

The Development of an Integrated System for Drought Monitoring and Prediction in South Africa

Report to the
Water Research Commission

by

SJL Mallory¹, P Wolski², B Abiodun², T Sawunyama³, J Tanner⁴

¹IWR Water Resources

²Climate System Analysis Group

³Inkomati-Usuthu Catchment Management Agency

⁴Institute for Water Research

WRC Report No. 2746/1/21

ISBN 978-0-6392-0290-7

October 2021



**WATER
RESEARCH
COMMISSION**

Obtainable from

Water Research Commission
Private Bag X03
Gezina, 0031
Pretoria
Republic of South Africa

orders@wrc.org.za or download from www.wrc.org.za

DISCLAIMER

This report has been reviewed by the Water Research Commission (WRC) and approved for publication. Approval does not signify that the contents necessarily reflect the views and policies of the WRC, nor does mention of trade names or commercial products constitute endorsement or recommendation for use.

EXECUTIVE SUMMARY

Droughts can have a devastating effect on the economy of a country as well as the livelihoods and quality of life of communities, especially those with limited resources or ability to adapt. South Africa is especially prone to drought due to our highly variable rainfall, both seasonally and over longer cycles of above and below average rainfall.

An early warning system would allow action to be taken timeously so as to mitigate the effects of drought where possible. The ultimate purpose of this project was to develop such an integrated system to monitor and predict droughts so that action can be taken timeously to manage the impact of droughts.

There are several components to drought, such as, rainfall, streamflow, reservoir storage and groundwater. These hydrological and climatic aspects cannot be viewed in isolation since they are linked within in a complex system. An understanding of the cause and effect and interaction between these aspects is essential to develop a comprehensive and integrated drought warning system.

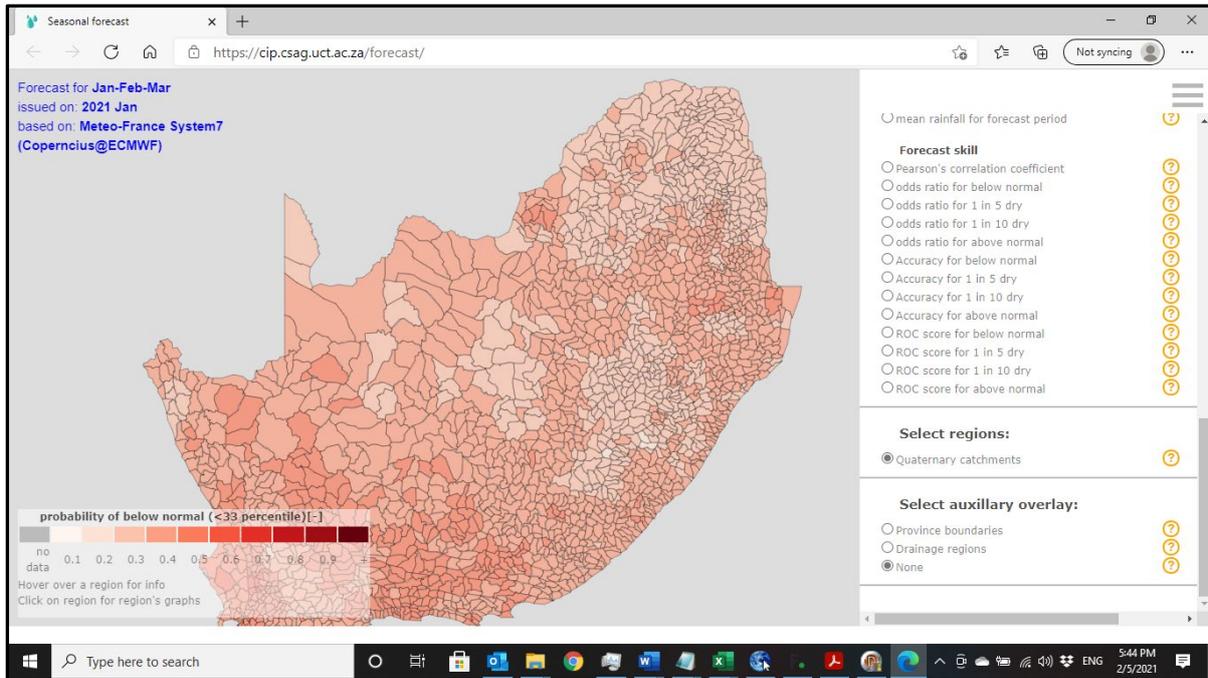
The driver behind the hydrological cycle is rainfall, and hence monitoring and prediction of rainfall is key to predicting streamflow, reservoir storage, groundwater and soil moisture.

Two different approaches to rainfall forecast at seasonal time scale were developed and are reported on, namely:

- Numerical climate forecast based on climate models (section 4.3)
- Statistical forecast based on rainfall monitoring and persistence of current rainfall anomaly

Numerical climate forecast data from 10 General Circulation Models (GCMs) has been used as input into hydrological models to simulate and predict streamflow and soil moisture. The forecast skill of these forecasts was evaluated and it was found that there is significant skill for forecasts up to three months into the future. Due to the chaotic nature of weather systems, forecasts beyond three months does not have significant skill. The GCM climate data is statistically downscaled to the spatial scale of Rainfall Zones based on hydrological and climate datasets created within the Water Resources 2012 study (Bailey & Pitman, 2015, thereafter WR2012). In this study, the monthly rainfall obtained from WR2012 for the period of 1983-2009 was used.

The website <http://cip.csag.uct.ac.za/forecast/> was developed as part of this project to spatially present the GCM data and the statistical analysis of the data sets (see Figure A).



Streamflow is modelled with the aid of a Pitman type hydrological model which has been modified to carry out multiple runs to cater for rainfall input from GCMs. Every month, rainfall forecast values are replaced with observed rainfall from the CHIRPS satellite rainfall database. Hence the Pitman simulation provide simulated naturalised flow that is updated every month and forecast three months into the future.

Soil moisture is a key component of the Pitman model and hence soil moisture forecasts are an intrinsic component of the output from the hydrological model. The Sami Groundwater model has been incorporated into the modified Pitman model in order to forecast groundwater storage.

Storage in dams is modelled with a Reservoir Simulation Model using the forecast naturalised flows, catchment water use and abstractions from the dam as input to forecast storage three months into the future.

Four pilot studies were carried out to test the four components of the forecast, namely streamflow, reservoir storage, groundwater and soil moisture. It was concluded that forecasts of streamflow and storage are useful in that they are reducing the uncertainty relating to a stochastic analysis in which a much wider range of flows and hence storage is possible. The flow and storage forecasts can mostly be compared to observed flow and/or storage available from the Department of Water and Sanitation website while equivalent observed data for soil moisture and groundwater is not available. Soil moisture status and prediction are however available from the website <http://wxmaps.org/pix/soil10>. The use of the Gravity Recovery and Climate Experiment (GRACE) model was investigated as tool for monitoring groundwater but the resolution of the data is too coarse for quaternary scale monitoring. Borehole water level monitoring, on the other hand, only provides an estimate of the groundwater storage over the

small area representative of the borehole water level and is too limited to represent quaternary scale groundwater storage, typically not more than a few square kilometres.

A modelling User Interface was developed, referred to as the Integrated Drought Monitoring and Prediction System (see Figure B).

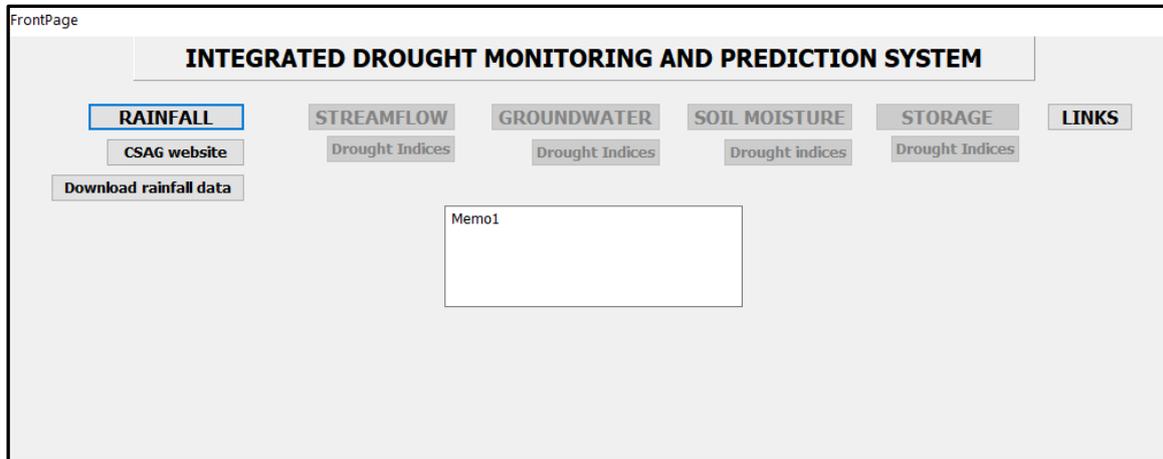


Figure B: Integrated Drought Monitoring and Prediction System

This application links to the CSAG rainfall website to assess and download rainfall forecasts. Streamflow, Groundwater and Soil moisture components are modelled using the recoded Pitman Model while storage is modelled using the Water Resources Modelling Platform. For each forecasting component, a monthly time step time series is produced from which drought indices are calculated.

The application also provided links to other useful forecasting websites.

The way forward with this forecasting and monitoring system is to implement it within the catchment operated by the Inkomati Catchment Management Agency (IUCMA) and within selected catchments operated by the Department of Water and Sanitation's Stand Alone Dams project.

Acknowledgements

The input from the Reference Group members is gratefully acknowledged.

Mr Wandile Nomquphu	: Water Research Commission (Chair)
Dr Tendai Sawunyama	: Inkomati-Usuthu Catchment Management Agency
Dr Eddie Riddell	: South African National Parks
Dr Beason Mwaka	: Department of Water and Sanitation
Ms Celiwe Ntuli	: Department of Water and Sanitation
Dr Willem Landman	: University of Pretoria
Prof John Ndiritu	: University of the Witwatersrand
Prof Chris Reason	: University of Cape Town
Prof Babatunde Abiodun	: University of Cape Town
Ms Tsepang Makholela	: Department of Forestry, Fisheries and the Environment
Prof Coleen Vogel	: University of the Witwatersrand
Prof Andries J. Jordaan	: University of the Free State
Mr Mxolisi Mukhawana	: Department of Water and Sanitation
Dr Chris Lennard	: University of Cape Town
Mrs Andriette Combrink	: Zutari Africa

The authors wish to acknowledge Ms Wendy Williams for her proof reading and editing of the final report.

Capacity Building

A PhD student (Mr Ridick Roland Takong), registered at the University of Cape Town, has benefitted from the involvement in this project. As part of his PhD thesis, Mr Takong studied the capability of some climate simulation datasets in replicating rainfall characteristics in Southern Africa, with a focus on precipitation over the Drakensberg Mountain Range, which is the source of rivers that support socio-economic activities in Lesotho, South Africa, and Namibia. He identified the climate simulation datasets that give a realistic representation of rainfall characteristics over Southern Africa, especially over the Drakensberg Mountain Range. Mr Takong has been working on the thesis for the past four years. He will submit the thesis for examination in May 2021.

Ms Kopano Mokoena, registered at Rhodes University for an MSc in Geohydrology, contributed to this study by investigating the use of GRACE, a remote-sensing tool, for monitoring groundwater storage. While she concluded that this tool does not provide sufficient resolution for monitoring groundwater at quaternary catchment scale, she benefitted from her involvement in this study by gaining practical experience in groundwater monitoring and modelling. Ms Mokoena hopes to complete her MSc in 2021.

TABLE OF CONTENTS

Executive Summary	i
Acknowledgements.....	iv
Capacity Building.....	iv
List of Figures	viii
List of Tables	xi
1 INTRODUCTION	1
2 EXISTING DROUGHT MONITORING AND PREDICTION SYSTEMS	2
2.1 Introduction	2
2.2 Surface Water Monitor	3
2.3 African Flood and Drought Monitor.....	4
2.4 Global Integrated Drought Monitoring and Prediction System	5
3 RAINFALL PREDICTABILITY	7
3.1 Introduction	7
3.2 Drivers of southern Africa rainfall	8
4 SEASONAL RAINFALL FORECASTING APPROACHES FOR SA IN THE CONTEXT OF DROUGHT FORECAST.....	13
4.1 Introduction	13
4.2 Measures of forecast skill.....	14
4.2.1 Deterministic forecast of a continuous variable	14
4.2.2 Probabilistic forecast of a (binomial) categorical variable	15
4.2.3 Deterministic forecast of a (binomial) categorical variable	16
4.3 Numerical dynamical forecasting based on climate models	18
4.3.1 Statistically-calibrated (Model Output Statistics or MOS) dynamical seasonal forecast	19
4.3.2 Multi-model ensemble of dynamical seasonal forecast models	25
4.4 Statistical forecast of end-of-season anomalies based on anomaly persistence and teleconnections	34
4.4.1 Introduction	34
4.4.2 Forecast of rainfall anomaly based on linear regression between with current rainfall anomaly	36
5 STREAMFLOW AND STORAGE MONITORING AND PREDICTABILITY	45
5.1 Introduction	45
5.2 Short-term streamflow predictability	45
5.2.1 Literature review	45
5.2.2 Characterization of hydrological droughts	47
5.2.3 Application of the Standardised Runoff Index (SRI)	50
5.3 Long term predictability	51
5.3.1 Literature review	51
5.4 Storage prediction.....	52
6 SOIL MOISTURE AND GROUNDWATER PREDICTABILITY	53
6.1 Introduction	53
6.2 GRACE Mission Data.....	54

6.3	Groundwater modelling	57
6.3.1	Groundwater recharge	58
6.3.2	Flow to downstream catchments	58
6.3.3	Evaporation from the riparian zone	59
6.3.4	Discharge into streams (baseflow)	59
6.4	Conclusions and recommendations	61
7	INTEGRATION OF PREDICTION METHOLOGIES TOWARDS AN INTEGRATED DROUGHT MONITORING AND PREDICTION SYSTEM	62
7.1	General approach	62
7.2	Implementation and functionality of the four information pathways.....	64
7.2.1	Current drought status	64
7.2.2	Statistical forecast of the end of the season rainfall anomaly	65
7.2.3	Numerical rainfall forecast.....	65
7.2.4	Hydrological model implementation and structure	66
8	PILOT STUDIES	69
8.1	Introduction	69
8.1.1	Theewaterskloof.....	69
8.1.2	White River System.....	79
8.1.3	Karatara River.....	92
8.1.4	Gutshwa River	98
9	RESULTS AND DISCUSSION.....	105
9.1	Rainfall forecasting	105
9.2	Streamflow, soil moisture and groundwater forecasts	106
9.3	Storage forecasts.....	107
9.4	Pilot studies	107
9.4.1	Theewaterskloof Dam	107
9.4.2	White River Systems.....	107
9.4.3	Karatara River catchment	108
9.4.4	Gutshwa River catchment	108
10	CONCLUSIONS AND RECOMMENDATIONS.....	109
10.1	Rainfall prediction	109
10.2	Streamflow forecasts.....	109
10.3	Storage forecasts.....	110
10.4	Groundwater/soil moisture	110
10.5	Way forward.....	110
10.6	General recommendations	111
11	REFERENCES	112
APPENDIX A: USER MANUAL		120
1	Objectives of the website	122
2	Data sources.....	122
3	Products.....	122
3.1	Dynamical seasonal forecast.....	122
3.1.1	Background.....	122
3.1.2	Rainfall Forecast Data in Pitman model format	123
3.1.3	Forecast of aggregated 3-month rainfall.....	124

3.2	Current rainfall information based on monitoring data.....	128
3.3	Statistical forecast of rainfall anomaly at the end of the rainy season	129
3.3.1	Background.....	129
3.3.2	Forecast of end-of-season rainfall anomaly based on linear regression of current rainfall anomaly and additional explanatory variables	131
4	Basic functionality of the website.....	132
4.1	Contextual Menu	132
4.2	Interacting with maps.....	137
4.3	Graphs and figures	138
4.3.1	Graphs for dynamical forecast	138
4.3.2	Graphs for monitoring data	141
4.3.3	Graphs for statistical forecast.....	143

LIST OF FIGURES

Figure 2.1: Surface Water Monitor: Home Page.....	3
Figure 2.2: Surface Water Monitor: Rainfall forecast.....	3
Figure 2.3: African Flood and Drought Monitor Home Page.....	4
Figure 2.4: Overview of the Drought Monitoring and Forecasting components.....	5
Figure 2.5: GIDMaPS website.....	6
Figure 4.1: <i>Example ROC curve for probabilistic event forecast</i>	16
Figure 4.2: <i>Schematic of contingency table for binomial event forecast</i> ,.....	17
Figure 4.3: Summary of results of calibration and validation for a selected quaternary catchment using a split-sample approach.....	22
Figure 4.4: Skill measures for seasonal forecasts based on PCR MOS downscaling of SAWS (ECHAM 4.5) forecast to WR2012 rainfall data in four climatological seasons.....	23
Figure 4.5: Skill measures for seasonal forecasts based on PCR MOS downscaling of CFSv2 forecast to WR2012 rainfall data in four climatological seasons.....	24
Figure 4.6: Skill measures for seasonal forecasts based on PCR MOS downscaling of ECMWF forecast to WR2012 rainfall data. DJF season.....	25
Figure 4.7: Deterministic forecast skill.....	29
Figure 4.8: Probabilistic forecast skill (ROC score).....	30
Figure 4.9: As for Figure 4.8 but for forecast of 1 in 5 year drought.....	30
Figure 4.10: As for Figure 4.8 but for forecast of 1 in 10 year drought.....	31
Figure 4.11: As for Figure 4.7 but for the Western Cape only.....	31
Figure 4.12: As for Figure 4.7 but for the Limpopo only.....	32
Figure 4.13: Deterministic forecast skill (Pearson's correlation).....	32
Figure 4.14: Probabilistic forecast skill (ROC score).....	33
Figure 4.15: Illustration of relationship between end of season accumulated rainfall and current rainfall anomaly.....	34
Figure 4.16: Simple categorical forecast of end of the season anomaly of total rainfall based on the current anomaly of accumulated rainfall.....	36
Figure 4.17: Correlation between the end-of-season and current anomaly of accumulated rainfall in a sub-catchment in the winter rainfall region.....	36
Figure 4.18: Separation of variance explained by basic linear model in end of the season anomaly into components.....	38
Figure 4.19: Separation of variance in end of the season anomaly into components.....	39
Figure 4.20a: Regions of similar separation of variance explained by a regression model (with ENSO).....	40
Figure 4.20b: Regions of similar separation of variance explained by a regression model (with ENSO).....	41
Figure 4.21: As Figure.4.20, but for the winter rainfall regime.....	42
Figure 4.22: Skill of the statistical forecast of the end of the season rainfall anomaly expressed in terms of ROC score.....	43
Figure 4.23: As for Figure 4.22, but for winter rainfall regime.....	44
Figure 4.24: CSAG rainfall forecasting website.....	44
Figure 5.1: Illustration of the streamflow drought indices (Fundel et al., 2013).....	48
Figure 5.2: SDI Analysis at Crocodile River at Karino River flow station.....	51

Figure 5.3: SDI Analysis at the Berg River at Misverstand River flow station	51
Figure 5.4: Storage projections using an ARMA stochastic model.....	52
Figure 6.1: Grace model results from Houberg <i>et al.</i> (2012).....	56
Figure 6.2: Principles of groundwater flow.....	57
Figure 6.3: Simulated soil moisture derived from the Pitman model	60
Figure 6.4: Simulated groundwater storage.....	61
Figure 7.1: Schematic of the integrated drought warning system	63
Figure 8.1: Location of the Theewaterskloof system.....	70
Figure 8.2: Natural flow time series into the Theewaterskloof Dam (based on WR2012 hydrology)	71
Figure 8.3: Monthly distribution of natural flow	71
Figure 8.4: Transfers from Theewaterskloof Dam to Cape Town	72
Figure 8.5: Modelled natural flow using projected rainfall (January to April 3 month projections)	73
Figure 8.6: Modelled natural flow using projected rainfall (May to August 3 month projections)	74
Figure 8.7: Modelled natural flow using projected rainfall (September to December 3 month projections)	75
Figure 8.8: Modelled storage using projected natural flow (January to April 3 month projections)	76
Figure 8.9: Modelled storage using projected natural flow (May to August 3 month projections)	77
Figure 8.10: Modelled storage using projected natural flow (September to December 3 month projections).....	78
Figure 8.11: Location of the White River system	79
Figure 8.12: Natural flow time series into the Longmere Dam (based on IUCMA 2019)	81
Figure 8.13: Monthly distribution of natural flow	81
Figure 8.14: Modelled natural flow using projected rainfall (January to April 3 month projections)	83
Figure 8.15: Modelled natural flow using projected rainfall (May to August 3 month projections)	84
Figure 8.16: Modelled natural flow using projected rainfall (September to December 3 month projections)	85
Figure 8.17: Modelled storage using projected natural flow (January to April 3 month projections)	86
Figure 8.18: Modelled storage using projected natural flow (May to August 3 month projections)	87
Figure 8.19: Modelled storage using projected natural flow (September to December 3 month projections).....	88
Figure 8.20: Rainfall prediction versus observed rainfall (January to April 3 month projections)	89
Figure 8.21: Rainfall prediction versus observed rainfall (May to August 3 month projections)	90
Figure 8.22: Rainfall prediction versus observed rainfall (September to December 3 month projections)	91
Figure 8.23: Location of the Karatara River catchment.....	92

Figure 8.24: Natural flow time series at Gauge K4H002, Katara River	93
Figure 8.25: Monthly distribution of natural flow	93
Figure 8.26: Land use in the Karatara catchment.....	94
Figure 8.27: Modelled natural flow using projected rainfall (January to April 3 month projections)	95
Figure 8.28: Modelled natural flow using projected rainfall (May to August 3 month projections)	96
Figure 8.29: Modelled natural flow using projected rainfall (September to December 3 month projections)	97
Figure 8.30: Catchment of the Gutshwa River, Mpumalanga	98
Figure 8.31: Natural flow time series at out of the Gutshwa River catchment	99
Figure 8.32: Monthly distribution of natural flow	99
Figure 8.33: Modelled versus observed groundwater storage, expressed in dimensionless terms.....	100
Figure 8.34: Modelled groundwater storage using projected rainfall (January to April 3 month projection)	101
Figure 8.35: Modelled groundwater storage using projected rainfall (May to August 3 month projection)	102
Figure 8.36: Modelled groundwater storage using projected rainfall (September to December 3 month projection)	103
Figure 8.37: Rainfall in the Gutshwa catchment (10 year moving average).....	104

LIST OF TABLES

Table 2.1: SPI drought categories (Mckee et al., 1993).....	2
Table 3.1: Drivers of southern Africa rainfall variability at various spatial and temporal scales	9
Table 5.1: Crocodile River at Karino – river flow thresholds for drought monitoring.....	49
Table 5.2: Runoff Anomaly categories based on their anomaly percent	49
Table 5.3: Z-Index denoted drought	50
Table 8.1: Summary of climate and hydrology information for the upper Theewaterskloof catchment.....	70
Table 8.2: Water use in the Theewaterskloof catchment.....	71
Table 8.3: Dams in the White River System	80
Table 8.4: Summary of climate and hydrology information for the upper Theewaterskloof catchment.....	80
Table 8.5: Water use within the White River system.....	82
Table 8.6: Summary of climate and hydrology information for the K40C quaternary catchment.....	93
Table 8.7: Summary of climate and hydrology information of the Gutshwa River	98

ABBREVIATIONS AND ACRONYMS

AAO	Antarctic Oscillation
ARC	Agricultural Research Council
ARMA	Autoregressive Moving Average
AUC	Area Under Curve
CFS	Coupled Forecast System
CHIRPS	Climate Hazards Group Infrared Precipitation
CSAG	Climate System Analysis Group
DAS	Data Assimilation Scheme
DJF	December January February
DWS	Department of Water and Sanitisation
ECMWF	European Centre for Medium range Weather Forecasts
ENSO	El Niño-Southern Oscillation
GCM	General Circulation Model
GDI	Groundwater Drought Indices
GIDMaPS	Global Integrated Drought Monitoring and Prediction System
GPCC	Global Precipitation Climatology Centre
GRACE	Gravity Recovery And Climate Experiment
GWI	Groundwater Indices
IDMaPS	Integrated Drought Monitoring and Prediction System
ILRI	International Livestock Research Institute
ITCZ	Inter-Tropical Convergence Zone
IUCMA	Inkomati-Usuthu Catchment Management Agency
JJA	June July August
MAM	March April May
MCC	Mesoscale Convection Complexes
MJO	Madden Julian Oscillation
MOS	Model Output Statistics
MSDI	Multivariate Standardised Drought Index
NMME	North American Multi Model Ensemble
NOAA	National Oceanic and Atmospheric Administration
ORSS	Odds Ratio Skill Score
PCA	Principal Component Analysis
PCR	Principal Component Regression
PDO	Pacific Decadal Oscillation
PDSI	Palmer Drought Stress Index
ROC	Receiver Operating Characteristic
SAA	South Atlantic Anticyclone
SAM	Southern Annular Mode
SAWS	South African Weather Service
SDI	Streamflow Drought Index
SGI	Standardised Groundwater Index
SIA	South Indian Anticyclone

SICZ	South Indian Convergence Zone
SIOD	Subtropical Indian Ocean Dipole
SON	September October November
SPI	Standardised Precipitation Index
SPI	Standardised Precipitation Index
SRI	Standardised Runoff Index
SSI	Standardised Streamflow Index
SST	Sea Surface Temperature
SWM	Surface Water Monitor
TIOD	Tropical Indian Ocean Dipole
TTT	Tropical Temperate Trough
TWS	Total Water Storage
WR2012	Water Resources 2012 Study
WReMP	Water Resources Modelling Platform
WRVCB	White River Valley Conservation Board

This page was intentionally left blank

1 INTRODUCTION

Droughts can have a devastating effect on the economy of a country as well as the livelihoods and quality of life of communities, especially those with limited resources or ability to adapt. South Africa is especially prone to drought due to our highly variable rainfall, both seasonally and over longer cycles of above and below average rainfall. An early warning system would allow action to be taken timeously so as to mitigate the effects of drought where possible. The ultimate purpose of this project is to develop such an integrated early warning system taking all aspects of drought into account, such as, rainfall, streamflow, storage and groundwater. These hydrological and climatic aspects cannot be viewed in isolation since they are linked within in a complex system. An understanding of the cause and effect and interaction between these aspects is essential to develop a comprehensive and integrated drought warning system.

The focus of this first report is on hydrological predictability. With the exception of desalination, all runoff and groundwater sources in southern Africa are derived from rainfall, hence the starting point in developing a drought warning system is to attempt to predict or at least reduce uncertainty with regard to rainfall prediction. Streamflow is strongly correlated to rainfall while the state of groundwater and evapotranspiration potential also influence streamflow. Groundwater as a water resource is also prone to droughts and hence must form part of any drought early warning system.

This report highlights various aspects of hydrological predictability and makes a recommendation on the way forward with the development of an integrated drought warning system.

- Rainfall
- Streamflow
- Reservoir storage
- Groundwater
- Soil moisture

2 EXISTING DROUGHT MONITORING AND PREDICTION SYSTEMS

2.1 Introduction

A literature review of drought monitoring and prediction systems revealed that there several regional drought monitoring systems that have been developed recently while there is also one global drought monitoring system. The three systems which are well documented and which were reviewed are:

- Surface Water Monitor System (USA)
- African Drought Monitor, and
- GIDMaPS (Global)

These systems all utilise the same general approach for monitoring, which is to use rainfall, streamflow and soil moisture indices, these are defined as the difference in the parameter (e.g. rainfall, streamflow, soil moisture) from the mean over a specified time period. (McKee et al. (1993)). Expressed mathematically for runoff (as an example):

Standardised Precipitation Index (SPI)

This index, $SPI_{i,k}$, requires rainfall values R_{ij} where i denotes the hydrological year and j th month within a hydrological year. The cumulative rainfall, $V_{i,k}$, for the i -th Hydrological year and k -th reference period can be obtained from:

$$V_{i,k} = \sum_{j=1}^k R_{t,j}, \quad j=1,2,\dots, \quad k=1,2,3,4$$

$$SPI_{i,k} = (V_{i,k} - V_{mean_k})/S_k$$

Where V_{mean_k} and S_k are respectively the mean and standard deviation of the cumulative rainfall for the k -th reference period.

The reference periods used for runoff and rainfall would typically vary from 3, 6, 12, 24, or 48 months while reference periods for soil moisture would be much shorter.

The definitions of states of drought, according to McKee et al. (1993) SPI are as indicated in Table 2.1.

Table 2.1: SPI drought categories (Mckee et al., 1993)

SPI value	Drought category	Time in category
0 to -0.99	Mild drought	24%
-1.00 to -1.48	Moderate drought	9.2%
-1.50 to -1.99	Severe drought	4.4%
< -1.99	Extreme drought	2.3%

2.2 Surface Water Monitor

The Surface Water Monitor (SWM) is described as a real-time hydrological monitoring and prediction system for the USA. It generates information on a daily basis for rainfall, runoff and soil moisture conditions in the USA and make forecasts for 1 to and 3 months into the future. See Figures 2.1 and 2.2.

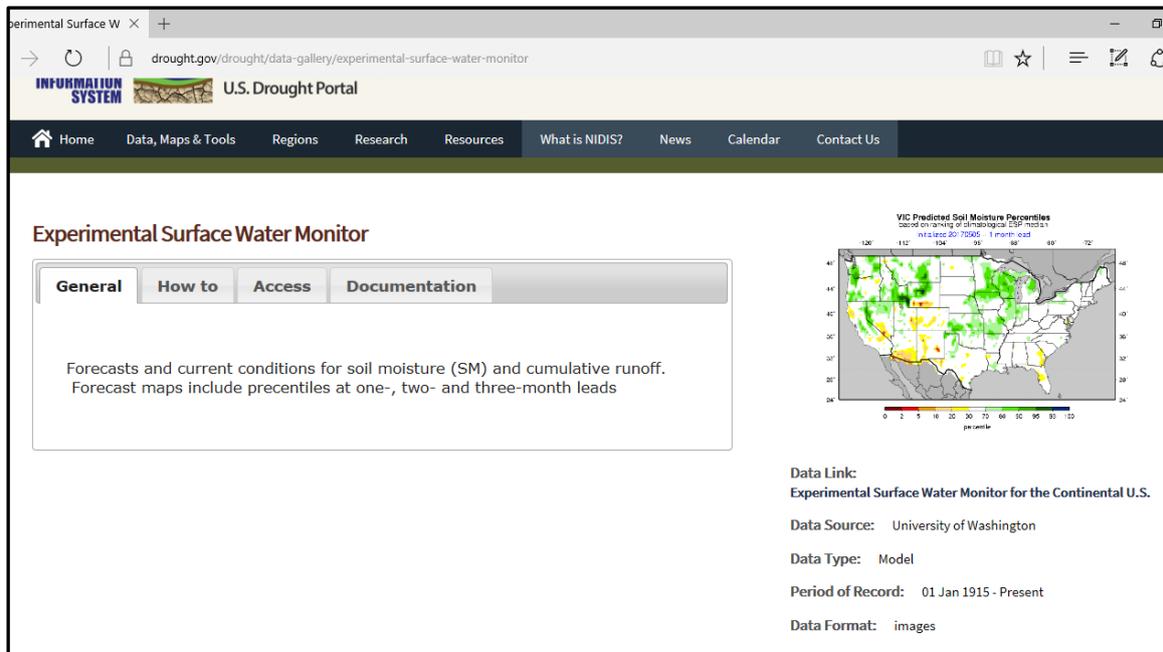


Figure 2.1: Surface Water Monitor: Home Page

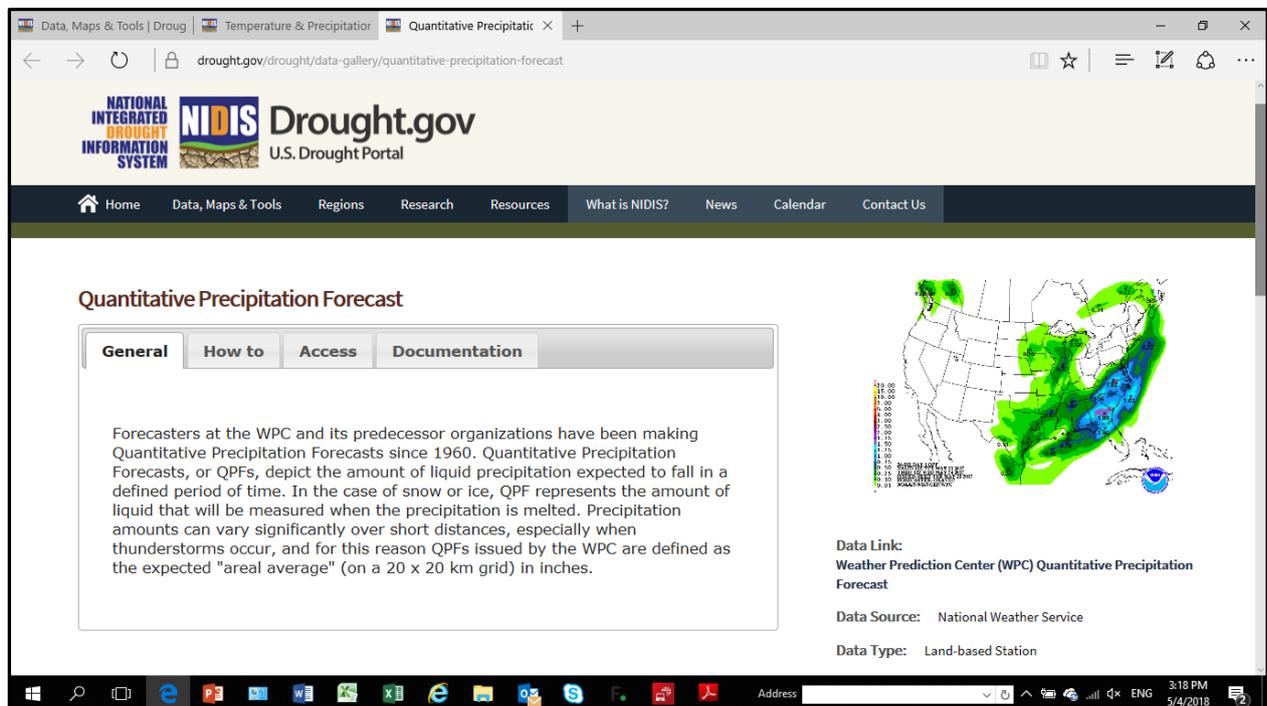


Figure 2.2: Surface Water Monitor: Rainfall forecast

SWM carries out forecasts using the Ensemble Streamflow Prediction Method (Twedt et al., 1977). This method entails setting antecedent conditions such as soil moisture and snow water equivalent within a hydrological model which then models a range of possible future states using past weather observations. An improvement on Twedt's original method is to only use ensemble members that match the current ENSO state (La Niña, Neutral, El Niño).

2.3 African Flood and Drought Monitor

In collaboration with the International Hydrological Programme, Princeton University has developed an experimental drought monitoring and forecast system for sub-Saharan Africa. The system aims to provide timely and useful information on drought by integrating climate predictions, hydrological models and remote sensing data. The main components of the system include the provision of near real time evaluations of the terrestrial water cycle and an assessment of drought conditions. The system has been designed for data scarce regions and uses macro scale hydrological modelling. Hydrological and drought forecasts are provided for 6 months. The predictive skill of the system has been evaluated for 30 years of historic hind casts and shows potential for providing useful forecasts of developing drought conditions, particularly for the first month. The web interface is shown in Figure 2.3.

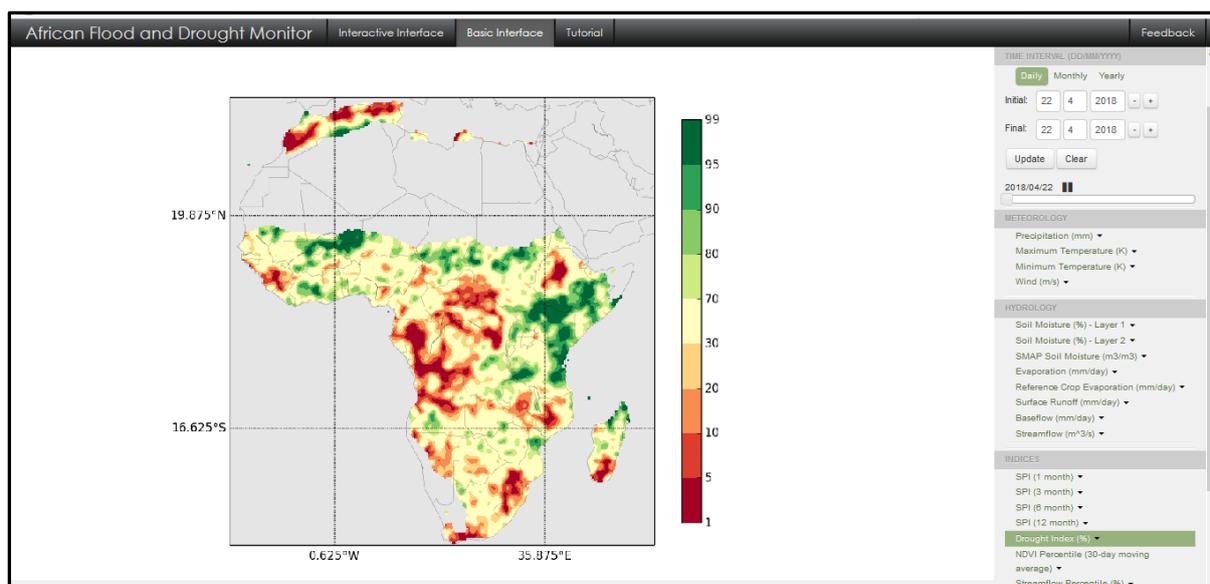


Figure 2.3: African Flood and Drought Monitor Home Page

The African Drought Monitor and Forecast system explanation

(<http://unesdoc.unesco.org/images/0023/002319/231937e.pdf>) outlines the monitor's three parts (see Figure 2.4 below):

- a) Historic reconstructions of the terrestrial hydrological cycle that is derived from simulations of the VIC land surface model forced by a hybrid reanalysis observational meteorological dataset. The datasets are used for a variety of applications including analysis of historic drought events, estimation of trends and variability, and investigation of drought mechanisms.

- b) Real-time monitoring component that updates the model run to 2-3 days from real time forced by bias corrected and downscaled TMPA satellite precipitation and GFS analysis fields of temperature and wind speed. There is also potential to force other impact models such as crop models.
- c) Seasonal hydrological forecast component that uses bias corrected and downscaled CFSv2 climate forecasts of precipitation and temperature to drive the model and provide ensemble predictions of drought conditions, for precipitation, soil moisture and streamflow. The figure below shows existing components in normal font, while potential future components are in italic font.

CFSv2 is an acronym for version 2 of the Coupled Forecast System which is a fully coupled General Circulation Model representing the interaction between the Earth's atmosphere, oceans, land and sea ice.

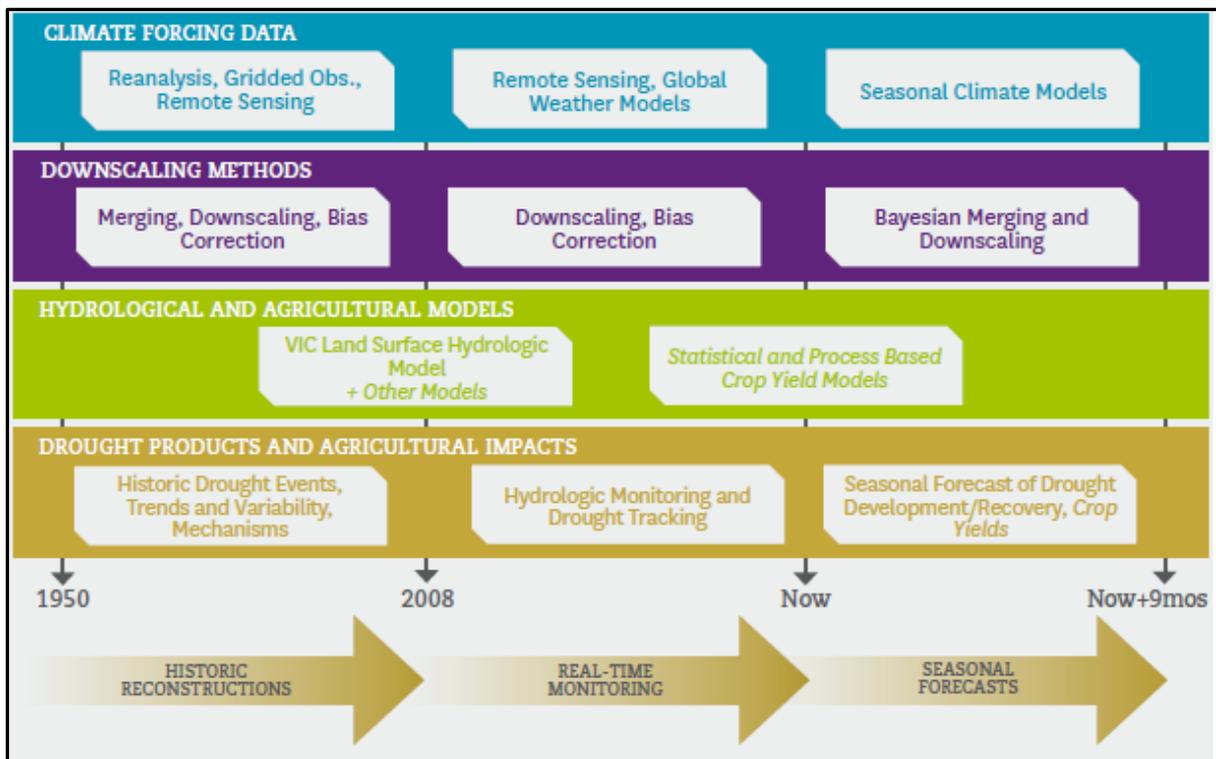


Figure 2.4: Overview of the Drought Monitoring and Forecasting components

2.4 Global Integrated Drought Monitoring and Prediction System

The Global Integrated Drought Monitoring and Prediction System (GIDMaPS) integrates precipitation and soil moisture data from several remote sensing platforms to produce drought indices for the howl globe (Hao Z et al., 2014). See Figure 2.5 which is an extract from the GIDMaPS website. GIDMaPS adds a Multivariate Standardized Drought Index (MSDI) (Hao and AghaKouchak, 2013b) to the SPI and SMI used by other drought monitoring systems. MSDI is the joint cumulative probability of rainfall and soil moisture.

