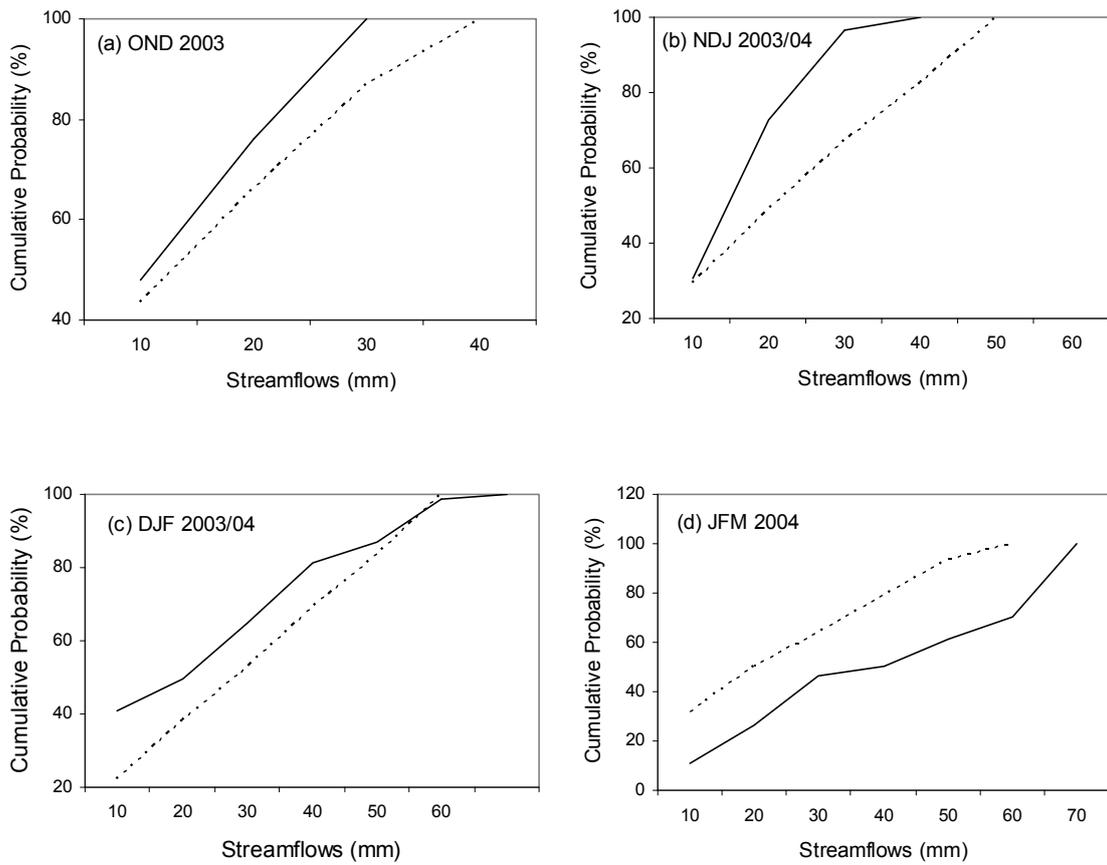
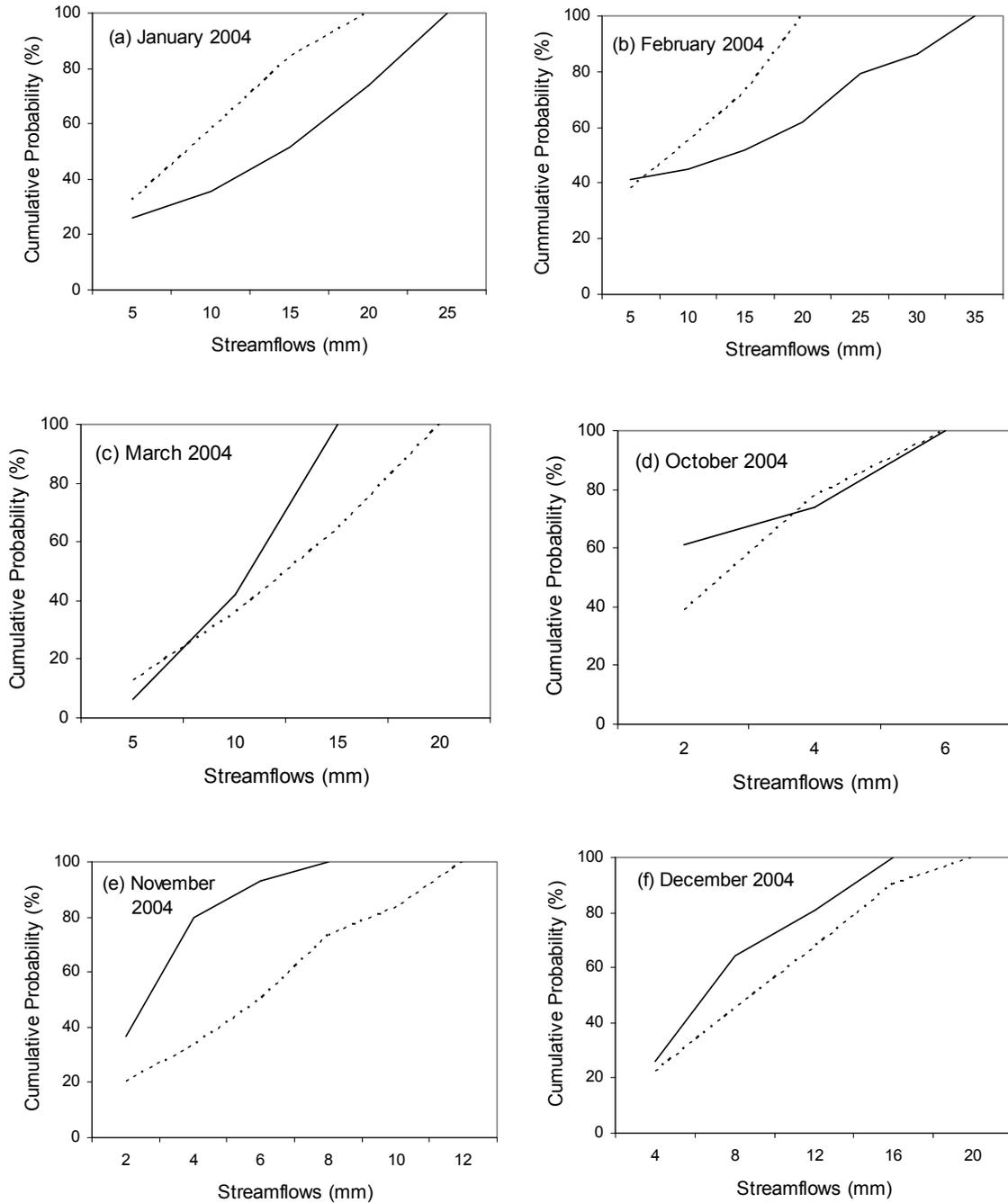Figure 10.10 Cumulative probabilities of accumulated daily flows simulated with the *ACRU* model versus observed flows at the mouth of the Mgeni catchment for OND (a), NDJ (b), DJF (c) and JFM (d) of 2003/04, using the ERBM. The dashed line represents the cumulative probability of observed streamflows, and the solid line represents the cumulative probability of forecasted streamflows
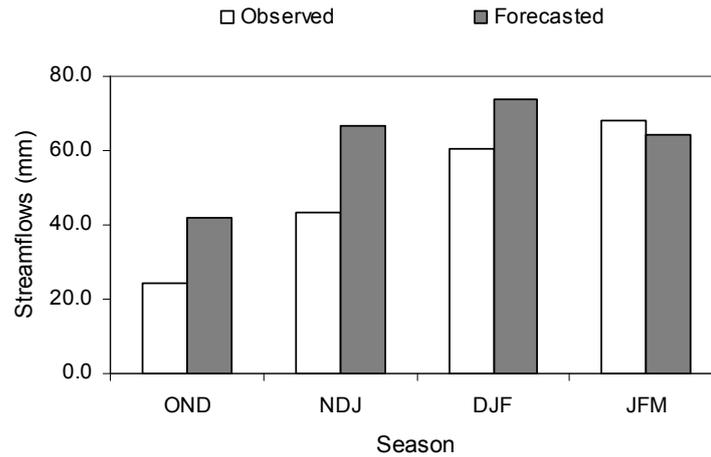
Figure 10.11 Three month totals of accumulated daily *ACRU* simulated versus observed flows (mm) at the mouth of the Mgeni catchment for 2003/04, using the ERBM

## 10.4  Conclusions

Seasonal climate forecasts for southern Africa are generally issued in a discrete, tercile format (Klopper and Landman, 2003). Such probabilistic forecasts cannot be applied in their published form in a hydrological/crop yield model that operates on a daily time step. Two approaches have been proposed to translate the triplet of probabilities into daily quantitative values. The first approach is to sample daily rainfall values from the same dates in selected analogue years, as described in Chapter 9. The other alternative, as demonstrated in this chapter, is to employ a conditional rainfall generator model. The model is designed to randomly generate ensembles of 10 members from selected analogue years for each forecast day, and uses the "Ensemble Re-ordering" (Clark *et al.,* 2004) as a post-processing step to reconstruct the temporal persistence of the synthetically generated daily rainfall data. The use of the ERBM is quite effective as a surrogate approach to generating stochastic realisations of daily rainfall series that resemble actual rainfall data with respect to a range of relevant statistics. The model thus has the property of reproducing the transitional probabilities of rain days and dry days as well as the persistence of dry and wet spells within forecast cycles, all of which are important in the evaluation and forecasting of streamflows and crop yields, as well as of droughts and floods.  The ensembles of simulated rainfall were used as input

into the *ACRU* model to generate an ensemble of simulated streamflows at monthly and seasonal time scales in the Mgeni catchment. Reasonably good results were obtained for most of the selected periods when simulating with the *ACRU* model, which indicates that the ERBM derived rainfalls are useful for various agrohydrological applications in South Africa, and possibly elsewhere.