GIDMaPS uses the Ensemble Streamflow Prediction method (Twedt et al., 1977) to perform seasonal forecasts.

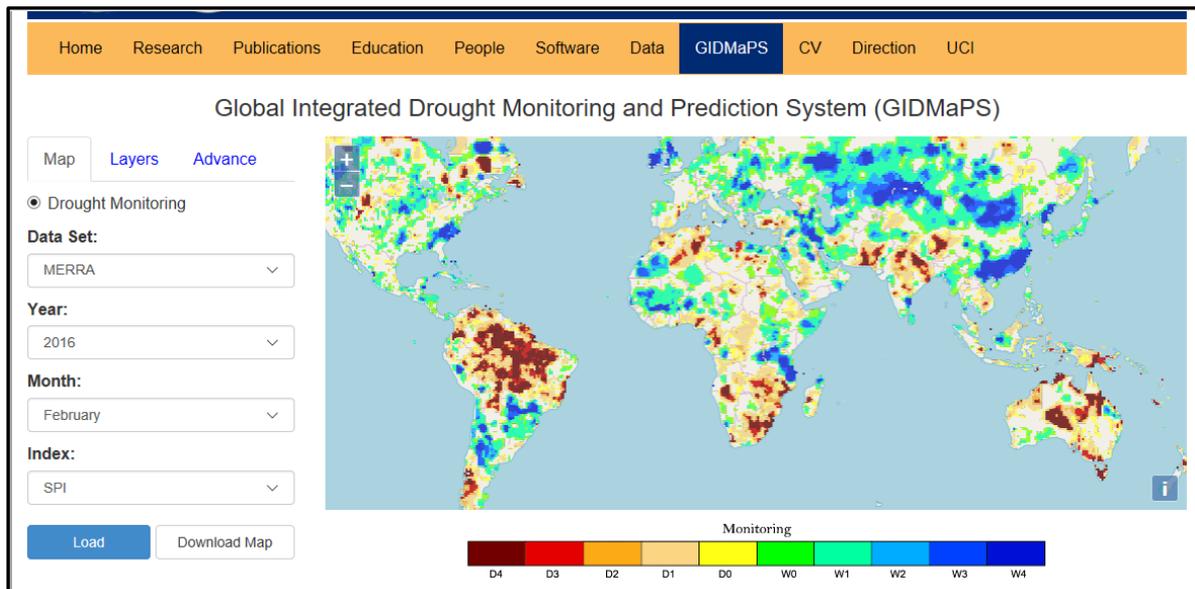


Figure 2.5: GIDMaPS website

3 RAINFALL PREDICTABILITY

3.1 Introduction

There is a limit to the predictability of the climate system because of the chaotic nature of that system. This means that the climate evolution is sensitive to small perturbations in the initial state, and that sensitivity manifests in the dynamical models of the climate system. Since the initial state of the climate system at the time of the forecast can never be determined precisely, that sensitivity limits ability to predict the future evolution of the climate system with dynamical models. In fact, the dynamic equations that govern the atmospheric motions are theoretically deterministic, but the chaotic nature of the system means that errors and uncertainty in the starting point amplify through the forecast time (Palmer, 2001). As a consequence, forecast systems that rely on the initial conditions allow for reliable forecast only at short, the so-called meteorological time scales – in the order of 1-5 days ahead.

Forecasting at longer, the so-called seasonal time scales (1-3 months ahead) relies on boundary conditions. The atmosphere variability is strongly influenced by energy fluxes at the ocean and continental surface. The characteristics of these surfaces change relatively slower than the timing of internal atmospheric variability – sea surface temperature anomalies, soil moisture or snow cover can persist for periods of time longer than a couple of weeks. Under some conditions, the association between these anomalies and local climate can extend predictability of the climatic responses into the seasonal time scale. A good example of such a longer-term predictability is the El Niño-Southern Oscillation (ENSO) phenomenon. ENSO varies slowly, and is predictable at least six months in advance, and there are relatively strong association between rainfall in various parts of the world and the state of ENSO. ENSO is the strongest natural fluctuation of climate on inter-annual time-scales. But weaker ENSO-like fluctuations, for example, Pacific Decadal Oscillation (PDO) may manifest at decadal time-scales. Importantly, predictability related to ENSO does not manifest universally, but only in some regions of the world and in some seasons. In southern Africa, for example, ENSO-rainfall association supports predictability of summer rainfall mostly in the north-east part of the country, in the beginning of the rainy season (Nov-Dec) and under strong ENSO forcing, i.e. during strong El Niño or La Niña, but not in neutral ENSO state.

Apart from the factors that are dependent on the nature of the forecasting tools (models), i.e. the presence of the relevant boundary forcing as described above, skill of the prediction also depends on models' ability to accurately represent the significant climate processes. This is affected by the choice of forecast scheme, discretization of the system equations in time and space, limited understanding of atmospheric processes, inaccurate parameterization of small-scale sub-grid processes, and the propagation of errors in their interaction.

All the above factors cause that in practice, a detailed prediction of atmospheric features is limited to about two weeks, and at longer time scales, only in certain regions and in certain seasons, and even then, prediction will be associated with substantial uncertainty.

The primary objective of this chapter is to describe various rainfall drivers over southern Africa which may influence the predictability of the seasonal rainfall over the regions, and to evaluate

the skill of three climate forecast systems, in the context of hydrological early warning system at seasonal time scales.

3.2 Drivers of southern Africa rainfall

Several studies have identified and discussed various atmospheric features and processes that modulate the spatial and temporal distribution of rainfall in southern Africa. Table 3.1 provides a brief description on the most important rainfall drivers and short review on how they influence southern Africa rainfall. A more comprehensive review can be found in Reason (2018).

Southern Africa rainfall has a strong seasonal cycle, but most areas in the region receive their rainfall in summer; the exception is the south-western Cape that experiences winter rainfall and the southern Cape coastal region receives rainfall year-round. Although the majority of summer rainfall is associated with the Tropical Temperate Trough (TTT) cloud-bands, various atmospheric features interact to contribute to the summer rainfall (Table 3.1). The regional atmospheric circulation converges moisture from the surrounding oceans into the South Indian Convergence Zone (SICZ) (Streten, 1973; Cook, 2000) which sustains TTTs. In that system, the westerlies bring moisture from the tropical Atlantic around the Angolan Low and the easterlies supply moisture from the southwest Indian Ocean through the South Indian Anticyclone (SIA). The north-easterlies bring in moisture from the equatorial western Indian Ocean into the boundary of the SICZ (Cook, 2000; Todd and Washington, 1999). Hence, a change in any of the regional circulations, may have ripple effects on other circulations and hence on the southern Africa rainfall.

El Niño-Southern Oscillation (ENSO) is the teleconnection with the greatest impact on southern Africa summer rainfall (Lindesay, 1988; Rocha and Simmonds, 1997; Reason et al., 2000; Cook, 2001; Reason and Jagadheesa, 2005). A link between El Niño and summer droughts over Southern Africa was found to be strongest in the south-eastern part of the sub-continent and over northeast South Africa (Richard et al., 2000). ENSO has the greatest influence on rainfall over southern Africa during its mature phase between the austral summer months of December and March. Meque and Abiodun (2014) showed that El Niño events favour widespread dry conditions in the southern Africa region, except over northern Mozambique, southern Tanzania, north-eastern Zambia and eastern DRC, where El Niño events produce wet conditions. The impact of ENSO on southern Africa rainfall come directly through changes in regional circulation. El-Niño events impose regional circulation changes that reduce moisture convergence, uplift and instability needed for tropical-extratropical cloud-band development over the region (Cook, 2001; Mulenga et al., 2003; Ratnam et al., 2014). Meque and Abiodun (2014) showed that during El Niño years, the SICZ (and TTTs) shift north-eastwards due to weakening of the SIA, consequently inhibiting precipitation over southern Africa. The impact of ENSO on southern African rainfall also indirectly through changes in other atmospheric teleconnections, like Pacific South American pattern (PSA; Karoly, 1989; Ghil and Mo, 1991; Mo and Higgin, 1998; Mo, 2000). The relationship between ENSO and southern Africa rainfall is characterized by a degree of uncertainty. For example, the 1997/1998 El Niño event was one of the strongest in the last century, but it did not lead to the

expected severe drought over subtropical southern Africa. Other phenomena shown to modulate the rainfall variability in southern Africa include the Tropical Indian Ocean Dipole (TIOD) (Saji et al., 1999; Yamagata et al., 2004), the Subtropical Indian Ocean Dipole (SIOD; Behera and Yagamata, 2001;), The Southern Annular Mode (SAM) (Gillett et al., 2006), the Benguela Niño (Rouault et al., 2009).

Table 3.1: Drivers of southern Africa rainfall variability at various spatial and temporal scales

System	Description and influence on rainfall
Regional	
Inter-Tropical Convergence Zone (ITCZ)	ITCZ is the dividing line between the two main air masses (the south-easterly trade winds and north-easterly trade winds) that influence southern Africa climate. ITCZ has a clear seasonal cycle, as it follows the apparent movement of the sun. Over southern Africa, it starts to move south of the equator in October reaching the its southernmost position over central Madagascar and Mozambique in February, after which it starts to retreat northward. The southward movement of the ITCZ ushers in the rainfall season over southern Africa while the northward retrieval terminates the rainfall season. An anomalous southward shift and strengthening of the ITCZ over tropical south-eastern Africa produces wet summers in South Africa (Cook et al., 2004).
South Indian Anti-cyclone (SIA)	SIA is a region of semi-permanent high atmospheric pressure over the Indian Ocean between 20°S and 35°S. It is one of the centres of anticyclonic activity in the southern hemisphere subtropical belt. A stronger (weaker) SIA increases (decrease) rainfall across southern Africa (Cook <i>et al.</i> , 2004; Reason and Jagadheesha, 2005; Reason <i>et al.</i> , 2006; Manhique <i>et al.</i> , 2011; Munday and Washington, 2017).
South Atlantic Anticyclone (SAA)	SAA is a semi-permanent high-pressure system in the southern part of the Atlantic Ocean. It has seasonal shift of about 6° latitudinally and 13° zonally. An anomalously strong SAA is associated with wet conditions over some parts of southern Africa (Walker, 1990; Tyson, 1986).
Botswana High	This is a tropical upper-level anticyclone that forms at 500 hPa level over central Namibia and western Botswana during the late austral summer. Botswana High is said be part of atmospheric circulation induced by heat release from heavy precipitation Congo basin. The relative strength and position of the Botswana high influence rainfall over the region. For example, a stronger than usual Botswana high has be shown to produce normal rainfall over Zimbabwe (Ratna et al., 2013).

System	Description and influence on rainfall
Regional	
Angola Low	This is a shallow heat low located over southern Angola and northern Namibia. It starts developing around October and strengthens during January and February, acting as the tropical source region for the tropical-extratropical cloud bands that bring most of the summer rainfall over southern Africa south. A stronger Angola low results in higher of rainfall over the region (Reason et al., 2006b; Mason and Jury, 1997; Cook et al., 2004; Hart et al., 2010)
Tropical Temperate Trough (TTT)	TTT is a tropical-extratropical cloud-band development that often extend NW-SE over southern Africa, from the Angolan Low region out into the southwest Indian (Harrison, 1984; Todd and Washington, 1999; Washington and Todd, 1999; Fauchereau <i>et al.</i> , 2009; Hart <i>et al.</i> , 2010; Manhique <i>et al.</i> , 2011; Hart <i>et al.</i> , 2013). Ocean. The TTT cloud-bands constitute the South Indian Convergence Zone (SICZ). TTTs bring a large portion of the summer rainfall south of about 15°S (Todd and Washington 1998; Washington and Todd, 1999; Manhique et al., 2011; Hart et al 2010, 2013; Ratna et al., 2012; Tozuka et al., 2014). For example, Hart et al. (2013) found that TTTs can contribute 30-60% of the mean summer rainfall over South Africa.
Mid-latitude frontal systems	These are frontal systems that form over the southern part of the Atlantic Ocean and approach the southwestern part of South Africa from the west. They occur mainly during the winter months in the Southern Hemisphere (May to July) although development can occur throughout the year. In winter months, the mid-latitude frontal systems frequently impact southwestern Africa, the only region in the sub-continent that receives mainly winter rainfall.
Tropical cyclones	Tropical cyclones that form over the warm South West Indian Ocean usually move with easterly mid-level flow towards southern Africa and make landfall over the eastern Madagascar and the western Indian Ocean islands of Mauritius and Reunion (Xie et al., 2002; Malan et al., 2013)
Easterly waves or lows	These are easterly waves or lows that pass across southern Africa in summer without TTT development. These waves tend to be semi-stationary rather than propagating. During their passage, the easterly waves or lows produce widespread heavy rainfall north of about 25-30°S (Dyson and Van Heerden, 2001; Reason and Keibel, 2004)
Cut-off lows	Cut-off lows are cold cored westerly systems that form in the mid- and upper troposphere on the equatorward side of the subtropical jet. They are formed from Rossby wave breaking that occurs or before the day the cut-off low forms. Occasional cut-off lows can contribute substantially to the winter rainfall in some years (Favre et al., 2012; Molekwa et al., 2014) and to spring rainfall over the south coast of South Africa (Weldon and Reason, 2015; Engelbrecht et al., 2015).

System	Description and influence on rainfall
Regional	
Atmospheric rivers	Atmospheric rivers are narrow plumes of enhanced winds that transport large amounts of tropical-sourced moisture into the higher latitudes over large distances. Heavy rainfall events in many subtropical and mid-latitude regions have been associated with atmospheric rivers (Zhu and Newell, 1994, 1998; Gimeno et al., 2016).
Mesoscale or local	
Mesoscale Convective Complexes (MCC)	MCCs are clustered along the eastern regions of southern Africa, adjacent to the warm waters of the Mozambique Channel and Agulhas Current. A few infrequent systems are found to be developing in Namibia and Botswana. The systems are found to predominantly occur during the months of November-February, with maximum activity occurring in November and December. Blamey and Reason (2013) showed MCCs can contribute up to 20% of the summer rainfall over north-eastern South Africa and southern Mozambique.
Air mass thunderstorms	These are deep convection systems that develop when the growth of cumulus clouds are sustained by a combination of diurnal heating and low-level moisture convergence from either mesoscale or synoptic forcing. Air mass thunderstorms bring a substantial portion of the summer rainfall over much of southern Africa, especially over the Highveld of South Africa where they are accompanied by hails (Smith et al., 1998).
Tornadoes	Tornadoes are violent rotating columns of air extending from a thunderstorm. Tornadoes occur over eastern South Africa and can lead to heavy rainfall, severe damage and sometimes loss of life (de Coning and Adam, 2000).
Terrain-induced circulations	The complex terrain of southern Africa leads to a variety of other local / mesoscale circulation systems that are important to local rainfall. For example, Nocturnal low-level jets induce night-time rainfall over Botswana and Namibia (Monahan et al., 2010).
Teleconnection	
Madden Julian Oscillation (MJO)	The Madden-Julian oscillation (MJO), an intra-seasonal (30- to 90-day) variability in the tropical atmosphere, is a traveling pattern that propagates eastward at about 4 to 8 m/s through the atmosphere over the warm parts of the Indian and Pacific oceans. It forms and is sustained by a large-scale coupling between atmospheric circulation and tropical deep convection. A strong 30-60 day periodicity has been identified in southern Africa rainfall. MJO activities have been shown to influence rainfall and wind oscillations over Tanzania (Kijazi and Reason, 2005; Mapande and Reason, 2005), TTTs over southern Africa (Hart et al., 2013), tropical cyclone tracks over the western Indian Ocean (Xie et al, 2002; Annamalai et al., 2005; Malan et al., 2013), and convection over southern Angola / northern Namibia (Hermes and Reason, 2009b).

System	Description and influence on rainfall
Regional	
El Niño – Southern Oscillation (ENSO)	ENSO is an irregularly periodic variation in winds and sea surface temperatures over the tropical eastern Pacific Ocean. The warming phase of ENSO is called El Niño and the cooling phase is known as La Niña. ENSO, which remains the dominant mode of inter-annual climate variability globally, is the mode with the greatest impacts on southern African summer rainfall (Lindesay, 1988; Rocha and Simmonds, 1997; Reason et al., 2000; Cook, 2001; Reason and Jagadeesha, 2005). Although the relationship between ENSO and rainfall in Southern Africa is not linear, La Niña years are generally associated with wet conditions, whereas El Niño years are associated with dry conditions over the region (e.g. Ropelewski and Halpert, 1987; Kruger, 1999; McHugh and Rogers, 2000; Reason and Jagadeesha, 2005).
Benguela Niño	The Benguela upwelling system is caused by surface water located near the western coast of southern Africa being forced offshore by strong southerly winds and deeper cold water rising to the surface and replacing the warmer water (Rouault et al., 2009). Although there is strong link between Benguela Niño and southwestern African rainfall (Hirst and Hastenrath, 1983; Nicholson and Entekhabi, 1986; Rouault et al., 2003b), but the relationship is strongly nonlinear and it also sensitive to magnitude of SST anomalies in the South West Indian Ocean (Hansingo and Reason, 2009; Reason and Smart, 2015).
South Indian Ocean subtropical dipole (SIOD)	SIOD is an inter-annual dipole event in the subtropical Indian Ocean. It has a positive phase characterised by unusually warm SSTs in the southwest Indian Ocean south of Madagascar and anomalously cool SSTs in the southeast Indian Ocean off Australia. SIOD is associated with high pressure anomalies in the mid-latitude South Indian Ocean. Positive IOD result in above normal rainfall over parts of southern Africa in the austral summer (Behera and Yagamata, 2001; Reason, 2001, 2002).
Southern Annular Mode (SAM), also known as Antarctic Oscillation (AAO)	SAM, the principal mode of atmospheric variability between the extra-tropics and high latitudes in the southern hemisphere, consists of an oscillation in atmospheric pressure between the Antarctic region and the southern mid-latitudes. A positive (negative) event is characterized by anomalously low (high) pressure over Antarctica and anomalously high (low) pressure over the mid-latitudes of the southern hemisphere. Gillett et al. (2006) found a positive association between the SAM and rainfall over south-eastern southern Africa while Rouault (2005) also found a negative association between SAM and winter rainfall over the western Cape region of South Africa.

4 SEASONAL RAINFALL FORECASTING APPROACHES FOR SA IN THE CONTEXT OF DROUGHT FORECAST

4.1 Introduction

From the point of view of objectives of this project, i.e. early warning system for hydrological drought at seasonal time scales, in a typical situation, the forecasted climate (i.e. variables such as rainfall and air temperature at surface) needs to be translated into hydrological responses (variables such as runoff or streamflow) or impacts with the aid of a hydrological model.

The process of generating hydrological responses from forecast climate introduces a range of uncertainties arising due to the nature, assumptions and simplifications of hydrological models and intermediate steps such as downscaling and/or bias correction. Those uncertainties may dilute the skill of the hydrological forecast. Thus, in spite of the attention of the project being focused on hydrological responses, climatic variables have to be considered as a starting point for hydrological analyses, as only in this way one can understand and quantify the “cascade of uncertainty” involved in data processing and serial linking of a number of models.

Considering the above, the following two broad framings to the forecast of hydrological drought as a basis for drought warning can be adopted:

- Forecasts of rainfall as an input to a hydrological model, where the results of hydrological simulations are interpreted in terms of a drought warning.
- Forecast of rainfall as a variable underlying the drought, and where forecast rainfall anomaly is interpreted in terms of a drought warning.

This chapter considers the second of the above, and two (sub-categorized into three) different approaches to rainfall forecast at seasonal time scale have been analysed:

- Numerical climate forecast based on climate models (section 4.3)
 - with output of these models taken “as is” (note that this does not preclude bias correction of climate model forecast) – this is based on NMME and Copernicus@ECMWF multi-model seasonal forecast ensembles (section 4.3.1).
 - with output of these models statistically calibrated or downscaled to observations of rainfall (this process is different to the bias correction) – this is based on data from SAWS, SEAS5@ECMWF and CFSv2 single model seasonal forecast ensembles (section 4.3.2)
- Statistical forecast based on rainfall monitoring and persistence of current rainfall anomaly (section 4.4)

Forecast skill is the ability of that forecast to correctly forecast future conditions. Forecast skill is a characteristic of the forecasting system, but it also reflects predictability of a particular variable at a particular location at a particular time. Forecast skill is a key piece of information necessary to interpret results of a forecast.

Since forecasting systems differ in skill, this section is constructed around exploration and evaluation of skill of the three different forecasting systems in the context of their use to satisfy objectives of this project. This is, of course, in addition to the explanation of how each of the forecasting systems work. Prior to that, the skill measures adopted in this study (section 4.2) are described in more detail.

It is worth noting at this stage, that the three approaches to seasonal rainfall forecast have been evaluated (and in case of statistical forecast based on persistent anomaly – developed) as this project was progressing and were somewhat guided by the availability of forecast data at the computing system of CSAG, one of the project partners. Data availability was influenced by activities external to this project. In particular, the full NMME and Copernicus@ECMWF forecast ensembles were only available to use in the last year of this project – 2019/2020. Prior to that, only data from NOAA CFS model contributing to the NMME was publicly available, while access to the ECMWF IFS model through a research licence. This project was not originally formulated to use these ensembles, and intended to use only dynamical forecasts from SAWS, CFS and SEAS5/ECMWF systems. These three systems were thus the basis for development of the statistically calibrated forecast. The NMME and Copernicus@ECMWF ensembles could not be statistically calibrated during this project, as this task exceeds time and funding available here. These forecasts were therefore use “as is”.

4.2. Measures of forecast skill

There are a large number of measures of forecast skills that are suitable for different types of forecasts (e.g. deterministic vs. probabilistic) and different types of forecast variables (e.g. categorical, i.e. drought-no drought vs. continuous – i.e. rainfall amount). The skill measures adopted in this project are presented below.

4.2.1. Deterministic forecast of a continuous variable

A numerical ensemble forecast is not deterministic, but probabilistic in nature, as it is based on an ensemble of simulations that capture forecast uncertainty. Ensemble forecast can, however, be expressed in a deterministic way, when considering the ensemble median rather than all ensemble members. Deterministic skill is then assessed through correlation (Pearson’s or Spearman’s) between the observed variable and the median of the ensemble forecast of that variable.

While deterministic forecasts are not directly useful in the context of generating a drought warning, the deterministic approach to forecast and to skill evaluation is the most transparent and the easiest to understand, compared to the probabilistic forecast and skill measures. Deterministic forecast were therefore adopted as one of the measures of forecast skill.

4.2.2. Probabilistic forecast of a (binomial) categorical variable

Forecast in a categorical form, i.e. forecast a category, or a range of values, might be expressed in probabilistic terms, i.e. forecast of a category occurring will have a probability associated with it. A categorical forecast is the traditional way of presenting seasonal climate forecast, with three categories, or terciles. Tercile forecast distinguishes: the above normal conditions (e.g. rainfall higher than the 66th percentile of its historical distribution), the below normal conditions (falling below 33rd percentile) and the normal conditions (falling between the 33rd and 66th percentile of historical variability), and provides a probability of each category, e.g. 45%, 30% and 25% (these probabilities obviously have to sum to 100%). While the tercile forecast contains three categories, skill of such forecast is evaluated separately for each of the categories, treating them as a binomial events, i.e. the forecast skill is assessed for three cases: below-normal vs. “non-below-normal”, normal vs. “non-normal” and above-normal vs. “non-above-normal”.

From the perspective of this project, a forecast of below normal conditions is in fact a forecast of drought conditions, or a drought event that occurs on average once in 3 years. By analogy, one can consider a forecast of a 1 in 10 year drought event that would indicate rainfall lower than the 10th percentile of historical distribution. Such a forecast would be a binomial in a sense that it would consider a “non-event”, i.e. rainfall higher than the 10th percentile of historical distribution, without detailing how high the actual value will be.

One of measures of skill of a binomial probabilistic forecast is Receiver Operating Characteristic (ROC) score. ROC score describes the ability of the forecast to discriminate between events and non-events. In a ROC curve the true positive rate is plotted in function of the false positive rate for different cut-off points of probability used to separate events from non-events (Figure 4.1). Each point on the ROC curve represents a pair of true and false positives corresponding to a particular probability decision threshold. The area under the ROC curve (AUC) is a measure of how well a parameter can distinguish between two diagnostic groups (e.g. first tercile vs. non-first tercile). ROC skill score relates ROC AUC for given forecast to the ROC AUC obtained under random forecast. ROC skill score value of 1 denotes a perfect forecast, ROC skill score of 0.5 indicates that forecast is not better than a random guess.

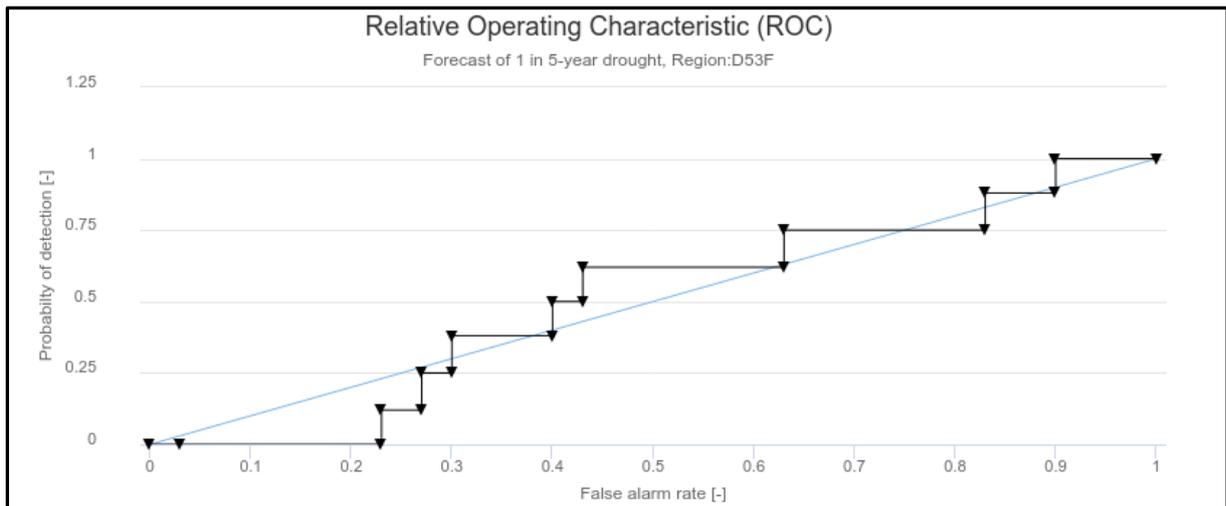


Figure 4.1: Example ROC curve for probabilistic event forecast

4.2.3 Deterministic forecast of a (binomial) categorical variable

While a probabilistic forecast of a drought event might have its merit in many contexts, as it allows for a context specific interpretation of probability associated with an event (e.g. in some context, drought preparing action might be triggered by when there is a 30% probability of a drought occurring, while others, less risk averse ones, might require a higher probability, say, 50%). In spite of this, a probabilistic event forecast remains difficult to communicate and often to understand. An alternative to such forecast is a deterministic event forecast, i.e. a statement – a drought will or will not occur. Similarly to the deterministic forecast of a continuous variable, such a forecast statement can be obtained from a probabilistic forecast assuming a certain threshold probability that allows converting the probabilistic forecast into an (apparently) definitive statement. In this case the producer of the forecast imposes their idea of what the probability cut off threshold, and thus risk (or uncertainty) tolerance of the user is.

A deterministic binomial forecast yields results that are relatively easy to interpret in terms of skill – one can relatively easily evaluate number of hits (event occurred and was forecast), correct negatives (event did not occur and was not forecast), misses (event occurred but was not forecast) and false positives (event was forecast, but did not occur). That is often present in the form of a contingency table (Figure 4.2).

Contingency Table

		Observed		Total
		yes	no	
Forecast	yes	hits	false alarms	forecast yes
	no	misses	correct negatives	forecast no
Total		observed yes	observed no	total

a.

		Observed		Total
		yes	no	
Forecast	yes	82	38	120
	no	23	222	245
Total		105	260	365

b.

Source: <https://www.cawcr.gov.au/projects/verification/>

Figure 4.2: Schematic of contingency table for binomial event forecast,

Expressing skill of a deterministic binomial forecast in a single, numerical value, is, however, surprisingly difficult. While an intuitive measure of skill accuracy would be for example “percent correctly forecast”, i.e. ratio of hits to total events, such a measure does not take into account such a factor as the number of false positives. To illustrate the problem with the “percent correctly forecast” as a skill measure, a forecast that issues a warning every time, would correctly forecast 100% of events, but it would obviously be very poor. Other skill measures suffer similar deficiencies, and these deficiencies magnify if forecast is of rare events (i.e. if there are considerably more non-events than events as is the case in severe droughts).

There is a plethora of skill measures that express skill of deterministic binomial forecast – see for example.

https://www.cawcr.gov.au/projects/verification/#Methods_for_dichotomous_forecasts.

Most of the commonly used skill measures are sensitive to the climatological frequency of the forecast events and are thus not applicable to rare events (as is the case with droughts).

The most universal and robust skill measure that is applicable to forecast of rare events is Odds Ratio Skill Score (ORSS). Its main drawback is that ORSS is not determined when any of the rows or columns in the contingency table are completely zero, which might happen in

operational forecast post-processed to derive rare events. This simply indicates that the forecast is not realistic.

$$ORSS = \frac{hits * correct\ negatives - misses * false\ alarms}{hits * correct\ negatives + misses * false\ alarms}$$

4.3. Numerical dynamical forecasting based on climate models

Dynamical numerical modelling of the global climate system is currently the most advanced and sophisticated way of generating climate forecast at the seasonal time scale. Numerical forecasting models are similar in nature to the global climate models used in generating climate change projections, although they have several major conceptual differences.

The main characteristics of forecast models are as follows:

- Historically forecasting models were atmosphere only and relied on an independent modelling of expected future evolution of sea surface temperatures to define their boundary conditions. That modelling could be relatively simple – e.g. assume persistence of relative anomalies observed at the time of issuing the forecast. However, currently, forecast models are mostly coupled, i.e. they include fully comprehensive sub-models simulating oceanic circulation. As such, coupled models rely on “internally” generated evolution of sea surface temperatures initialized on the state observed at the time of issuing the forecast.
- Seasonal forecast models are gridded and have relatively coarse grid – in the order of 0.25 to 1.5 degrees (~25km to 150km) in size.
- Forecast models generate a comprehensive set of data on variables reflecting the state of the global climate system (i.e. distribution of pressures, humidity, winds and temperatures at various levels in the atmospheric column), as well as variables at the earth’s surface, such as rainfall, air temperature, soil moisture.
- Forecast model output is, similarly to the output of global climate models, often characterized by biases, i.e. there are systematic differences between modelled values of atmospheric variables and observations. The implication is that output of forecast models has to be either bias-corrected, or subject to further process of downscaling or calibration, or used in a relative way, i.e. as % departure from the climatological mean.
- Forecast models are typically initialized from observed state of the climate system, i.e. the values of the physical variables (such as air temperature, humidity, pressure) in a model at the start of the forecast are generated from the actual observed data through a process of interpolation or data assimilation.
- Forecast models are run as an initial condition ensemble, i.e. the initial conditions are perturbed, so that individual simulations start with minimally different values, with differences between individual simulations growing with simulation time. That process is intended to allow capturing the so-called initial condition uncertainty that reflects the

chaotic character of the climate system. Most forecast systems generate a 10 member ensemble, but some run a 25- or even 50-member ensemble.

The implication of the ensemble approach is that there is no single forecast of a particular variable over a particular time, but rather a range of variables is forecasted. It is not possible to select “the best” ensemble member, as each ensemble member represents a future that is possible under given forcing, and all of these futures are equally probable. This pertains, however, only to the “model space”. There might be (and are) differences between “model space” and the real world. These differences reflect errors of a particular model, but also the effects arising from the model-independent levels of climate predictability in a particular location and season (section 3.1).

4.3.1. Statistically-calibrated (Model Output Statistics or MOS) dynamical seasonal forecast

Forecast climate models, as mentioned above, generate information (such as rainfall) as an average over their large computational unit (grid) – typically of spatial scale between 100 and 250 km. This does not take into account that surface climate varies at much smaller spatial scales, in some situations as low as 10s of metres. Hydrological models, in turn, require information at spatial scales that are “in-between”, i.e. in the order of 1km to tens of km.

Also, as mentioned above, data output by climate forecast models are often biased, i.e. systematically offset compared to observations. This is an intrinsic deficiency of climate models that is not simply reducible.

In order to account for climate model biases and allow derivation of climate data at smaller spatial scales output of climate models is often downscaled. It is worth noting that downscaling is not the same as interpolation, and also conceptually different from bias correction.

The process of downscaling relies on the relationship between large scale atmospheric circulation features and locally specific responses of surface variables such as rainfall. That process can be formalized through a nested, small scale climate model (dynamical downscaling), or establishing a statistical relationship between drivers and responses (statistical downscaling). In both situations, inputs to downscaling are large scale atmospheric variables, the output is a climate variable at local scale. As such, the process in principle does not involve using the actual surface variable values (at that model’s grid) generated by the forecast climate model.

In contrast, the process of bias correction involves the actual surface variable generated by the forecast climate model, and is simply a process of changing some statistical properties of either that model-generated variable, or of that variable observed at local scale.

In this study, the so-called Model Output Statistics (MOS) approach (Wilks, 2001; Paeth et al., 2011) has been adopted. This is a *sensu stricto* statistical downscaling procedure, and such a process is often termed “statistical forecast calibration”.

4.3.1.1. Implementation of the MOS downscaling/calibration of dynamical forecasts

It was decided to focus on bespoke statistical downscaling of several available seasonal forecast products and used as a reference the hydrological and climate datasets created within the Water Resources 2012 study (Bailey and Pitman, 2016, thereafter WR2012). In the study, the monthly rainfall obtained from WR2012 for the period of 1983-2009 were used.

Historical observational data

The analyses are based on aggregate rainfall dataset covering South Africa that is a part of WR2012. That study consolidated observational rainfall and streamflow data in a hydrological modelling framework to derive regionally- and time-consistent description of surface water resources in South Africa. In that, the area of South Africa was divided into 1944 quaternary (fourth-order in a hierarchical classification system) catchments. Station-based observations of rainfall from over 1700 stations were interpolated to generate monthly areal rainfall for 456 rainfall regions each of which encompass one or several quaternary catchments. Rainfall extracted for these rainfall regions was used. The WR2012 dataset spans the period of 1920-2009, but only the 1983-2009 data was used in this study.

Seasonal forecast data

Three operational seasonal climate forecasts were used:

- A forecast generated by South African Weather Service (SAWS) using a fully coupled ECHAM4.5-MOM3-SA global climate model (Beraki et al., 2014).
- NOAA's Climate Forecasting System v. 2 (CFSv2, Saha et al., 2014, freely available through cfs.ncep.noaa.gov)
- ECMWF forecast based on IFS atmospheric model, described in details here: <https://www.ecmwf.int/en/forecasts/documentation-and-support/long-range>,
- and available from <https://www.ecmwf.int>. At the time of implementation (~2018) the ECMWF IFS model results were not publicly available, and this project had access to these data through research licence only.

Adopted downscaling methodology procedure

The downscaling procedure adopted in this study is one of MOS approaches – Principal Component Regression (PCR). The use of that procedure follows earlier downscaling work carried out in southern Africa by Landman et al., 2001, Landman et al., 2012, and Muchuru et al., 2014, and Archer et al., 2019.

In the PCR MOS, the predictor is based on 850mb geopotential height fields (z_{850}) for the 0-55E, 0-45S domain simulated by each of the forecast systems as a three-months forecast initialized in the beginning of each of the DJF, MAM, JJA and SON seasons. That large scale z_{850} field captures the configuration of main rain-influencing large-scale circulation features affecting rainfall over southern Africa, namely the southern African thermal low, Angola Low, South Indian Ocean high, South Atlantic high and mid latitude lows. For summer rainfall, these features affect rainfall through influencing the characteristics of the so called easterly wave in general, and the low-level tropical and South Indian Ocean easterlies in particular. For winter rainfall, these features express the main driving forces of the westerly wave and mid-latitude

cyclones. That large-scale field has been shown to explain a considerable proportion of variance of local rainfall in the southern Africa region (e.g. Landman, 2001; Wolski et al., 2017).

The forecast ensemble mean of z850 field is subject to dimensionality reduction using principal component analysis (PCA), and significant principal components are selected using n-rule (Peres-Neto et al., 2005). A MOS multiple regression model is calibrated individually for each of the quaternary catchments, using the season's rainfall total as predictand, and the scores of up to 10 significant PCs as predictors.

4.3.1.2. Forecast skill

To assess the level of predictive skill of the forecast, a cross-validation approach was used, i.e. the level of skill attained when the calibrated model is applied to data that is not used in calibration was assessed. There are several ways of implementing the cross-validation procedure. Typically, one would use a "split-sample" approach, when the available data is split into two groups, and one is used to calibrate the model, while the other one is used for validation and calculation of skill measure. The "leave one out" approach was adopted, i.e. one year was left out of the available time series and the MOS model calibrated on that sub-sample. The calibrated model was then used to predict the left out year. The process was repeated for each year in the available dataset and the skill measure constructed based on the prediction of each of the left-out years.

Results of a split-sample calibration-validation procedure for an individual rainfall region are illustrated in Figure 4.3.

Results of assessment of skill of downscaled seasonal forecasts for SAWS forecast are presented in Figure 4.4, for CFSv2 forecast in Figure 4.5 and for ECMWF forecast in Figure 4.6 (DJF season only).

The results can be summarized as follows:

- The skill of the downscaled forecast varies strongly between seasons, with DJF and JJA displaying some levels of skill, and the transitional seasons of MAM and SON showing skill only locally, skill levels are relatively low in general, at best reaching 0.6 in terms of Pearson's correlation, and 0.7 in terms of ROC scores,
- There is some consistency between various forecast systems, i.e. the levels of skill and spatial regions with skill are relatively similar across the three forecasting systems, although there are minor differences too. In all analysed forecasts, levels of skill occur in DJF in the summer rainfall region, but there is virtually no skill in JJA in the winter rainfall region.

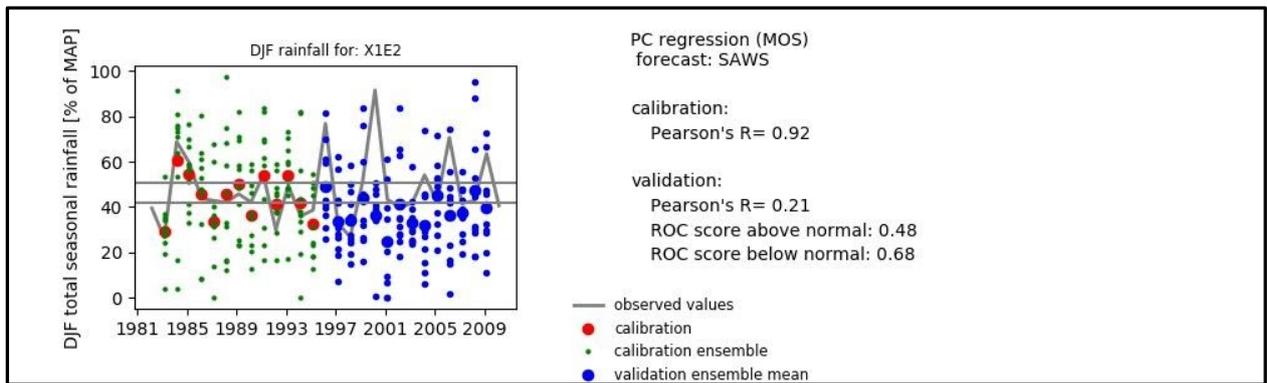
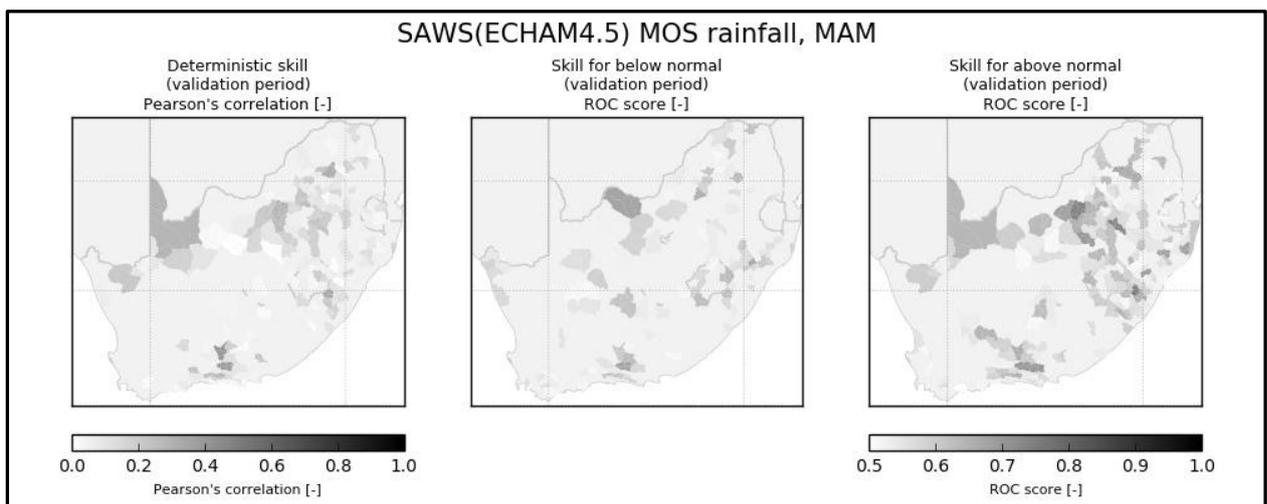
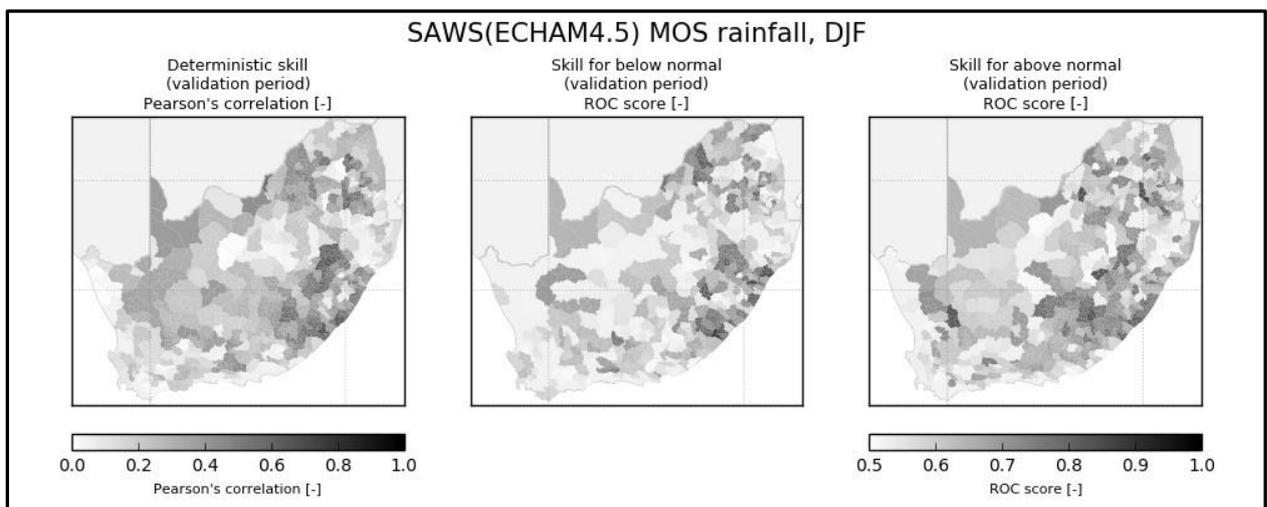


Figure 4.3: Summary of results of calibration and validation for a selected quaternary catchment using a split-sample approach



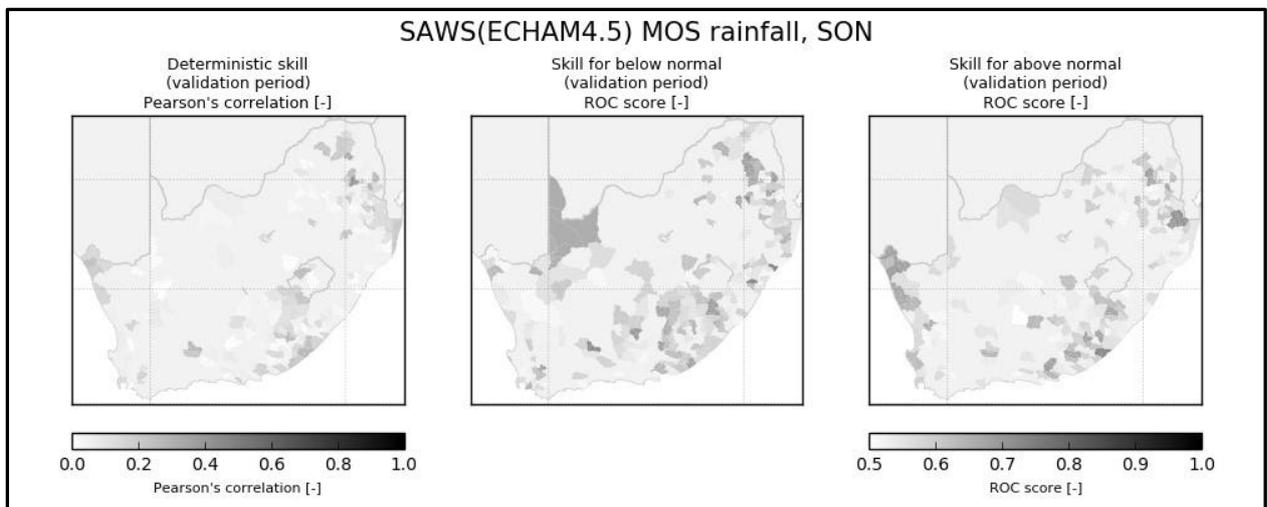
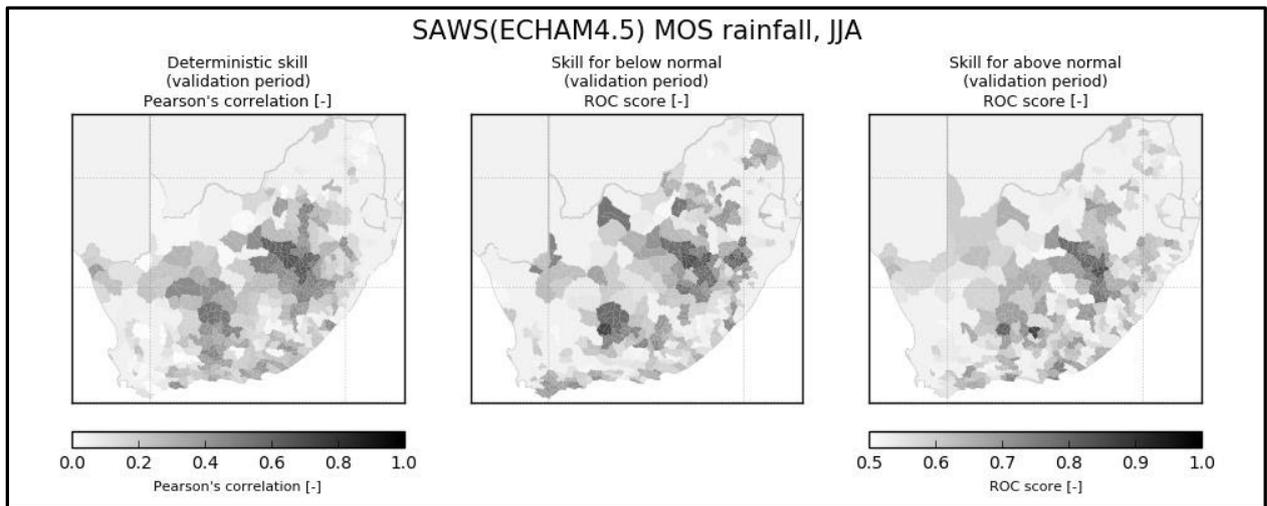
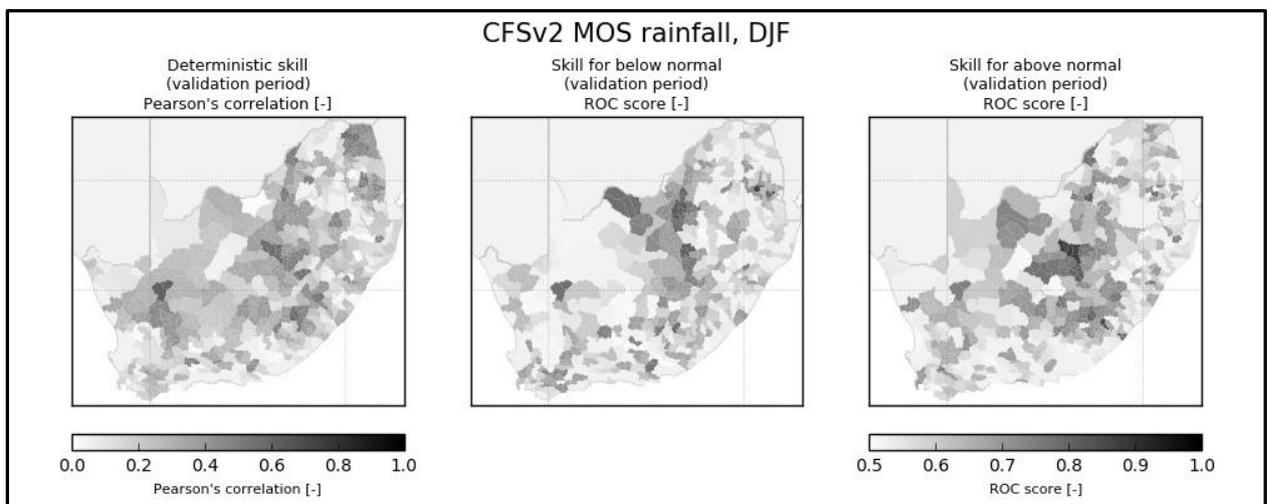


Figure 4.4: Skill measures for seasonal forecasts based on PCR MOS downscaling of SAWS (ECHAM 4.5) forecast to WR2012 rainfall data in four climatological seasons



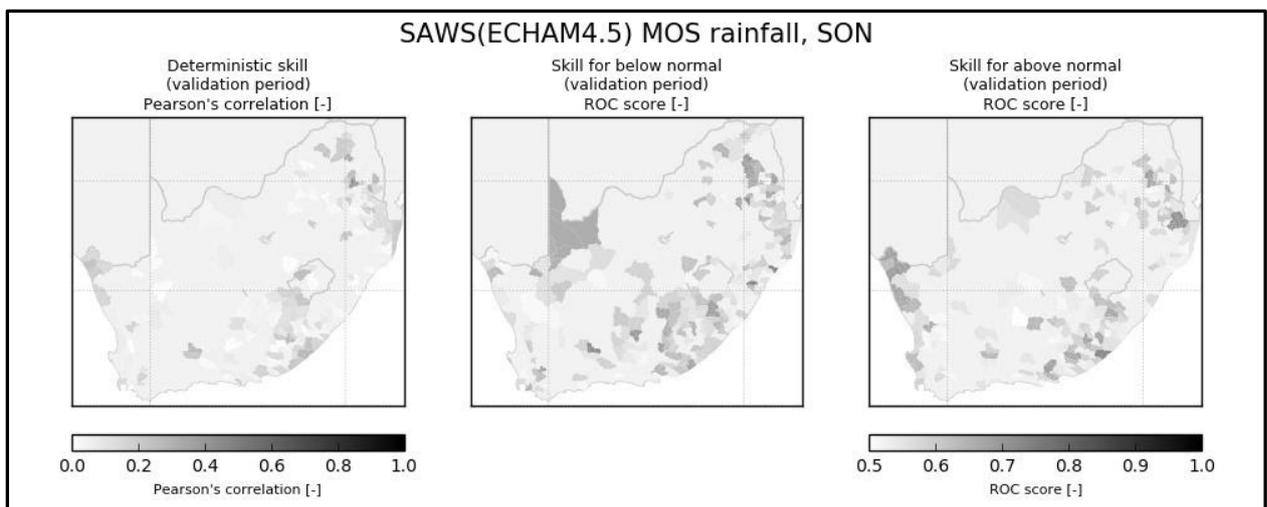
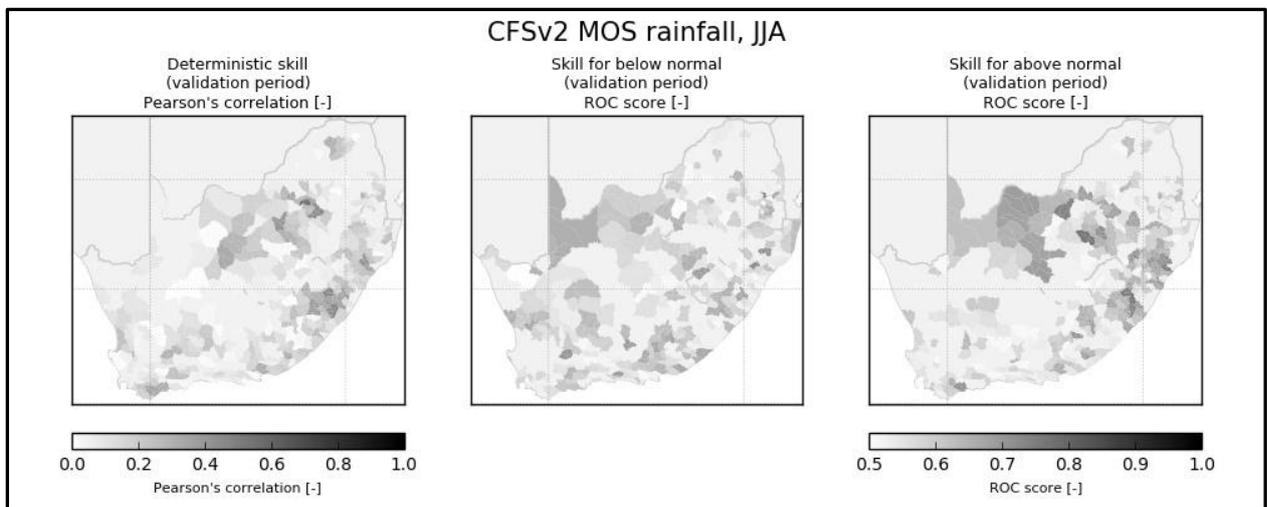
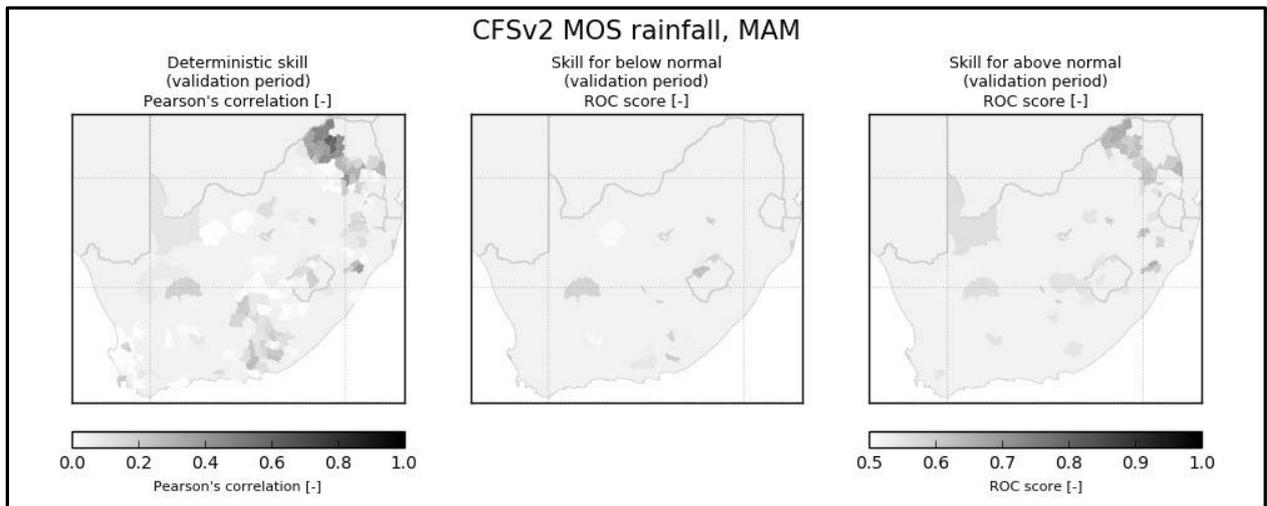


Figure 4.5: Skill measures for seasonal forecasts based on PCR MOS downscaling of CFSv2 forecast to WR2012 rainfall data in four climatological seasons

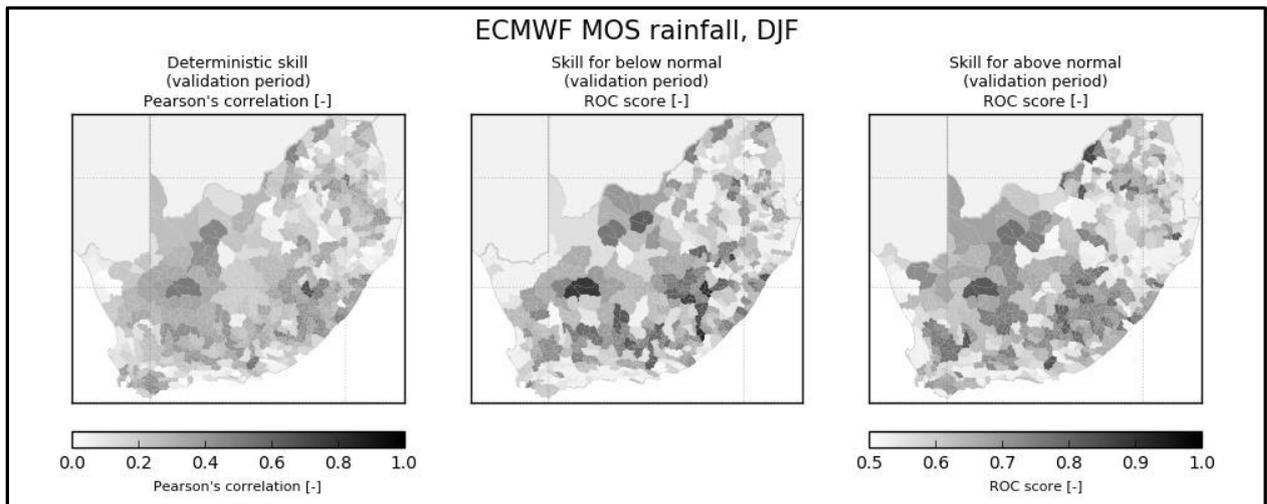


Figure 4.6: Skill measures for seasonal forecasts based on PCR MOS downscaling of ECMWF forecast to WR2012 rainfall data. DJF season

4.3.2. Multi-model ensemble of dynamical seasonal forecast models

In the course of this project, two seasonal forecast multi-model ensembles have become publicly available – the North America Multi Model Ensemble (NMME) (Kirtman et al., 2014) and Copernicus@ECMWF ensemble. These ensembles include two of the models used in section 4.3 – CFS v.2, and ECMWF IFS (called here SEAS5). It is worth noting that a “multi-model ensemble” encompasses a group of forecast models, each of which generates an ensemble forecast (the so-called initial condition ensemble), i.e. it generates multiple possible time series of future climate variables.

These two multi-model forecast ensembles provide forecast data regularly – every month, and also provide a set of historical forecasts (the so-called retrospective forecasts) that allow evaluation of the quality of the each of the forecasting model. All models in both ensembles provide at least surface temperature and rainfall data on the daily time scale and extending 3-9 months ahead of each forecast date. The forecast models are of spatial resolution (1 deg, ~100 km) that is low enough to allow a direct use of forecast rainfall and temperature data on the monthly basis in further analyses.

These ensembles are the basis for operational analyses of seasonal forecast globally issued by the hosting institutions – ILRI and ECMWF. In that both institutions rely on the strength of the multiple models, and, at the highest level of aggregation of information, synthesize all individual model forecasts into a single, multi-model statement.

As such, these ensembles constitute a state-of-the-art of seasonal forecasting globally. In this project, it was therefore decided to capitalize on the availability of forecast data from these multi-model ensembles and undertook synthesis of the forecast data into the specific context of this project. It is recommended that these ensembles should become the primary source of information towards the forecast of the prospects of drought.

It is suggested that the ultimate strength of the multi-model forecast can be revealed through:

- Calibration of the forecast to observations (i.e. MOS post-processed – see section 4.3.1), and,
- Analyses of the forecast models as an ensemble rather than individually.

Achieving this was not possible within the framework of this project. In this study the focus was on the main features of the ensembles, processing of forecast data into the data flow enabling the establishing an early warning system for hydrological drought, and on evaluation of their skill of individual forecast models.

Multi-model ensemble seasonal forecast data

NMME forecast data are available through the University of Columbia (<http://iridl.ldeo.columbia.edu/SOURCES/.Models/.NMME/>) and included forecasts from seven models coded:

- GFDL-FLOR
- GFDL
- NCEP-CFSv2
- NCAR-CESM1
- CanCM4i
- GEM-NEMO
- NASA-GEOSS2S

Copernicus@ECMWF ensemble is available through Copernicus Climate Data Store: <http://cds.climate.copernicus.eu>, and includes forecast from four forecast models (available in mid-2019, but seven are available in Nov 2020):

- ECMWF IFS 43.r1 – coded here SEAS5
- Meteo-France ARPEGE 6.4 model – coded here System7
- DWD ECHAM 6.3 model coded here GCFS v.2.0
- CMCC CESM-CAM model, coded here SPS3

Apart from the above, Copernicus@ECMWF ensemble includes CFSv2 (the same as NMME), UK Met Office HadGEM3 system (at the time these analyses were conducted it did not have comprehensive dataset allowing its incorporation here) and JMA MRI-CPS2 model (available only in November 2020).

Individual models in each of the ensembles differ in the atmospheric, land surface and ocean sub-models used, in the process of initialization (staggered, or perturbed), in spatial resolution (1-2 deg), duration of the simulation (up to 12 months) and size of the ensemble (3-50 members). Details of each of the Copernicus@ECMWF models can be found at <https://confluence.ecmwf.int/display/CKB/Description+of+the+C3S+seasonal+multi-system>, while details of the models can be found through links on the data source web page (<http://iridl.ldeo.columbia.edu/SOURCES/.Models/.NMME/>). Models contributing to each of

the ensembles are subject to periodic updates to newer versions, and as a result model names change from time to time.

Both ensemble forecasts are issued every month, with NMME ensemble forecast available on the 5th of each month, while the Copernicus@ECMWF ensemble available on the 13th.

As mentioned above, forecast data from both ensembles are generated every month by the modelling centres generating individual forecasts. Retrospective forecasts are available for all the models for the period of 1993-present for the Copernicus@ECMWF ensemble, and 1983-present for the NMME ensemble.

Pre-processing of data for skill assessment

The following, semi-automated procedure has been implemented on the computing system at CSAG, UCT:

- A subset of monthly forecast data (rainfall and air temperatures) for each of the forecast models was downloaded for all available retrospective forecasts. That subset covers the area of southern Africa (Lat: 45S-0S, Lon: 0E-55E), and the first three months of the forecast, starting on the month on which a forecast (or retrospective forecast) is issued (so for example Jan, Feb, Mar for forecast issued in January).
- All available ensemble members are downloaded.
- The data were organized into a uniform file and variable naming convention.
- Data were converted from gridded format to area-average over the WR2012 rainfall regions and WR2012 quaternary catchments. This was done by intersecting catchment/region polygons with the gridded data and calculating area-average of grid cell values over the polygons.
- Data at the polygon level were then bias-corrected using a month-specific quantile mapping function. That function is determined based on the available retrospective forecasts. In that, it is considered that long-term distribution of a variable in all forecast ensemble members should correspond to the long-term distribution of that variable in the observed data. This approach allows for the individual forecasts to diverge from climatology both in terms of ensemble median and ensemble members. The GPCC 2018 data was used as the historical observed dataset.

Forecast skill

In order to calculate, the following indices were calculated from the retrospective forecast data:

- Ensemble median of three-month mean rainfall for each model ensemble.
- Ensemble median of mean monthly rainfall for the three months of the forecast for each model ensemble.
- Probability of the three-month mean rainfall in each model ensemble to fall:
- Within each of the three terciles (below-, above- and normal) of the observed three-month mean rainfall. The probability of the below-normal rainfall is equivalent to the

probability of 1 in 3 year drought below 20th percentile of the observed three-month mean rainfall, which is equivalent to 1 in 5 year drought, and below 10th percentile of the observed three-month mean rainfall, which is equivalent to 1 in 10 year drought.

The probabilities were calculated by fitting a normal distribution function to the distribution of the individual ensemble members values. Considering that most of the models in the multi-model ensembles have only 10 ensemble members, more extreme drought levels were not used.

These indices were then used to evaluate the following forecast skill measures:

- Pearson's correlation between ensemble media and observed rainfall (as an expression of deterministic skill of the forecast)
- ROC skill score for above-, normal- and below-normal rainfall, as well as for 1 in 5 year and 1 in 10 year drought.

Considering that the "grand ensemble" consists of 11 models, and 6 skill measures for each of the 12 calendar months and for each of the 1948 quaternary catchments were evaluated, the output of the process is not easy to present in a digestible form.

A visualisation of the two sets is presented below.

- Maps showing monthly Pearson's correlation and ROC skill scores for 1 in 3, 1 in 5 and 1 in 10 year drought for one of the models – ECMWF SEAS5, for the entire country (Figures 4.7 to 4.10), but also for the Western Cape and for Limpopo (Figures 4.11 to 4.12).
- A summary of the skill scores for each model and each month of the forecast for quaternary catchments falling within the Inkomati basin – quaternaries within region X of WR2012 (Figures 4.13 and 4.14).

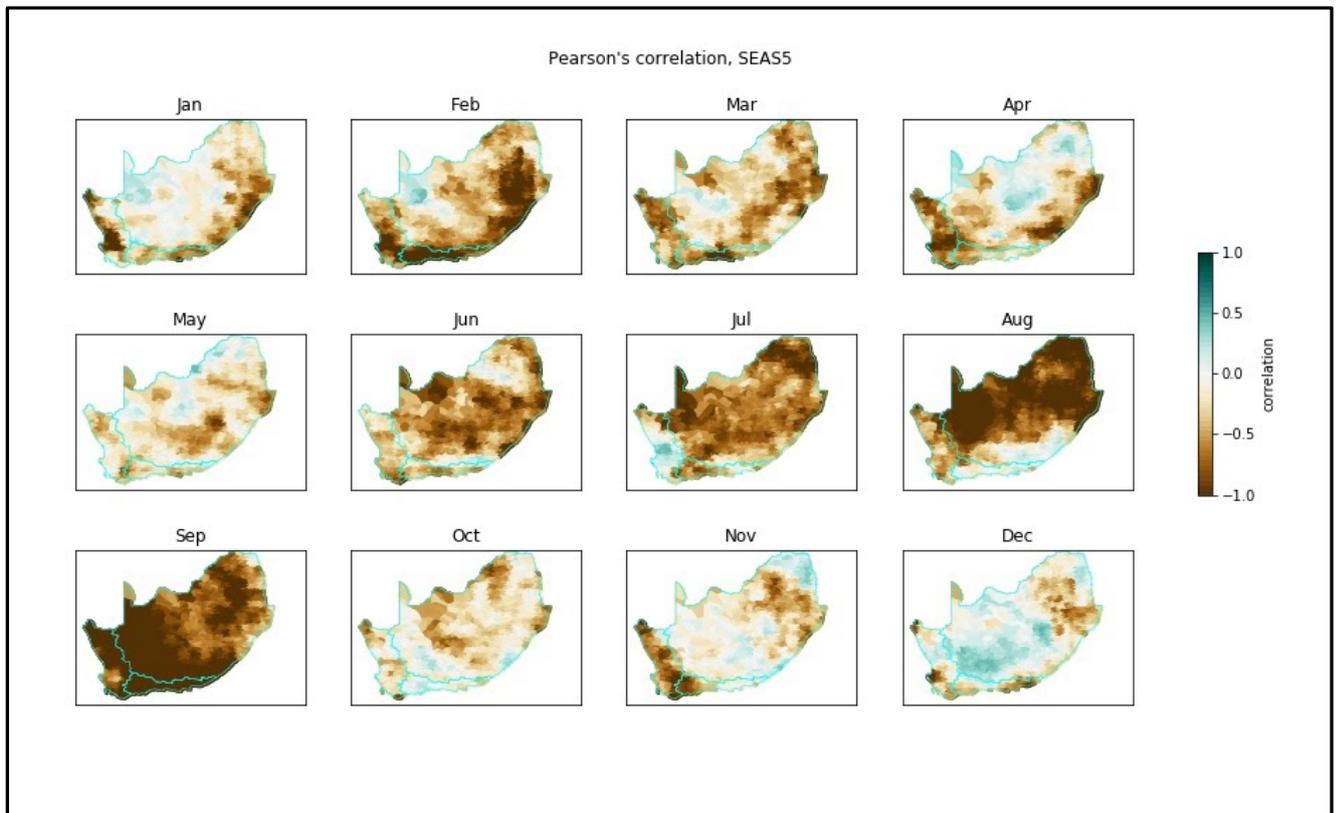


Figure 4.7: Deterministic forecast skill

This Figure presents correlations between observations and ensemble median at the scale of quaternary catchments, for forecast of three month rainfall mean, issued on each of the calendar months, ECMWF SEAS5 system

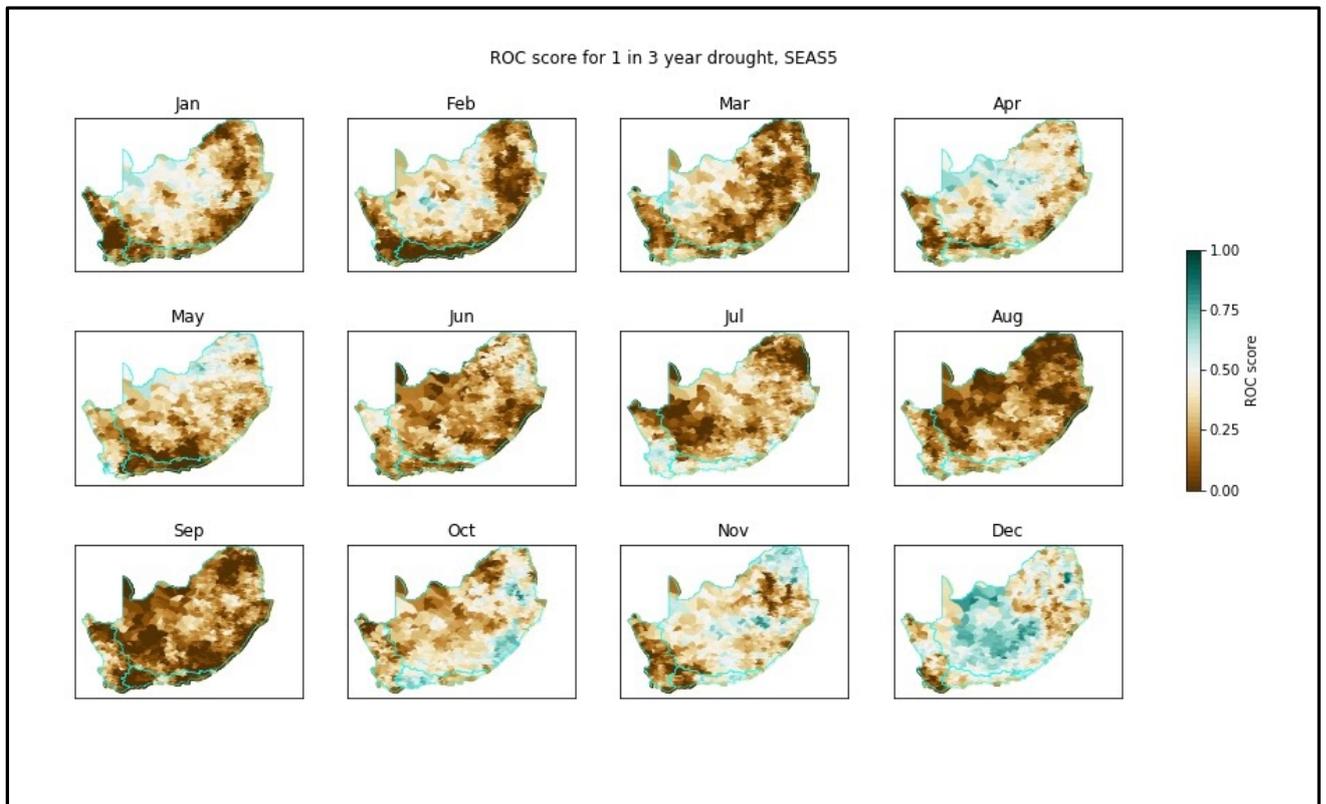


Figure 4.8: Probabilistic forecast skill (ROC score)

This Figure presents forecast skill at the scale of quaternary catchments, for forecast of below-normal rainfall (1 in 3 year drought), issued on each of the calendar months, ECMWF SEAS5 system. Forecast is better than random guess when ROC score is > 0.5

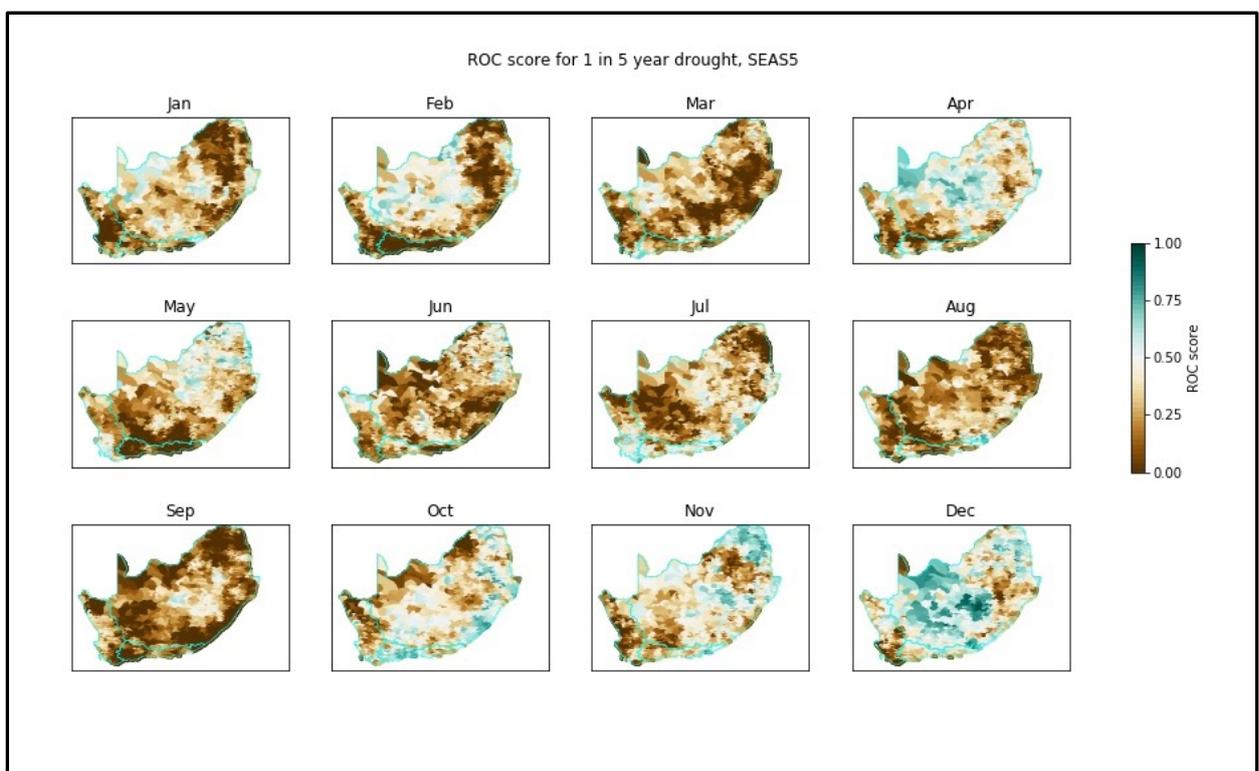


Figure 4.9: As for Figure 4.8 but for forecast of 1 in 5 year drought

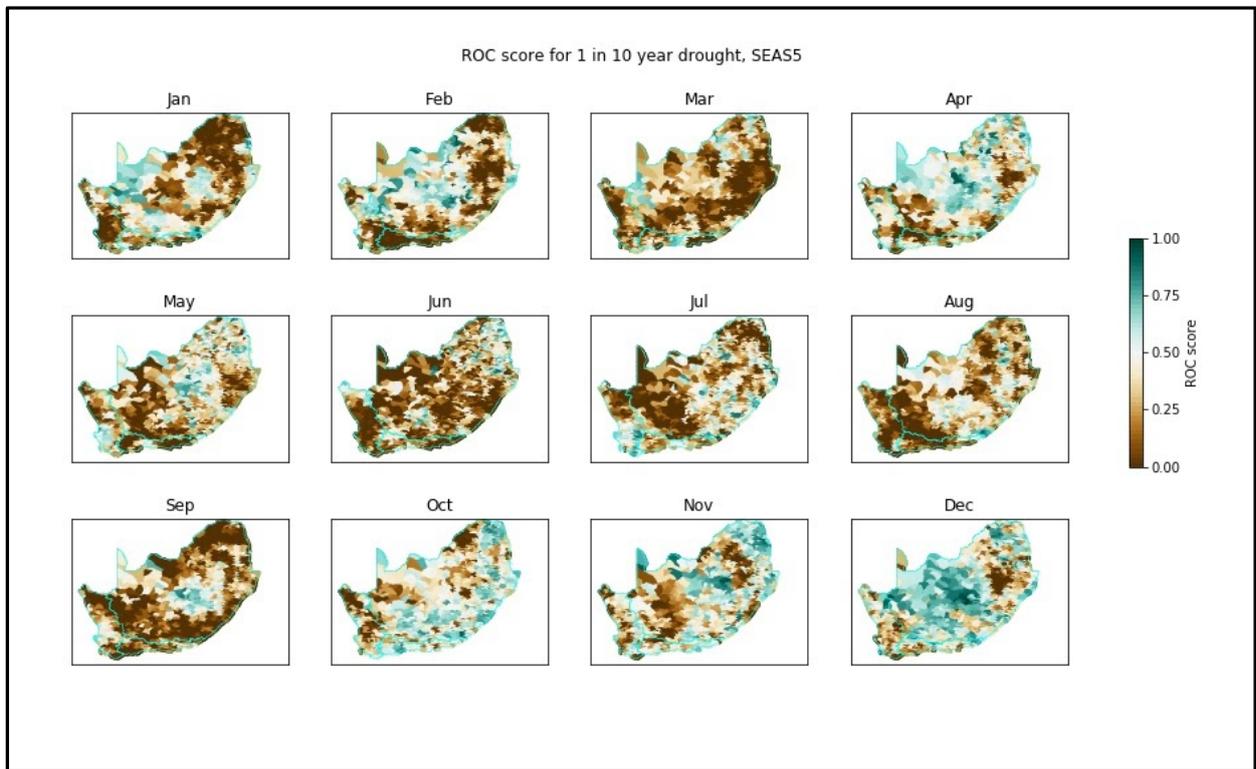


Figure 4.10: As for Figure 4.8 but for forecast of 1 in 10 year drought

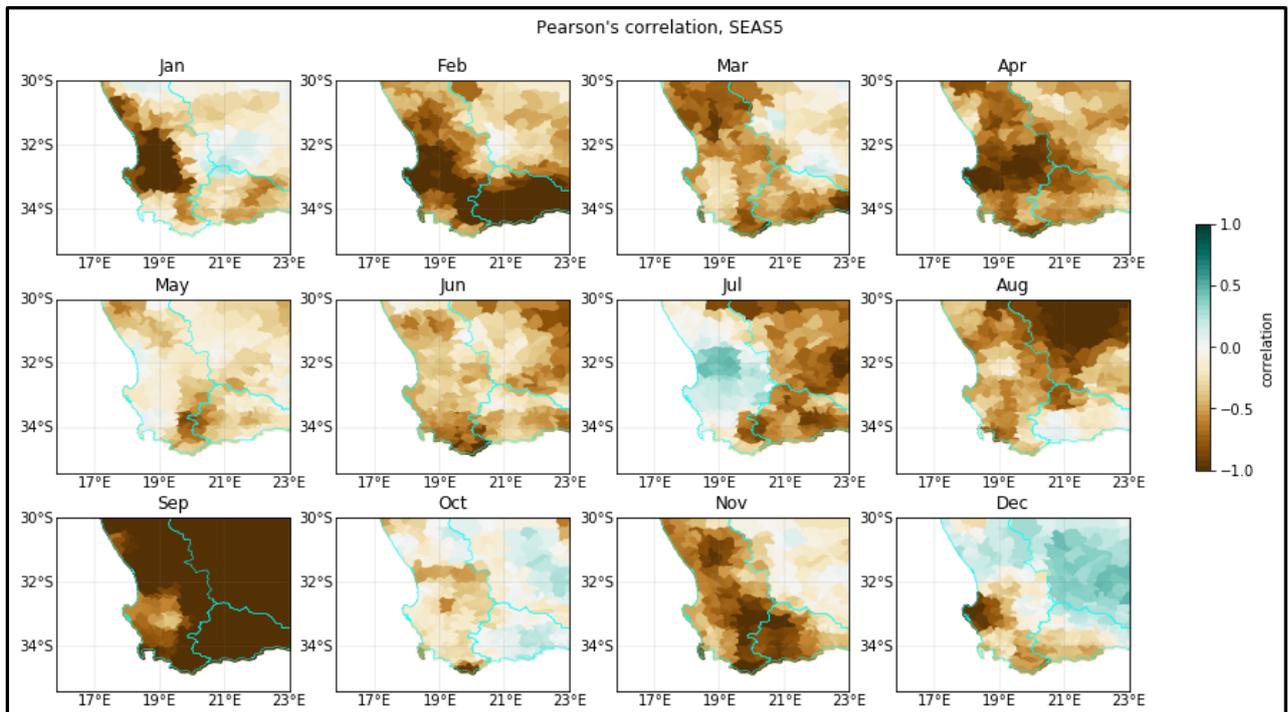


Figure 4.11: As for Figure 4.7 but for the Western Cape only

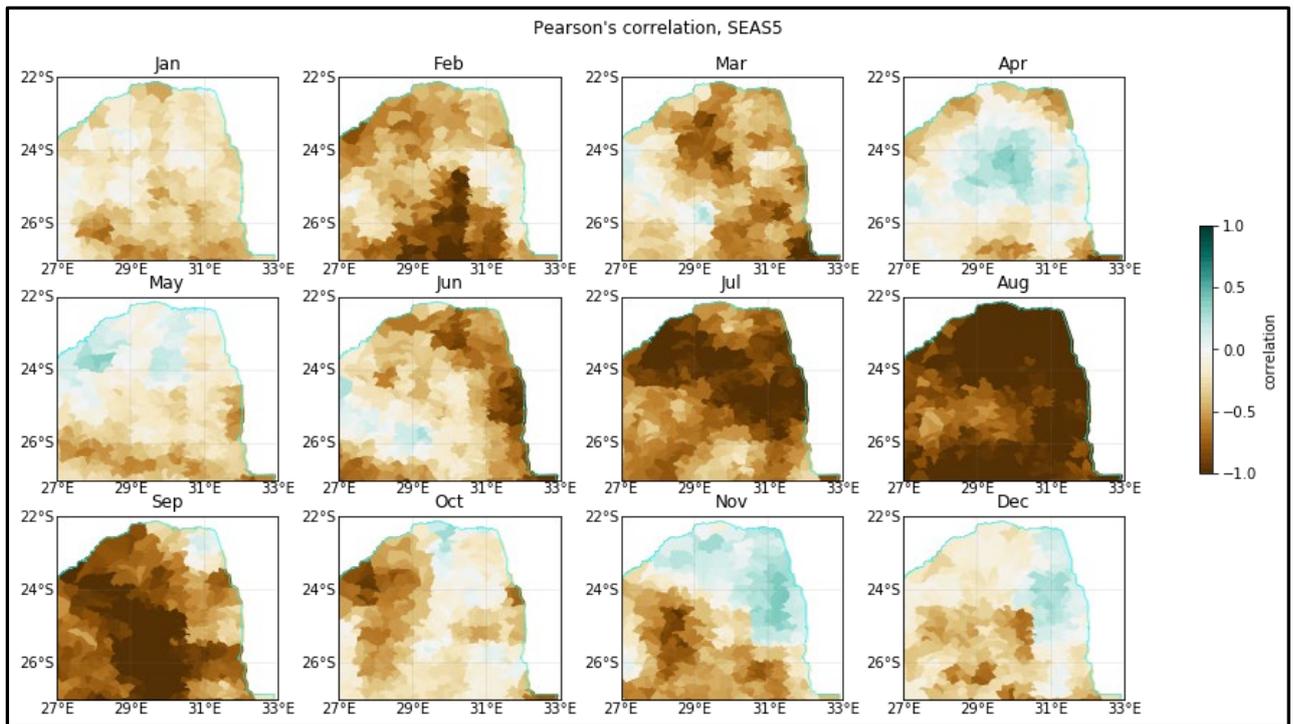


Figure 4.12: As for Figure 4.7 but for the Limpopo only

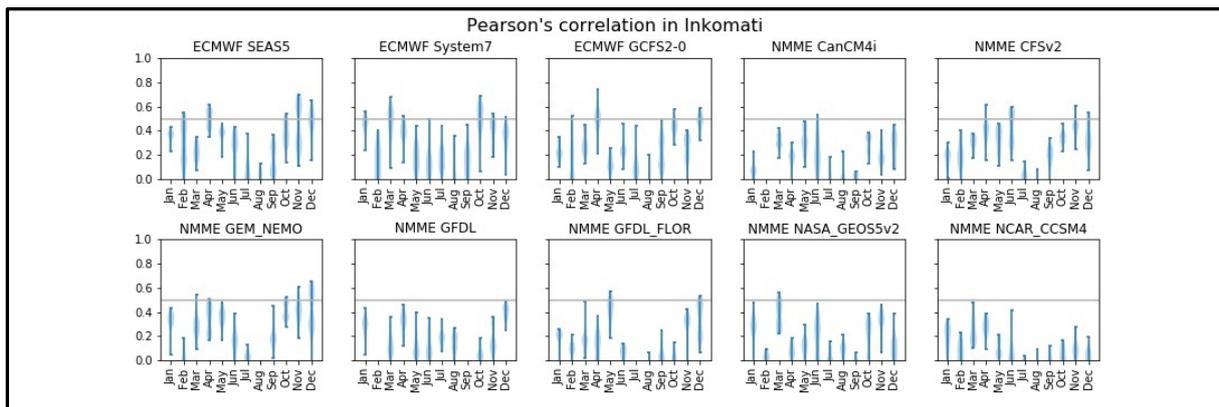


Figure 4.13: Deterministic forecast skill (Pearson's correlation)

Figure 4.13 shows the forecast skill of three month rainfall mean, issued on each of the calendar months by each of the ECMWF and NMME ensemble models for the Inkomati basin. Each distribution illustrates a range of forecast skill in quaternaries of the Inkomati basin.