Although the focus in this chapter was only on generating ensembles of conditional rainfall sequences, the ERBM can be extended to a multivariate weather generator for simulating of maximum and minimum temperatures or solar radiation.

# 11. SUMMARY, CONCLUSIONS AND RECOMMENDATIONS

## 11.1 Summary and Conclusions

The development of a framework for the application of near real time, plus daily to seasonal rainfall forecasts as a nested input to agrohydrological models, thereby enabling the forecasting of agrohydrological variables across a range of time scales and lead times, is a new concept in southern Africa. The framework which has been developed is a flexible one and is designed to include generic windows which facilitate the selection of near real time remotely sensed observations, as well as short, medium term and longer term forecasts supplied by various weather and climate models from different institutions across a range of spatial scales. The framework is capable of importing ASCII outputs of individual and merged rainfall fields estimated from a network of daily reporting raingauges, radars and satellite images, as well as rainfall forecasts provided by three Numerical Weather Prediction models, into GIS for spatial disaggregation, reformatting, data joining and finally for extracting *ACRU* formatted daily rainfall values to 1 946 hydrologically inter-linked Quaternary Catchments.

Furthermore, generic algorithms have been developed within the framework to translate categorical monthly and three month seasonal forecasts into a daily time series values suitable for use with agrohydrological models. Two types of methods can be used in translating the categorical forecasts. The first is the *Historical Sequence Method*, which is designed to sample daily rainfall values from the same dates in selected analogue years. The other is the *Ensemble Re-ordering Based Method*, which randomly generates ensembles of 10 members from selected analogue years for each forecast day, and uses the method as a post-processing step to reconstruct the temporal persistence of the synthetically generated daily rainfall data. The latter method is also designed to generate the transitional probabilities of rain days and dry days, as well as the persistence of dry and wet spells within forecast cycles, all of which are important in forecasting streamflows and crop yields, as well as droughts and floods.

The framework has been developed for applications on a real time basis. In this study, however, it was applied at the Mgeni catchment with an archive of historical

forecasts in order to verify the output of weather and climate models as well as the two temporal downscaling methodologies mentioned above. In the evaluation process, the key questions were:

- How accurate are the SIMAR products when compared to the observed reference? How did the uncertainty in these products translate into the streamflow uncertainty? Can they be used reliably for operational agrohydrological applications?

- What is the skill of the C-CAM and UM rainfall forecasts with respect to location, rainfall magnitude, lead times and when they are transformed to streamflows? What is the role of the initial hydrological conditions in affecting the skill of streamflows forecasts? How skilful is the combined use of the C-CAM and UM models when compared to the individual runs? To what extent does the "ensemble approach" explain the uncertainty with a particular NCEP-MRF forecast? Is an ensemble mean more skilful than individual members? Does the skill increase with decreasing lead time when compared to the reference?

- How effective are the *Historical Sequence Method* and the *Ensemble Re-ordering Based Method* in translating the categorical monthly and seasonal rainfall forecasts to daily values? How can hydrological ensemble forecasts obtained from ensembles of analogue years be verified? How skilful are the seasonal forecasts when compared to climatology? Do they preserve the statistical moments of historical time series from which the rainfall sequences are populated? What is the skill of the daily dry and wet spells generated by the *Ensemble Re-ordering Based Method*?

In an attempt to address these questions, a range of verification techniques has been employed and the following conclusions are drawn, based on the results obtained for each of the models embedded within the forecasting framework:

### 11.1.1 Conclusions on SIMAR Rainfall Fields

- The use of daily reporting raingauges as well as merged fields of satellite/raingauges and satellite/radars/raingauges provides relatively realistic rainfall results, without much difference in their hydrological

outputs, whereas the radar and raw satellite information by themselves cannot be used in operational hydrological application in their current status.

- The rainfall-streamflow transformation by the *ACRU* model had a dampening effect on the accumulated flows. As a result, the influence of rainfall errors on the streamflows was not comprehensively investigated in the SIMAR products. The reason for the dampening effect was the absence of high runoff producing rainfall events during the study period.

## 11.1.2 Conclusions on C-CAM, UM and NCEP-MRF  Rainfall Forecasts

- The C-CAM model is capable of identifying a rainfall event, but with a tendency of under-estimating its magnitude. For lead times up to 2 days there is an acceptable skill in the C-CAM forecasts, but for the 3 and 4 day lead times the skill is low and unreliable.

- The UM model is capable of identifying rainy days from non-rainy days, but with a significant over-estimation of rainfall amount. There is no significant difference between the 1 and 2 day lead time UM forecasts.

- The role of the initial hydrological conditions in affecting the skill of C-CAM and UM streamflows forecasts was significant. The results show that the under-estimation of the C-CAM forecasts was reduced from 34% to 10%, while the over-estimation in the UM forecasts was reduced from 291% to only 59% when the *ACRU* model was initialised with observed rainfalls up to the previous day at each forecast run within the study period.

- Owing to the spatial scale gap between the 2.5º gridded NCEP-MRF forecasts and Quaternary Catchments, these forecasts were not applied with the *ACRU* model. Results obtained for a continuous period of 92 days showed that the rainfall forecasts were significantly over-predicted and that the quality of the NCEP-MRF forecasts was seen to slightly decrease with increasing lead time. However, taking into account the scale gap and discrepancies between the forecast and observation times, a correlation coefficient of 0.46 is considered a fair result.

- The "ensemble approach" was successful in capturing the observation for only two out of the four selected events and there is no guarantee that the

ensemble mean is better than that of any single ensemble member. However, ensemble forecasts are very important in calculating Probabilistic Quantitative Precipitation Forecasts, which are often required in risk based decision making.

- The combined use of the C-CAM and UM models by a "weighted averaging" had an important effect in smoothing the respective extreme under- and over-estimations of rainfalls. However, the combined use is overshadowed more by the over-estimation of the UM forecasts than the under-estimation of the C-CAM forecasts.

### 11.1.3  Conclusions on Monthly and Seasonal Forecasts

- Reasonably good results were obtained in regard to $r^2$, bias, RMSE and MAE from both the *Historical Sequence Method* and the *Ensemble Re-ordering Based Method* for most of the selected periods, which indicates the soundness of applying either method in transforming skilful categorical seasonal forecasts into ensembles of daily quantitative rainfall values for application in hydrological/crop yield models.

- For the seasonal streamflow forecasts, the observed mean, standard deviation, skewness coefficient and coefficient of variation were simulated fairy well by both methods in most of the selected periods. Monthly streamflow forecasts for the year 2004 using the *Ensemble Re-ordering Based Method* also indicated the model to capture the observed statistics well for most of the months of the year except for the dry months from June to September.

- The *Ensemble Re-ordering Based Method* is capable of reproducing statistically acceptable transitional probabilities of rain days and dry days as well as their persistence within a given month and this is of great importance in agricultural operations.

It should be noted that the NWP models and methodologies conceptualised in this study were only tested on one catchment. The developments presented in this thesis may, therefore, be considered as a pilot study, laying the foundation for further research to be undertaken in making the framework more operational and more useful to decision makers responsible for water resources and agricultural

operations. The areas that need further research are outlined in the sub-section which follows.

## 11.2  Recommendations for Future Research

There are many scientific and social problems that might need to be addressed in the process of practical implementation of hydro-climatic forecasts. Based on the experiences in undertaking this study, the following recommendations for future research are made. These recommendations are outlined in the context of three broad categories:

### 11.2.1  Recommendations on Issues of Verification

- In this study, data obtained from raingauge networks was used for verifying the forecasts and it was implicitly assumed that the reference data were perfect. In reality, however, upscaling point observations to a resolution of a grid box is often subjected to spatial errors, even if there is dense raingauge coverage over an area, as was the case in the Mgeni catchment. For a fairer verification, therefore, it is highly recommended to use remotely sensed observations with a correction factor from raingauge measurements as an alternative for reference rainfall. Furthermore, observational uncertainty should be accounted for in the verification process.

- In order to assess the quality and reliability of the various forecasts, more research needs to be undertaken in extending the verification process to other parts of South Africa with different hydrological regimes to that of the Mgeni catchment. Verification against observed streamflows should also be explored by including the actual activities within the Mgeni catchment in the hydrological modelling. These include land uses and their influences on hydrological responses, irrigation demand and supply, abstractions from dams and return flows from urban and industrial areas as well as inter-catchment transfers.

- In order to gain more confidence, users need to assess the quality of forecasts, but care must be taken to avoid misinterpretation of valuable information.

### 11.2.2 Recommendations for Model Improvements

- Rainfall estimates derived from SIMAR are of great importance for defining the "now state" of a catchment at the beginning of a forecast period. However, more work is still required in improving the time of delivery of the daily rainfall maps in near real time. More research and development is also needed in improving the quality of the products, especially for rainfall estimates derived from satellite and radar measurements.

- Although rainfall forecasts from the C-CAM, UM and NCEP-MRF models have scored some successes, there is still a room for improvement with respect to their skill, especially for longer lead times and for distributing the information at a finer spatial scale that is comparable to relatively homogeneous hydrological response units and their respective dominant processes.

- A hybrid modelling strategy for Quantitative Precipitation Forecasts from two or more operational NWP models needs to be considered in order to eliminate the problems of systematic errors that often occur when a single model is used. The use of multi model ensembles for seasonal forecasting (e.g. Doblas-Reyes *et al.,* 2005) should also be considered in the future.

- The input file menu in *ACRU* model needs to be reconfigured in order to accommodate rainfall forecasts for rapid simulation and self-updating (i.e. "hot starting") schemes.

- Further research is needed in the framework development to accommodate other weather variables (e.g. temperature, solar radiation) that are often required, especially for crop yield forecasting.

### 11.2.3 Recommendations for Practical Applications

- Updating the Quaternary Catchment rainfall database is required for the application of the *Historical Sequence Method* and the *Ensemble Re-ordering Based Method* on catchments other than the Mgeni,

- Research is needed in identifying the potential end users and into how the agrohydrological forecasts can be translated and distributed to them, as different users need different types of forecasts, formats and levels of

accuracy. Water managers, for example, may be more interested in the amount of runoff that flows into their reservoirs, while farmers may be more interested in the onset, timing and distribution of rainfall within a growing season.

- As an operational utility, one of the major and challenging problems is the speed of delivery from a forecasting centre to end users. A web based interface is therefore highly recommended for effective and real time transfer of information to decision makers. However, end users need to first understand the skill and limitation of the forecasts. Research is therefore needed on how to communicate the forecasts to end users so as they can decide for themselves whether to take the risk of using the forecast information or not.

- Finally, development of a simple decision support tool is recommended for potential users to assess their potential economic benefits and losses associated with the use of particular forecast information.

The author believes that addressing the above-mentioned recommendations would enhance this framework for short to longer term climate forecasts and could potentially maximises opportunities to improve management of climate related risks in agricultural and water related decision making.

In final conclusion, the primary scientific contribution of this study has been the development of a flexible, easily updatable and user friendly forecasting framework that is capable of bridging the gaps that exist between outputs of weather and climate models and their practical application in agrohydrological models over southern Africa. Findings obtained from the evaluation of the various weather/climate models and temporal downscaling methodologies in this study have also added new knowledge to the science of forecasting on issues related to the uncertainties that cascades through the translation of weather/climate forecasts into streamflow forecasts and forecast verification strategies. The scene is now set for operationalising this framework for integrated, time-varying agrohydrological forecasts over southern Africa.

# 12. REFERENCES

Acocks, J.P.H., 1988. Veld types of South Africa. Botanical Research Institute, Pretoria, RSA. pp 146.

Adiku, S.G.K., Dayananda, P.W.A., Rose, C.W. and Dowuona, G.N.N., 1997. An analysis of the within-season rainfall characteristics and simulation of the daily rainfall in two savanna zones in Ghana. *Agricultural and Forest Meteorology*, 86: 51-62.

Ahrens, B. and Jaun, S., 2007. On evaluation of ensemble precipitation forecasts with observation based ensembles. *Advances in Geosciences*, 10:139-144.

Anctil, F., Perrin, C. and Andreassian, V., 2003. ANN output updating of lumped conceptual rainfall\runoff forecasting models. *Journal of the American Water Resources Assciation*, 39: 1269-1279.

Anstee, S., 2004. Unpublished material. Application of numerical weather prediction to rapid environmental assessment. DSTO-GD-0403. DSTO Systems Sciences Laboratory, Department of Defence, Australia.

Arnell, N.W., 1999. The effect of climate change on hydrological regimes in Europe: A continental perspective. *Global Environmental Change*, 9: 5-23.

Baethgen, W.E., Meinke, H. and Gimenez, A., 2004. Unpublished material. Adaptation of agricultural production systems to climate variability and climate change: Lessons learned and proposed research approach. In: *Insights and Tools for Adaptation: Learning from Climate Variability,* NOAA-OGP, Washington DC, USA.

Banitz, E., 2001. Evaluation of short-term weather forecasts in South Africa. *Water SA,* 11: 489-498.

Barnston, A.G., Van den Dool, H.M., Zebiak, S.E., Barnett, T.P., Ji, M., Rodenhuis, D.R., Cane, M.A., Leetmaa, A., Graham, N.E., Ropelewski, C.R., Kousky, V.E., O'Lenic, E.A. and Liverzey, R.E., 1994. Long-lead seasonal forecasts – where do we stand? *Bulletin of the American Meteorological Society*, 66: 2097-2114.

Bartman, A.G., Landman, W.A. and Rautenbach, C.J. de W., 2003. Recalibration of general circulation model output to austral summer rainfall over southern Africa. *International Journal of Climatology*, 23: 1407-1419.

Basson, M.S., 1997. *Overview of Water Resources Availability and Utilisation in South Africa*. CPT Book Printers, Ltd., Cape Town, RSA.

Bekele, F., 1992. Unpublished material. Ethiopian use of ENSO information in its seasonal forecasts. National Meteorological Services Agency, Addis Abeba, Ethiopia.

Bezuidenhout, C.A., 2005. Development and evaluation of model-based operational yield forecasts in the South African sugar industry. Unpublished PhD Thesis, School of Bioresources Engineering and Environmental Hydrology, University of KwaZulu-Natal, Pietermaritzburg, RSA.

Blaikie, P., Cannon, T., Davies, I. and Wisner, B., 1994. Unpublished material. At risk-natural hazards, people's vulnerability, and disasters. Routledge, London, UK.

Bocchiola, D. and Rosso, R., 2006. The use of scale recursive estimation for short term quantitative precipitation forecast. *Physics and Chemistry of the Earth*, 31: 1228-1239.

Bogardi, J.J, Villagran, J.C., Birkmann, J., Renaud, F., Sakulski, D., Chen, X., Affeltranger, B., Mensa, A. and Kaplan, M., 2005. Vulnerability in the context of climate change. In: *An International Workshop Human Security and Climate Change*, Holmen Fjord Hotel, Asker, Norway. 21-23 June 2005.

Bohle, H., Dowing, T.E. and Watts, M., 1994. Climate change and social vulnerability: the sociology and geography of food insecurity. *Global Environmental Change*, 4: 38-48.

Budhakooncharoen, S., 2003. Unpublished material. Rainfall estimate for flood management using meteorological data from satellite imagery. Mahanakorn University of Technology, Bangkok, Thailand.

Chen, Z., Grasby, S.E. and Osadetz, K.G., 2002. Prediction of average annual groundwater levels from climate variables: an empirical model. *Journal of Hydrology*, 260: 102-117.

Chen, Z., Grasby, S.E. and Osadetz, K.G., 2004. Relation between climate variability and groundwater levels in the upper carbonate aquifer, southern Manitoba, Canada. *Journal of Hydrology*, 290: 43-62.

Chiew, F.H.S., Srikanthan, R., Frost, A.J. and Payne, E.G.I., 2005. Reliability of daily and annual stochastic rainfall data generated from different data lengths and data characteristics. In: eds. Zerger, A. and Argent, R.M., *MODSIM 2005 International Congress on Modelling and Simulation.* Modelling and Simulation Society of Australia and New Zealand, December 2005, pp. 1223-1229. ISBN: 0-9758400-2-9.

Chiew, F.H.S., Zhou, S.L. and McMahon, T.A., 2003. Use of seasonal streamflow forecasts in water resources management. *Journal of Hydrology*, 270: 135-144.