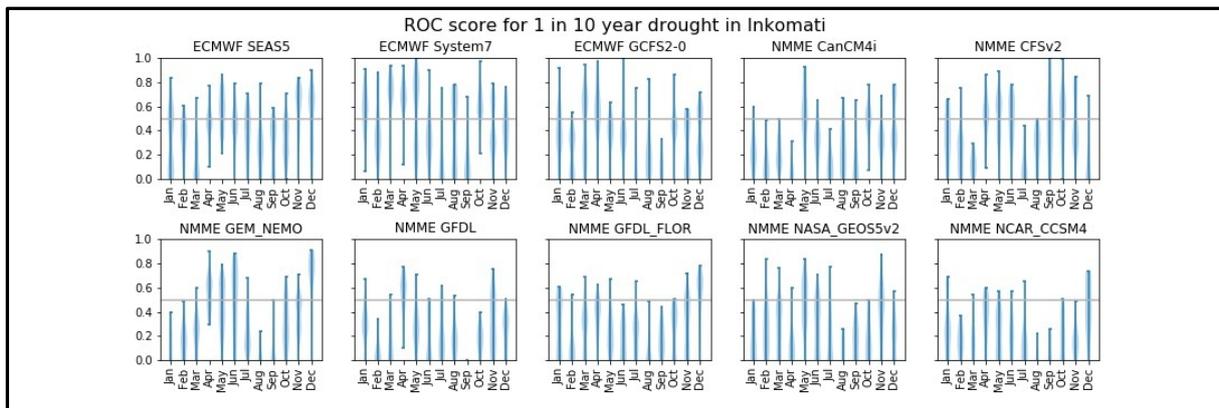
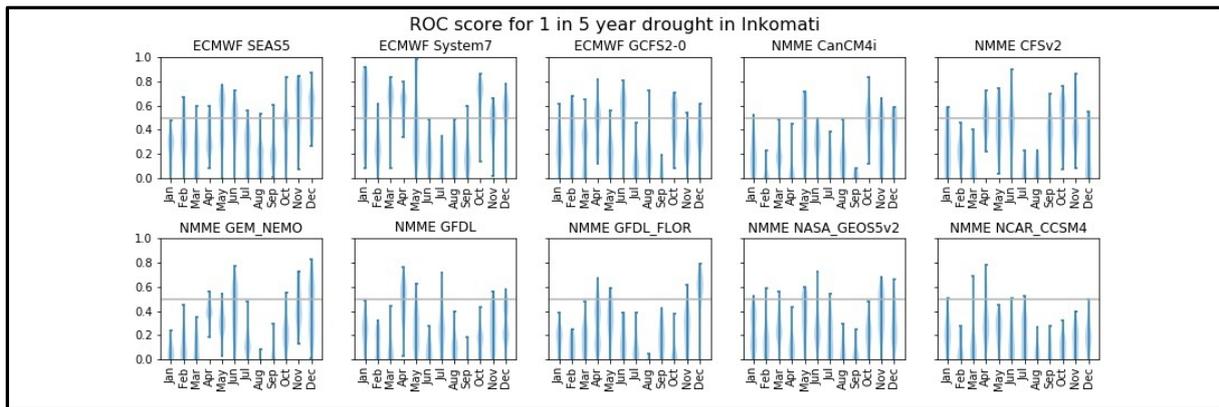
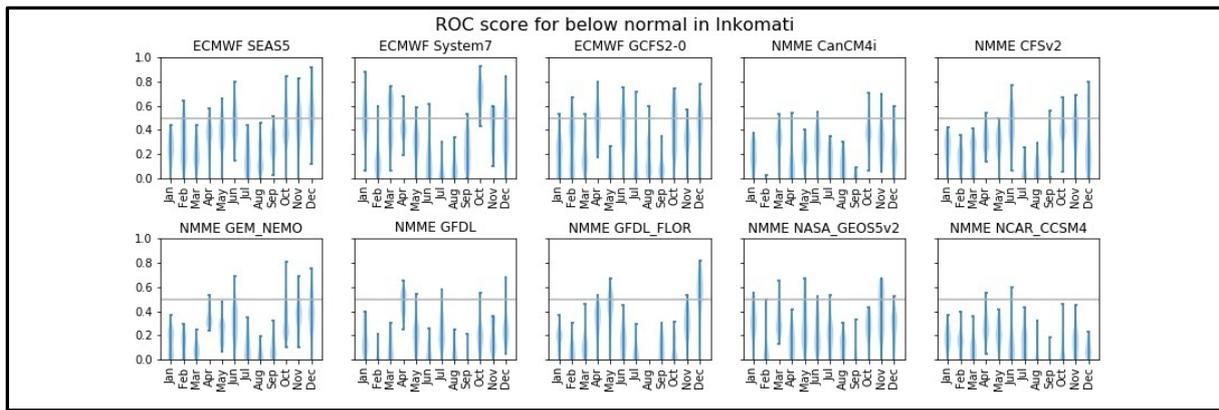


Figure 4.14: Probabilistic forecast skill (ROC score)

Figure 4.14 shows the forecast skill at the scale of quaternary catchments, for forecast of below-normal rainfall (1 in 3 year drought), 1 in 5 year drought, and 1 in 10 year drought, issued on each of the calendar months by each of the ECMWF and NMME ensemble models for the Inkomati basin. Each distribution illustrates a range of forecast skill in quaternaries of the Inkomati basin. Forecast is better than a random guess when the ROC score is > 0.5.

4.4. Statistical forecast of end-of-season anomalies based on anomaly persistence and teleconnections

4.4.1. Introduction

The approach presented here was originally developed by P. Wolski and P. Johnston for the forecast of rainfall anomaly during the 2015-2017 Cape Town Drought, and published as a popular science blog <https://www.groundup.org.za/article/will-there-be-more-rain-winter/> and <http://www.csag.uct.ac.za/2018/03/15/will-there-be-more-rain-this-winter/>

The approach is based on observations illustrated in Figure 4.15, that rainfall anomaly observed in the beginning of the season tends to persist throughout the season, particularly when considered in terms of accumulated, or total season's rainfall.

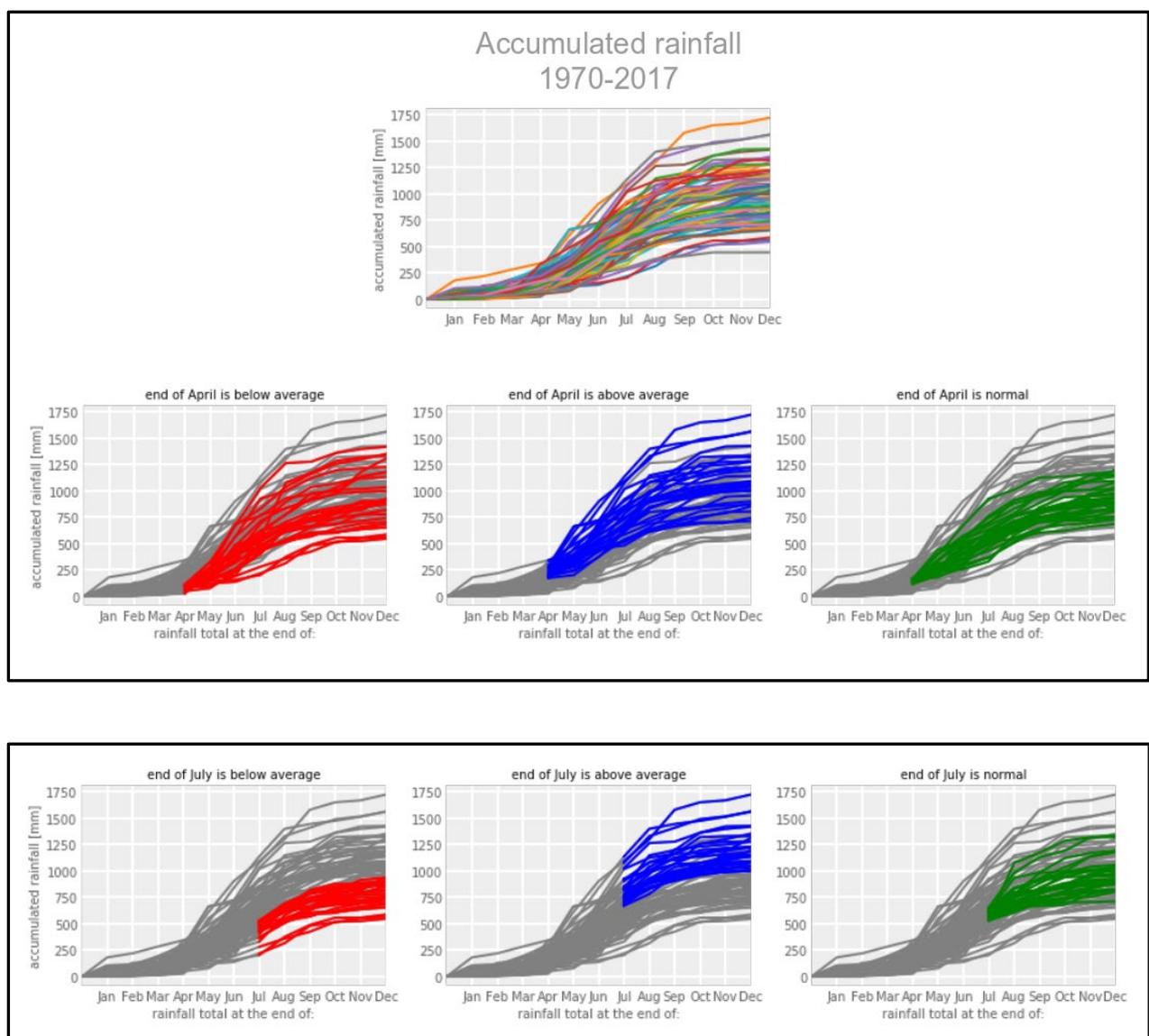


Figure 4.15: Illustration of relationship between end of season accumulated rainfall and current rainfall anomaly

The top row of Figure 4.15 illustrates monthly accumulated rainfall for each year in a 47 year period. Middle row from top illustrated how (categorical) anomaly in the end of April diversifies accumulated rainfall “trajectories”. The bottom row shows the same but for the end of July.

There could be an underlying climate factor that causes lower (or higher) rainfall, and that persists throughout seasons. In this way, the amount of rainfall in the beginning of the rainy season is an indicator of the amount of rainfall the rest of the season receives. However, the majority of the effect arises due the fact that this considers the accumulated rainfall figures. As a result, an anomaly occurring earlier in the season has bearing to the total rainfall at the end of the season. The role of the actual anomaly increases as the season progresses – thus anomaly in the beginning of the rainy season has little implications to the total rainfall that year, but anomaly towards the end of the season is not likely to be reflected in the annual total.

This observation lends itself to the formulation of a probabilistic forecast of end of the season anomaly as a function of the anomaly in the month of forecast in two forms.

- Categorical forecast – probability of above-, below- and normal rainfall total based on current tercile. This forecast is simply formulated by creating contingency tables of number of cases when in historical observations an association between current tercile and end of the season tercile occurred. This contingency table can then be presented in terms of probabilities. A series of contingency tables leads to a schematic as in Figure 4.16.
- Forecast of rainfall anomaly based on linear regression between current and end of the season anomaly. The basis for this forecast is illustrated in Figure 4.17. In the forecast, a linear regression is constructed based on historical data and parameters of the regression equation are used for prediction based on a given data. Since prediction using linear regression have an associated prediction error, that can be used to formulate prediction probabilities for different categories of anomalies, i.e. below-, above- or normal, or other, more relevant from the point of view of drought – e.g. 1 in 10 year drought.

Only the latter is described here in detail.

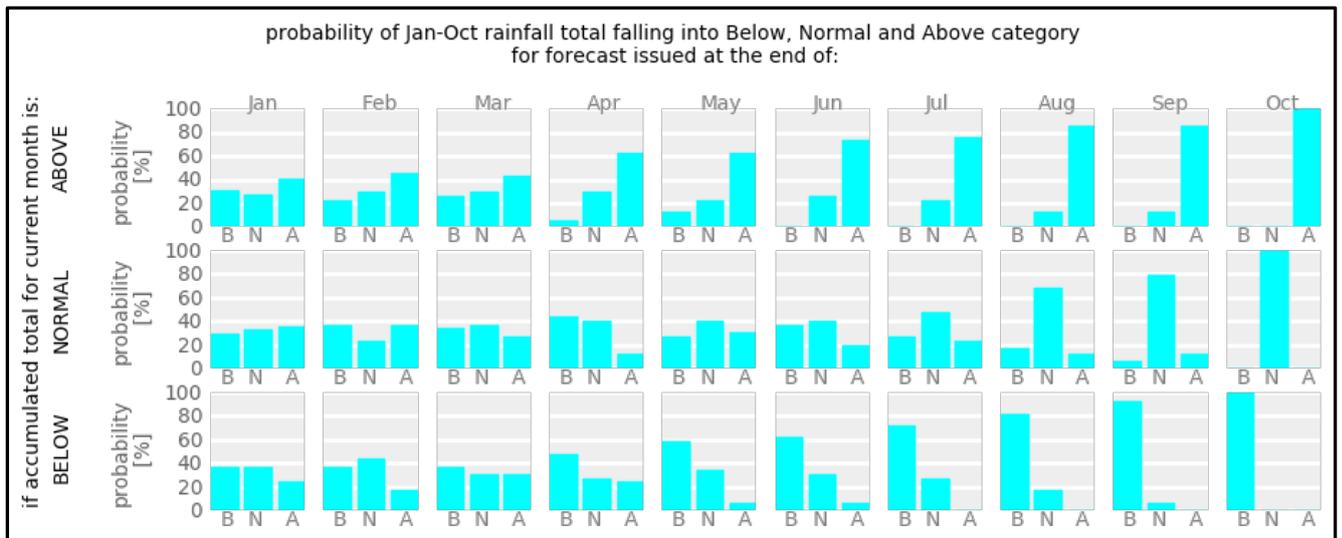


Figure 4.16: Simple categorical forecast of end of the season anomaly of total rainfall based on the current anomaly of accumulated rainfall

This particular forecast shown in Figure 4.16 is for a sub-catchment located in the winter rainfall region, with the extended rainy season spanning April-September. End-of-the season drought can be predicted with a considerable probability already in April.

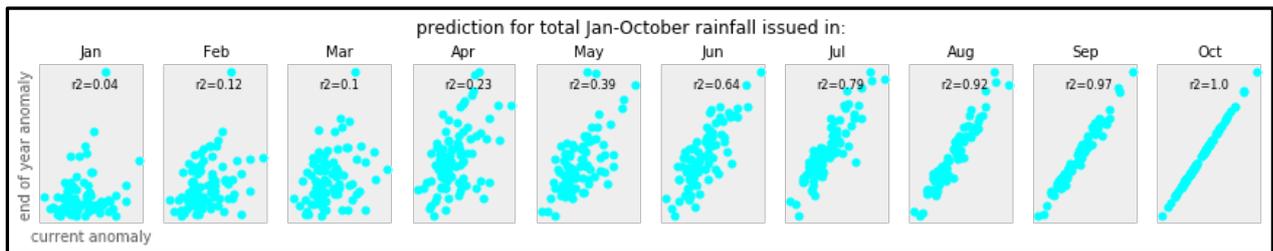


Figure 4.17: Correlation between the end-of-season and current anomaly of accumulated rainfall in a sub-catchment in the winter rainfall region

4.4.2. Forecast of rainfall anomaly based on linear regression between with current rainfall anomaly

In the basic version of the forecast, a linear regression is constructed based on historical data and parameters of the regression equation are used for prediction based on a given data obtained from monitoring.

$$PA_{season} = a_m PA_m + b_m + \epsilon_m$$

where PA_{season} is the end of season anomaly, PA_m is the anomaly of accumulated rainfall in a given month, a_m and b_m are parameters of the regression model that are specific to the month (and obviously location), and ϵ_m is the model error.

The forecast is then simply based on

$$PA_{season,fcst} = amPA_{m,obs} + bm + m$$

Where index *fcst* indicates the value that is forecast, and index *obs* indicates an observed value of anomaly in a given calendar month.

Since the prediction using linear regression have an associated prediction error (in the simplest linear regression case considered to be normally-distributed), that can be used to formulate prediction probabilities for different categories of anomalies, i.e. below-, above- or normal, or other, more relevant from the point of view of drought – e.g. 1 in 10 year drought.

The extended version of the forecast includes additional variables. Since it is known that seasonal rainfall anomalies in South Africa are associated with the state of global modes of variability, such as ENSO, AAO and IOD, and that these modes drive seasonal predictability of rainfall (see Chapter 3), they were included as additional explanatory variables in the regression model.

$$PA_{season} = amPA_m + cENSO_m + bm + m$$

where $ENSO_m$ is the value of ENSO index in month *m*.

Implementation of the model

The forecast model has been implemented at the WR2012 quaternary catchments using gridded blended satellite-station rainfall product – CHIRPS v. 2.0 (Funk et al., 2014). The monthly gridded data over the 1970-2018 period were interpolated to the level of WR2012 sub catchments. The regression model was developed for each individual catchment. Because South African rainfall is characterized by three seasonality regimes – summer, winter and all-year-round, the model was set up considering local season for each of the quaternaries. That differentiation was obtained by hierarchical clustering of standardized rainfall climatology into three classes. The Jan-Dec season for winter rainfall regime was used, and the July-June rainfall season for summer and all-year-round rainfall regime.

Interrogating the forecast model

The performance of the statistical forecast model for various months can be interrogated by considering the amount of variance the regression model explains in the end of the season accumulated rainfall anomaly. In simple terms:

$$VPA_{season,obs} = VPA_{season,model} + V$$

where VPA is the variance of observations, regression fitted values and regression residuals respectively.

Furthermore $VPA_{season,model}$ can be disaggregated into $VPA_{season,committed}$ and $VPA_{season,persistent}$,

Considering that:

$$PA_{season} = PA_m + PA_{anon-committed}$$

and

$$PA_{anon-committed} = dPA_m + PA_{residual}$$

and that $VPA_{committed}$ is the variance of PA_m , and variance of $PA_{residual}$ is V_ϵ .

The variances can then be calculated and plotted as in Figure 4.18. It is clear from this figure that the majority of the explanatory power arises from what is called the “committed” anomaly, i.e. anomaly that arose in a given month.

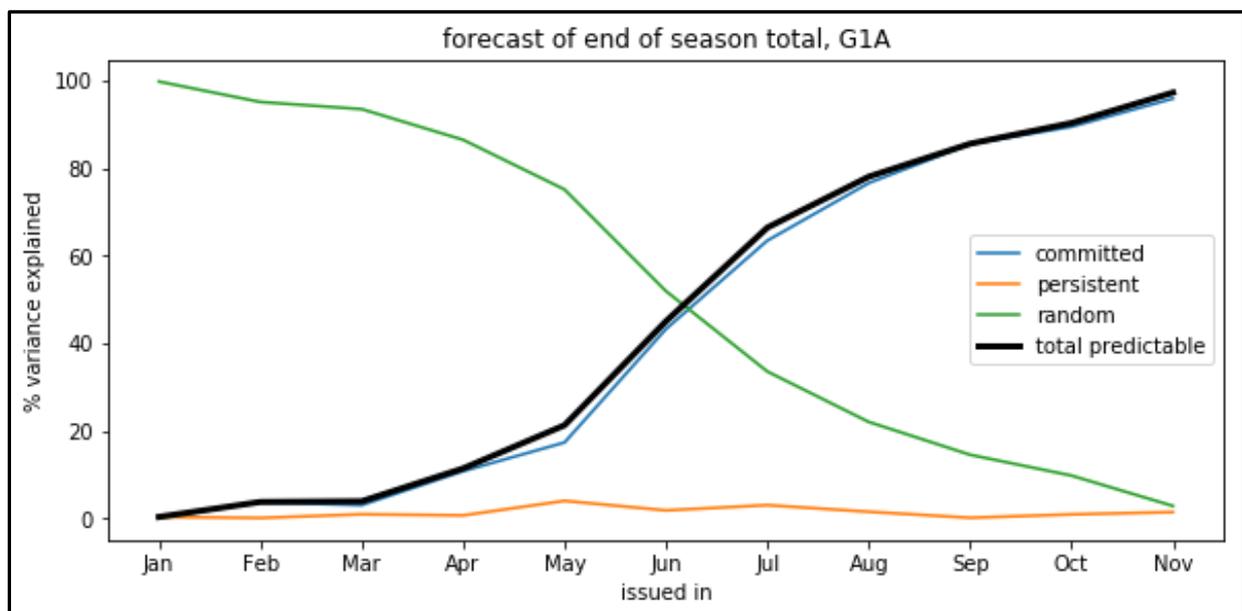


Figure 4.18: Separation of variance explained by basic linear model in end of the season anomaly into components

The approach can be applied to the extended model that includes ENSO, and the results are shown in Figure 4.19. That figure illustrates that in the case of this particular quaternary catchment, the role of ENSO is limited – including it in the regression model increases explanatory power of that model by ~4-5% in May and June only.

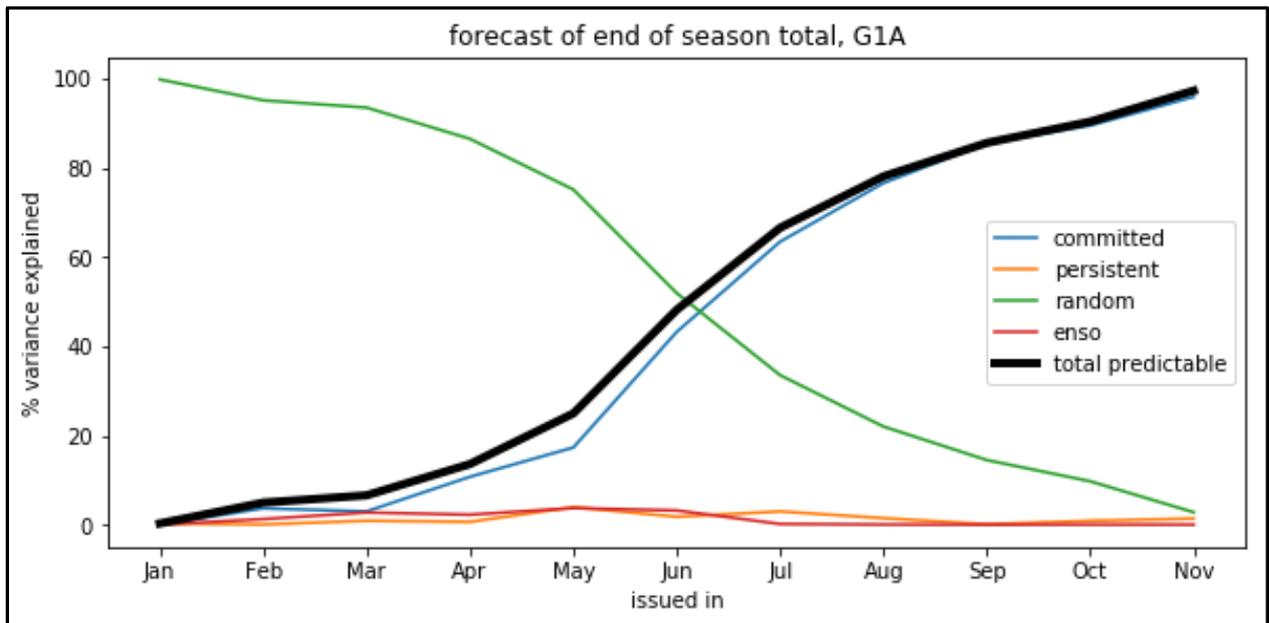


Figure 4.19: Separation of variance in end of the season anomaly into components

The components in Figure 4.19 are explained by an extended linear model that includes ENSO as an explanatory variable

A comprehensive mapping of the role of various variance components in the regression model for the entire country is illustrated in Figures 4.20 and 4.21. Identified homogeneous regions differ relatively little in the partitioning of variance, with the differences highlighting minor differences in rainfall seasonality within each of the rainfall regimes. There are also some differences in terms of the role of persistence and ENSO, but in general, neither ENSO, nor persistence of anomaly do not explain more than 10% of variance in the end-of-the-season anomaly.

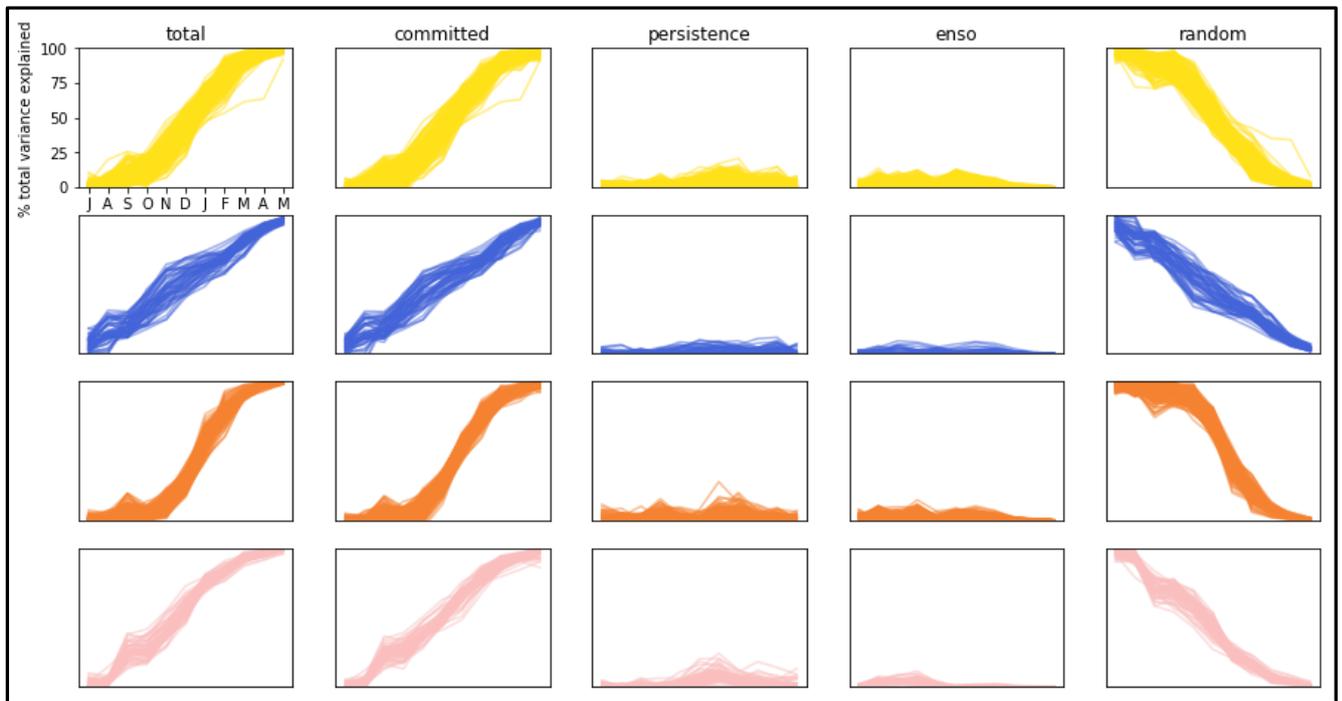


Figure 4.20a: Regions of similar separation of variance explained by a regression model (with ENSO)

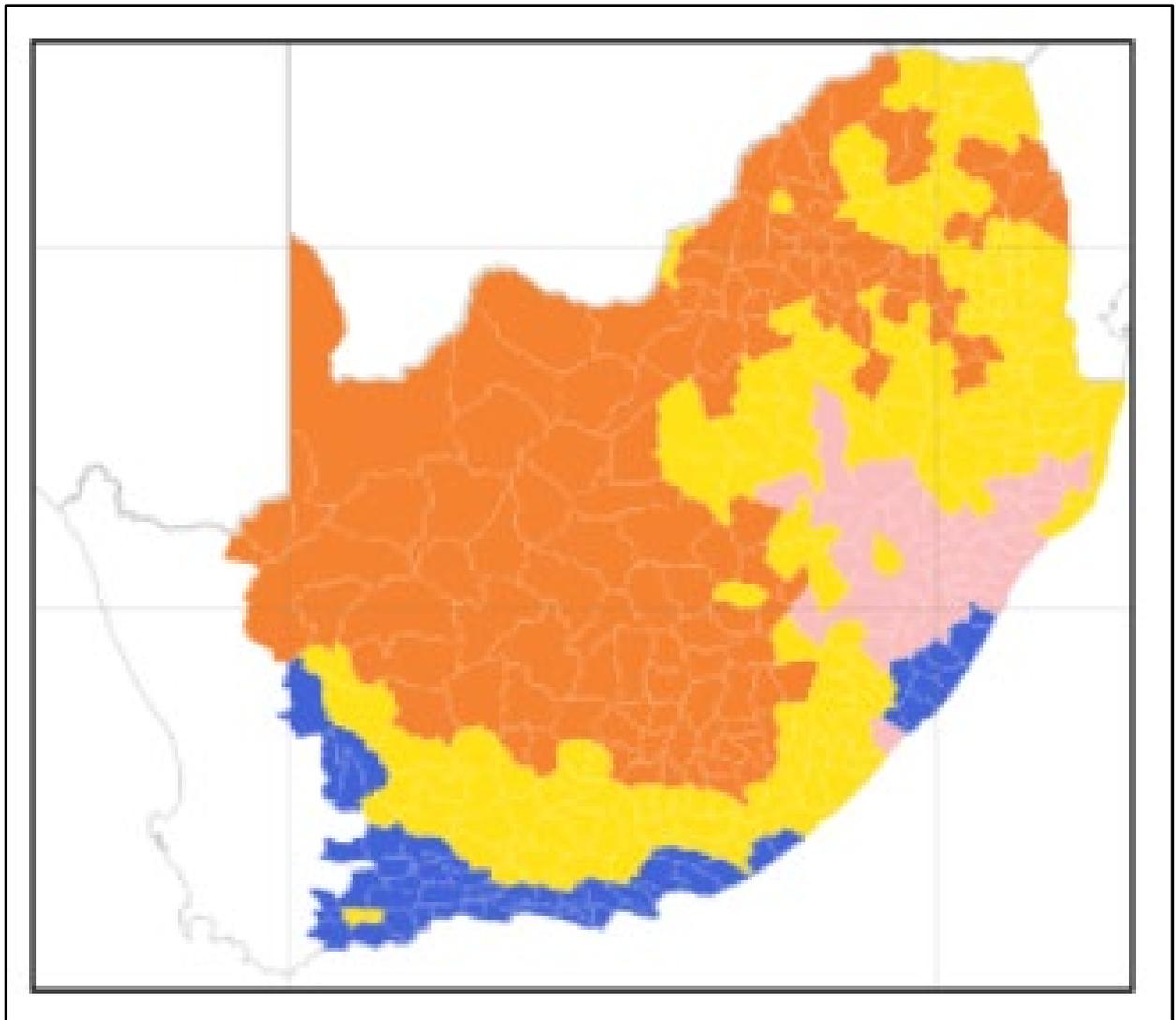


Figure 4.21b: Regions of similar separation of variance explained by a regression model (with ENSO)

Figure 4.20a and 4.20b show the end of the season rainfall anomaly in the summer and all-year-round rainfall regimes. Regions were defined by hierarchical clustering of variance component curves. The colours in the map (Figure 4.20b) correspond to the colours in the summary graphs (Figure 4.20a). Each line in the summary graphs represents single quaternary catchment.

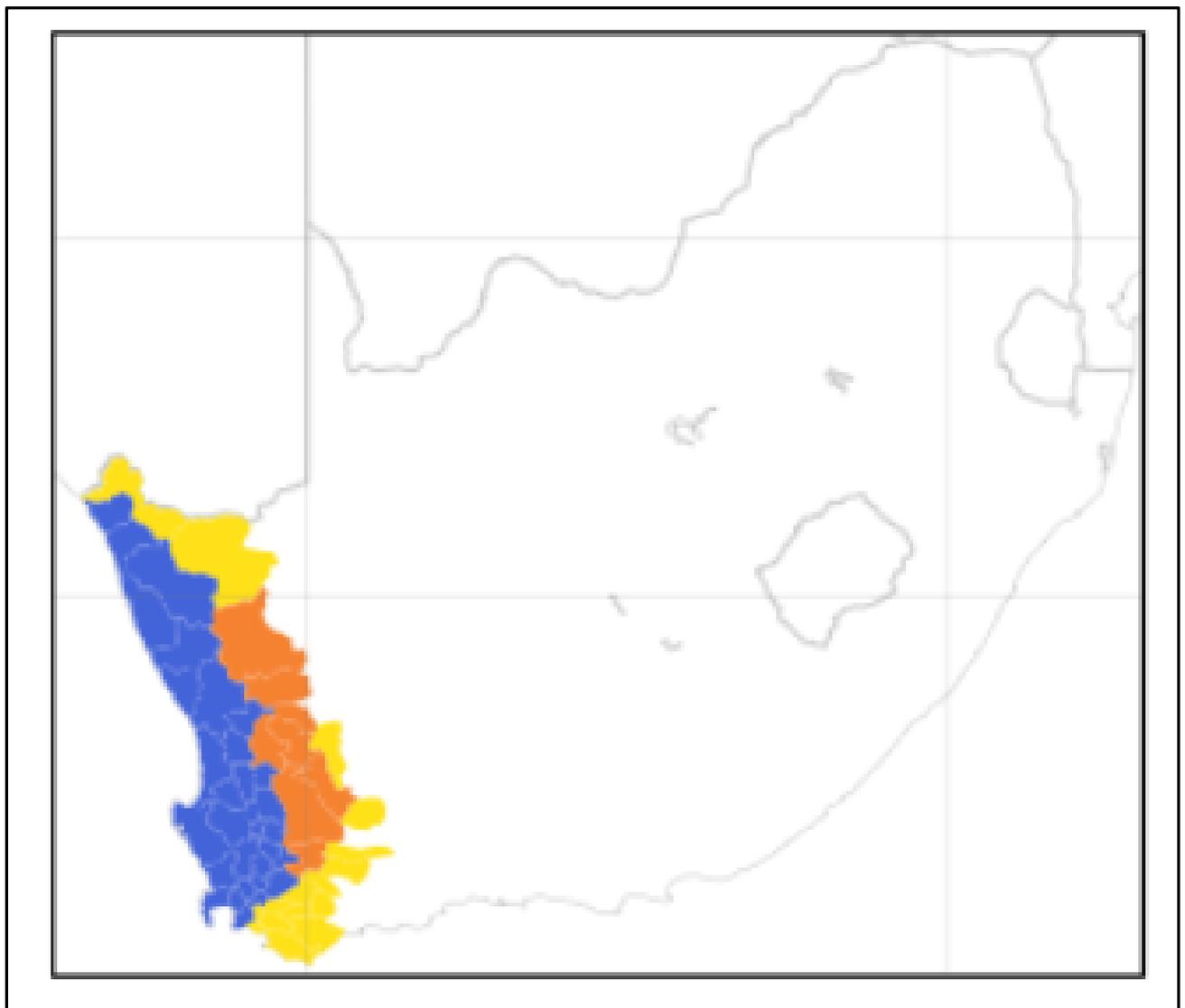
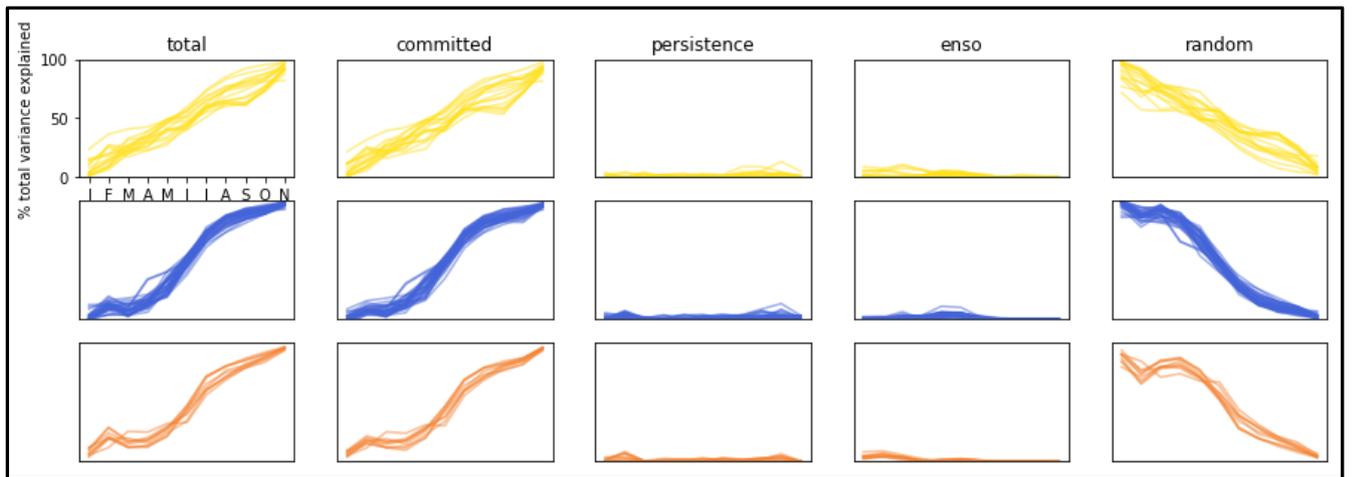


Figure 4.22: As Figure.4.20, but for the winter rainfall regime

Skill of the forecast model

The skill of the forecast was assessed formally using ROC skill score. Figures 4.22 and 4.23 show skill of the forecast (in the version including ENSO) in predicting the below-normal conditions, or in other words a 1 in 3 year drought.