Clark, M.P., Gangopadhyay, S., Brandon, D., Werner, K., Hay, L., Rajagopalan, B. and Yates, B., 2004. A resampling procedure for generating conditioned daily weather sequences. *Water Resources Research*, 40: W04304, doi: 10.1029/2003WR002747.

Coburn, A.W., Sspence, R.J.S. and Pomonis, A., 1994. Unpublished material. Vulnerability and Risk Assessment. Cambridge Architectural Research Limited, Cambridge, UK.

Collischonn, W., Haas, R., Andereolli, I. and Tucci, C.E.M., 2005. Forecasting river Uruguay flow using rainfall forecasts from a regional weather-prediction model. *Journal of Hydrology*, 305: 87-98.

CRED, 2002. Online emergency events database: Available from: <http://www.cred.be/emdat/intro.html> [Accessed 10 June 2005].

Cunha, L.V., Oliveira, R.P., Nascimeto, J. and Ribeiro, L., 2005. Impacts of climate change on water resources: A case study on Portugal. In: *The Fourth Inter-Celtic Colloquium on Hydrology and Management of Water Resources*, Guimaraes, Portugal, July 11-14, 2005.

Curtis, D.C. and Humphery, J.H., 1995. Use of radar-rainfall estimates to model the January 9-10, 1995 floods in Sacramento, CA. Paper presented. In: *The 1995 Southwest Association of ALERT Systems Conference*, Tulsa, OK, USA.

DEAT, 2001. South African estuaries: Generalised land cover for the Mgeni catchment. Available from http://www.environment.gov.za [Accessed 13 August 2007].

Deyzel, I.T.H., Pegram, G.G.S., Visser, P.J.M. and Dicks, D., 2004. Spatial Interpolation and Mapping of Rainfall (SIMAR). Volume 2: Radar and Satellite Products. *WRC Report* No. 1152/1/04. Water Research Commission, Pretoria, RSA.

Doblas-Reyes, F.J., Hagedorn, R. and Palmer, T.N., 2005. The rationale behind the success of multi-model ensembles in seasonal forecasting-II. Calibration and combination. *Tellus*, 57A: 234-252.

Downing, T.E. and Ludeke, M., 2002. Social geographies of vulnerability and adaptation. In: eds. Reynolds, J.F. and Stafford, D.M.*, Global Desertification: Do Humans Cause Deserts?*, Chapter 14: 233-252, Dahlem University Press, Germany.

Downing, T.E., Patwardhan, A., Klien, R., Mukhala, E., Stephen, L., Winograd, M. and Ziervogel, G., 2003. Vulnerability assessment for climate adaptation. Technical Paper nN. 3. UNDP, New York, USA.

Dunsmore, S.J., Schulze, R.E. and Schmidt, E.J., 1986. Antecedent soil moisture in design stormflow estimation. *ACRU* Report No. 23. Department of Agricultural Engineering, University of Natal, Pietermaritzburg, RSA.

DWAF, 2001. Appendix 1: Mvoti to Mzimkulu Water Management Area Situation Assessment (February 2001): A report by A.J Wilson and Associates International CC Management Consultants, Pretoria, RSA.

DWAF, 2004. Mooi-Mgeni River Transfer Scheme Phase 2: Feasibility Study Environmental Impact Assessment. *DWAF Report No. P WMA 07/V20/00/0504*, Landscape Dynamics in association with Eco-Agent and Isquare, Pretoria, RSA.

Dyson, L.L. and Van Heerden, J., 2002. A model for the identification of tropical weather systems over South Africa. *Water SA*, 28: 249-258.

Dyson, L.L., Van Heerden, J. and Marx, H.G., 2002. Short term weather forecasting techniques for heavy rainfall. *WRC Report* No. 1011/1/02. Water Research Commission, Pretoria, RSA.

Ebert, E.E., 2001. Ability of a poor man's ensemble to predict the probability and distribution of precipitation. *Monthly Weather Review*, 129: 2461-2480.

Engelbrecht, F., 2005. Simulations of climate and climate change over southern and tropical Africa with the Conformal-Cubic Atmospheric model. In: ed. Schulze, R.E., *Climate Change and Water Resources in Southern Africa: Studies on Scenarios, Impacts, Vulnerabilities and Adaptation. WRC Report* 1430/01/06, Chapter 4, 57-74. Water Research Commission, Pretoria, RSA.

Engelbrecht, F., 2007. Personal communication. Meteorological Group, University of Pretoria, Pretoria, RSA. 14 August 2007.

ESRI, 2005. GIS and Mapping Software. ESRI® ArcMAP™ 8.2, Copyright © 1995-2007 ESRI, USA.

Fair, K.A, 1999. Unpublished material. Review of the operating rules of the Mgeni river system, *BKS Report No.* P671431, Pretoria, RSA.

Faniran, J.A., Ngceba, F.S., Bhat, R.B. and Oche, C.Y., 2001. An assessment of the water quality of the Isinuka springs in the Transkei region of Eastern Cape, republic of South Africa. *Water SA*, 27: 241-250.

FAO, 2004. Drought impact mitigation and prevention in the Limpopo River Basin. Land and Water Discussion Paper 4. Viale delle Terme di Careacalla, Rome, Italy.

Feddersen H, and Andersen, U., 2005. A method for statistical downscaling of seasonal ensemble predictions. *Tellus* A, 57: 398-408.

Federico, S., Avolio, E., Bellecci, C. and Colacino, M., 2006. The application of LEPS technique for quantitative precipitation forecast (QPF) in southern Italy. *Advances in Geosciences*, 7: 1-8.

Fischer, G., Shah, M. and van Velthuizen, H., 2002. Climate change and agricultural vulnerability. A special report by the International Institute for Applied Systems Analysis as a contribution to World Summit on Sustainable Development, Johannesburg 2002. Remaprint, Vienna, Austria.

Franchini, M., 1996. Use of a genetic algorithm combined with a local search method for the automatic calibration of conceptual rainfall-runoff models. *Hydrological Science*, 4: 21-39.

FSIEWS, 2000. Handbook for defining and setting up a food security information and early warning system. FAO, UN, Rome, Italy.

Ganguly, A.R and Bras, R.L., 2001. Distributed quantitative precipitation forecasting using information from radar and numerical weather prediction models. *Journal of Hydrometeorology*, 4: 1168-1180.

Garbrecht, J.D. and Schneider, J.M., 2004. Long-term variability of Oklahoma precipitation and water resources availability. USDA, Agricultural Research Service, Grazinglands Research Laboratory, El Reno, OK, USA.

Ghile, Y.B., 2004. An adaptation of the SCS-*ACRU* hydrograph generating technique for application in Eritrea. Unpublished MSc Dissertation, School of Bioresources Engineering and Environmental Hydrology, University of KwaZulu-Natal, Pietermaritzburg, RSA.

Glantz, M.H., 1996. Forecasting El Nino: Science's gift to the 21st century. In: *A Workshop on Reducing Climate-Related Vulnerability in Southern Africa*, Victoria Falls, Zimbabwe. 1-4 October 1996.

Gleick, P.H., 2000. Water: The potential consequences of climate variability and change for the water resources of the United States. Available from: <http://pacinst.org> [Accessed 26 March 2005].

Golding, B.W., 2000. Quantitative precipitation forecasting in the UK. *Journal of Hydrology*, 239: 286-305.

Goswami, M. and O' Connor, K.M., 2006. Real time flow forecasting in the absence of quantitative precipitation forecasts: A multi-model approach. *Journal of Hydrology*, 334: 125-140.

Goswami, M., O' Connor, K.M., Bhattarai, K.P. and Shamseldin, A.Y., 2005. Assessing the performance of eight real time updating models and procedures for the Brosna river. *Hydrology and Earth System Sciences*, 9: 394-411.

Grigg, N.S., 2000. Risk management in water utilities. Available from: <http://www.engr.colostate.edu> [Accessed 07 August 2005].

Habets, F., Lemoigne, P. and Noilhan, J., 2004. On the utility of operational precipitation forecasts to served as input for streamflow forecasting. *Journal of Hydrology*, 293: 270-288.

Haines, A. T., Finlayson, B.L. and McMahon, T.A., 1988. A global classification of river regimes. *Applied Geography*, 8: 255-272.

Hallowes, J.S., 2002. Evaluation of a methodology to translate rainfall forecasts into runoff forecasts for South Africa. Unpublished MSc Dissertation, School of Bioresources Engineering and Environmental Hydrology, University of Natal, Pietermaritzburg, RSA.

Hammer, G.L. and Nicholls, N.N., 1996. Managing for climate variability – the role of seasonal climate forecasting in improving agricultural systems. In: *Proceedings 2nd Australian Conference on Agricultural Meteorology.* Bureau of Meteorology, Melbourne, Australia. pp 19-27.