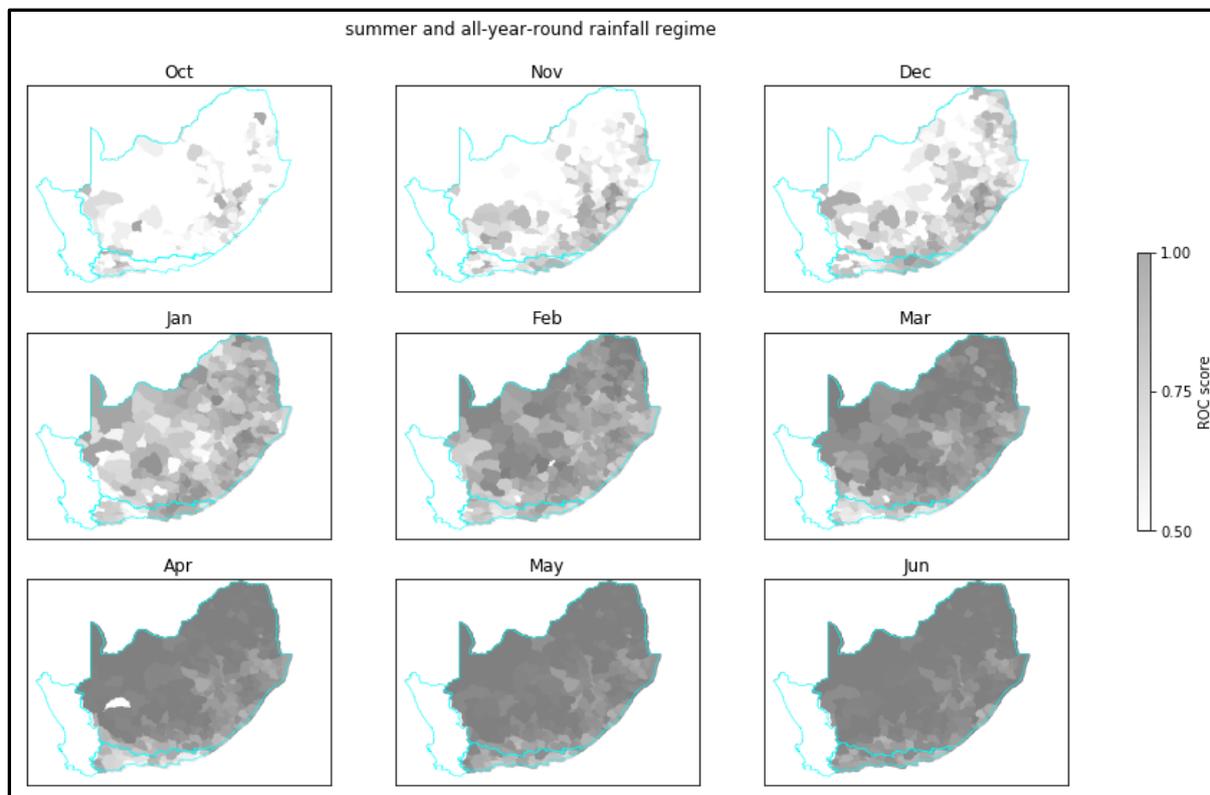


Figure 4.23: Skill of the statistical forecast of the end of the season rainfall anomaly expressed in terms of ROC score

The skill shown in Figure 4.22 is the ROC score for the below-normal category (or 1 in 3 year dry conditions), for summer and all-year-round rainfall regime

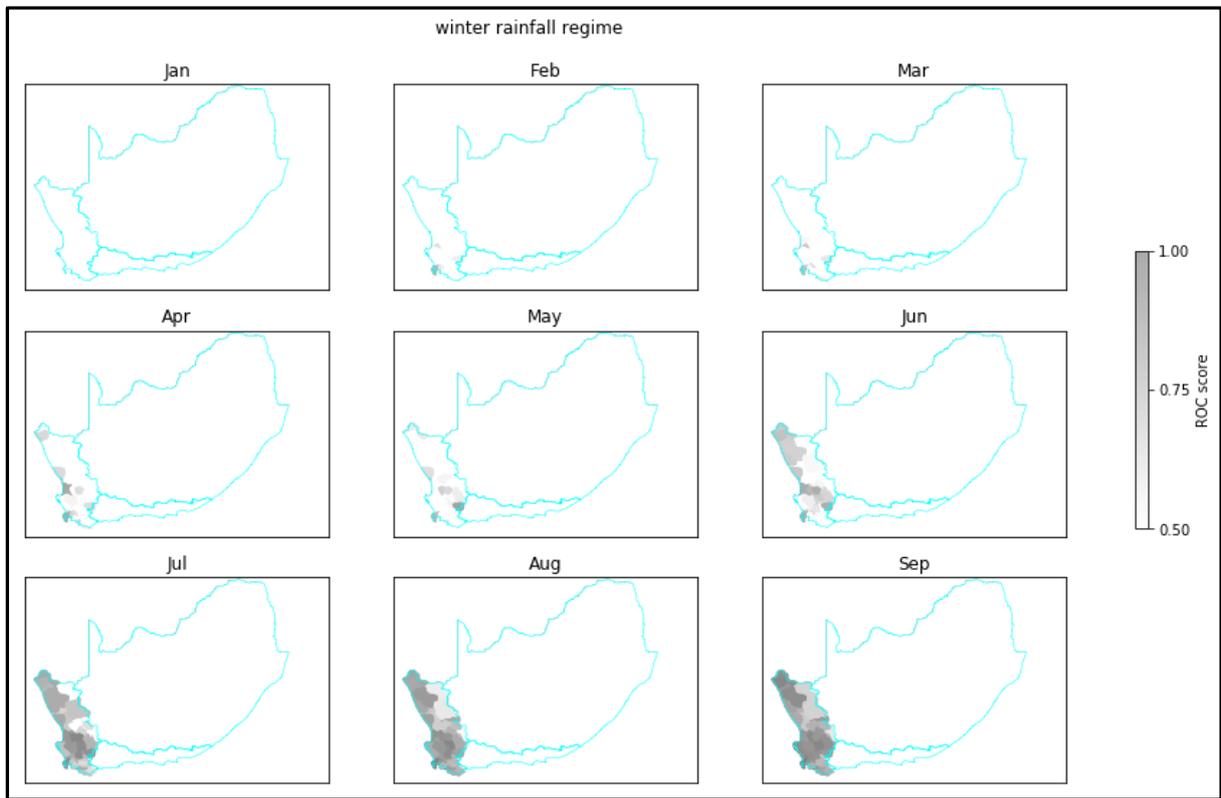


Figure 4.24: As for Figure 4.22, but for winter rainfall regime

The website <http://cip.csag.uct.ac.za/forecast/> was developed as part of this project to spatially present the GCM data and the statistical analysis of all data sets described above. See Figure 4.24.

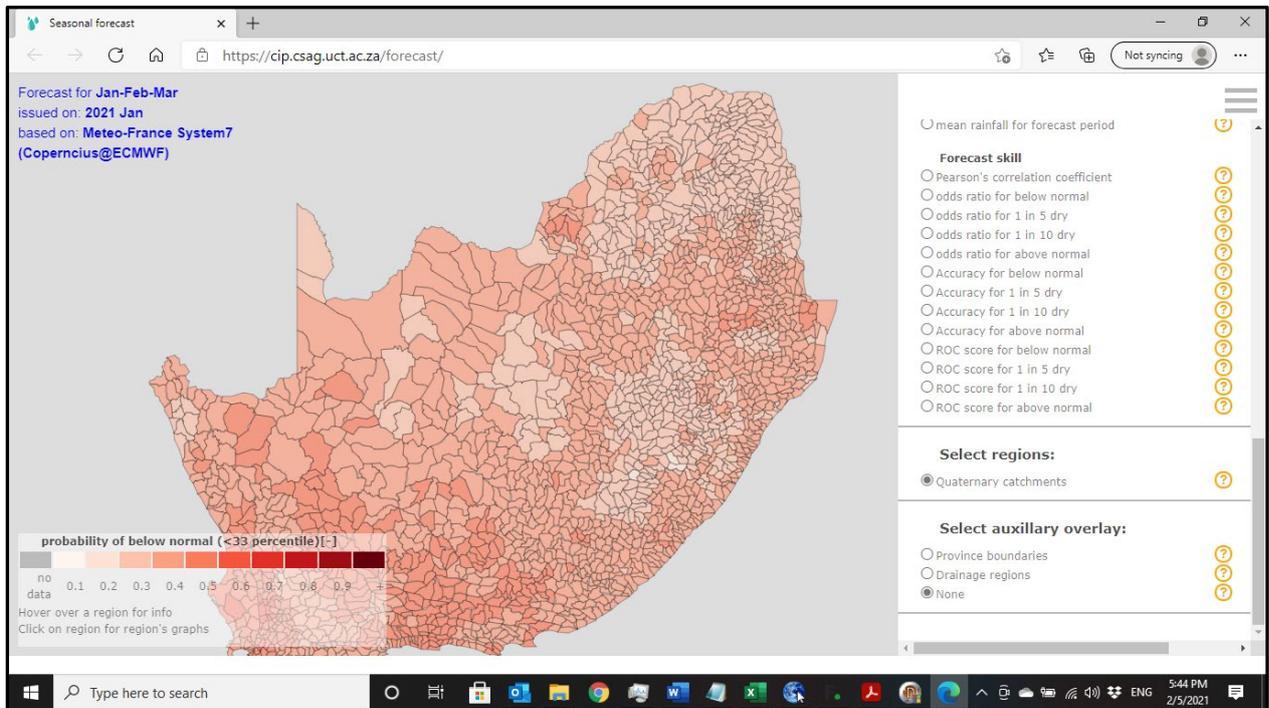


Figure 4.25: CSAG rainfall forecasting website

5 STREAMFLOW AND STORAGE MONITORING AND PREDICTABILITY

5.1 Introduction

While streamflow is a popular indicator of drought since it is easily observed by hydrologists, water managers and even the general public with the advent of websites which make this information readily available, it is of limited practical application in South Africa. The reason for this is that South Africa experiences highly variable and spatially skewed distribution of streamflow and hence relies on numerous dams to supply bulk of water to cities, industries and irrigators. The exception is rural communities who rely directly on streamflow and are therefore very vulnerable to drought. Predictability of streamflow is therefore addressed in this report within the context of run-of-river abstraction, which due to the nature of South African river entails short-term predictability.

5.2 Short-term streamflow predictability

5.2.1 Literature review

It is common to study the characteristics of streamflow-related droughts on seasonal time scales (Vidal et al., 2010). Hydrological predictability at seasonal lead times (1-6 months) comes from knowledge of initial hydrological conditions (soil moisture, groundwater and streamflow) and seasonal climate forecast skill of meteorological variables (Shukla et al., 2013). However, streamflow droughts also appear on shorter time scales (Tallaksen et al., 1997). Some of the short, consecutive events might be connected to one prolonged drought, nevertheless they can cause damage to certain sectors, especially rural communities dependant on river flow, and are worthwhile to consider individually. Streamflow drought forecasts on time scales up to one month are potentially useful for hydropower generation, irrigated agriculture, water quality, navigation and tourism, in general, all sectors that can use the information for upcoming streamflow drought events to take preventative action. For these sectors, skilful forecasts of streamflow droughts could help to prevent or mitigate the consequences of water shortage (Steinemann, 2006).

In the past, numerous studies have investigated the contributions of the initial hydrologic conditions in seasonal hydrologic predictability in different regions in the world. Maurer and Lettenmaier (2003) used multiple regression to identify the sources of hydrologic predictability in the Mississippi River basin and found that soil moisture was the primary source of runoff predictability at one (1) month lead time in all seasons except the summer months over western mountainous region, where snow dominated the runoff predictability. In a similar study using Principal Component Analysis, Maurer et al. (2004) investigated the controlling factors to the runoff predictability over all of North America and concluded that soil moisture and snow water content could provide useful levels of seasonal hydrologic predictability beyond what is available via climate only. Berg and Mulroy (2006) utilized a residual analysis approach and found that a statistically significant number of stations in the Nelson River basin in Canada even macroscale estimates of initial soil moisture could be used to improve streamflow

predictability at one (1) to three (3) months lead time. Likewise, Mahanama et al. (2008) showed that in the tropical island country of Sri Lanka, initial soil moisture and monthly runoff contribute to the seasonal hydrologic predictability up to 3 months lead time. They found the correlation between initial soil moisture and monthly runoff to be highest at 1 month lead time mainly during April-May June and July-August-September. Based on their results, they concluded that improving estimate of initial soil moisture is far more achievable than improvement of seasonal precipitation forecast skill. More recently, Koster et al. (2010) and Mahanama et al. (2011) used suites of hydrological models to evaluate the contributions of soil moisture and snow water content to streamflow predictability across United States. These studies indicated that the contribution of initial hydrologic conditions to seasonal hydrologic predictability was consistent among hydrologic models. All the studies cited above have addressed sources of seasonal hydrologic predictability and their relative influence. However, as far as can be ascertained, only Shukla et al.(2013) attempted to address this globally, where they found out that in the southern hemisphere, initial hydrologic conditions mainly dominate during forecasts periods starting 1 April and 1 July over arid regions and temperature dry winter regions. On the other hand, the study demonstrated that over equatorial humid and monsoonal climate regions the contribution of climate forecast skill is higher than the initial hydrologic conditions throughout most of the year.

In the past, numerous studies have investigated the contributions of the initial hydrologic conditions in seasonal hydrologic predictability in different regions in the world. Maurer and Lettenmaier (2003) used multiple regression to identify the sources of hydrologic predictability in the Mississippi River basin and found that soil moisture was the primary source of runoff predictability at 1 month lead time in all seasons except the summer months over western mountainous region, where snow dominated the runoff predictability. In similar study using Principal Component Analysis (PCA), Maurer et al. (2004) investigated the controlling factors to the runoff predictability over all of North America and concluded that soil moisture and snow water content could provide useful levels of seasonal hydrologic predictability beyond what is available via climate only. Berg and Mulroy (2006) utilized a residual analysis approach and found that a statistically significant number of stations in Nelson River basin in Canada even macroscale estimates of initial soil moisture could be used to improve streamflow predictability at 1 to 3 months lead time. Likewise, Mahanama et al. (2008) showed that in the tropical island country of Sri Lanka, initial soil moisture and monthly runoff contribute to the seasonal hydrologic predictability up to 3 months lead time. They found the correlation between initial soil moisture and monthly runoff to be highest at 1 month lead time mainly during April-May June and July-August-September. Based on their results, they concluded that improving estimate of initial soil moisture is far more achievable than improvement of seasonal precipitation forecast skill. More recently, Koster et al. (2010) and Mahanama et al. (2011) used suite of hydrological models to evaluate the contributions of soil moisture and snow water content to streamflow predictability across United States. These studies indicated that the contribution of initial hydrologic conditions to seasonal hydrologic predictability was consistent among hydrologic models. All the studies cited above have addressed sources of seasonal hydrologic predictability and their relative influence. However, as far as can be ascertained, only Shukla et al. (2013) attempted to address this globally, where they found out that in the southern Hemisphere, where South Africa is located, initial hydrologic conditions mainly dominate during forecasts periods starting 1 April and 1 July over arid regions and temperature

dry winter regions. On the other hand, the study demonstrated that over equatorial humid and monsoonal climate regions the contribution of climate forecast skill is higher than the initial hydrologic conditions throughout most of the year.

5.2.2 Characterization of hydrological droughts

Low-flow can be defined as a flow of water in a stream during prolonged dry weather. Droughts include low-flow periods, but a continuous seasonal low-flow event does not necessarily constitute a drought, although many researchers refer to a continuous low-flow period in one year as an annual drought. In general, two approaches to predict properties of low streamflow events in the long range can be discerned. First, stochastic approaches that relate the current state of a catchment and potential predictors to what has been observed in the past, to infer the likelihood of low streamflow within the prediction period. These include regression techniques (Moreira *et al.*, 2008), time series models (Bordi and Sutera, 2007) and neural network techniques (Morid *et al.*, 2007). Hwang and Carbone (2009) used autoregressive models to predict drought indices and to quantify the uncertainty in the prediction. The second, less common approach for the long-range prediction of droughts involves a coupled atmospheric-hydrological model. Wood *et al.* (2002) employ monthly forecasts from a global atmospheric model to drive a grid-based hydrological model that produces reasonable predictions of streamflow up to several months in advance. The refined systems of Li *et al.* (2008) and Luo and Wood (2007) were able to predict average monthly drought conditions up to three (3) months ahead.

A number of consecutive time intervals where the selected flow variable (a discharge or flow volume) has lower values than a reference flow level indicates the *duration of a drought event*.

- For each such event, the sum of deviations of a flow variable from the reference level represents the cumulative flow-deficit amount (*drought severity*).
- This deficit divided by the duration is the measure of *drought intensity*.

Streamflow droughts are generally characterized by the indices duration (time between onset and offset), severity (cumulative water deficit) and magnitude (severity/duration) (Tallaksen *et al.*, 1997; Yoo *et al.*, 2011).

There are various drought indices applicable to define streamflow drought and the commonly used are:

- Standardised Runoff Index (SRI) (Shukla and Wood, 2007)
- Indices based on low flows
- Indices based on runoff anomaly.

Defining hydrological droughts solely by considering these indices requires the assignment of thresholds. Whenever the predicted or observed runoff falls below a threshold, this counts as an event of streamflow drought. Figure 5.1 illustrates the streamflow drought indices drawn from an observed or forecast hydrograph, with indices being dependent on choice of threshold. The solid line is the observed or forecast runoff that is in certain periods, above or below

streamflow detection threshold (dashed line). For a forecast member of each 32-day forecast and the corresponding observation, the longest consecutive period below the threshold (streamflow drought duration) is evaluated. The deficit during this period (severity, shaded area) is the cumulative difference between threshold and runoff. The quotient of severity and duration, called magnitude, is evaluated as well. Timing is defined as the moment when half of the event has happened.

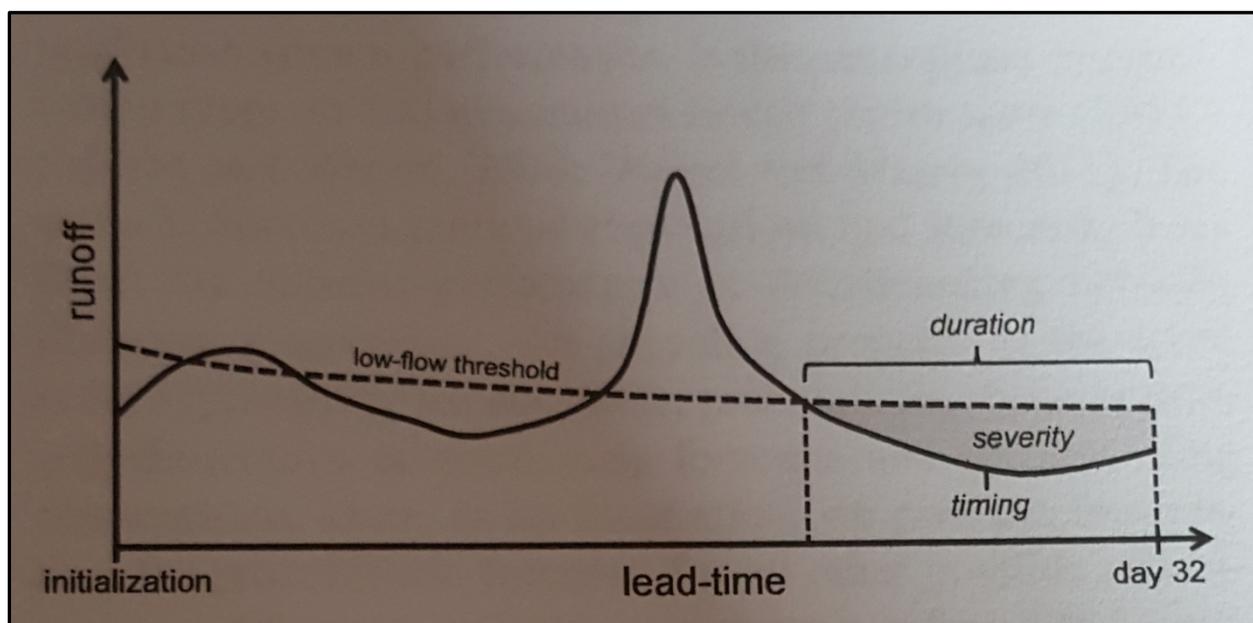


Figure 5.1: Illustration of the streamflow drought indices (Fundel et al., 2013)

The time of drought occurrence has been given different definitions, for instance the starting date of the drought, the mean of the onset and the termination date, or the date of the minimum flow. Often another drought deficit characteristic, the drought intensity, is introduced as the ratio between drought deficit volume and drought duration. Based on the time series of drought deficit characteristics it is possible to derive drought deficit indices (Tallaksen and Van Lannen, 2004).

The threshold might be chosen in many different ways, amongst other things, a function of the type of water deficit to be used. In some applications the threshold is a well-defined flow quantity. It is also possible to apply low flow indices, e.g. percentile from the flow duration curve See Table 5.1. The threshold might be fixed or vary over the year.

Table 5.1 Crocodile River at Karino – river flow thresholds for drought monitoring

X2H006_Crocodile River at Karino													
Percentile	Thresholds	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	June	July	Aug	Sept
0.05	Very Low Flow	4.6	4.6	4.6	5.8	6.7	7.7	4.6	4.6	4.6	4.6	4.6	4.6
0.10	Low Flow	4.8	4.8	4.8	7.4	9.6	8.3	4.8	4.8	4.8	4.8	4.8	4.8
0.25	Below Normal Flow	6.6	7.2	10.1	13.9	12.0	15.2	8.7	6.6	6.6	6.6	6.6	6.6
0.50	Normal Flow	7.7	10.2	15.8	18.2	25.5	24.9	15.9	10.9	7.4	7.4	7.4	7.4
0.75	Above Normal Flow	9.7	13.3	22.2	31.7	43.7	33.2	21.8	14.6	9.5	8.3	8.3	8.3
0.95	High Flow	12.5	31.3	65.2	43.0	86.8	83.9	48.7	24.3	13.3	10.4	9.0	9.9
Max	Extreme High Flow	16.1	44.8	77.3	71.3	109.8	103.0	94.8	47.0	29.4	19.1	12.8	10.8

Runoff Anomaly

The annual runoff variation is used as drought indicator based on anomaly percent. The categories of runoff are separated into 5 according to their percentage anomalies in Table 5.2 which are denoted by digits -2, -1, 0, +1 and +2, where digit 2 refers to the anomaly percent $\Delta < -30\%$ for a low-runoff year, digit -1 refers to $-30\% \leq \Delta \leq -10\%$ for a relatively lower-runoff year; digit 0 represents $-10\% \leq \Delta < 10\%$ for a year of normal runoff; digit +1 denotes $10\% \leq \Delta < 30\%$ for a relatively higher-runoff year; digit +2 refers to $\Delta > 30\%$ for a high-runoff year.

Table 5.2: Runoff Anomaly categories based on their anomaly percent

Runoff anomaly	Corresponding Category	Runoff status
$\Delta < -30\%$	-2	Very Low
$-30\% \leq \Delta \leq -10\%$	-1	Low
$-10\% \leq \Delta < 10\%$	0	Normal
$10\% \leq \Delta < 30\%$	+1	High
$\Delta > 30\%$	+2	Very high

From the runoff anomaly assessment of Wang et al. (2008) it was noted that there are common features of high and low runoff for different rivers and hence set of standards are developed to classify runoff levels (water deficiency or abundance) in rivers to indicate associated drought or flood categories. To achieve this, runoff data should be fit in to follow a normal distribution or other type of distribution. Normalizing runoff would convert the probability density function of Pearson type III distribution into the standard normal distribution as function of Z. According to the properties of a normal distribution of variable Z, the Z values are divided into 5 levels and delimit their corresponding bounded domains as the drought index of each category, given in Table 5.3.

Table 5.3: Z-Index denoted drought

Category	AF	Z value	D/F	TFD
1	>95%	$Z > 1.6448$	Flood	5%
2	70 to 95%	$0.5244 < Z \leq 1.6448$	Light flood	25%
3	30 to 70%	$-0.5244 \leq Z \leq 0.5244$	Normal	40%
4	5 to 30%	$-1.6448 \leq Z < -0.5244$	Light drought	25%
5	< 5%	$Z < -1.6448$	Drought	5%

AF = accumulation frequency, D/F = drought or flood, and TFD = theoretical frequency distribution.

5.2.3 Application of the Standardised Runoff Index (SRI)

Drought, being complex in nature, is difficult to define, quantify and monitor. Streamflow data is generally used for the analysis of hydrological drought (Shukla and Wood, 2008) based on a Streamflow Drought Index (SDI).

This index, $SRI_{i,k}$, requires streamflow volume values $Q_{i,j}$ where i denotes the hydrological year and j -th month within a hydrological year. The cumulative streamflow volume, $V_{i,k}$, for the i -th hydrological year and k -th reference period can be obtained from:

$$V_{i,k} = \sum_{j=1}^k Q_{i,j} \quad i=1,2,\dots, \quad j=1,2,\dots, \quad k=1,2,3,4$$

$$SRI_{i,k} = (V_{i,k} - V_{mean_k}) / S_k$$

Where V_{mean_k} and S_k are respectively the mean and standard deviation of the cumulative streamflow volumes for the k -th reference period.

The definitions of states of drought with SRI are:

Non drought: $SRI \geq 0.0$

Mild drought: $-1.0 \leq SRI < 0.0$

Moderate drought: $-1.5 \leq SRI < -1.0$

Severe drought: $-2.0 \leq SRI < -1.5$

Extreme drought: $SRI < -2.0$

As an illustration, two streamflow stations were selected for calculating and testing the performance of the SRI, one from the Crocodile River at Karino, in the Inkomati-Usuthu Water Management Area and another from the Berg River at Misverstand in the Western Cape. The analysis results shown in Figure 5.2 has shown that the SRI can well discover the main droughts known to have occurred in the Crocodile River sub-catchment: 1982-1984, 1992-1996, and 2015-2016. This sub-catchment is characterized by large spatial variability in the river regimes and flow magnitudes (long dry winter months and short wet months). Similarly, Figure 5.3 indicates similar pattern of droughts, with extreme drought experienced in 2015-2016. From the results below the SRI method can easily be used in an early drought warning system, but good quality streamflow data is required.

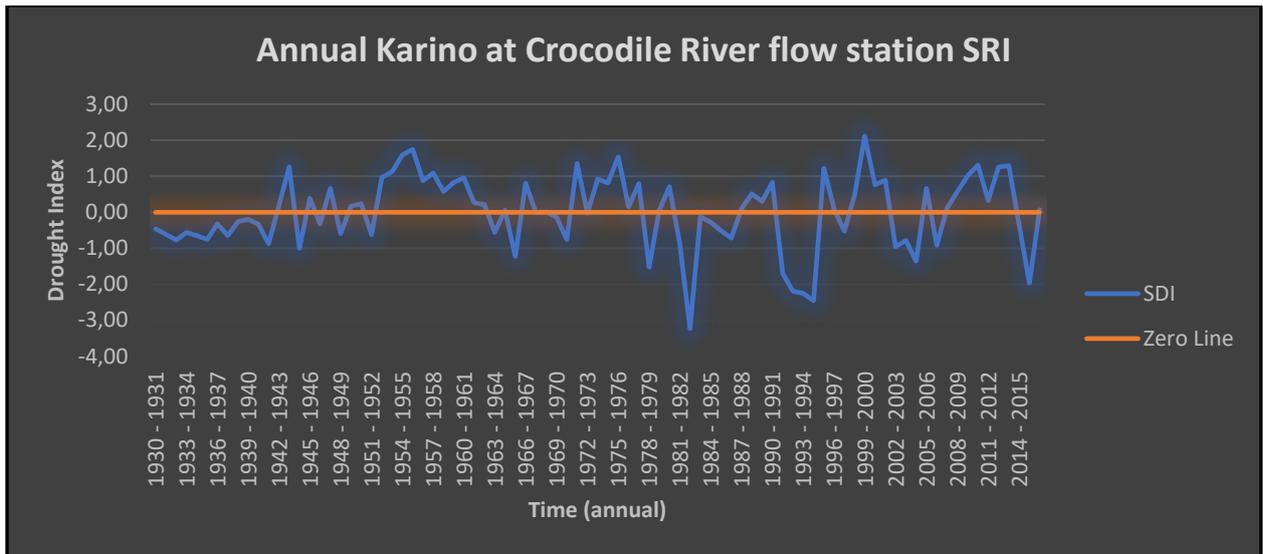


Figure 5.2: SDI Analysis at Crocodile River at Karino River flow station

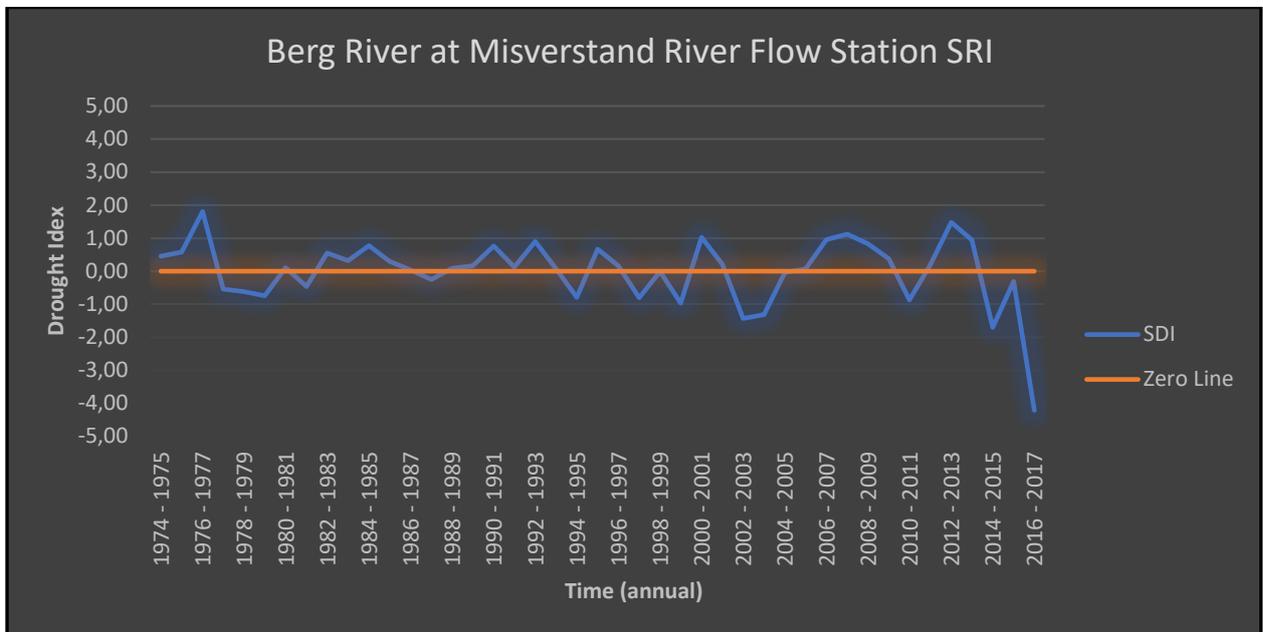


Figure 5.3: SDI Analysis at the Berg River at Misverstand River flow station

5.3 Long term predictability

5.3.1 Literature review

There is a dearth of literature on long-term predictability. Recent papers on this topic refer to 3 to 6 months as 'long-term' and use methods incorporated into the models presented in Section 2, that is, predictions linked to weather phenomenon such as ENSO. 'Long-term' within the context of this project implies greater than six months (the limit of GCM rainfall prediction) and up to the ten years, as required for long term water resources planning. Long-term predictability would rely on identifying and quantifying cycles in streamflow or rainfall.

5.4 Storage prediction

Storage prediction or forecasting has played a major role in the management of South Africa's large dams since the 1980s. This was brought about through the development of a stochastic model by Pegram (DWAf, 2006) which was initially used to manage the Vaal System and later applied to all large bulk water supply systems within South Africa. The method uses an Autoregressive Moving Average model (ARMA) to generate statistically plausible natural flow time series, typically up to 200 time series for so-called long term simulation and 500 to 1 000 time series for short term simulations. A typical outcome from such a process is shown in Figure 5.4. It is important to note, however, that this process does not predict or forecast the future flows but rather indicates the probability of future storage given a widest range of possible future flow events (from extreme drought to very wet periods). This has however proven to be a very useful tool to timeously introduce water restrictions to avoid system failure.

The ARMA model described above assumes that all generated time series are stationary. Hence this method is not applicable to climate change where there is likely to be a trend of decreasing or increasing streamflow. Also, the ARMA does not take into account antecedent conditions but assumes any event is possible. Mallory et al. (2009) used a method described by Sellick and Bonthuys (Sellick, 2008) to incorporate antecedent condition into stochastic streamflow modelling. This concept was applied to Crocodile and Sabie River systems (DWAf, 2009; DWAf 2013). The latter model referred to as the Cai Squared model is a short-term predictive model that works well over the period of April to November in the aforementioned catchments.

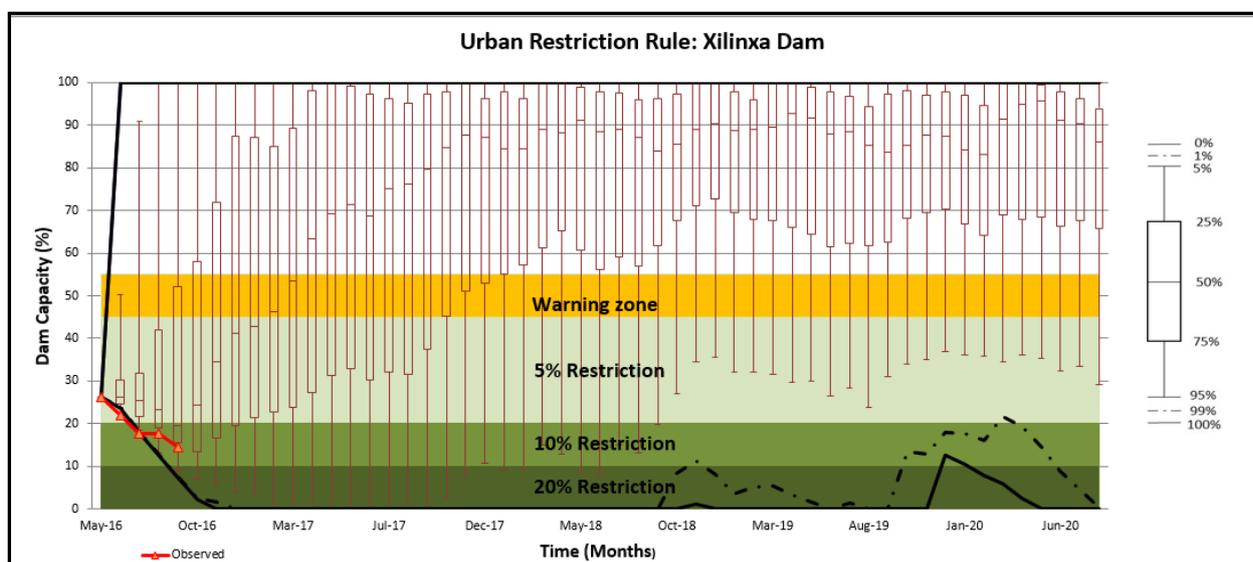


Figure 5.4 : Storage projections using an ARMA stochastic model

6 SOIL MOISTURE AND GROUNDWATER PREDICTABILITY

6.1 Introduction

Different groundwater systems (in terms of aquifer type as well as the local environment) perform differently with respect to drought. According to Peters et al. (2005), this performance is identified by three indicators: reliability, resilience and vulnerability. Reliability is the frequency or probability that a system is in a satisfactory state. Resilience describes how quickly a system is likely to recover once a failure has occurred. The magnitude of failure was assumed to be the drought deficit as defined using the threshold level approach outlined in Peters et al. (2005). Thresholds would need to be defined if this approach is taken forward. Vulnerability is the severity of a failure. Different authors define vulnerability differently, it has been defined as both the average drought deficit (Kjeldsen and Rosbjerg, 2001), and as the maximum drought (Moy et al., 1986). These indicators also do not reflect the dependency between the occurrence, frequency and severity of droughts. Mishra and Singh (2010) point out that due to the shortcomings in the conventional concept of a groundwater drought, increasingly groundwater drought is identified on the basis of the temporal variability of the weather.

Performance indicators have mostly been applied to streamflow or surface water reservoirs. The few studies that investigate the performance of groundwater or baseflow, use predefined indicators to compare the sensitivity or susceptibility to drought of different catchments. These studies often use very different indicators. Currently there is no consensus about how to evaluate or define the performance of groundwater during drought (Peters et al., 2005).

The difficulty in defining drought characteristics has led to compartmentalised and quite specific drought indices, such as the Palmer Drought Stress Index (PDSI, for hydrological drought), the Standardised Precipitation Index (SPI, for meteorological drought), and the Standardised Groundwater level Index (SGI, for groundwater drought). Recently attempts have been made to integrate drought indices (Ma et al., 2014). Thomas et al. (2014) developed a framework to evaluate a holistic drought characterisation using data from the GRACE satellites, focusing on the total water storage deficits to characterise drought occurrence.

The development of groundwater drought indicators has employed water budget approaches (Mendicino et al., 2008), statistical applications using in situ groundwater observations (Bloomfield and Marchant, 2013) or hydrological model simulations (Houberg et al., 2012; Li and Rodell, 2015). Thomas et al. (2017) used a water budget approach to derive a groundwater drought index, and integrated anthropogenic impacts and natural groundwater drought. Both Li and Rodell (2015) and Thomas et al. (2017) compared GRACE derived groundwater storage changes with in-situ observation of changes using a groundwater drought index (originally introduced in Li and Rodell, 2015). Thomas et al. (2017) compared their GRACE derived Groundwater Drought Index (GDI) with in-situ based Groundwater Indices (GWI) and their results documented a strong correlation. In addition, they compared the GDI with various other drought indices such as the PDSI and the SPI and identified a temporal offset of 5 months, suggesting that traditional drought indicators demonstrated

drought conditions 5 months prior to evidence of groundwater drought (for their case study area which was in the Central Valley of California).

Peters et al. (2005) list the overall performance indicators they utilised which are based on the three single performance indicators originally defined by Hashimoto et al. (1982) for a surface water reservoir, and mentioned above (reliability, resilience and vulnerability). *Reliability* α is the frequency or probability that a system is in a satisfactory state. It is calculated as the period of time when the recharge or groundwater discharge is above the threshold level divided by the total duration. *Resilience* γ describes how quickly a system is likely to recover to a satisfactory state once a failure has occurred. It is calculated as the inverse of the expected drought duration. *Vulnerability* v is the likely magnitude of a failure if one occurs. The magnitude of failure was assumed to be the drought deficit as defined using the threshold level approach. Vulnerability is both calculated as the average deficit per drought and as the maximum deficit. Using outputs (groundwater recharge, gradient and discharge) from the modified Pitman Model together with defined thresholds, Peters et al. (2005) indicators could be applied. An overview of the performance indicators used by Peters et al. (2005) is given below:

- α Reliability: the probability that the system is not in a drought,
- γ Resilience: the inverse of the expected duration of a drought,
- v Vulnerability: expected severity of a drought if one occurs, expressed both by the average drought deficit and the deficit of a drought with a 50-year return period,
- SL Sustainability index (defined by Loucks, 1997), which combines α , γ and v ,
- Pd.yr Overall performance indicator defined in Peters et al. (2005) based on the average drought deficit per year,
- PD.10 Overall performance indicator defined in Peters et al. (2005) based on the drought deficit for droughts with return periods larger than 10 years,
- PD.Cor Overall performance indicator defined in Peters et al. (2005) based on the loss function defined by Correia et al. (1986).