Hammer, G.L., Hansen, J.W., Philips, J.G., Mjelde, J.W., Hill, H., Love, A. and Potgieter, A., 2001. Advances in application of climate prediction in agriculture. *Agricultural Systems*, 70: 515-553.

Hammer, G.L., Holzworth, D.P. and Stone, R., 1996. The value of skill in seasonal forecasting to wheat crop management in a region with high climatic variability. *Australian Journal of Agricultural Research*, 47: 717-737.

Hansen, J.W., 2002. Realizing the potential benefits of climate prediction to agriculture: Issues, approaches, challenges. *Agricultural Systems*, 74: 309-330.

Hartmann, H.C., Bales, R. and Sorooshian, S., 1999. Weather, climate, and hydrologic forecasting for the Southwest U.S. CLIMAS Report Series, CL2-99, Institute for the Study of Planet Earth. The University of Arizona, Tucson, AZ 85721, USA.

Hartmann, H.C.,Pagano, T.C., Sorooshian, S. and Bales, R., 2002. Confidence builders: Evaluating seasonal climate forecasts from user perspectives. *Bulletin of American Meteorological Society*, 83: 683-698.

Hastenrath, S., Greischar, L. and Van Heerden, J., 1995. Prediction of the summer rainfall over South Africa. *Journal of Climate*, 8: 1511-1518.

Hay, L.E. and Clark, M.P., 2003. Use of statistically and dynamically downscaled atmospheric model output for hydrologic simulations in three mountainous basins in the western United States. *Journal of Hydrology*, 282: 56-75.

Hobbs, J.E., 1980. *Applied Climatology: A Study of Atmospheric Resources*. Dawson Westview press, Folestone, Kent, England.

Hobson, A.N., 1997. Use of a stochastic weather generator in a watershed model for streamflow simulation. Unpublished MSc Dissertation, Department of Civil, Environmental and Architectural Engineering, University of Colorado, CO, USA.

Hossain, A.N.H., 2003. Climate forecast applications in Bangladesh for water related disaster mitigation. In: *International Conference on Total Disaster Risk Management*, Bangladesh. 23 December 2003.

Hossain, M.S.N., 2001. Assessing human vulnerability due to environmental change: Concepts and assessment methodologies. Unpublished MSc Dissertation, Department of Civil and Environmental Engineering, Royal Institute of Technology, Stockholm, Sweden.

Hoyt, W.G. and Langbein, W.B., 1995. *Floods*. Princeton University Press, Princeton, USA.

Hurrell, J., 1995. Decadal trends in the North-Atlantic oscillation: Regional temperatures and precipitation. *Science*, 269: 676-679.

IPCC, 1996. Climate change 1995: In: ed. Watson, R.T., Zinyowera, M.C. and Moss, R.H., *Impacts, Adaptations and Mitigations of Climate Change*: *Scientific-Technical Analyses*. Cambridge University Press, New York, USA.

IPCC, 2001. Third Assessment Report of the Intergovernmental Panel on Climate Change. Parts 1, 2, and 3. Synthesis Report and Policy Makers Summaries. Cambridge University Press, Cambridge, UK.

ISDR, 2004. Living with risk, a global review of disaster reduction initiatives. Available from: <http://www.unisdr.org> [Accessed 23 November 2005].

Jolliffe, I.T. and Stephenson, D.B., 2003. Categorical events. In Ed. Jolliffe I.T. and Stephenson, D.B., *Forecast Verification: A Practitioner's Guide in Atmospheric Science.* Ch. 1, John wiley and Sons Ltd., Chichester, England.

Jones, J.W., Hansen, J.W., Royce, F.S. and Messina, C.D., 2000. Potential benefits of climate forecasting to agriculture. *Agriculture, Ecosystems and Environment*, 82: 169-184.

Jordan, P., Seed, A., May, P. and Keenan, T., 2004. Evaluation of dual polarization radar for rainfall-runoff modelling –A case study in Sydney, Australia. Paper presented. In: *Sixth International Symposium on Hydrological Applications of Weather Radar*, Melbourne, Australia. 2-4 February 2004.

Jordan, P., Seed, A. and Weinmann, P.E., 2000. Errors in radar measurements of rainfall. In: *Proceedings of Hydro 2000 – 3rd International Hydrology and Water Resources Symposium*, Perth, Australia. pp 421-426.

Jose, A.M., Francisco, R.V. and Cruz, N.A., 1996. A study of impact of climate variability/change on water resources in the Philippines. *Chemosphere*, 33: 1687-1704.

Kabat, P. and Bates, B., 2002. The evidence In: Ed. Appleton, B*., Climate Changes the Water Rules: How Water Managers can Cope with Today's Climate Variability and Tomorrow's Climate Change.* Ch.1, 2 to 11, Printfine Ltd., Liverpool, UK.

Kamarianakis, Y., Chrysoulakis, N. and Feidas, H., 2006. Comparing rainfall estimates derived from raingauges and satellite images at the eastern Mediterranean region. Paper Presented. In: *9th AGILE Conference on Geographic Information Science*, Visegràd, Hungary.

Katzfey, J.J and McGregor, J.L., 2003. Unpublished material. High-resolution weather predictions for America's Cup in Auckland: A blend of model forecasts, observations and interpretation. CSIRO Atmospheric Research, Aspendale, Vic, Australia.

Kershaw, R., 2006. Parameterization of atmospheric processes in the Unified Model. In: *18th BMRC Modelling Workshop 2006:* Presentations *The Australian Community Climate And Earth System Simulator (ACCESS) Challenges and Opportunities*. Bureau of Meteorology, Melbourne, Australia. 28 November - 1 December 2006.

Kienzle, S.W., Lorentz, S.A. and Schulze, R.E., 1997. Hydrology and Water Quality of the Mgeni Catchment. *WRC Report* No. TT87/97. Water Research Commission Pretoria, RSA.

Kinuthia, J.H., 1999. Unpublished material. Global warming and climate impacts in southern Africa: How might things change? Kenya Meteorological Department, Nairobi, Kenya.

Kishore, K., 2002. An integrated climate risk management approach to disaster reduction and adaptation to climate change. *Paper presented at UNDP Expert Group Meeting on Integrated Disaster Reduction with Adaptation to Climate Change*, Havana, Cuba. 19-21 June, 2002.

Kjeldsen, T.R. and Rosbjerg, D., 2001. A framework for assessing the sustainability of a water resources system. Environment and Resources, Technical University of Denmark, DK-2800 Kongens Lyngby, Denmark.

Klopper, E. and Landman, W.A., 2003. A simple approach for combining seasonal forecasts for southern Africa. *Meteorology Applications*, 10: 319-327

Klopper, E., 1999. The use of seasonal forecasts in South Africa during the 1977/98 rainfall season. *Water SA,* 25: 311-314.

Kroese, N.J., 2004. Spatial Interpolation and Mapping of Rainfall (SIMAR). Volume 1: Maintenance and Upgrading of Radar and Raingauge Infrastructure. *WRC Report* No. 1151/1/04. Water Research Commission, Pretoria, RSA.

Kundzewicz, Z., 2001. Coping capcity for extreme events. Paper presented. In: *An International Conference on Freshwater,* Bonn, Germany. 3-7 December, 2001.

Kunz, R.P., 1993. Techniques to assess possible impacts of climate change in southern Africa. Unpublished MSc Dissertation, Department of Agricultural Engineering, University of Natal, Pietermaritzburg, RSA.

Lall, U., Rajagopalan, B. and Tarboton, D.G., 1996. A nonparameteric wet/dry spell model for resampling daily precipitation. *Water Resources Research*, 32: 2803-2823.

Landman, W. and Goddard, L., 2005. Predicting southern African summer rainfall using a combination of MOS and perfect prognosis. *Geophysical Research Letters*, 32: L15809, doi: 10.1029/2005GL022910.

Landman, W.A. and Klopper, E., 1998. 15-year simulation of the December to March rainfall season of the 1980s and 1990s using canonical correlation analysis (CCA). *Water SA,* 24: 281-286.

Landman, W.A. and Mason, S.J., 1999. Operational long-lead prediction of South African rainfall using canonical correlation analysis. *International Journal of Climatology*, 19: 1073-1090.

Landman, W.A., Botes, S., Goddard, L. and Shongwe, M., 2005. Assessing the predictability of extreme rainfall seasons over southern Africa. *Geophysical Research Letters*, 32: L23818, doi: 10.1029/2005GL023965.

Landman, W.A., Mason, S.J., Tyson, P.D. and Tennant, W.J., 2001. Statistical downscaling of GCM simulations to streamflow. *Journal of Hydrology*, 252: 221-236.

Laurent, H., Jobard, I. and Toma, A., 1998. Validation of satellite and ground based estimates of precipitation over the Sahel. *Atmospheric Research*, 47-48, 651-670.

Letcher, R.A., Chiew, F.H.S. and Jakeman, A.J., 2004. An assessment of the value of seasonal forecasts in Australian farming systems. Environmental Studies, the Australian National University, Australia.

Lettenmaier, D.P. and Wood, E.F., 1993. Hydrologic forecasting. In: Ed. Maidment, D.R., *Handbook of Hydrology*. McGraw-Hill, Inc. New York, USA.

Levitt, A.M.,1997. *Disaster Planning and Recovery: A Guide for Facility Professionals*. John Wiley & Sons. New York, USA.

Livezey, R.E., 2003. Categorical events. In Eds. Jolliffe I.T. and Stephenson, D.B., *Forecast Verification: A Practitioner's Guide in Atmospheric Science.* Ch. 4, John Wiley and Sons Ltd., Chichester, England.

Lumsden, T.J., 2000. Development and evaluation of a sugarcane yield forecasting system. Unpublished MSc Dissertation, School of Bioresources Engineering and Environmental Hydrology, University of Natal, Pietermaritzburg, RSA.

Madsen, H. and Jacobsen, T., 2001. Automatic calibration of the MIKE SHE integrated hydrological modelling system. In: *4th DHI Software Conference*, Scanticon Conference Centre, Helsingør, Denmark. 6-8 June, 2001.

Maidment, D.R., 1993. *Handbook of Hydrology*. McGraw-Hill, Inc., New York, USA.

Mailier, P.J., Jolliffe, I.T. and Stephensom, D.B., 2006. Unpublished material. Quality of weather forecasts: Review and recommendations. Royal Meteorological Society, Reading, UK.

Maini, P., Kumar, A., Singh, S.V. and Rathore, L.S., 2004. Operational model for forecasting location specific quantitative precipitation and probability of precipitation over India. *Journal of Hydrology*, 288: 170-188.

Mason, M., 2000. Definitions of technical terms in forecast verification examples of forecast verification scores. In: *Proceedings of the Workshop on Forecast Quality*, Palisades, New York, USA. October 10, 2000.

Mason, S.J., 1990. Temporal variability of sea surface temperatures around southern Africa: A possible forcing mechanism for the 18-year rainfall oscillation? *South African Journal Science*, 86: 243-252.

Mason, S.J., 1998. Seasonal forecasting of South African rainfall using a non-linear discriminant analysis model. *International Journal of Climatology*, 18: 147-164.

Mason, S.J., Joubert, A. M., Cosijn, C. and Crimp, S.J., 1996. Review of seasonal forecasting techniques and their applicability to southern Africa. *Water SA,* 22: 203-206.

Mecklenburg, S., Joss, J. and Schmid, W., 2000. Improving the nowcasting of precipitation in an alpine region with an enhanced radar echo tracking algorithm. *Journal of Hydrology*, 239:46-48.

Moura, A.D., Bengtsson, L., Buizer, J., Busalacchi, A., Cane, M.A., Lagos, P., Leetmaa, A., Matsuno, T., Mooney, K., Morel, P., Sarachik, E.S., Shukla, J., Sumi, A. and Patterson, M., 1992. International Research Institute for Climate Prediction: A proposal. 1100 Wayne Avenue, Suite 1225, Silver Spring, MD 20910, USA, pp 51 (cited by Landmann *et al.,* 2001).

Murphy, A.H., 1993. What is a good forecast? An essay on the nature of goodness in weather forecasting. *Weather Forecasting*, 8: 281-293.

Murphy, S.J., Washington, R., Downing, T.E., Martin, R.V., Ziervogel, G., Preston, A., Todd, M., Butterfield, R. and Briden, J., 2001. Seasonal forecasting for climate hazards: Prospects and responses. *Natural Hazards*, 23: 171-196.

Nurmi, P., 2003. Recommendations on the verification of local weather forecasts (at ECMWF Member States). Consultancy Report, Reading, UK.

NWRS, 2002. Unpublished material. South Africa's water situation, and strategies to balance supply and demand. Proposed First Edition National Water Resource Strategy, DWAF, Pretoria, RSA.

Ogallo, L.A., Boulahya, M.S. and Keane, T., 2000. Applications of seasonal to interannual climate prediction in agricultural planning and operations. *Agricultural and Forest Meteorology*, 103: 159-166.

Olesen, J.E. and Bindi, M., 2002. Consequences of climate change for European agricultural productivity, land use and policy. *European Journal of Agronomy*, 16: 239-262.

Pappenberger, F., Beven, K.J., Hunter, N.M., Bates, P.D., Gouweleeuw, B.T., Thielen, J. and De Roo, A.P.J., 2005. Cascading model uncertainty from medium range weather forecasts (10 days) through a rainfall-runoff model to flood inundation predictions within the European flood forecasting system (EFFS). *Hydrology and Earth System Sciences*, 9: 381-393.

Pegram, G.G.S, 2004. Spatial Interpolation and Mapping of Rainfall (SIMAR). Volume 3: Data Merging for Rainfall Map Production. *WRC Report* No. 1153/1/04. Water Research Commission, Pretoria, RSA.

Pegram G.G.S., Deyzel I.T.H., Sinclair S., Visser P., Terblanche D. and Green G.C., 2004. Daily mapping of 24 hr rainfall at pixel scale over South Africa using satellite, radar and raingauge data. In: *2nd International Precipitation Working Group (IPWG) Workshop*, Naval Research Laboratory, Monterey, CA, USA. 25-28 October 2004.

Pegram, G.G.S. and Sinclair, D.S., 2002. A linear catchment model for real time flood forecasting. *WRC Report* No. 1005/1/02. Water Research Commission, Pretoria, RSA.

Perks, L.A., 2001. Refinement of modelling tools to assess potential agrohydrological impacts of climate change in South Africa. Unpublished PhD Thesis, School of Bioresources Engineering and Environmental Hydrology, University of Natal, Pietermaritzburg, RSA.

Piccolo, F. and Chirico, G.B., 2005. Sampling errors in rainfall measurements by weather radar. *Advances in Geosciences*, 2: 151-155.

Piechota, T.C., Chiew, F.H.S., Dracup, J.A. and McMahon, T.A., 1998. Seasonal streamflow forecasting in eastern Australia and the El Niňo-Southern Oscillation. *Water Resources Research*, 34: 3035-3044.

Pielke, R.A., 2000. Policy responses to El Nino 1997-1998. In: Ed. Changnon, S.A., *El Niňo 1997/1998*. The Climate Event of the Century, Oxford University Press, New York. pp 172-196.

Plate, E.J., 2002. Risk management for hydraulic systems under hydrological loads. In: Eds. Bogardi, J.J. and Kundzewicz, Z.W.*, Risk, Reliability, Uncertainty and Robustness of Water Resources Systems*. International Hydrology Series, Chapter 23: 209-220, Cambridge University Press, UK.

Podestá, G., Letson, D., Messina, C., Royce, F., Ferreyra, R. A., Jones, J., Hansen, J.W., Llvoet, I., Grondona, M. and O'brien, J.J., 2002. Use of ENSO-related climate information in agricultural decision making in Argentina: A pilot experience. *Agricultural Systems*, 74: 371-392.

Potts, J.M, Folland, C.K., Jolliffe, I.T. and Sexton, D., 1996. Revised 'LEPS' scores for assessing climate model simulations and long-range forecasts. *Journal of Climate*, 9: 34-43.

Rajagopalan, B. and Lall, U., 1999. A k-nearest-neighbor simulator for daily precipitation and other weather variables. *Water Resources Research*, 35: 3089-3101.

Reason, C.J.C., Engelbrecht, F., Landman, W.A., Lutjeharms, J.R.E., Piketh, S., Rautenbach, H. and Hewitson, B.C., 2006. A review of South African research in atmospheric science and physical oceanography during 2000-2005. *South African Journal of Science*, 102: 35-45.

Rebora, N., Ferraris, L., Von Hardenberg, J. and Provenzale, A., 2005. Stochastic downscaling of LAM predictions: An example in the Mediterranean area. *Advances in Geosciences*, 2: 181-185.

Ritchie, J.W., Zammit, C. and Beal, D., 2004. Can seasonal climate forecasting assist in catchment water management decision making? A case study of the border rivers catchment in Australia. *Agriculture, Ecosystems and Environment,* 104: 553-565.

Roads, J., 2004. Experimental weekly to seasonal, global to regional US precipitation forecasts. *Journal of Hydrology*, 288: 153-169.

Rowlston, W., 2003. Water policy makers' needs on climate variability and climate change in South Africa. In: Ed. Schulze, R.E., *Managing Water Related Issues on Climate Variability and Climate Change in South Africa. ACRUcons Report* 44, pp. 53-56*,* School of Bioresources Engineering and Environmental Hydrology, University of KwaZulu-Natal, Pietermaritzburg, RSA.

Rural Development Services, 2002. Unpublished material. Agricultural assessment of the Mgeni municipality. Metroplan, Pietermaritzburg, RSA.