The time series of groundwater recharge and discharge generated by a model needs to be sufficiently long to estimate return periods of at least 50 years for estimating the performance indicators. The model will be run with climate data from WR2012 (Bailey and Pitman, 2017) (in addition to forecasted climate data) and therefore a record length of at least 90 years will ensure climatic extremes are included in the simulation. A monthly time step is an appropriate time step for the analysis of groundwater droughts as this time step avoids large numbers of minor droughts and decreases dependence among droughts but still shows sufficient detail.

6.2 GRACE Mission Data

Remote sensing has been established as a powerful tool to observe water storage dynamics at large scales (Thomas et al. 2017a) with the launch of the GRACE satellites. Observations from GRACE gravity anomalies may be converted into changes of water equivalent height thus tracking changes in total water storage around the world. Should other components of the water balance be available (e.g. surface water and soil moisture information), GRACE has

been able to isolate groundwater storage changes (Rodell et al., 2009; Scanlon et al., 2015) at a regional scale. Many studies have evaluated GRACE-derived groundwater storage changes as a response to drought (Famiglietti et al., 2011; Scanlon et al., 2012), while Thomas et al. (2017b) evaluated a GDI based on GRACE observations in an effort to understand and identify groundwater drought.

GRACE-based storage changes are in good agreement with those obtained from land surface model simulations (Syed et al., 2008) and in situ observations (Rodell et al., 2007), and the utility of GRACE for characterising extreme drought has been demonstrated in a number of recent studies (Houberg et al., 2012; Schumacher et al., 2018). Thus, the potential to use GRACE observations to fill the current need for subsurface water information in the drought mapping and prediction process is evident. While GRACE has supported many advances in water cycle science, the monthly production frequency and coarse spatial resolution (about 150 000 km²) limit the utility of GRACE observations for a majority of applications that require near real time input of much finer resolution earth observation data. In order to realise the full potential of GRACE for hydrological applications at the basin scale, column integrated Total Water Storage (TWS) anomalies from GRACE must be effectively downscaled in space, vertically stratified into moisture component anomalies (e.g. soil moisture, groundwater and surface water), and extrapolated to the present, thereby meeting the specificity, timeliness, and high spatial resolution requirements of most applications.

Examples of data assimilation, which has synthesized the advantages of observations and numerical land surface models have been used to disaggregate GRACE observations temporally, horizontally and vertically. Zaitchik et al. (2008) assimilated GRACE TWS anomalies into a catchment Land Surface Model using a novel implementation with an Ensemble Kalman Smoother. This GRACE Data Assimilation Scheme (DAS) was shown to improve model skill in the simulation of hydrological states and fluxes in sub-GRACE resolution in the Mississippi basin. Data assimilation essentially reduced uncertainties in the Land Surface Model simulation resulting from the input data used to force the model, simplifications in model parameterisation and limitations in the described physical realism of the model, by using observation data sets for constraining the model simulations of terrestrial hydrology. Houberg et al. (2012) furthered this work and extended GRACE DAS (Zaitchik et al., 2008) to the North American domain as part of a larger project aimed toward integrating enhanced (i.e. via data assimilation) GRACE TWS data into the U.S. and North American Drought Monitors. Beside the wider range of hydroclimatic conditions, the study went beyond Zaitchik et al. (2008) by assessing the potential of GRACE DAS for drought monitoring and by evaluating GRACE DAS simulations using soil moisture measurements, and groundwater observations beyond the Mississippi Basin. They determined that drought conditions could be identified more comprehensively and objectively by integrating GRACE based drought indicators into the short and long term objective blends (a mix of precipitation data, various standardized indices such as the PDSI and simple water budget estimates of soil moisture) that constitute the U.S. and North American Drought Monitor baselines. These Drought Monitor baselines currently lack valuable information on deep (root zone and below) soil moisture and groundwater storage changes (Figure 6.1).

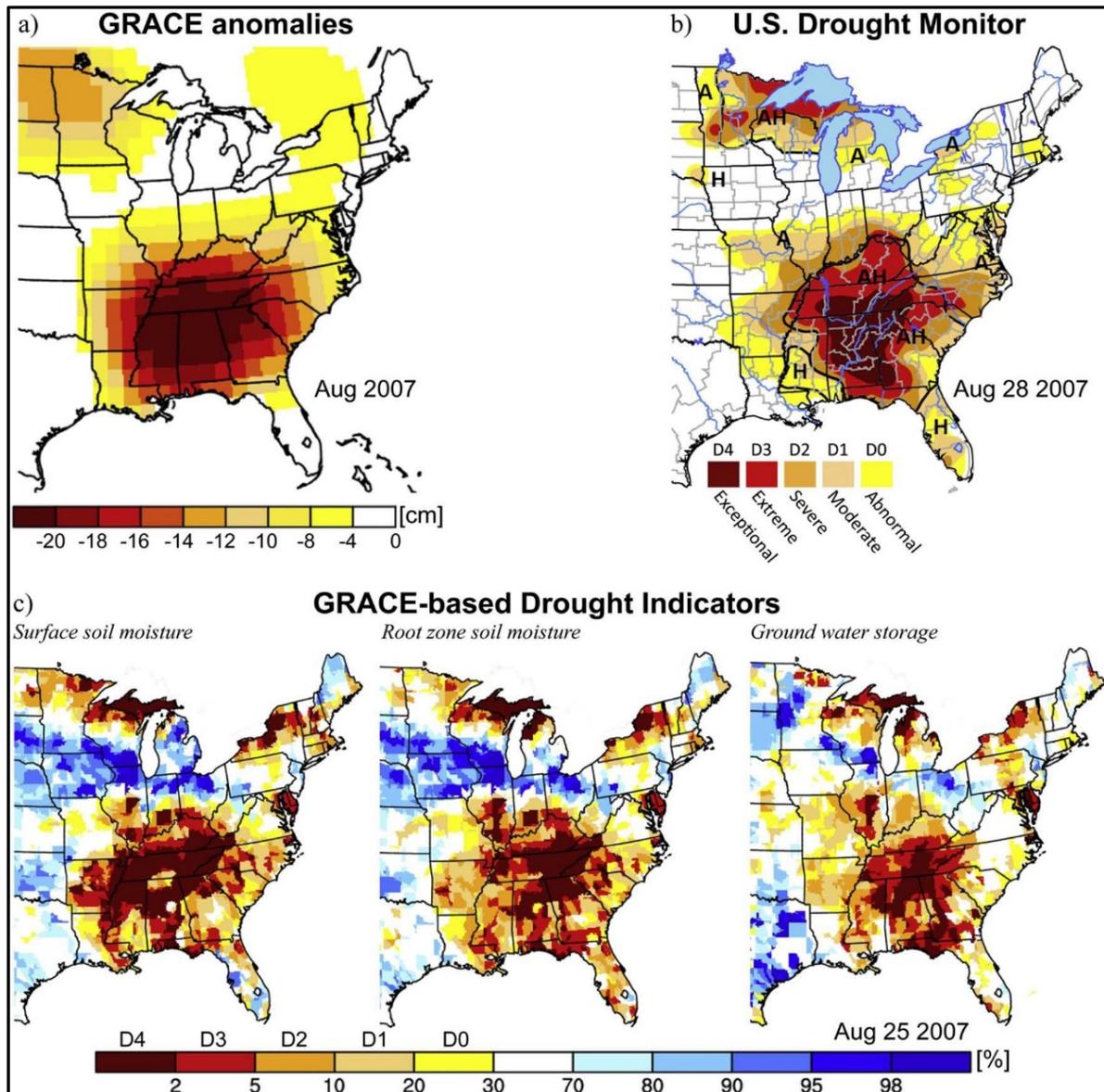


Figure 6.1: Grace model results from Houberg et al. (2012)

Figure 6.1 shows the correspondence between (a) the GRACE monthly water storage anomaly fields, (b) the U.S. Drought Monitor product, and (c) drought indicators based on model-assimilated GRACE TWS observations during the drought in the south eastern United States in August 2007

Schumacher et al. (2018) integrated GRACE mission into WaterGAP Global Hydrology Model which is a global water resources and use model for the semi-arid region of the Murray Darling Basin, Australia. In particular, they tested the ability of a parameter calibration and data assimilation approach to introduce long term trends into the Model, which were poorly represented due to errors in forcing, model structure and calibration. They also investigated the influence of selecting a specific GRACE data product and filtering method on the final parameter calibration and data assimilation results. They found that integrating the GRACE data into the Model did not only improve simulation of seasonality and trend of terrestrial water storage changes, but it also improved the simulation of individual water storage components.

They also importantly found that using solutions from different GRACE data providers produced slightly different outcomes and recommended that a rigorous evaluation of GRACE errors is required to realistically account for the spread of the differences in the results.

Clearly, there are numerous examples of authors applying GRACE mission data products in interesting and creative ways and with many different types of models. Many of these examples (Schumacher et al., 2018; Houberg et al., 2012; Li et al., 2012; Zaitchik et al., 2008; Thomas et al., 2017) are applied specifically for drought prediction and evaluation purposes. It is clear however, that a rigorous evaluation of GRACE errors is required, and that the assimilation of GRACE data into models is thoughtfully and carefully carried out.

6.3 Groundwater modelling

While the GRACE method could be used to give a broad catchment-based perspective of the state of groundwater, the data is not available at the resolution required to report at quaternary catchment scale. An alternative is to model the change in groundwater. There are two approaches readily available to do this, namely the Sami Groundwater Model, as described in the WRSM2000 Technical Guide (Bailey et al, 2007), and the Hughes Groundwater Model (Hughes et al, 2007). These two models are similar and the general principles are described below with the aid of Figure 6.2.

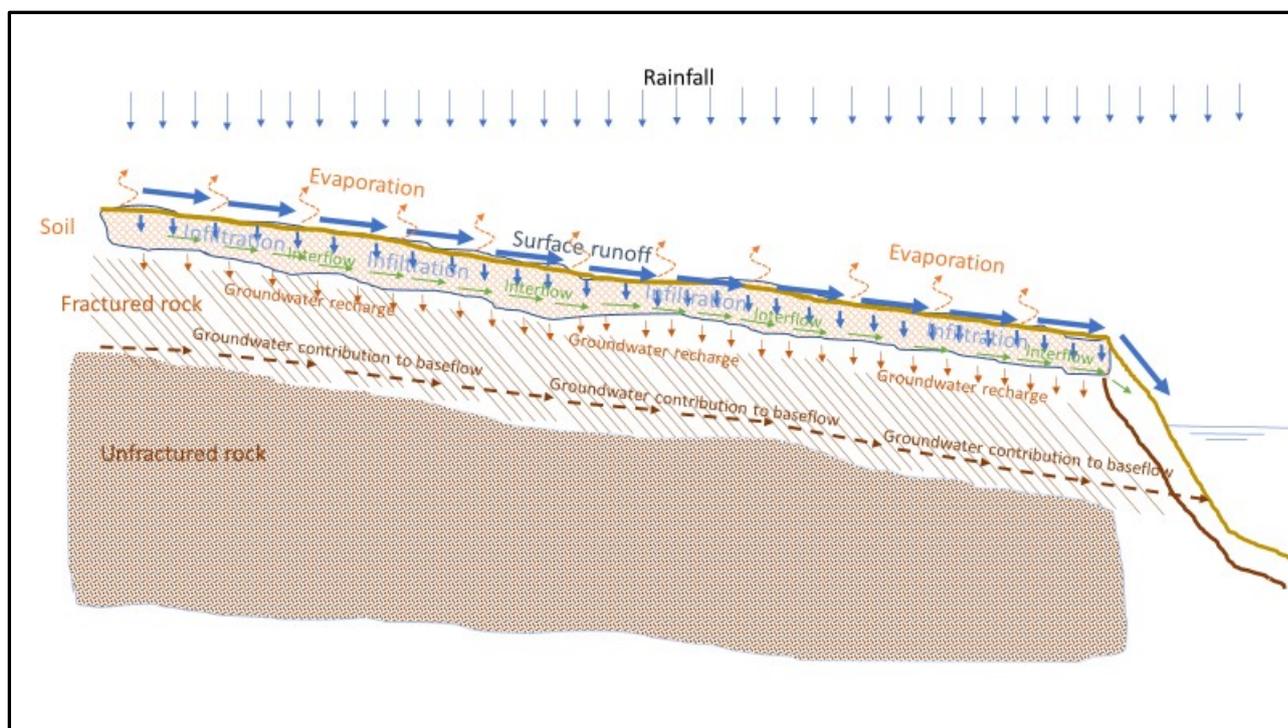


Figure 6.2: Principles of groundwater flow

Surface runoff, soil moisture and interflow are calculated by the Pitman model. The groundwater components which need to be estimated are:

- Recharge
- Flow to downstream catchments, and
- Groundwater discharge into the river
- Evaporation from the riparian zone

The above groundwater components can all be estimated from the soil moisture component 'S' which is an output of the Pitman model.

6.3.1 Groundwater recharge

Recharge is given by the following equation (Hughes 2007):

$$RE = GW \left(\frac{S - SL}{ST - SL} \right)^{GPOW}$$

Where

RE	=	variable of potential recharge (mm)
GW	=	parameter of maximum recharge in mm at maximum soil moisture (ST)
S	=	Input data of soil moisture in mm
SL	=	Parameter of soil moisture threshold below which there is no recharge
GPOW	=	Parameter of the storage-recharge relationship

Equation 6.1

6.3.2 Flow to downstream catchments

While most groundwater discharges into the streams within the catchment from which the recharge occurs, some of the groundwater flows sub-surface to the downstream catchment, also referred to as underflow.

This outflow can be estimated as the product of Transmissivity (T) and Hydraulic Gradient (HG) as given by Equation 6.2.

$$\text{Underflow} = \text{HG} \times \text{T}$$

Where:
T = Transmissivity
HG = Hydraulic gradient

Equation 6.2

The Hydraulic Gradient is not a constant but varies with storage as given by Equation 6.3 from the WRSM2000 Technical Guide (Bailey et al, 2007).

$$\mathbf{HG = HGRAD \times (STORE - SWL)/(CAP - SWL)}$$

Where:

HG = Hydraulic gradient at any point in time

HGRAD = maximum hydraulic gradient

STORE = Groundwater storage

SWL = Static water level

Equation 6.3

6.3.3 Evaporation from the riparian zone

While the Pitman model already caters for evaporation from soil moisture, there is an additional component to evaporation which relates to groundwater, and this is evapotranspiration from the riparian zone. Riparian vegetation transpires groundwater which enters the soil zone as it approaches the riparian zone. This is calculated simply as the area of the riparian zone multiplied by the net evaporation rate. However, the wetted area within the riparian zone is not constant and varies as a function of groundwater storage. The full equation from the WRSM2000 Technical Guide (Bailey et al, 2007) is as follows:

$$\mathbf{GWEVAP = (EVAPOTRANSPIRATION - RAINFALL) \times AREA \times (STORE - SWL)/(CAP - SWL)}$$

Where:

GWEVAP = Evaporation from groundwater recharge as it flows through the riparian zone

AREA = The estimated area of the riparian zone from which groundwater evaporation is occurring

Equation 6.4

Note that there are numerous methods for estimating evapotranspiration. These are not discussed in this report.

6.3.4 Discharge into streams (baseflow)

The flow out of groundwater into the stream or river, also referred to as baseflow, is estimated with equation 6.5.

$$\mathbf{GWBaseflow = (1 - e^{(HEAD \times BPOW)}) \times BFMAX}$$

Where:

GWBaseflow = Contribution of groundwater to surface water (stream, river or lake)

HEAD= Hydraulic head

BPOW = Power function which defines the relationship between hydraulic head and baseflow

Equation 6.5

Hydraulic head is not a constant but varies over time, as given by Equation 6.6.

$$\text{HEAD} = \text{STORE} - \text{SWL} - (\text{RUNOFF}/\text{CATCHMENT AREA})$$

Where:

HEAD = Hydraulic head at any point in time

RUNOFF = Streamflow at any point in time

CATCHMETN AREA = Catchment are from which the streamflow is derived

Equation 6.6:

Using equations 6.1 to 6.6 with a soil moisture time series (from the Pitman model) as input, the groundwater storage can be simulated. See Figures 6.3 and 6.4 which show simulated soil moisture and groundwater storage in the X21C catchment.

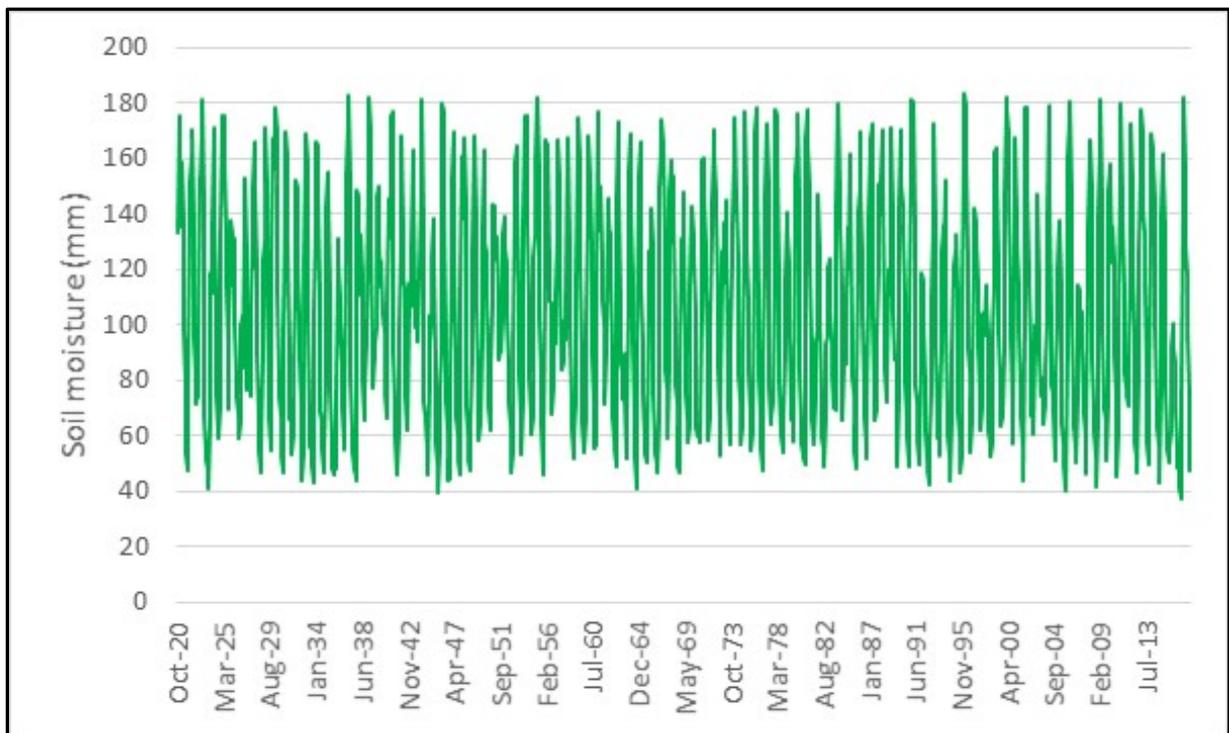


Figure 6.3: Simulated soil moisture derived from the Pitman model

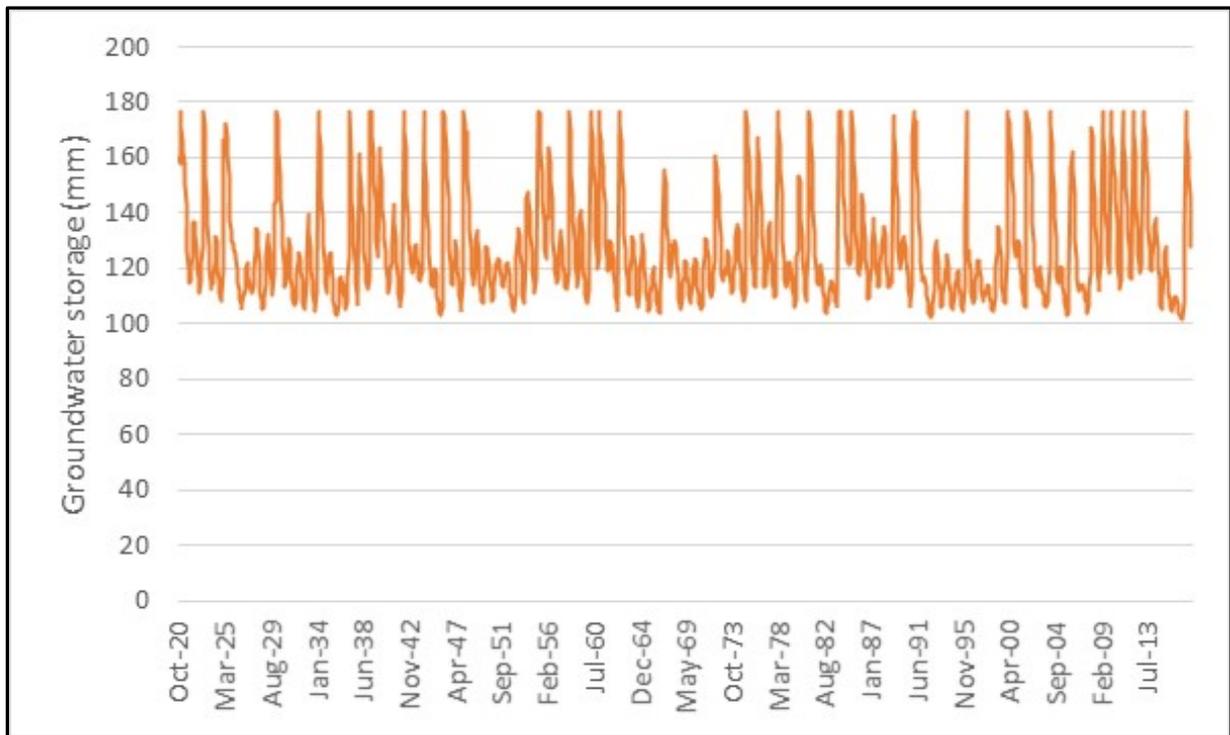


Figure 6.4: Simulated groundwater storage

6.4 Conclusions and recommendations

Monitoring groundwater storage is challenging, and while remote sensing techniques are constantly improving, these techniques do not provide observations at a fine enough resolution to be useful for managing catchments at quaternary scale. As an interim measure, it is recommended that modelled soil moisture and groundwater storage be used as drought indicators for these parameters.

7 INTEGRATION OF PREDICTION METHOLOGIES TOWARDS AN INTEGRATED DROUGHT MONITORING AND PREDICTION SYSTEM

7.1 General approach

Based on the analyses carried out in this project and described in section 4, the integrated drought warning system has been conceptualised as a combination of a number of sources of information:

- Near real time climate (rainfall and air temperature) data
- Indices of modes of climate variability (e.g. ENSO)
- Multi-model ensemble seasonal forecast data
- Near real time soil moisture monitoring data

In the above, the current climate variables obtained from the monitoring system are used within the drought warning system in three contexts:

- To evaluate the current status of drought (through drought indices)
- To derive implications of current rainfall anomaly to the end of season rainfall total using statistical model,
- To provide the input (initial condition) to hydrological modelling aimed at generating hydrological seasonal forecast.

Indices of modes of climate variability (and their forecast) provide information on drivers of climate variability that affect climate at seasonal time scale and underlie the predictability of climate at this time scale. Within the drought warning system, this information is used to construct messages about the future drought outlook through incorporating it as an explicit input to a statistical model.

The multi-model ensemble seasonal forecast data comprise the source of information about climate conditions at the seasonal time scale (3 months) that is, theoretically at least, the most conceptually defensible. This is because climate models allow for a physically consistent representation of the climate system, and thus for the most comprehensive linking of the current and future conditions, i.e. linking that makes fewer assumptions than any other type of statistical forecast. That forecast is, however, not devoid of uncertainties.

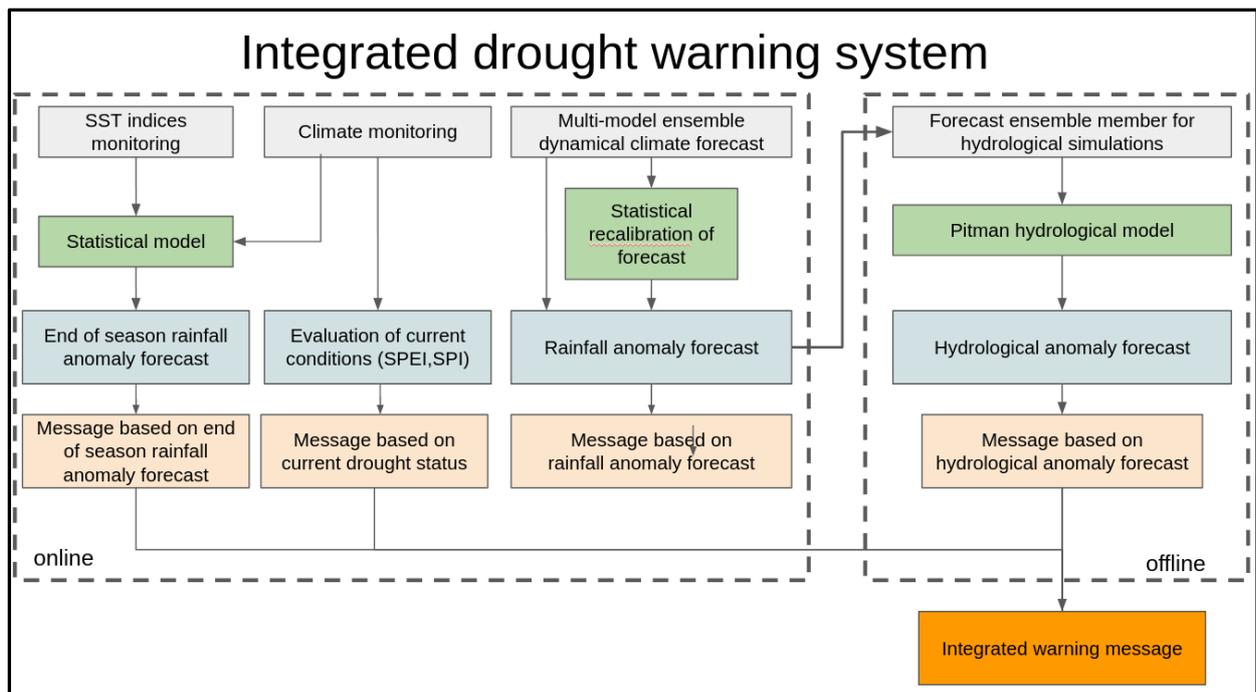


Figure 7.1: Schematic of the integrated drought warning system

The three sources of information (current climate, current state of modes of climate variability, results of dynamical seasonal climate forecast) allow for (direct or indirect) capturing three principal factors that affect the outlook of the drought:

- Current state of drought (or simply season’s anomaly), that is one of the main factors affecting future drought,
- Main drivers of predictability of climate – i.e. processes that affect climate and vary slowly enough to exert their influence at the seasonal time scale
- Combination of the knowledge about physical processes in the climate system and knowledge about the initial conditions of the climate system that is implicitly formalized in the form of initialized global forecast models.

In addition, through the incorporation of hydrological modelling that allows for translation of the seasonal climate forecast into seasonal hydrological forecast, the warning system allows for incorporation of information on the state of hydrological system, and capturing the influence of slow-varying processes within the earth surface (and sub-surface) on future status of various parts of water resource and hydrological systems – runoff, soil moisture, reservoir storage and groundwater storage.

The information sources listed above contribute to four “pathways” of generating messages relevant from the perspective of drought warning that encapsulate the different sources of drought forecast information:

- Information about the current drought,
- Information from statistical models based on assumption that the current drought affects future drought status,

- Information on the future climate anomalies based on climate models that comprehensively describe climate processes, and
- Information about the future state of elements of the hydrological/water resources system that merge information about future climate from climate models with that describing their current state.

The combination of these pathways amounts to a multi-method, multi-data sources approach that allows for derivation of robust messages in a situation of uncertainty of information sources and (cascading and accumulating) uncertainties and errors associated with each data processing path.

The system has two components – online (a website) and offline (software installed and operated on a PC). The online component focuses on analyses of climate data and derivation of information in the pathways a, b and c. The offline component focuses on derivation of hydrological information in the pathway (d).

While this way of approaching the functionality of the integrated system is recognised as not the most convenient, and makes the system available to only highly skilled users, there are three main motivations behind the adopted approach:

- The process of implementation of Pitman hydrological model in an online environment by far exceeded the scope of this project.
- The system is focused on water resources, and thus its primary audience is water resource managers who, with the access to Pitman model are able to interrogate its outputs in a more specific rather than generic way, the latter of which would obviously be implemented if Pitman model was to be run online with a generic configuration.
- The online component allows for interrogation of a very comprehensive range of seasonal climate forecasts, and selection of the forecast ensemble that is best suited to answer drought warning questions in a particular location. Only that forecast can then be used to generate hydrological forecast, again, providing targeted rather than generic information, with benefits to the final outcomes.

It is envisaged that the warning system is operated monthly and provides information about the period of three months ahead (including the current given month). Drought warning can be issued approximately in the middle of the month.

The details of the implementation of the four information “pathways” within the envisaged drought warning system are presented in the sections below.

7.2 Implementation and functionality of the four information pathways

7.2.1 Current drought status

Access to raw, original climate data sources is limited. Station data originating from SAWS or ARC are available for a fee, and thus cannot be utilized at this stage within this project. The current drought status is thus assessed based on readily available surrogate data sources –

gridded station data (GPCC) and satellite-based rainfall data (CHIRPS). As far as can be ascertained, there are no gridded air temperature data of sufficient resolution available in near real time, the analyses of rainfall has been limited to indices only. Two drought indices were used; a simple accumulated seasonal rainfall anomaly, and an SPI at three time scales – 12 months, 24 months and 36 months. These time scales are used because they are the most relevant from the perspective of water resources.

Information is presented at the level of quaternary catchments – with maps providing a regional overview, and time series available for each of the quaternary catchments, when selected.

Apart from the data-based product, links to drought information products of other institutions – notably the ARC Umlindi newsletter are provided (e.g. <https://www.arc.agric.za/ARC%20Newsletters/UMLINDI,%20Issue%202020-11,%20November%202020.pdf>), SAWS Drought Monitoring Desk (<https://www.weathersa.co.za/Documents/Climate/CLS-CI-Drought%20Monitoring.pdf>)

7.2.2 Statistical forecast of the end of the season rainfall anomaly

This pathway is based on the same datasets as the current drought status pathway and presents the implementation of the statistical model described in detail in section 4.4. Information is presented at the level of quaternary catchments – with maps providing a regional overview, and time series available for each of the quaternary catchments, when selected.

7.2.3 . Numerical rainfall forecast

This “pathway” relies on a rainfall forecast from a multi-model ensemble of global climate models.

A combination of two publicly available multi-model ensembles were used, namely, NMME and Copernicus@ECMWF. Those are described in detail in section 4.3.2.

This pathway is geared towards deriving forecast of three relevant indices: deterministic (ensemble median) rainfall anomaly, and probabilities of 1 in 3, 1 in 5 and 1 in 10 years rainfall anomaly in terms of 3-months rainfall mean, for each of the quaternary catchments (which are the basic units of water management in South Africa).

Three ways of interpreting the multi-model ensemble towards the future drought messages from the ensemble data are envisaged:

- Analysing all the individual models together as a “grand ensemble”. In that, the models forecast data are combined together into a single ensemble of 120-150 members, and forecast indices are calculated.
- Analysing an individual model that has the best skill for an area and period of concern.

- Analysing the level of agreement of models in the multi-model ensemble in terms of increased probability of drought as manifested by the drought indices.

In the operational setting, every month, once the forecast data are available: the following happens automatically in the back-end of the system, by execution of a number of scripts:

- Recent GPCC Monitoring and First Guess data are downloaded.
- GPCC 2018 data are merged with GPCC Monitoring data and GPCC First Guess data to generate time series of historical “observed” rainfall extending until the one month prior to the month of the forecast.
- Gridded GPCC rainfall data are converted to rainfall over the WR2012 catchments and rainfall regions.
- Forecast data for all models is downloaded for the forecast issued on the given month.
- Forecast data are converted to rainfall over the WR2012 catchments and rainfall regions
- The regions’ rainfall figures are then bias-corrected.
- Rainfall indices (ensemble median, probabilities of below normal, 1 in 5 and 1 in 10 years) are calculated on the forecast data.
- Continuous time series are generated that combine the historical monitoring data (merged GPCC time series) and the forecast data. These are later used to force the hydrological model.

Information is presented at the level of quaternary catchments – with maps providing a regional overview, and details of the forecast available for each of the quaternary catchments, when selected.

Maps present drought-oriented deterministic and probabilistic forecast indices (ensemble median rainfall amount, probability of below normal, probability of 1 in 5 year and probability of 1 in 10 year drought, as well as a set of skill measures for each index.

For each quaternary, upon selecting, detailed information on skill, including tabular and graphical information such as ROC curve, as well as time series allowing evaluation of the performance of given forecast historically are presented.

7.2.4 Hydrological model implementation and structure

- Rainfall prediction (3 months into the future) obtained from members of the “grand” multi-model ensemble of numerical seasonal forecasts and converted into % rainfall format consistent with the Pitman Model format. Other output includes:
 - Rainfall prediction map for the whole country at quaternary scale.
 - Precipitation Indexes.
- Pitman model upgraded on the WReMP platform to accept multiple rainfall scenarios. The current configuration reads 10 ensembles from SANS5. The Pitman Model produces:
 - Natural runoff ensemble consisting of 10 time series.
 - Time series of soil moisture.

- Time series of Groundwater storage.
- The runoff time series are then run through a WReMP (or WRYM) to produce:
 - Ensemble of streamflow.
 - Storage Ensemble in all dams.
 - Soil moisture prediction.
 - Groundwater storage prediction.

7.2.4.1 . Integration of hydrological model with numerical rainfall forecast

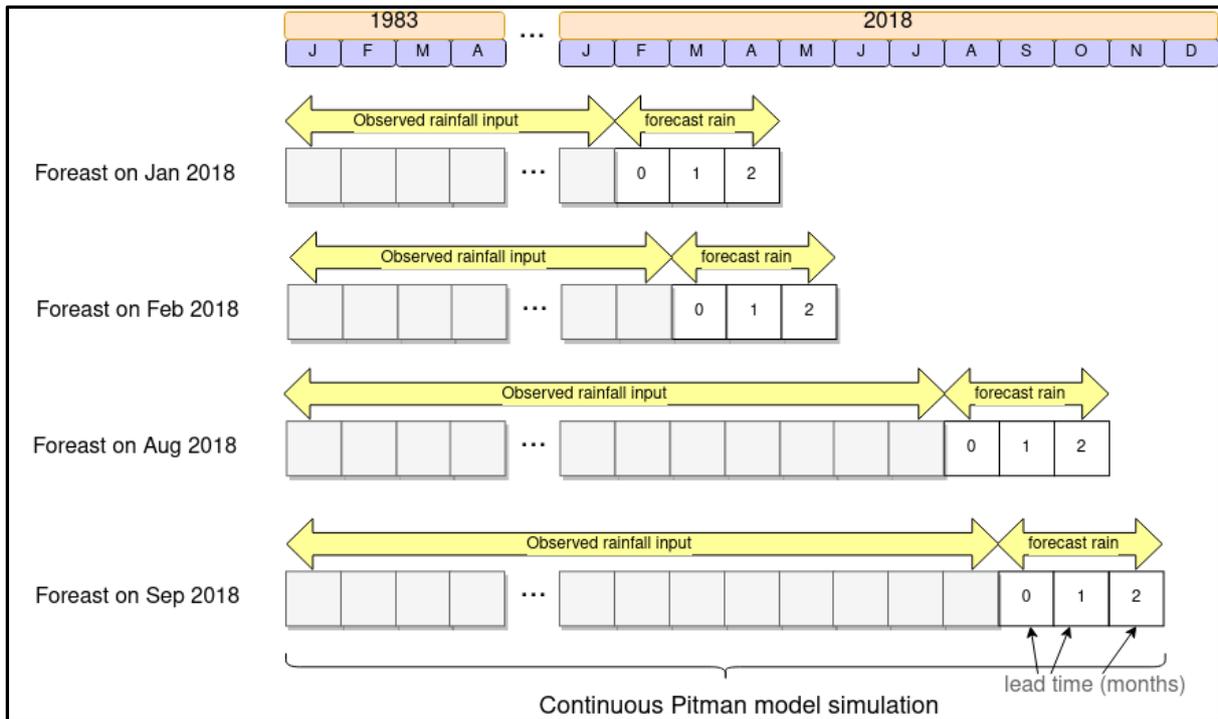


Figure 7.2: Schematic of operational implementation of a hydrological forecasting system

Figure 7.2 is based on continuous simulations with observed data and appended rainfall forecast data.

A Windows based application was developed to achieve the above drought monitoring and forecasting integration. See Figure 7.3.

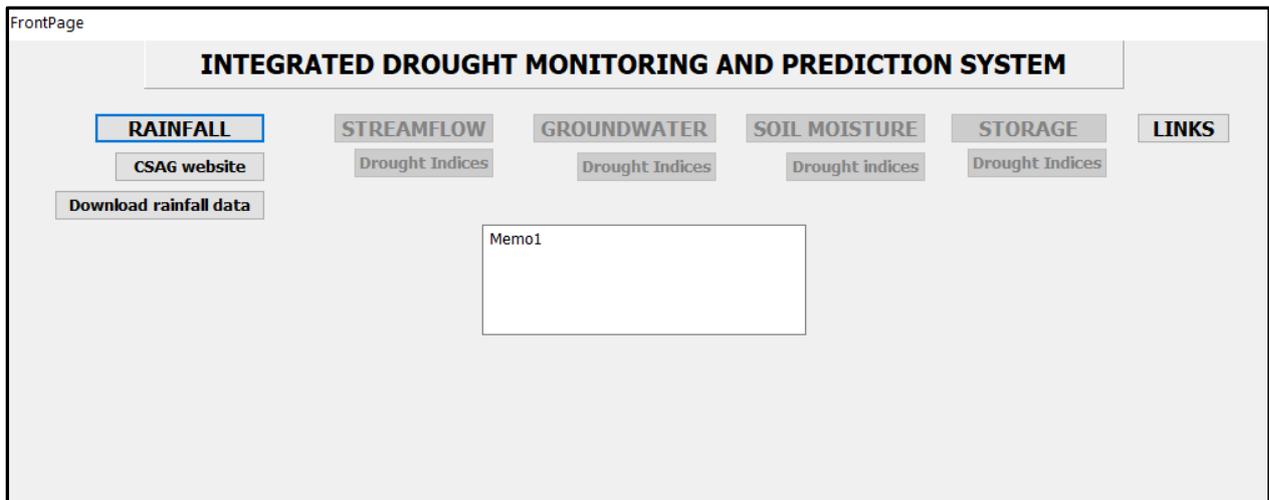


Figure 7.3: Integrated Drought Monitoring and Prediction System (IDMAPS)

The IDMAPS application links to the CSAG rainfall website (https://web.csag.uct.ac.za/wrc_earlywarning/) to rainfall forecasts. Streamflow, Groundwater and Soil moisture components are modelled using the recoded Pitman Model while storage is modelled using the Water Resources Modelling Platform. For each forecasting component, a monthly time step time series is produced from which drought indices are calculated.

The application also provided links to other useful forecasting websites.

8 PILOT STUDIES

8.1 Introduction

In order to test the integrates system, three catchments were selected to conduct pilot studies. The catchments were selected mainly to provide variation on climatic zones. Hence catchment in the winter and summer rainfall regions were selected. In addition, a catchment on the southern coast which experience both summer and winter rainfall was selected. Another criteria in selecting catchments was simplicity so as not to cloud the projections with unrelated operational issues.

The catchments selected were as follows:

- Theewaterskloof (Western Cape)
- White River system (Mpumalanga)
- Garden Route Dam (Southern Cape)

8.1.1 Theewaterskloof

The upper reaches of the Theewaterskloof, up to and including the Theewaterskloof Dam, is shown in Figure 8.1. The catchment comprises of quaternary catchments H60A, H60B and H60C.