Rutulis, M., 1989. Groundwater drought sensitivity of southern Manitoba. *Canadian Water Resources Journal*, 14: 18-33.

Sarachik, E.S., 1996. Climate Prediction and the Ocean. In: *A Workshop on Reducing Climate-Related Vulnerability in Southern Africa*, Victoria Falls, Zimbabwe. 1-4 October 1996.

Sauchyn, D.J., 2000. Climatic variability and its implications for sustainable agriculture. Agri-Food Innovation Fund Project. *Report No.* 96000473, Department of Geography, University of Regina, Saskatchewan, Canada.

SAWS, 2005. Seasonal forecast for southern Africa. Long Range Forecasting Group. Available from: <http://www.weathersa.co.za> [Accessed 10 June 2005].

Schäfer, N.W. and van Rooyen, P.G., 1993. Operational applications of the Mgeni systems analysis. In: Eds. Lorentz, S.A., Kienzle S.W. and Dent, M.C., *Proceedings of the Sixth South African National Hydrological Symposium,* 385-394. University of Natal, Pietermaritzburg, RSA. 8-10 September, 1993.

Schmidli, J., Frei, C. and Vidale, P.L., 2006. Downscaling from GCM precipitation: A benchmark for dynamical and statistical downscaling methods. *International Journal of Climatology*, 26: 679-689.

Schneider, J.M. and Garbrecht, J.D., 2003. Temporal disaggregation of probabilistic seasonal climate forecasts. In: *Proceedings American Meteorological Society, 14th Symposium on Global Change and Climate Variations*, Long Beach, CA, USA. February 9-13, 2003.

Schulz, R, Peall, S.K., Dabrowski, J.M. and Reinecke, A., 2001. Current-use insecticides, phosphates and suspended solids in the Lourens river, Western Cape during the first rainfall event of the wet season. *Water SA*, 27: 65-70.

Schulze, R.E., 1984. Hydrological models for application to small rural catchments in southern Africa: Refinement and development. *WRC Report* No. 63/2/84. Water Research Commission, Pretoria, RSA.

Schulze, R.E., 1995a*. Hydrology and Agrohydrology*. *A Text to Accompany the ACRU 3.00 Agrohydrological Modelling System*. *WRC Report* TT69/95, Ch.1, AT1-1 to AT1-13, Water Research Commission, Pretoria, RSA.

Schulze, R.E., 1995b. Streamflow. In: Ed. Schulze, R.E., *Hydrology and Agrohydrology: A Text to Accompany the ACRU 3.00 Agrohydrological Modelling System*. *WRC Report* TT69/95, Ch.10, AT10-1 to AT10-6, Water Research Commission, Pretoria, RSA.

Schulze, R.E., 1997. Impacts of global climate change in a hydrologically vulnerable region: Challenges to South African hydrologists. *Progress in Physical Geography*, 21: 113-116.

Schulze, R.E., 1998. Hydrological modelling concepts and practice. Unpublished lecture notes. School of Bioresources Engineering and Environmental Hydrology, University of Natal, Pietermaritzburg, RSA.

Schulze, R.E., 2001. Risk, hazards and vulnerability within a context of hydrological risk management. Example from South Africa. In: *Proceedings 10th South African National Hydrology Symposium*. School of Bioresources Engineering and Environmental Hydrology, University of Natal, Pietermaritzburg, RSA.

Schulze, R.E., 2002. Development of methodologies for operational scale hydrological forecasting for water resources management in the Thukela catchment, South Africa. *A Research Proposal to the World Meteorological Organisation*. School of Bioresources Engineering and Environmental Hydrology, University of Natal, Pietermaritzburg, RSA.

Schulze, R.E., 2003. Risk, hazards and vulnerability within a context of hydrological risk management: A conceptual framework and examples from South Africa In: Ed. Schulze, R.E., *Modelling as a Tool in Integrated Water Resources Management: Conceptual Issues and Case Study Applications. WRC Report* 749/1/02, Ch. 6, 109-137, Water Research Commission, Pretoria, RSA.

Schulze, R.E., 2005. Development of an integrated time-varying agrohydrological forecast system for southern Africa: A Research Proposal to the Water Research Commission. In: *A Workshop of Water Utilization in Agriculture.* Water Research Commission, Pretoria, RSA. 29 April, 2005.

Schulze, R.E., 2006. South African Atlas of Climatology and Agrohydrology. *WRC Report 1489/1/06,* Water Research Commission, Pretoria, RSA.

Schulze, R.E., Angus, G.R. and Guy, R.M., 1995b. Soils. In: Ed. Schulze, R.E.*, Hydrology and Agrohydrology: A Text to Accompany the ACRU 3.00 Agrohydrological Modelling System. WRC Report* TT69/95, Ch.5, AT5-1 to AT5-40, Water Research Commission, Pretoria, RSA.

Schulze, R.E., Dent, M.C., Lynch, S.D., Schäfter N.W., Kienzle, S.W. and Seed, A.W., 1995a. Rainfall. In: Ed. Schulze, R.E., *Hydrology and Agrohydrology: A Text to Accompany the ACRU 3.00 Agrohydrological Modelling System. WRC Report* TT69/95, Ch.3, AT3-1 to AT3-38, Water Research Commission, Pretoria, RSA.

Schulze, R.E., Hallowes, J., Lynch, S.D,. Perks, L.A. and Horan, M., 1998. Forecasting seasonal streamflow in South Africa: A preliminary investigation. In: Ed. Joubert, A.M., *Forecasting Rainfall and Streamflow over South and Southern Africa. Research Report No.* TRR/T98/046, ESKOM Technology Group, Rocheville, RSA.

Schulze, R.E., Kiker, G.A. and Kunz, R.P., 1993. Global climate change and agricultural productivity in southern Africa: Thought for food and food for thought. *Global Environmental Change*, 3: 330-349.

Schulze, R.E. and Kunz, R.P., 1995. Reference potential evaporation. In: Ed. Schulze, R.E., *Hydrology and Agrohydrology: A Text to Accompany the ACRU 3.00 Agrohydrological Modelling System. WRC Report* TT69/95, Ch.3, AT4-1 to AT3-38, Water Research Commission, Pretoria, RSA.

Schulze, R.E., Lorentz, S.A., Kienzle, S.W. and Perks, L.A., 2004. Case Study 3: Modelling the impacts of land-use and climate change on hydrological responses in the mixed underdeveloped/ developed Mgeni catchment South Africa. *In*: Eds. Kabat, P., Claussen, M., Dirmeyer, P.A., Gash, J.H.C., Bravo deGuenni, L., Meybeck, M., Pielke, R.A Snr., Vörösmarty, C.J., Hutejes, R.W.A. and Lűtkemeier, S. *Vegetation, Water, Humans and the Climate. A New Perspective on an Interactive System*. Springer-Verlag, Heidelberg, Germany. Chapter D7, 441-453.

Schulze, R.E. and Perks, L.A., 2000. Assessment of the Climate Change on Hydrology and Water Resources in South Africa. *ACRUcons Report* 33, pp. 92*,* School of Bioresources Engineering and Environmental Hydrology, University of KwaZulu-Natal, Pietermaritzburg, RSA.

Schulze, R.E. and Smithers, J.C., 2004. The *ACRU* agrohydrological modelling system as of 2002: Background, concepts, structure, output, typical applications and operations. *WRC Report No.* 749/1/04, Water Research Commission, Pretoria, RSA.

Scott, A.J. and Collopy, F., 1992. Error measures for generalizing about forecasting methods: Empirical comparisons. *International Journal of Forecasting*, 8: 69-80.

Semenov, M.A., Brooks, R.J., Barrow, E.M. and Richardson, C.W., 1998. Comparison of the WGEN and LARS-WG stochastic weather generators for diverse climates. *Climate Research*, 10: 95-107.

Seo, D.J., Breidenbach, J., Fulton, R. and Miller, D., 2000. Real time adjustment of range-dependent biases in WSR-88D rainfall estimates due to nonuniform vertical profile of reflectivity. *Journal of Hydrometeorology,* 1: 222-240.

Sharma, A. and Lall, U., 1999. A non parametric approach for daily rainfall simulation. *Mathematics and Computers in Simulation*, 48: 361-371.

Smith, K., 1996. *Environmental Hazards*. Routledge, London, UK.

Smithers, J.C. and Schulze, R.E., 1995. *ACRU Agrohydrological Modelling System: User Manual Version 3.00*. *WRC Report* TT70/95, Water Research Commission, Pretoria, RSA.

Smithers, J.C. and Schulze, R.E., 2004. *ACRU Agrohydrological Modelling System: User Manual Version 4.00*. School of Bioresources Engineering and Environmental Hydrology, University of KwaZulu-Natal, Pietermaritzburg, RSA.

Smithers, J.C. and Schulze, R.E., Lynch, S.D., Hallowes, L.A., Thornton-Dibb, S.L.C., Pike, A. and Rivers-Moore, N.A., 2004. In: Eds. Smithers, J.C. and Schulze, R.E., *ACRU Agrohydrological Modelling System: User Manual Version 4.00.* School of Bioresources Engineering and Environmental Hydrology, University of KwaZulu-Natal, Pietermaritzburg, RSA.

Smithers, J.C., Schulze, R.E., Pike, A. and Jewitt, G.P.W., 2001. A hydrological perspective of the February 2000 floods: A case study in the Sabie river catchment. *Water SA*, 27, 325-332.

Soveri, J. and Ahlberg, T., 1989. Multiannual variations of the groundwater level in Finland during the year 1962-89. Paper presented. In: *A Conference on Climate and Water,* Helsinki, Finland. 11-15 September, 1989.

Sun, X, Mein, R.G., Keenan, T.D. and Elliott, J.F., 2000. Flood estimation using radar and raingauge data. *Journal of Hydrology*, 239: 4-18.

Taljaard, J.J., 1994. Atmospheric circulation systems, synoptic climatology and weather phenomena of South Africa. Part 1: Controls of the weather and climate of South Africa. *South African Weather Bureau Technical Paper*, 27: 1-45.

Taljaard, J.J., 1996. Atmospheric circulation systems, synoptic climatology and weather phenomena of South Africa. Part 6: Rainfall in South Africa. *South African Weather Bureau Technical Paper*, 32: 1-98.

Tarboton, K.C. and Schulze, R.E., 2003. Hydrological consequences of development scenarios for the Mgeni catchment. In: Eds. Lorentz, S.A., Kienzle S.W. and Dent, M.C., *Proceedings of the Sixth South African National Hydrological Symposium,* University of Natal, Pietermaritzburg, RSA. pp 297-304.

Tennant, W.J., 2005. Personal communication. South African Weather Service, RSA. 05 June 2005.

Tennant, W.J., 2007. Personal communication. South African Weather Service, RSA. 14 July 2007.

Tennant, W.J., Toth, Z. and Rae, K.J., 2006. Application of the NCEP ensemble prediction system to medium-range forecasting in South Africa: New products, benefits and challenges. Submitted to *Weather and Forecasting*, Revised: June 9, 2006.

Toth, E., Brath, A. and Montanari, A., 2000. Comparison of short-term rainfall prediction models for real time flood forecasting. *Journal of Hydrology*, 239: 132-147.

Toth, Z., Kalnay, E., 1997. Ensemble forecasting at NCEP and the breeding method. *Monthly Weather Review*, 125: 3297-3319.

Toth, Z., Kalnay, E., Tracton, S. M., Wobus, R. and Irwin, J., 1997: A synoptic evaluation of the NCEP ensemble. *Weather and Forecasting*, 12: 140-153.

Toth, Z., Mendez, M.P. and Vintzileos, A., 2006. Bridging the gap between weather and climate forecasting: Research priorities for intra-seasonal prediction. In: *A Workshop*, 27 April 2006, Naval Weather Service, National Center for Environmental Prediction, MD, USA.

Toth, E., Montanari, A. and Brath, A., 1999. Real time flood forecasting via combined use of conceptual and stochastic models. *Physics and Chemistry of the Earth (B),* 24: 793-798.

Toth, Z., Talagrand, O. and Zhu, Y. 2005. The attributes of forecast systems: A framework for the evaluation and calibration of weather forecasts. In: Eds. Palmer, T.N. and Hagedorn, R., *Predictability of Weather and Climate*. Cambridge University Press, New York, USA.

Toth, Z., Zhu, Y., Marchok, T., Tracton, S.M. and Kalnay, E., 1998. Verification of the NCEP global ensemble forecasts. Paper presented. In: *12th Conference on Numerical Weather Prediction*, Phoenix, Arizona, USA. 11-16 January 1998.

Tychon, B, Balaghi, R. and Jlibene, M., 2003. Unpublished material. Risk management in agricultural water use. Foundation Universitaire Luxmbourgeoise, Belgium and Institute National de la Recherche Agronomique, Morocco.

Tyson, P.D., 1990. Modelling climatic change in southern Africa: A review of available methods. *South African Journal of Science*, 86: 318-330.

UK Met Office, 2007. The Unified Model. Submodels. Available from: < http://www.metoffice.gov.uk> [Accessed 17 July 2007].

UKCIP, 2003. Climate adaptation: Risk, uncertainty and decision making. In: Eds. Willows, R.I. and Connell, R,K., *UKCIP Technical Report*, Part 2, 41-52, UKCIP, Oxford, UK.

UNDP, 2002. A climate risk management approach to disaster reduction and adaptation to climate change. Havana, Cuba, 19-21 June 2002.

UNEP, 1992. *A Report on the 3rd workshop of the UNEP Working Group on ENSO and Climate Change*. Bangkok, Thailand. 4-7 November,1991.

UNEP, 1997. *Global Environmental Outlook*. United Nations. Oxford University Press, New York, USA.

Van Hemert, L., 2007. Personal communication. South African Weather Service, RSA. 23 September 2007.

Van Jaarsveld A.S. and Chown S.L., 2001. Climate change and its impacts in South Africa. *Trends in Ecology and Evolution*, 16: 13-14.

Van Kleef, H., 2003. Temporal variations in groundwater levels and rainfall in the upper aquifer unit of the Langebaan road aquifer. In: Ed. Schulze, R.E., *Managing Water Related Issues on Climate Variability and Climate Change in South Africa. ACRUcons Report No.* 44*,* pp 89-94, School of Bioresources Engineering and Environmental Hydrology, University of KwaZulu-Natal, Pietermaritzburg, RSA.

Van Zyl, D., 2003. *South African Weather and Atmospheric Phenomena*. Printed by Creda, Cape Town, RSA.

Webster, P.J. and Grossman, R., 2003. Unpublished material. Forecasting river discharge into Bangladesh on short, medium and long time scales. Georgia Institute of Technology, Atlanta, Georgia, USA.

Wei, M., Z. Toth, R. Wobus, Y. Zhu, C. Biship, 2005: Initial perturbations for NCEP ensemble forecast system. In: *Proceedings for the First THORPEX Internal Science Symposium*, Montreal, Canada. 6-10 December 2004.

Wilhite, D.A. and Glantz, M.H., 1985. Understanding the drought phenomenon: The role of definitions. *Water International*, 10: 111-120.

Wilhite, D.A., 2000. *Drought: A Global Assessment*. Natural Hazards and Disasters Series. Routledge Publishers, London, UK.

Wilhite, D.A., Hayes, M.J., Knutson, C. and Smith, K.H., 2000. Planning for drought: Moving from crisis to risk management. *Journal of American Water Resources Association*, 36: 697-710.

Wilks, D.S., 1995. *Statistical Methods in the Atmospheric Sciences.* San Diego, CA, USA: Academic Press.

Wilks, D.S., 2002. Realizations of daily weather in forecast seasonal climate. *Journal of Hydrometeorology*, 3: 195-207.

WMO, 1992. International Meteorological Vocabulary. Report No. 182, Geneva, Switzerland.

Wood, A.W., Leung, L.R., Sridhar, V. and Lettenmaier, D.P., 2004. Hydrological implications of dynamical and statistical approaches to downscaling climate model outputs. *Climatic Change,* 62: 189-216.

WWAP, 2003. Water for People Water for Life. The United Nations World Water Development Report. UNESCO-WWAP, Barcelona, Spain.

Xiong, L., O'Connor, M. and Guo, S., 2004. Comparison of three updating schemes using artificial neural networks in flow forecasting. *Hydrology and Earth System Sciences*, 8: 247-255.

Yapo P., Sorooshian, S. and Gupta, V., 1993. A Markov chain flow model for flood forecasting. *Water Resources Research*, 29: 2427-2436.

Yates, D. Gangopaddhyay, S., Rajagopalan, B. and Strezepek, K., 2003. A technique for generating regional climate scenarios using a nearest neighbor algorithm. *Water Resources Research*, 39, 1199, doi: 10.1029/2002Wr001769.

Yucel, I., Kuligowski, R.J. and Gochis D. J., 2004. Evaluating the performance of satellite rainfall estimates using data from NAME program. In: *Proceedings of the 29th Annual Climate Diagnostics and Prediction Workshop*, Madison, WI, USA. 20-24 October 2004.

Zhang, H. and Casey, T., 1999. Verification of categorical probability forecasts. *Weather and Forecasting*, 35: 80-89.

Zhu, Y., Toth, Z., Wobus, R., Richardson, D. and Mylne, K., 2002. The economic value of ensemble-based weather forecasts, *Bulletin of the American Meteorological Society*, 83:73-83.

Zorita, E. and von Storch, H., 1999. The analog method as a simple statistical downscaling technique: Comparison with more complicated methods. *Journal of Climate*, 12: 2474-2489.