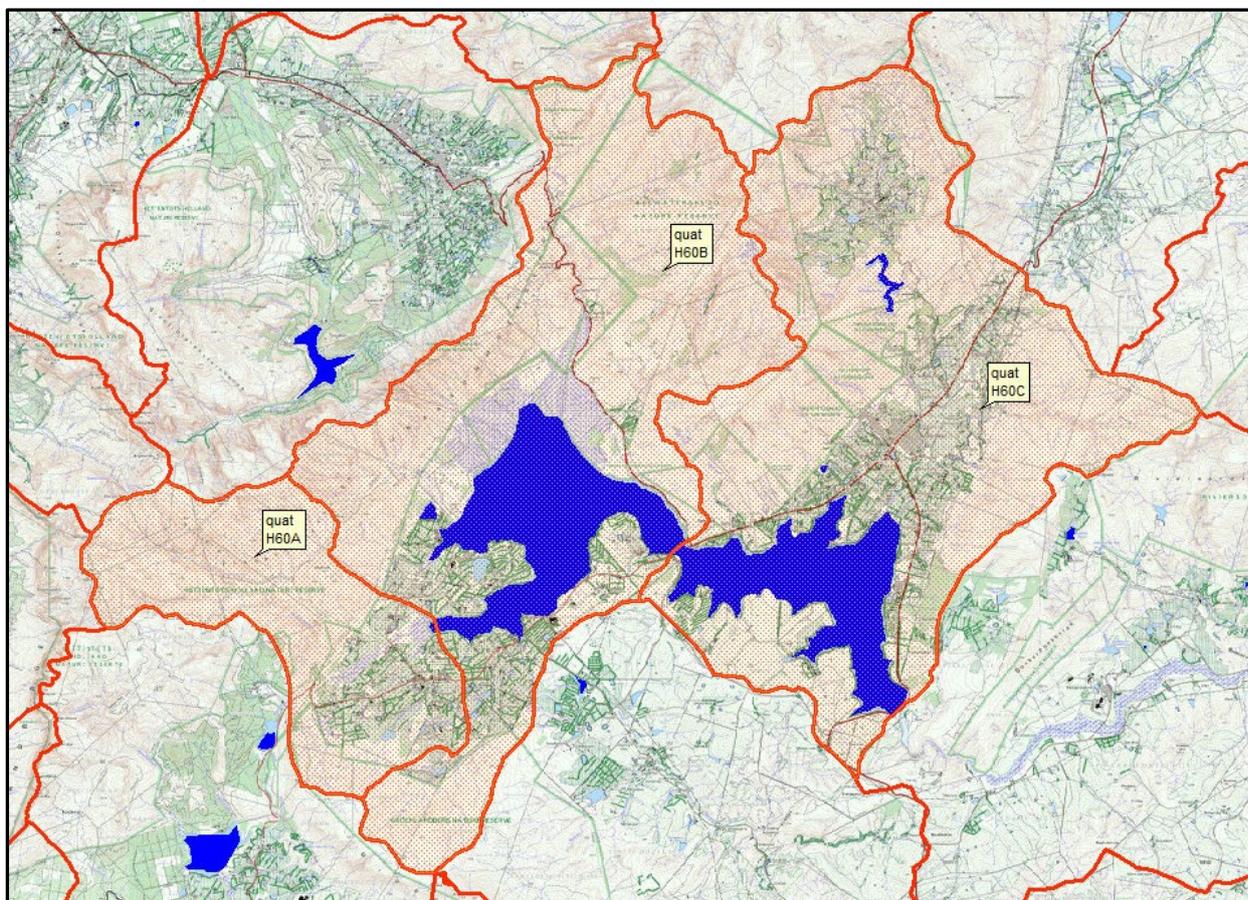


Figure 8.1: Location of the Theewaterskloof system

This system includes the Elandsdam Dam which supplies water to irrigators downstream of the dam and supplements the water supply to the town of Villiersdorp. The Theewaterskloof supplies water to Cape Town and well as irrigators downstream of the dam.

The hydrology of the system is summarised in Table 8.1 while estimated water use is summarized on Table 8.2.

Table 8.1: Summary of climate and hydrology information for the upper Theewaterskloof catchment

Catchment	Area (km ²)	Mean Annual Evaporation (mm)	Mean Annual Precipitation (mm)	Natural Mean Annual Runoff (million m ³ /annum)
H60A	73	1 440	2 141	112.1
H60B	210	1 465	1 241	79.9
H60C	217	1 470	994	62.9
Total	500			254.9

Source: WR2012

A time series of the natural flow into the dam is shown in Figure 8.2 while the monthly distribution of the natural flow is shown in Figure 8.3.

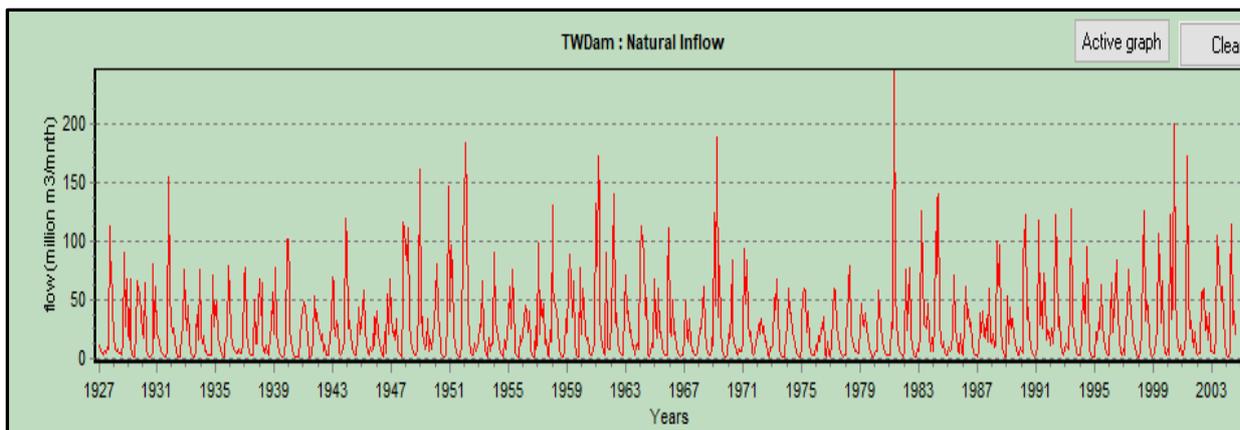


Figure 8.2: Natural flow time series into the Theewaterskloof Dam (based on WR2012 hydrology)

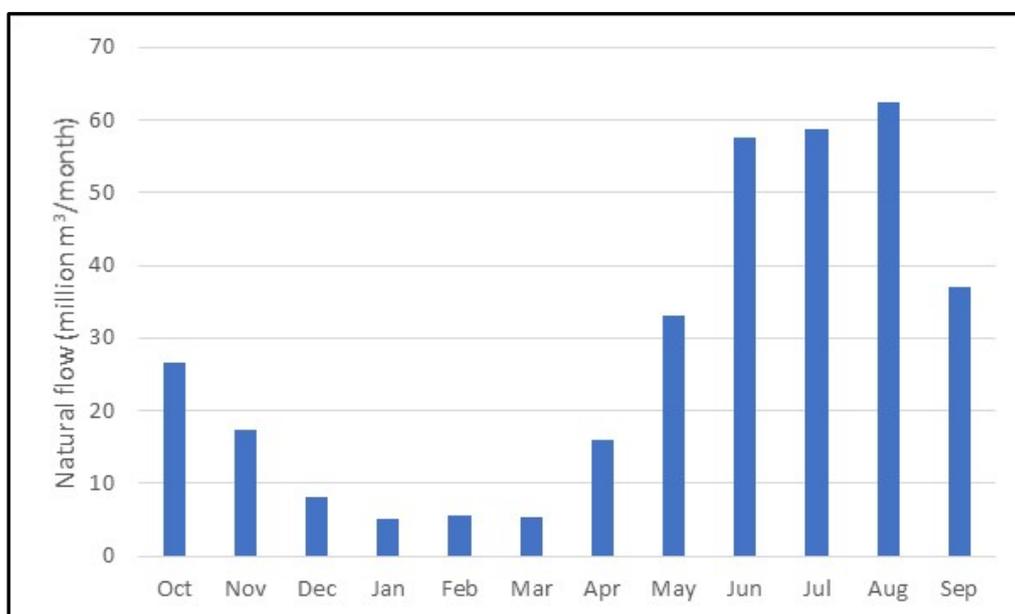


Figure 8.3: Monthly distribution of natural flow

Table 8.2: Water use in the Theewaterskloof catchment

User sector	Estimated water use/streamflow reduction
Irrigation	14.4
Domestic	0.7
Streamflow reduction due to commercial forestry	0.3
Transfers to Cape Town	220*
Release to downstream	37.2

* Varies

In addition to water use within the catchment, water is released from the dam to downstream users. This release, estimated at 37.2 million m³/annum was obtained from the flow record H6H012. Note that while the 1 in 50 yield of the Theewaterskloof Dam is approximately 220 million m /annum, the actual transfers from the dam to the Cape town water supply system varies according to a complex operating rule which takes into account the storage of the other dams in the system and restriction applied to the systems as a whole. For the purpose of this forecasting exercise, the recorded abstractions from the dam were requested from DWS but these were not yet available as 'verified data'. The abstraction was therefore estimated from the record of abstractions (Gauge G1H053). See Figure 8.4.

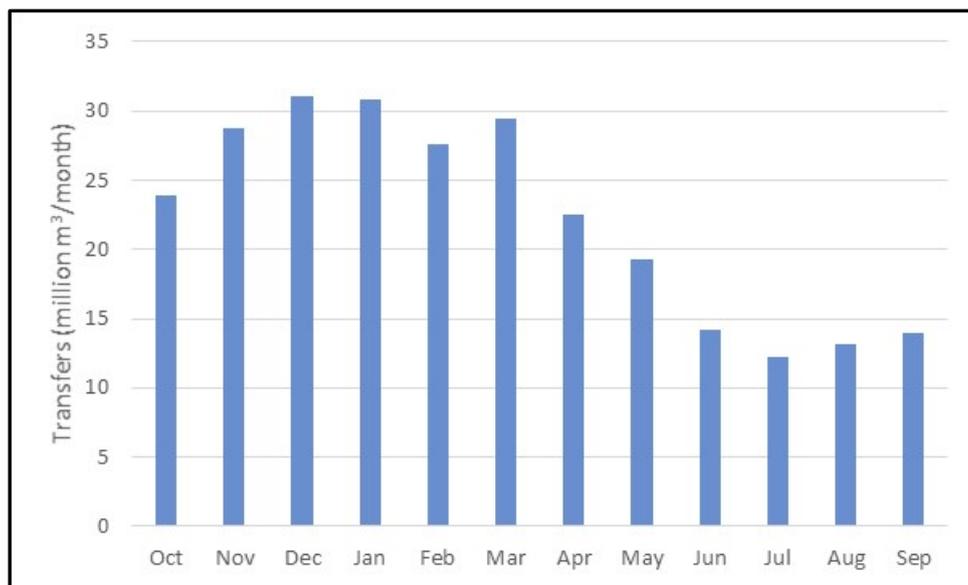


Figure 8.4: Transfers from Theewaterskloof Dam to Cape Town

The projected natural flows, commencing in January 2020, are shown in Figures 8.5, 8.6 and 8.7. These natural flows were derived using the Pitman model and the projected rainfall. Note that the rainfall projections are only for three months hence the limitation in the projected natural flows. The minimum and maximum natural flows in each month, as obtained from WR2012, are also shown on this graph.

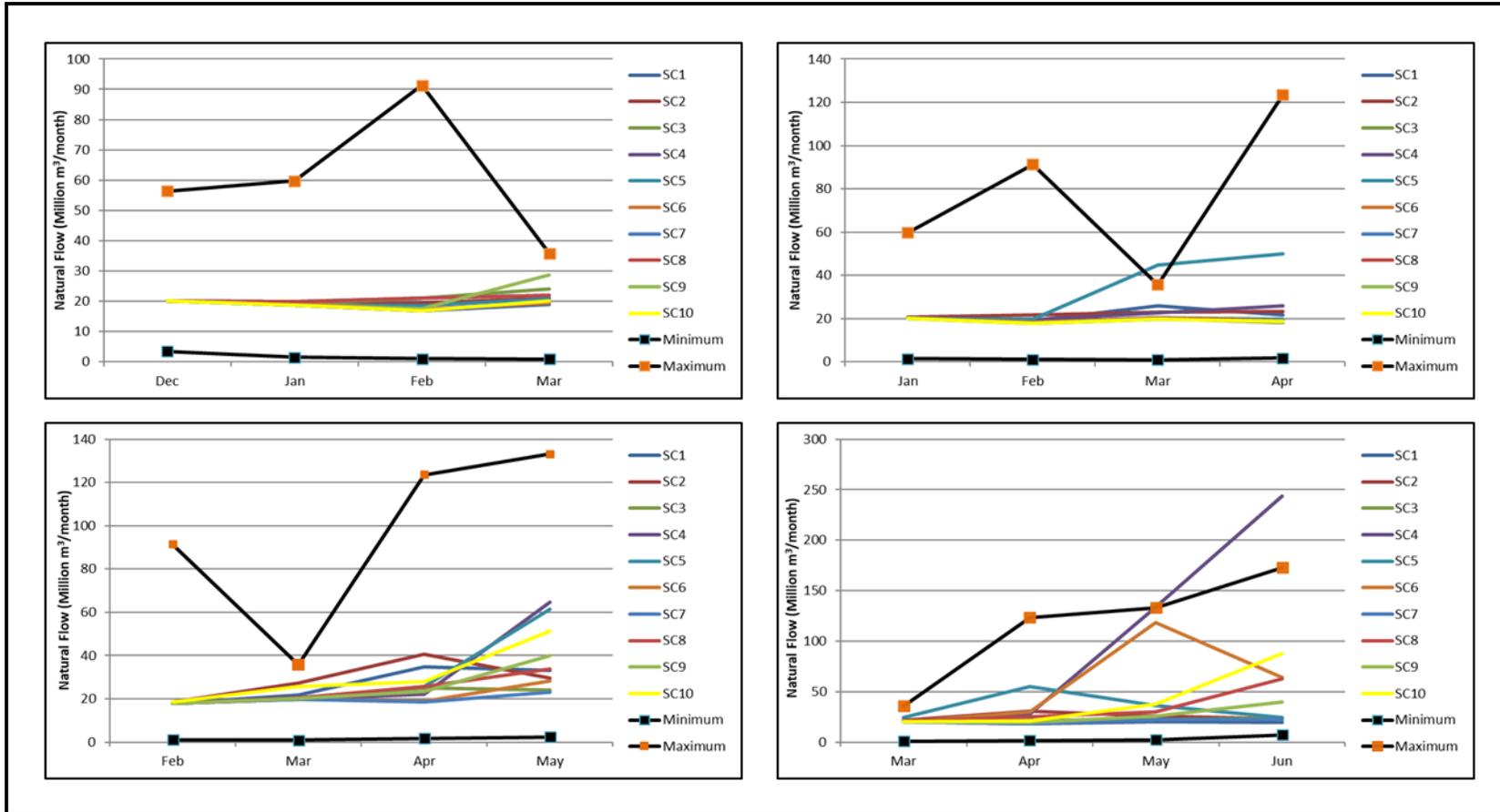


Figure 8.5: Modelled natural flow using projected rainfall (January to April 3 month projections)

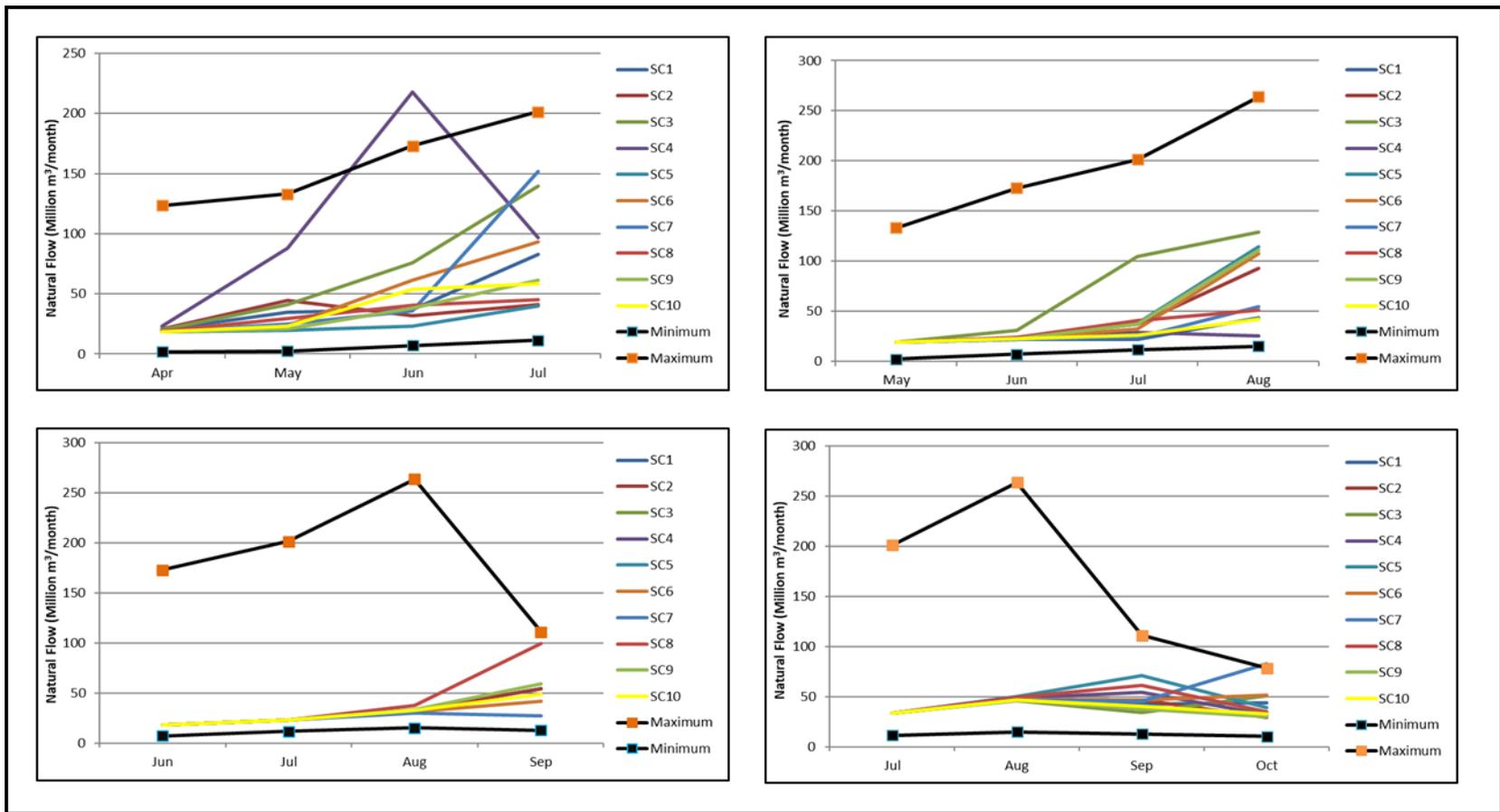


Figure 8.6: Modelled natural flow using projected rainfall (May to August 3 month projections)

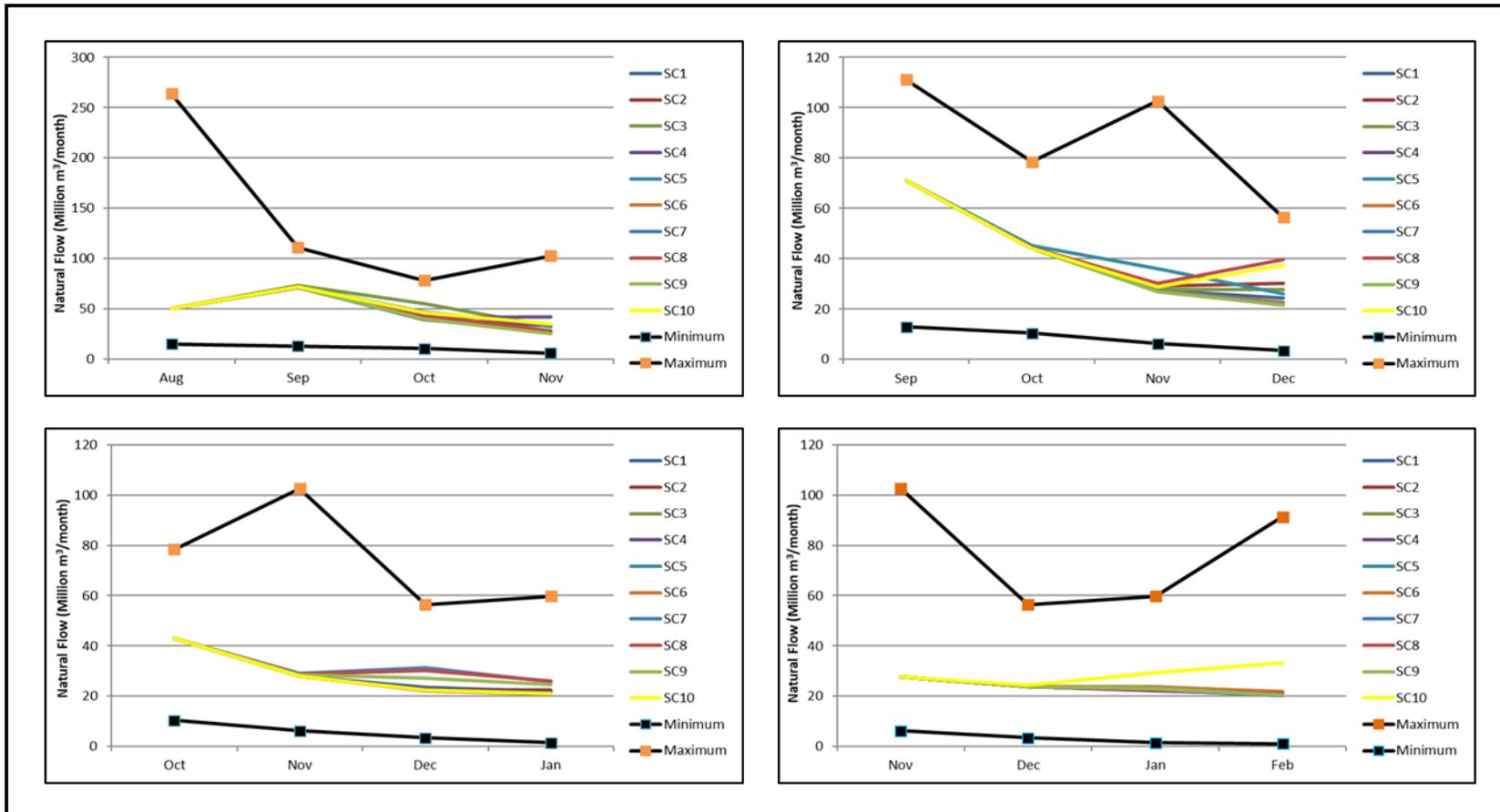


Figure 8.7: Modelled natural flow using projected rainfall (September to December 3 month projections)

The projected storage of the Theewaterskloof Dam is shown in Figures 8.8, 8.9 and 8.10. Note that the starting storage of the dam is set to the observed storage with each simulation.

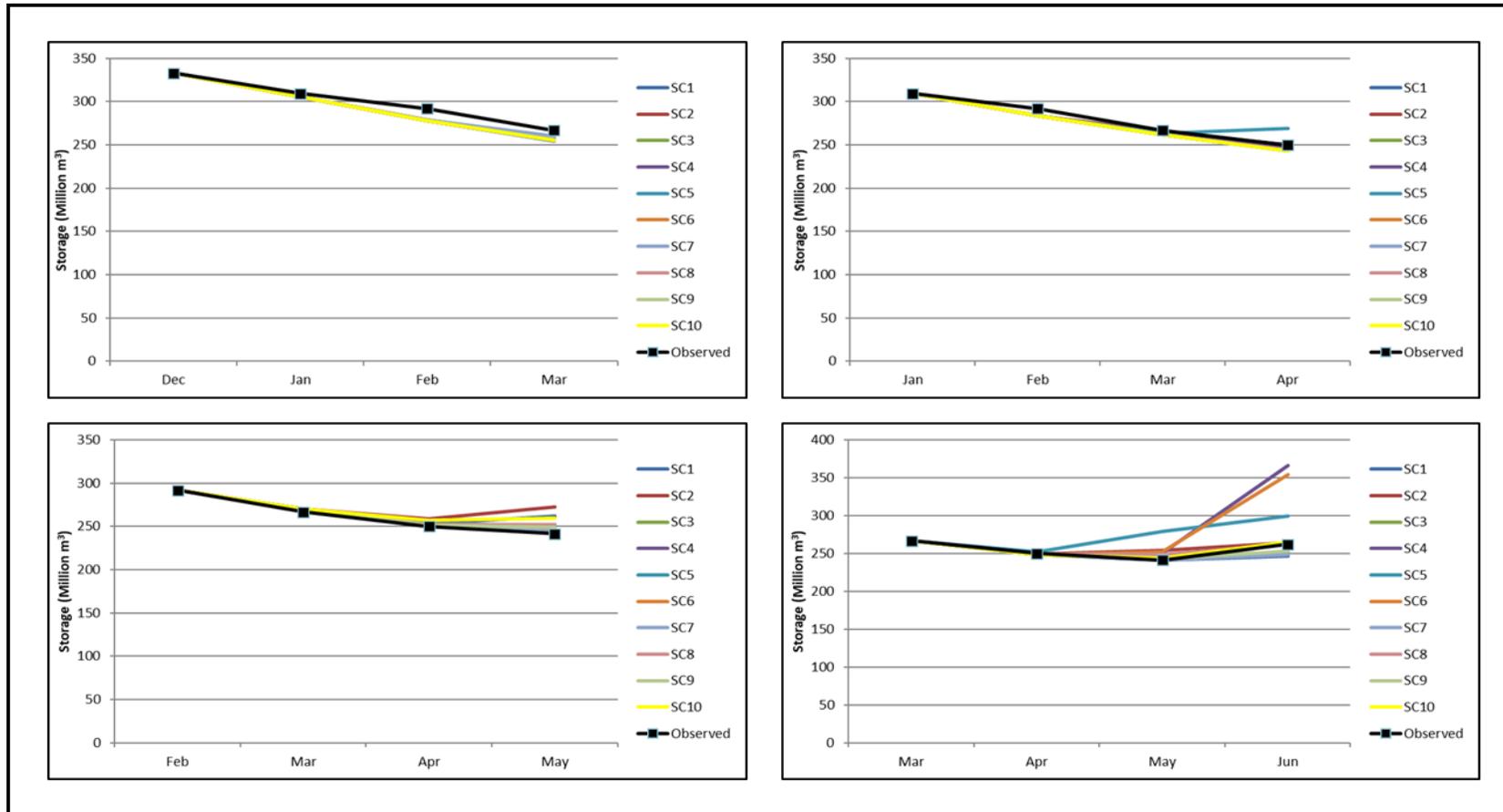


Figure 8.8: Modelled storage using projected natural flow (January to April 3 month projections)

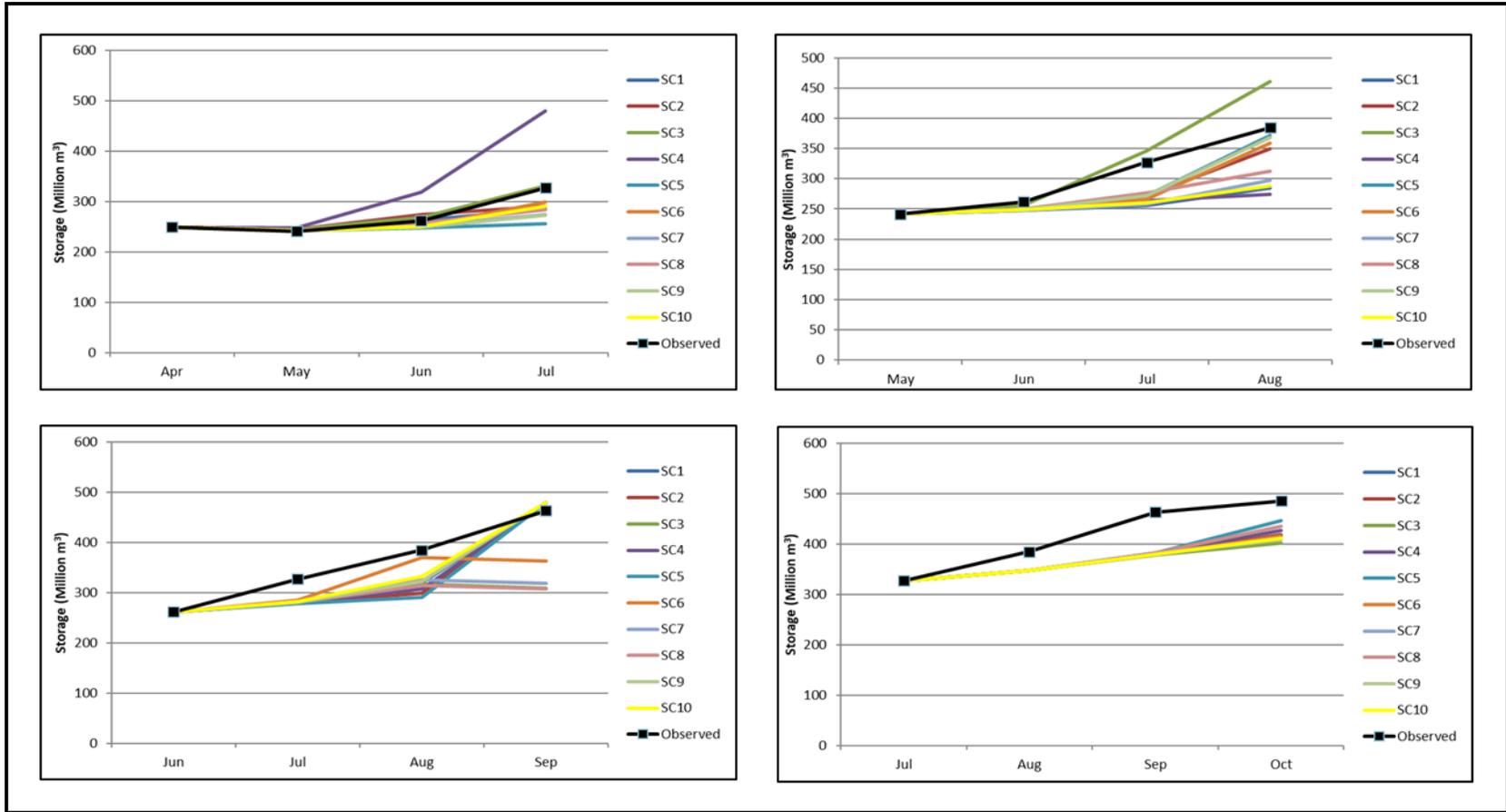


Figure 8.9: Modelled storage using projected natural flow (May to August 3 month projections)

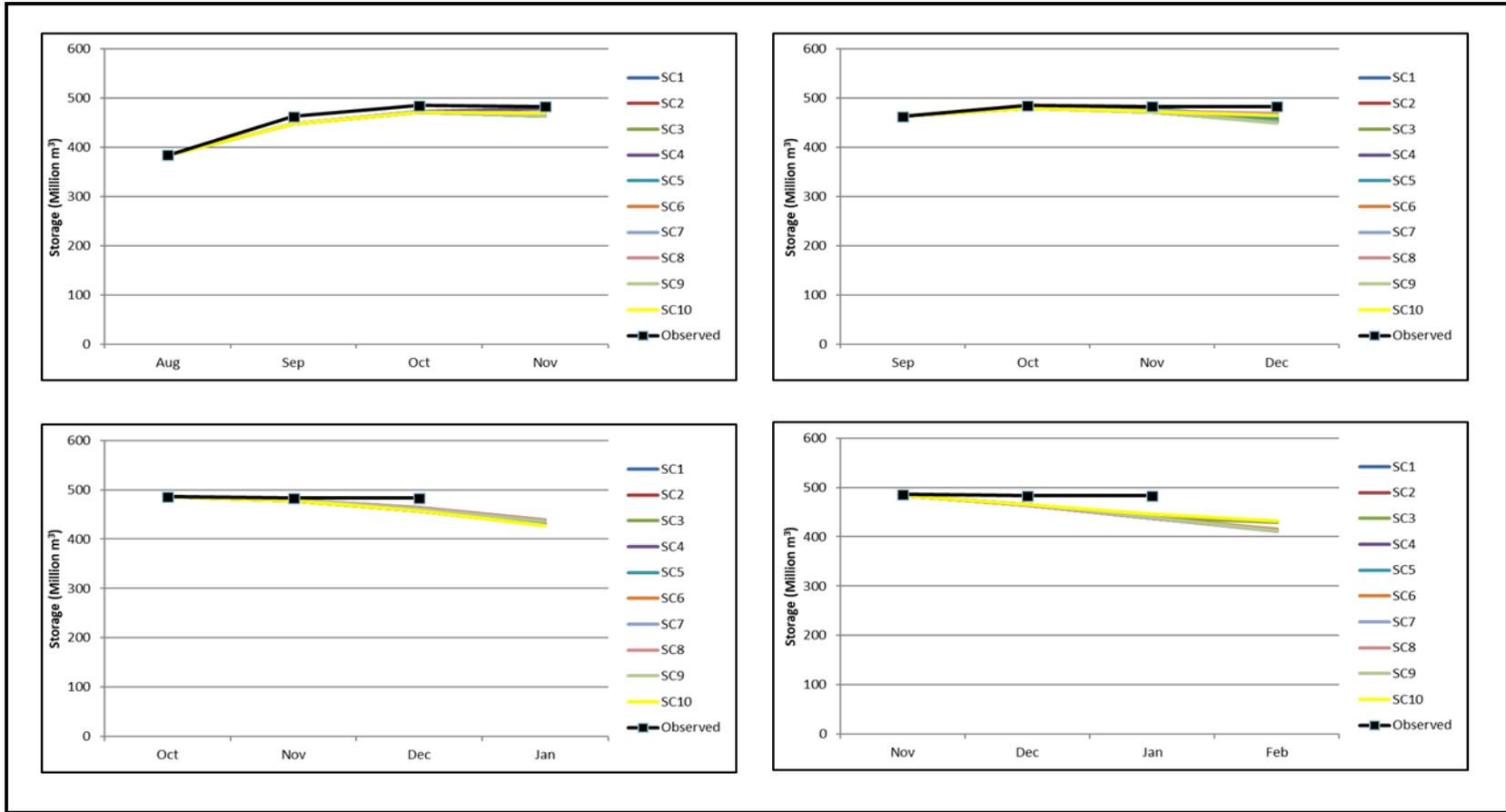


Figure 8.10: Modelled storage using projected natural flow (September to December 3 month projections)

8.1.2 White River System

The White River system is an important sub-catchment of the Crocodile catchment. It is a complex catchment containing four significant dams and transfers between two catchments, namely, the White River catchment and the sand River catchment. See Figure 8.11. The White River system, as modelled in this pilot study, includes quaternary catchments X22E, X22G and X22H.

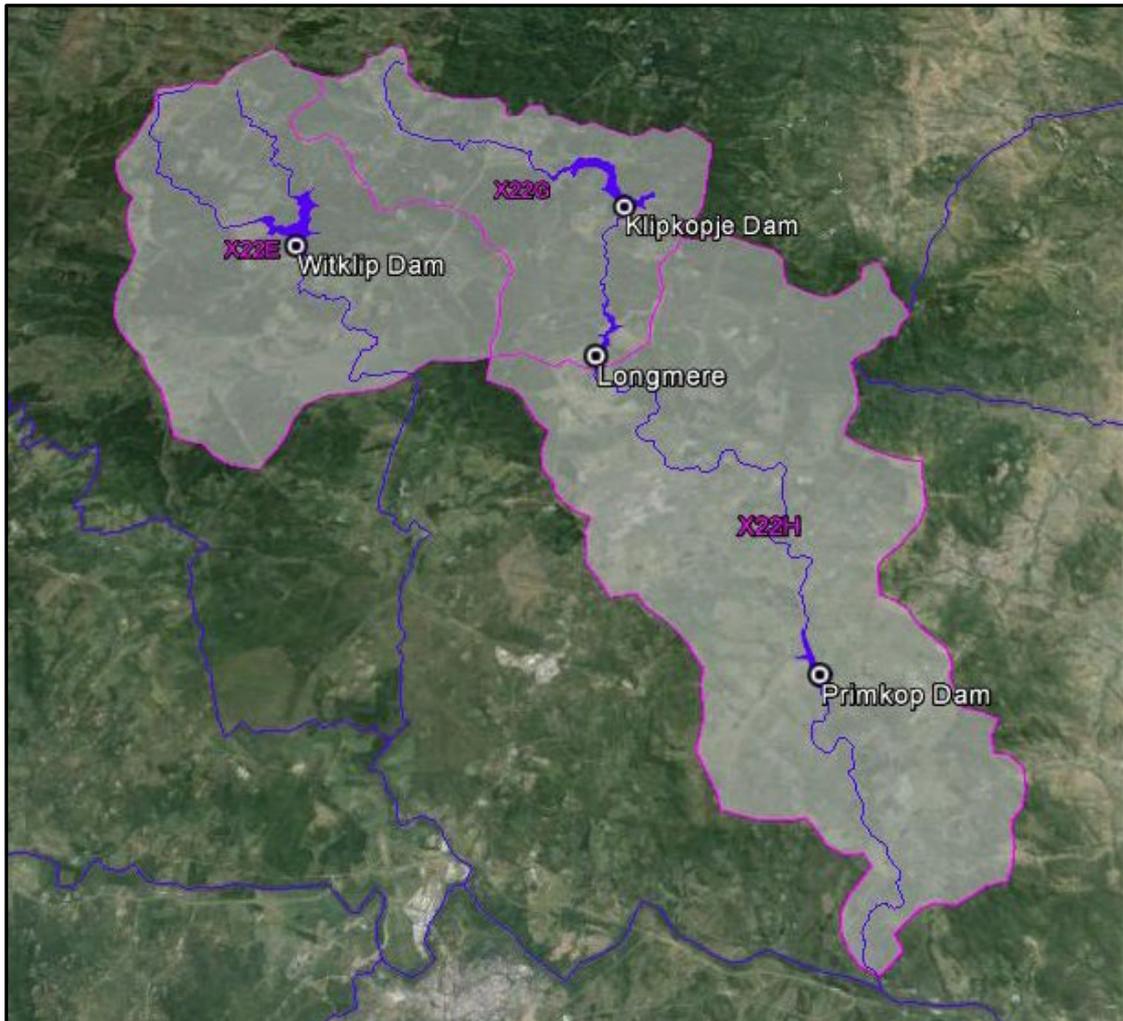


Figure 8.11: Location of the White River system

Details of the dams within this system are given in Table 8.3.

Table 8.3: Dams in the White River System

Dam	Full supply surface Area (ha)	Full supply capacity (million m³)	Owner	Operator
Klipkopje	235	12.30	WRVCB	WRVCB
Longmere	90	4.24	WRVCB	WRVCB
Primkop	41	1.93	WRVCB	WRVCB
Witklip	190	1.90	DWS	Sand River Irrigation Board

Table 8.4: Summary of climate and hydrology information for the upper Theewaterskloof catchment

Catchment	Area (km²)	Mean Annual Evaporation (mm)	Mean Annual Precipitation (mm)	Natural Mean Annual Runoff (million m³/annum)
X22E1	16.0	1 447	1 313	6.9
X22E2	48.3	1 453	1 192	16.5
X22E3	88.6	1 464	1 065	19.7
X22G1	77.0	1 458	1 116	13.6
X22G2	30.5	1 456	1 053	6.1
X22H1	66.2	1 453	1 006	11.8
X22H2	90.2	1 450	910	13.8
X22H3	43.8	1 472	806	5.1
Total				

Source: IUCMA, 2019

A time series of the natural flow into the Longmere Dam is shown in Figure 8.12 while the monthly distribution of the natural flow is shown in Figure 8.13.

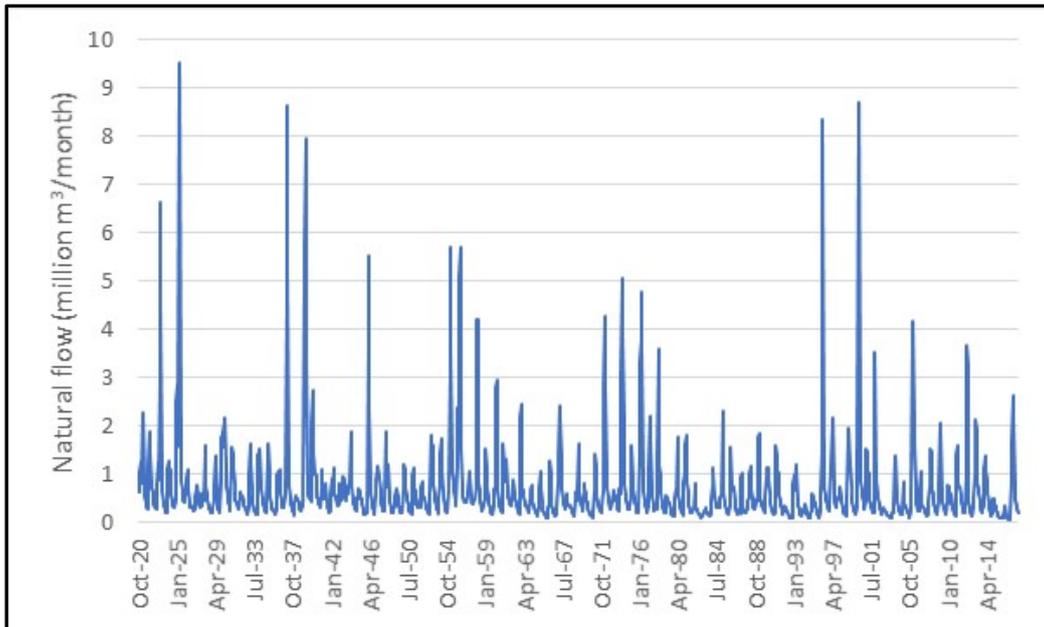


Figure 8.12: Natural flow time series into the Longmere Dam (based on IUCMA 2019)

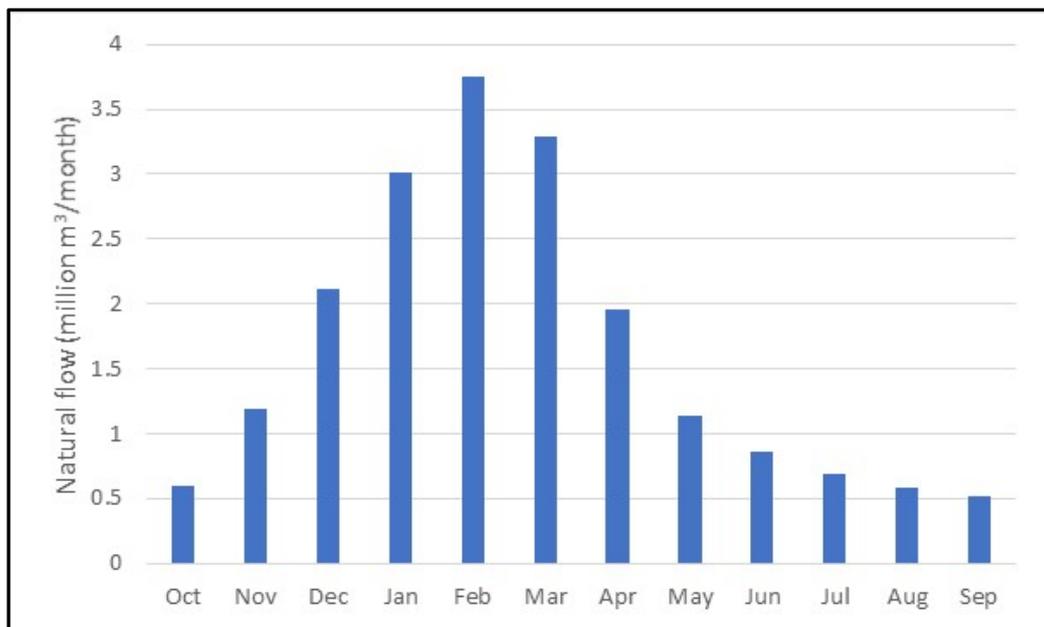


Figure 8.13: Monthly distribution of natural flow

The water use in the catchment is summarized in Table 8.5.

Table 8.5: Water use within the White River system

Water use sector	Water requirement (million m³/annum)
Municipal	5.0
Industrial	0.12
Irrigation	26.6
Streamflow reduction	15.1

The projected natural flows, commencing in January 2020, are shown in Figures 8.14, 8.15 and 8.16. These natural flows were derived using the Pitman model and the projected rainfall. Note that the rainfall projections are only for three months hence the limitation in the projected natural flows. The minimum and maximum natural flows in each month, as obtained from WR2012, are also shown on this graph.

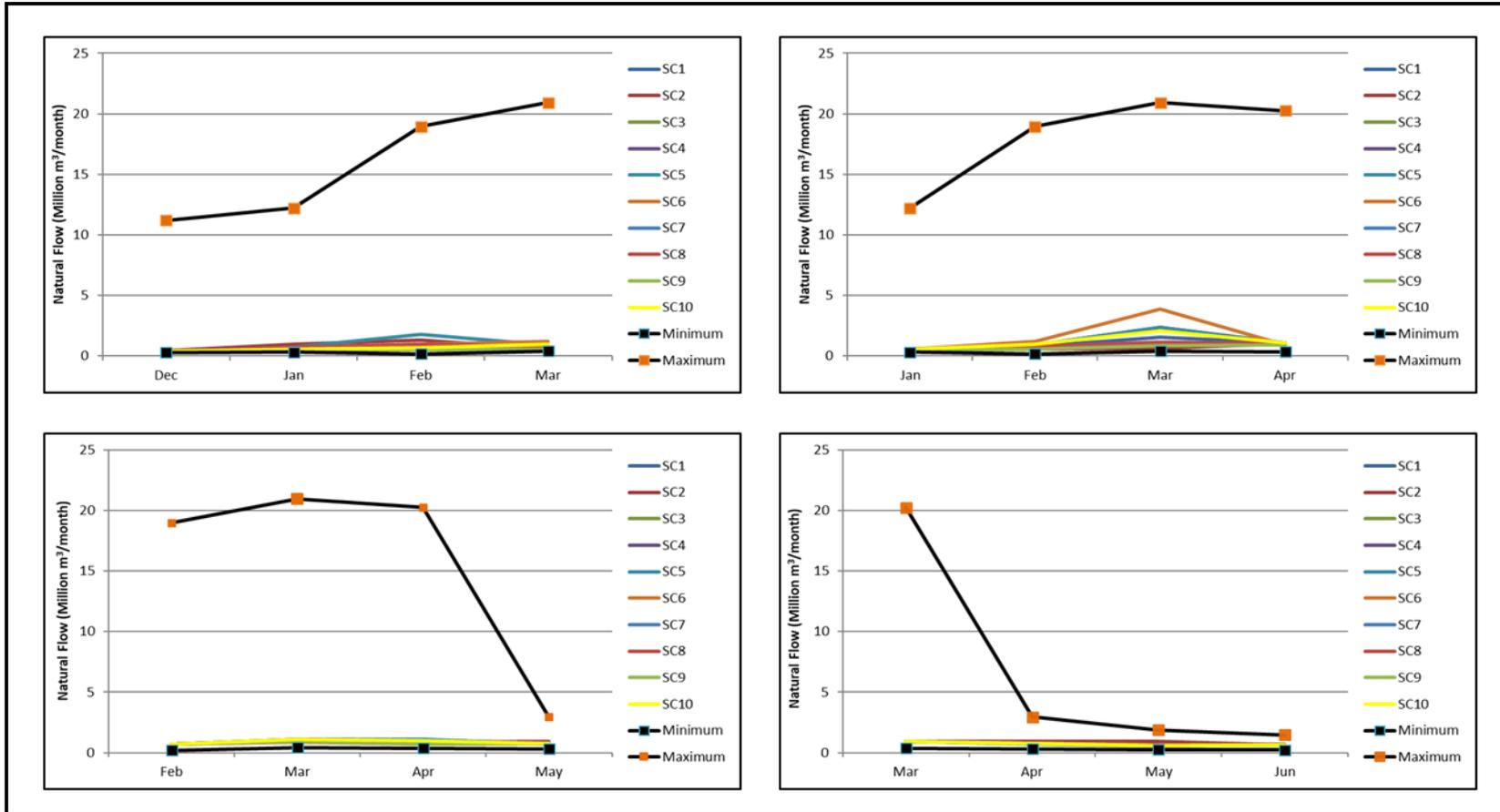


Figure 8.14: Modelled natural flow using projected rainfall (January to April 3 month projections)

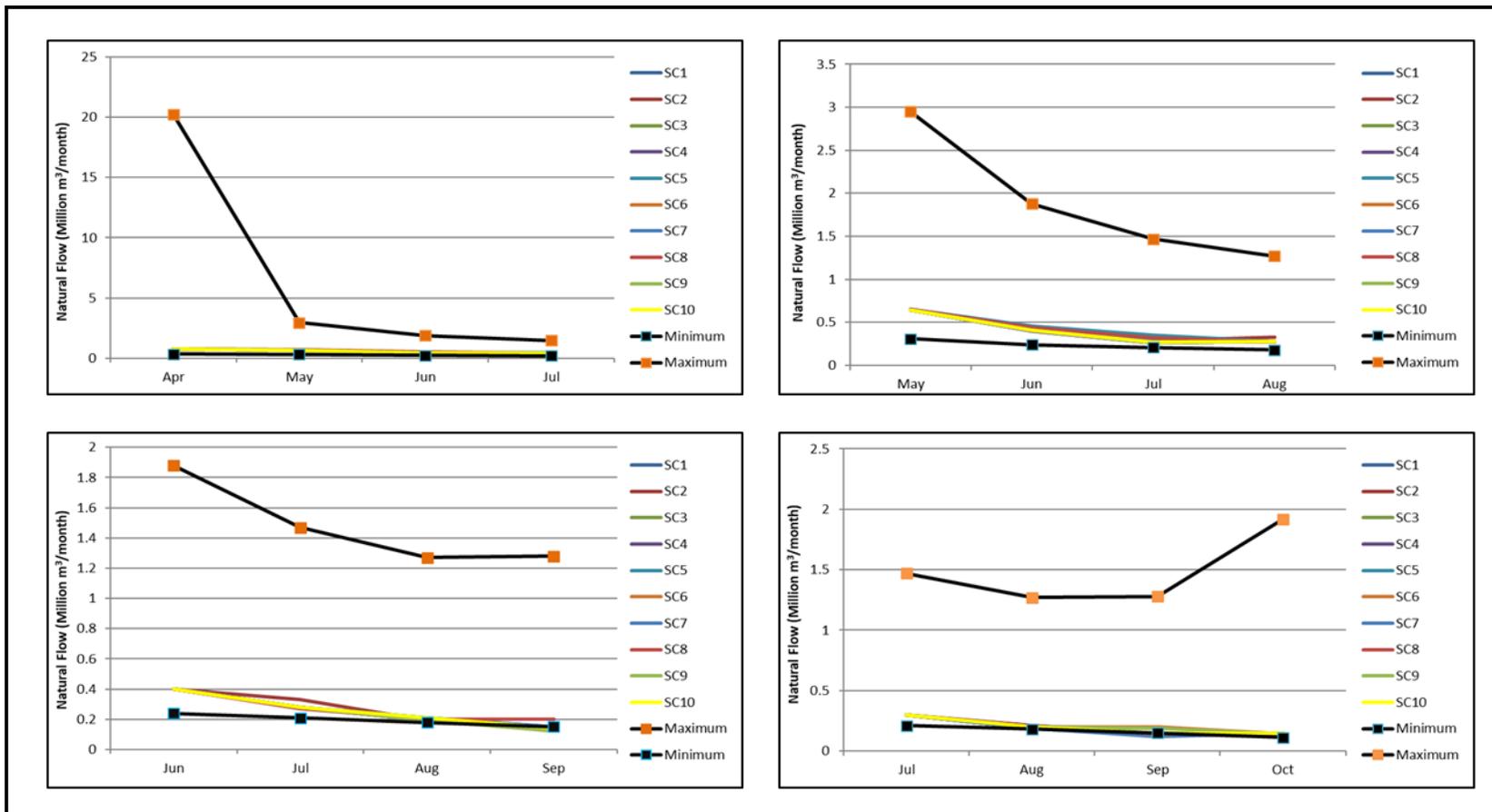


Figure 8.15: Modelled natural flow using projected rainfall (May to August 3 month projections)

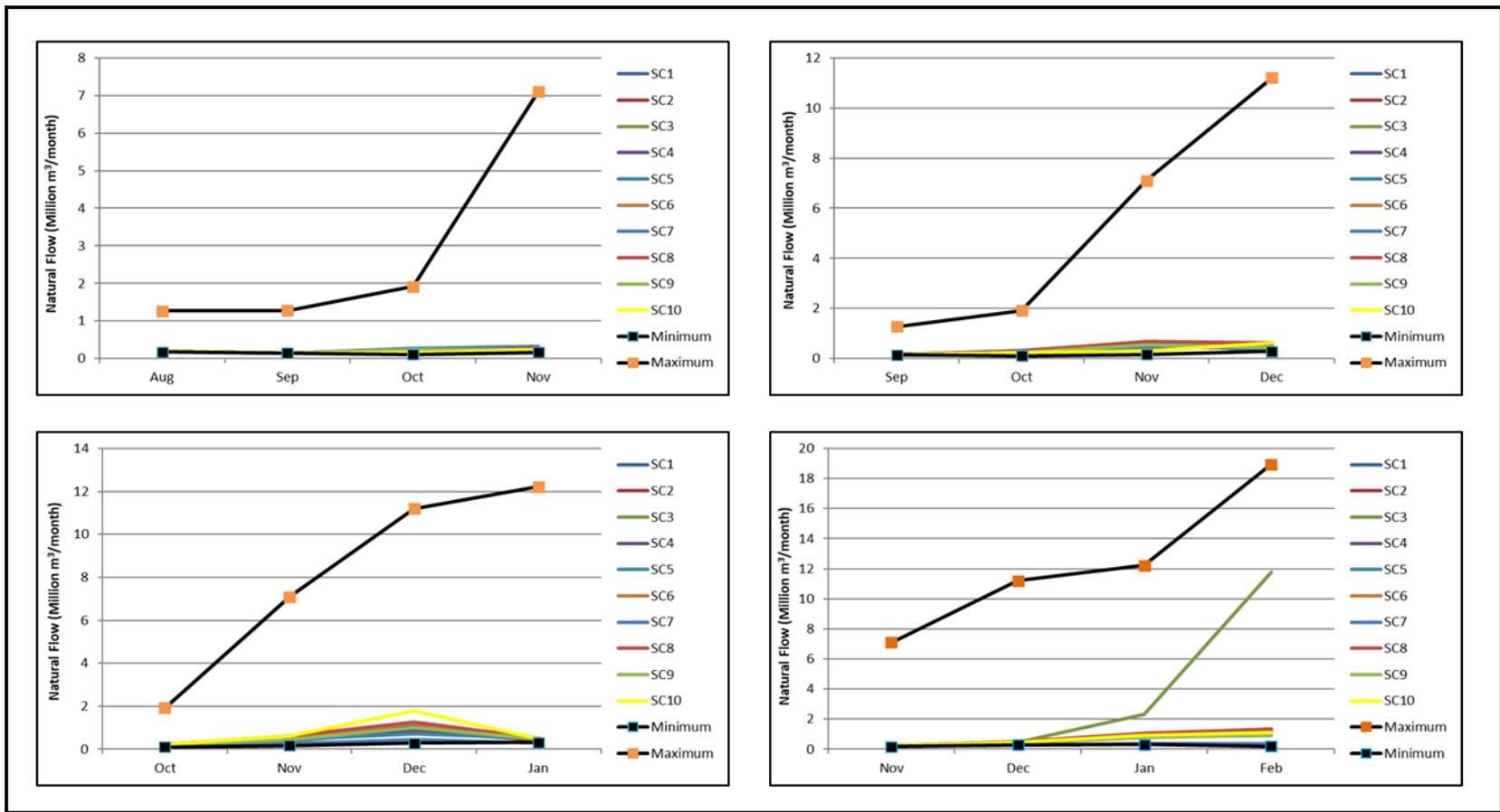


Figure 8.16: Modelled natural flow using projected rainfall (September to December 3 month projections)

The projected combined storage of the Klipkopje and Longmere dams are shown in Figures 8.17, 8.18 and 8.19. Note that the starting storage of the dam is set to the observed storage with each simulation.

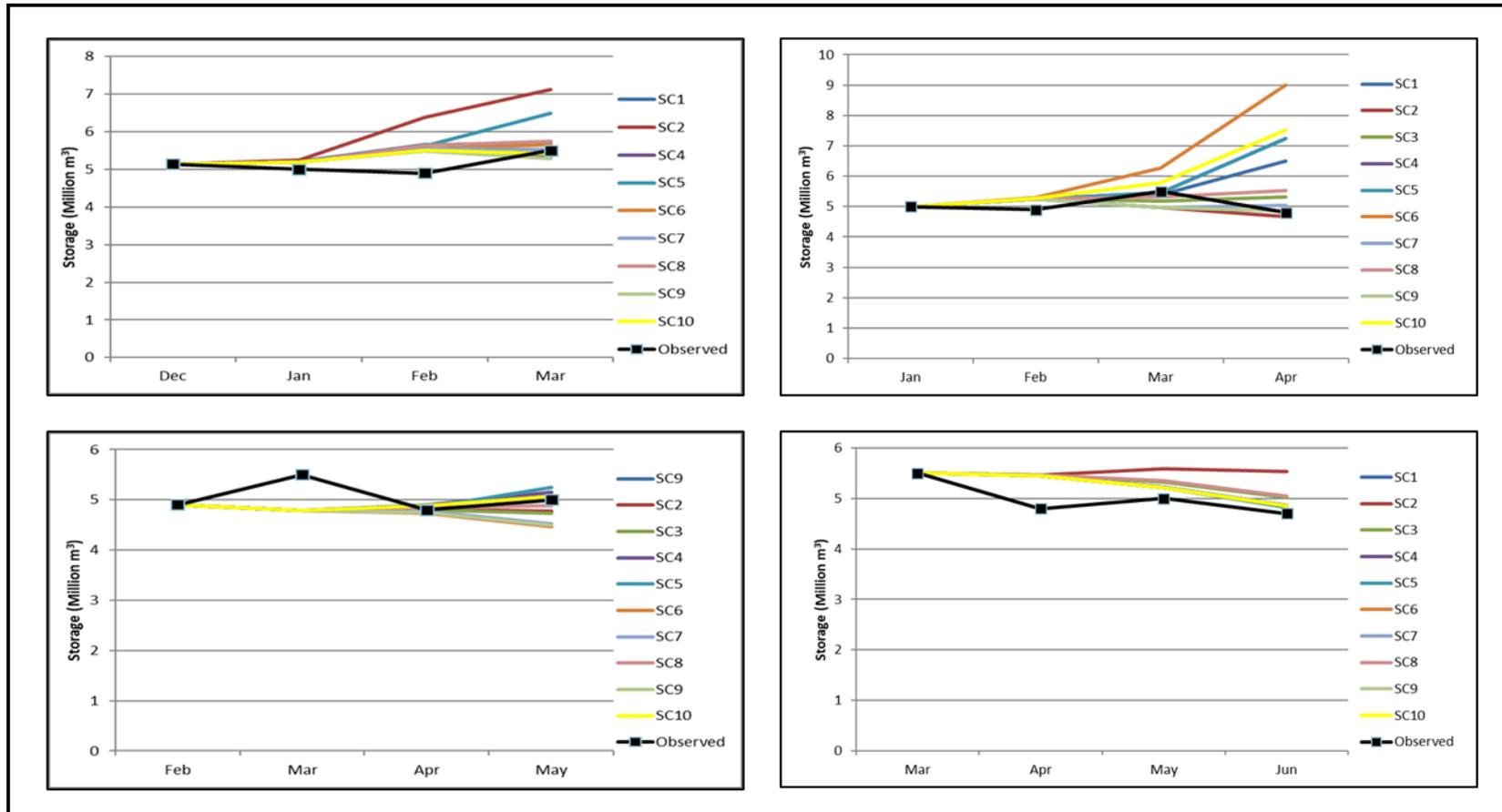


Figure 8.17: Modelled storage using projected natural flow (January to April 3 month projections)

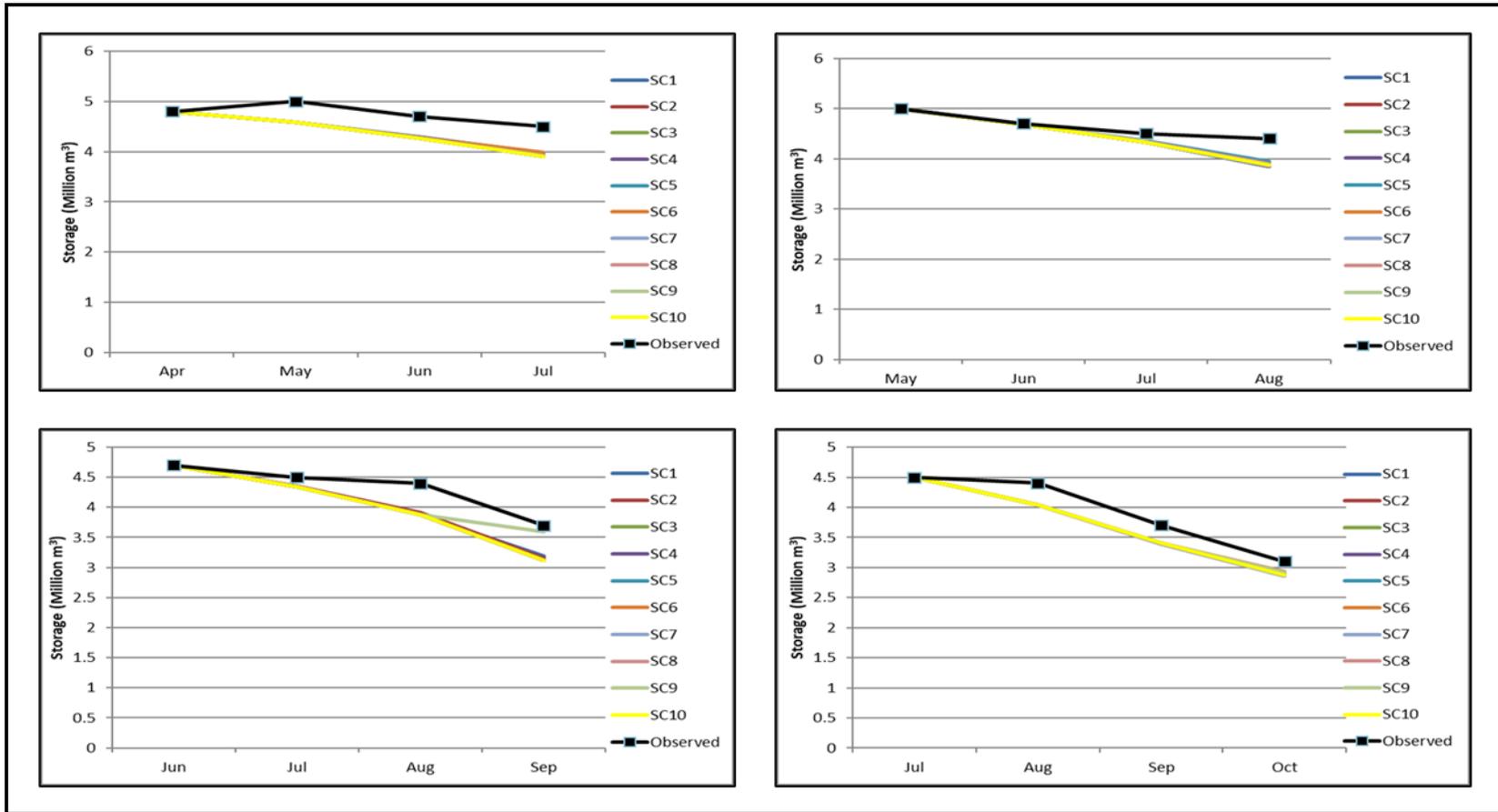


Figure 8.18: Modelled storage using projected natural flow (May to August 3 month projections)

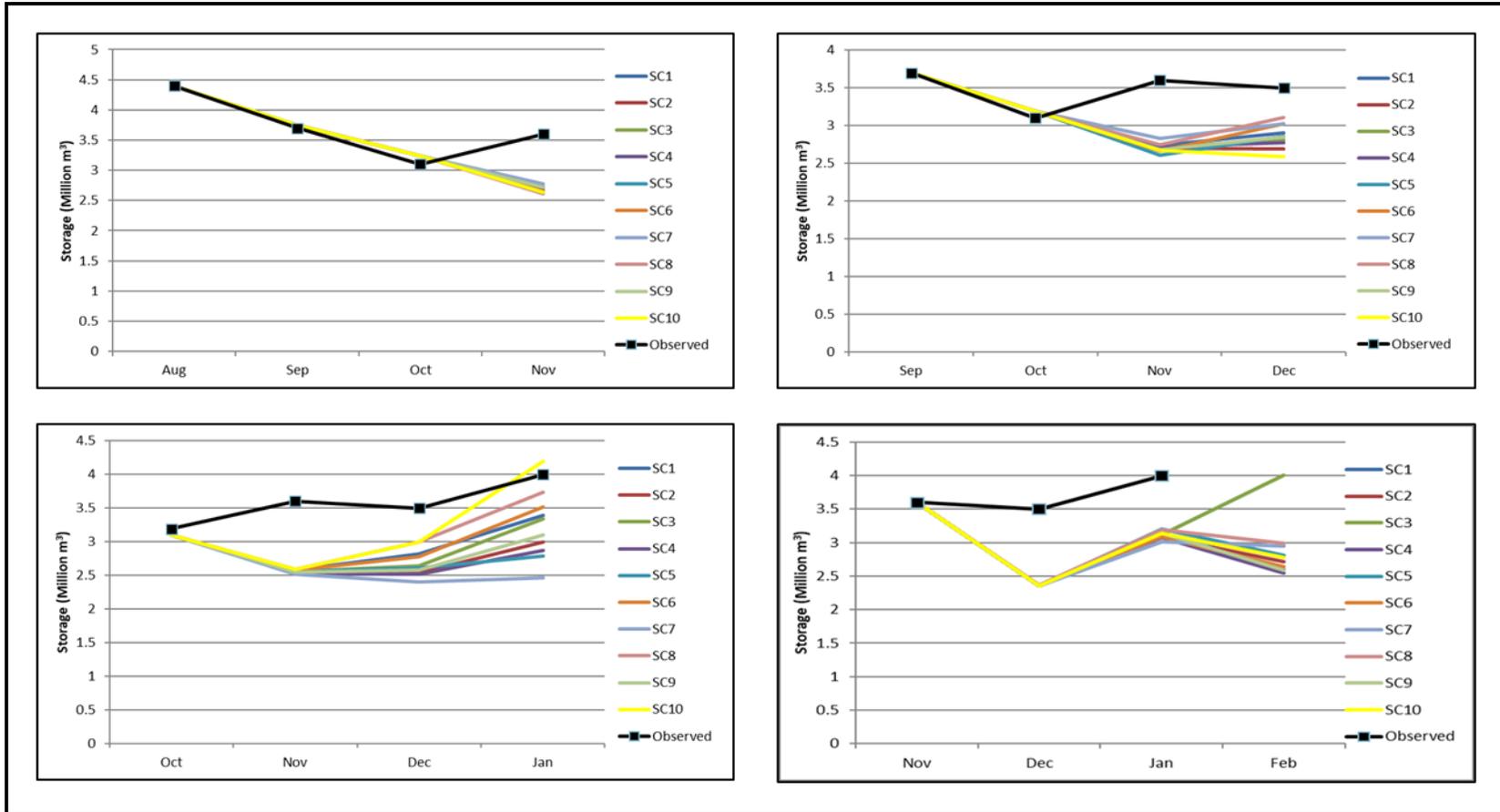


Figure 8.19: Modelled storage using projected natural flow (September to December 3 month projections)

The projected rainfall within the Klipkopje and Longmere dam catchments are shown in Figures 8.20, 8.21 and 8.22.

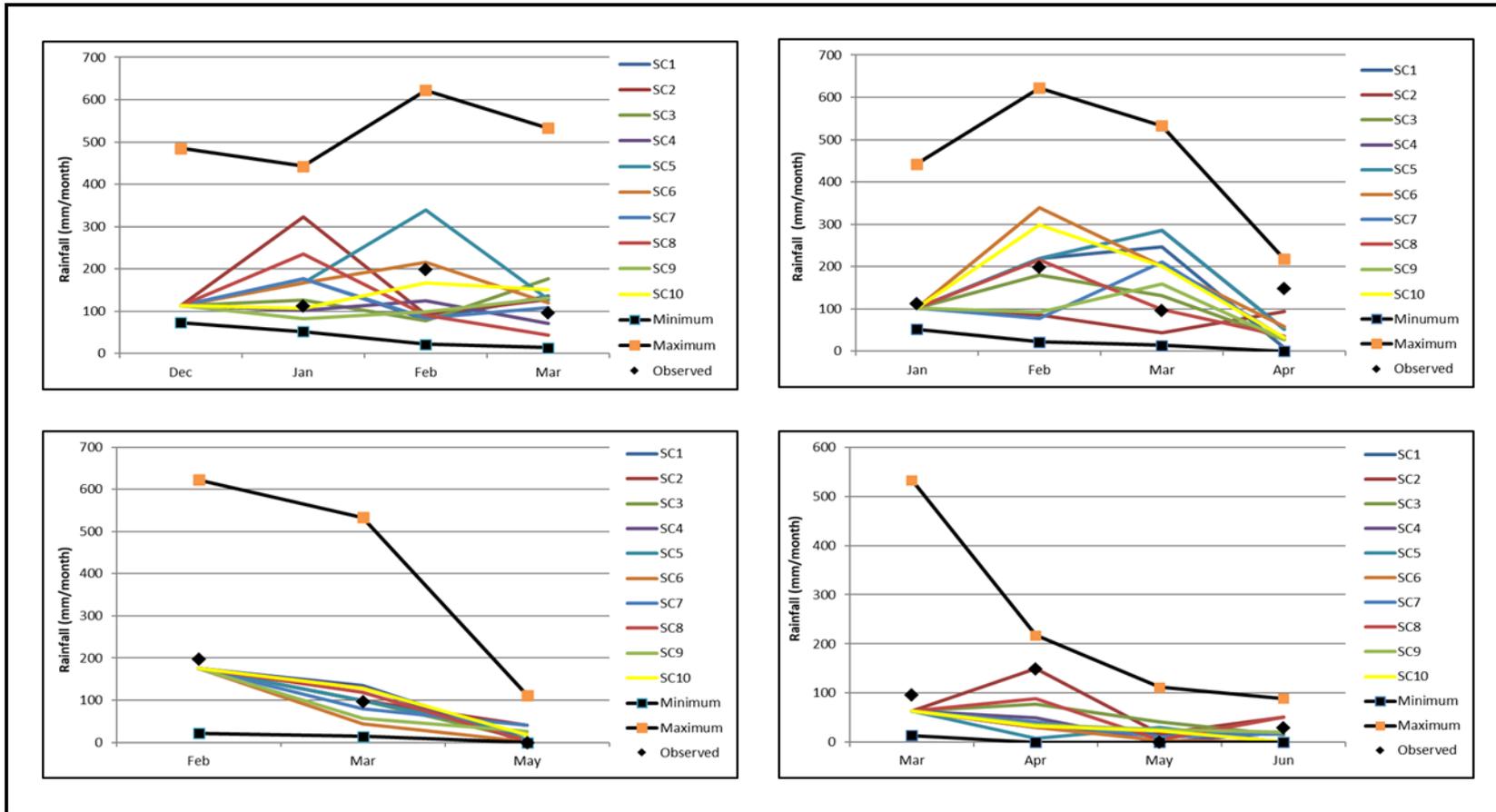


Figure 8.20: Rainfall prediction versus observed rainfall (January to April 3 month projections)

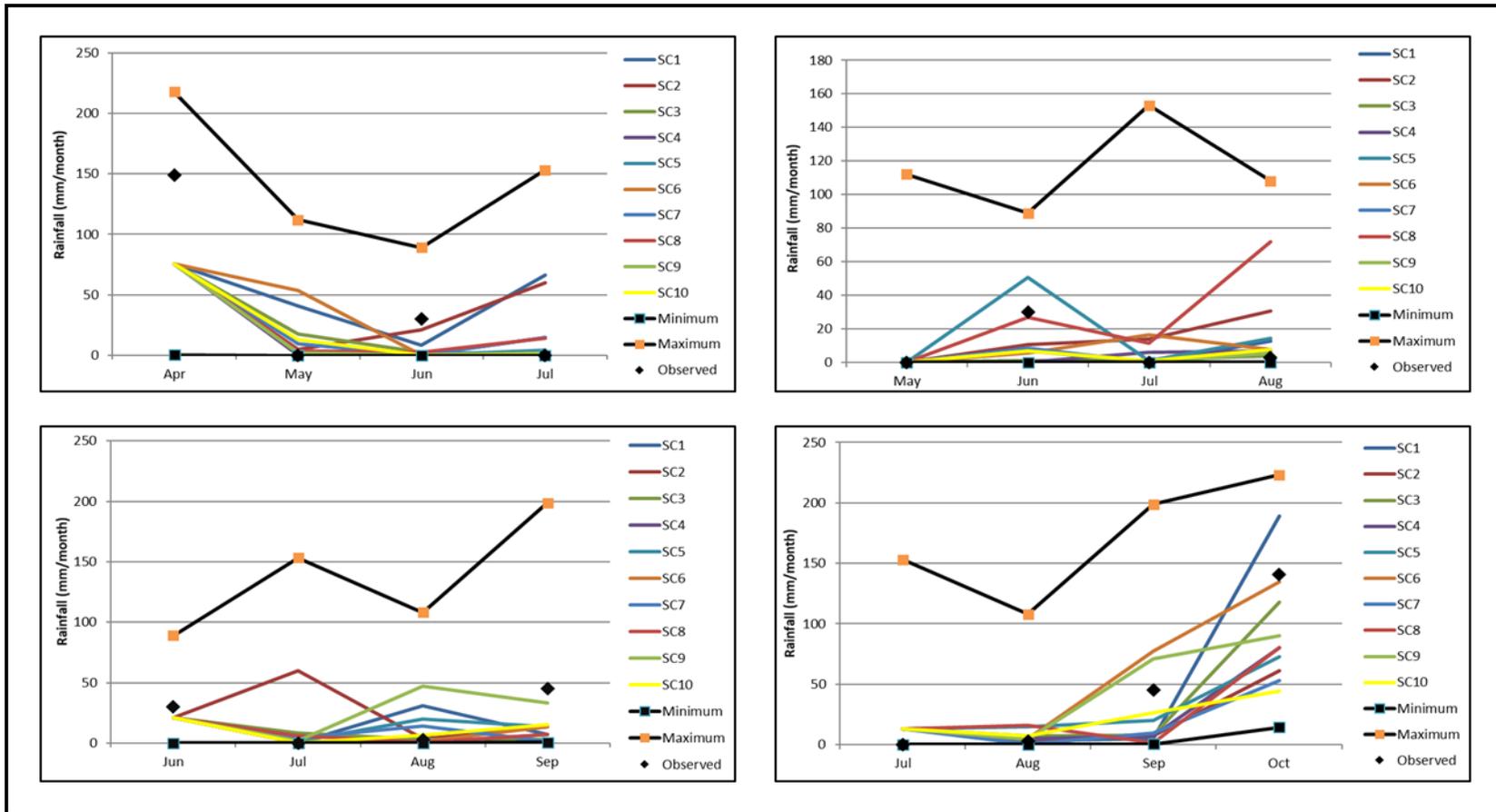


Figure 8.21: Rainfall prediction versus observed rainfall (May to August 3 month projections)

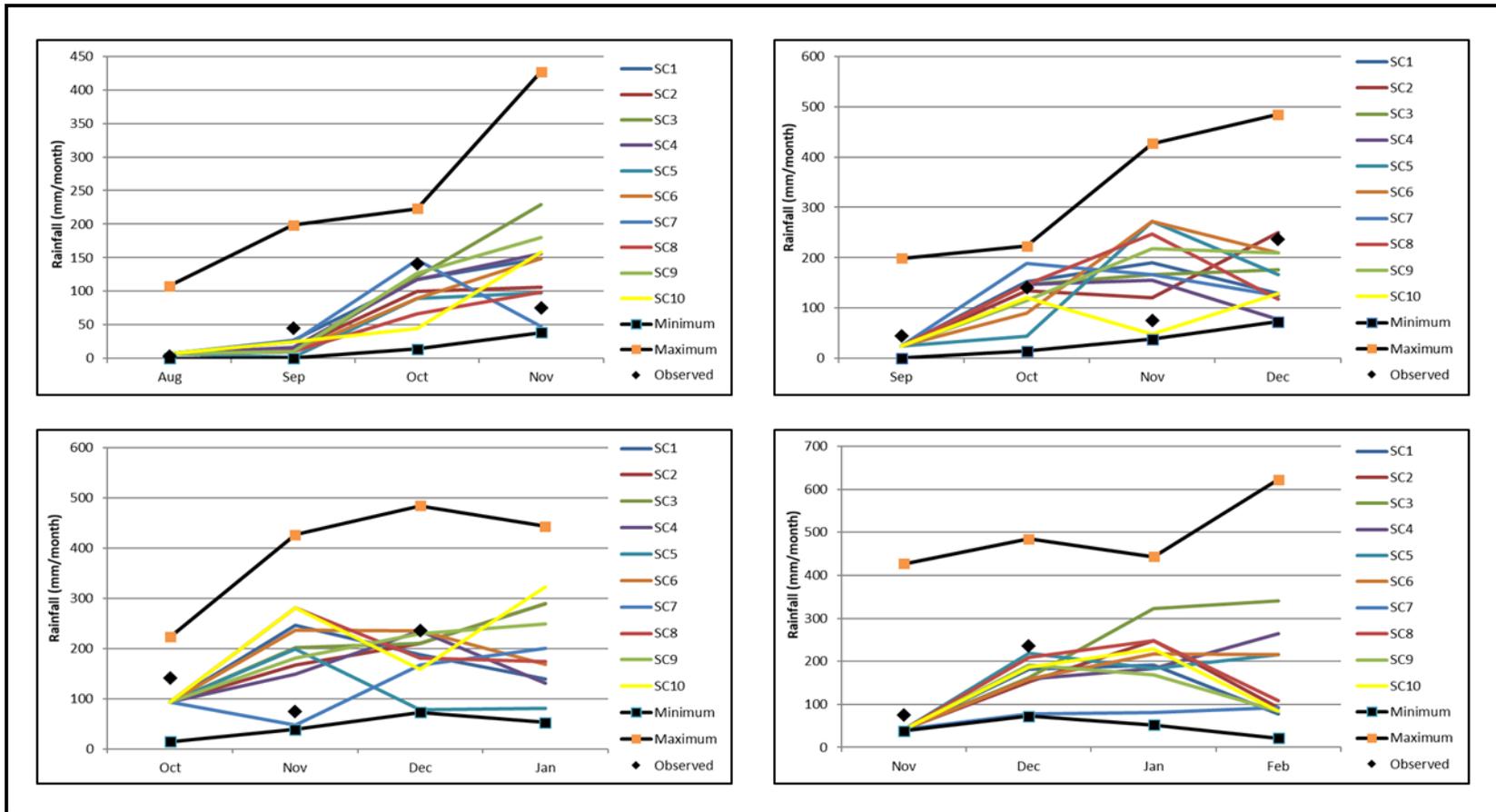


Figure 8.22: Rainfall prediction versus observed rainfall (September to December 3 month projections)

8.1.3 Karatara River

The Theewaterskloof and White River pilot studies made projections of storage in dams based on rainfall projections. In both cases the modelling procedure seems to underestimate the storage. Since there are many components to the integrates system it is no possible from the previous two pilot model setup to identify the source of the discrepancy. Therefore, a third pilot study was selected in a catchment without storage and a reliable flow gauge which is recording flow which is largely natural. The selected catchment is the Karatara River in the K40C quaternary catchment. See Figure 8.23. In this pilot model setup the project catchment runoff is compared with the observed runoff.

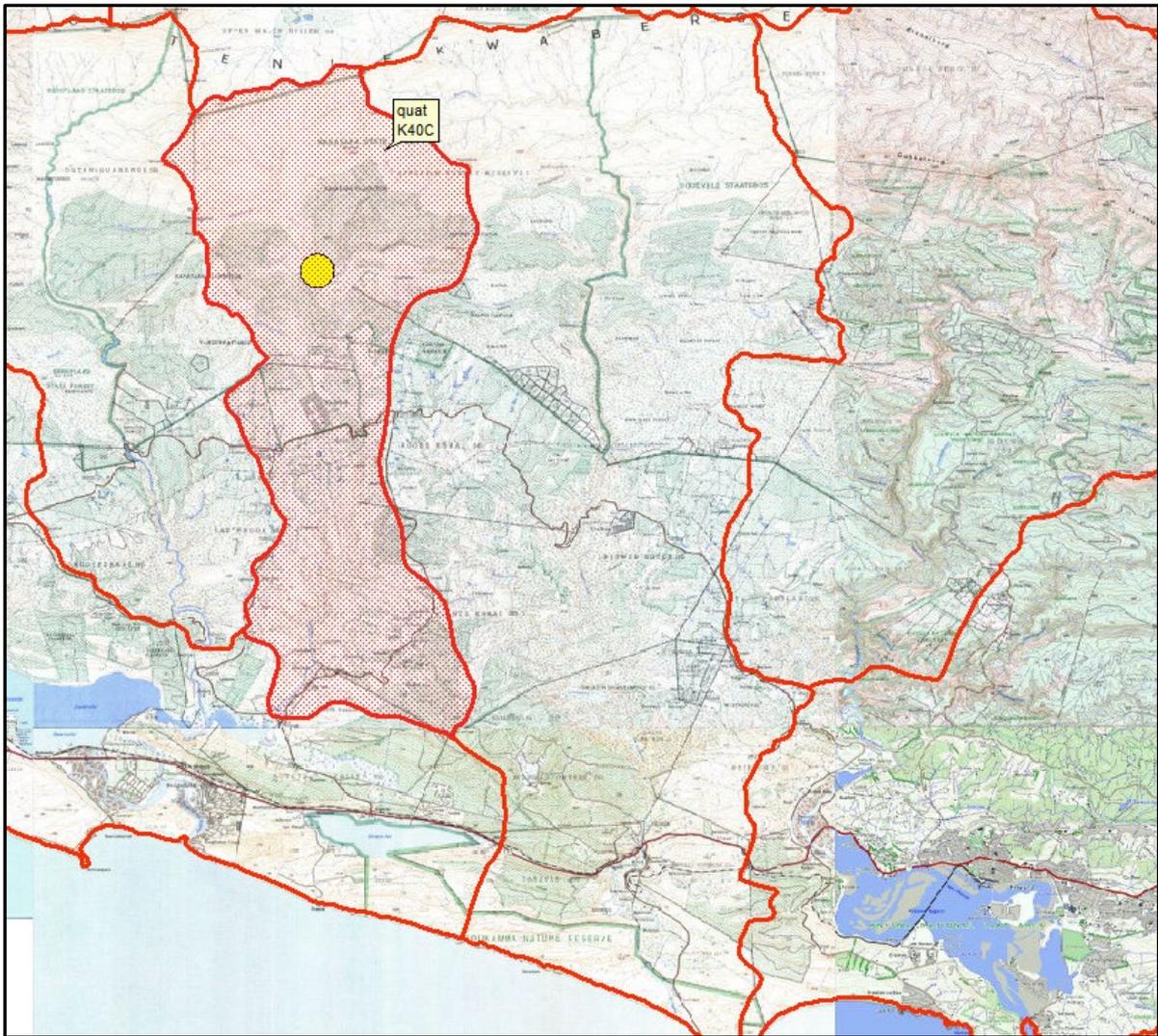


Figure 8.23: Location of the Karatara River catchment

Table 8.6: Summary of climate and hydrology information for the K40C quaternary catchment

Catchment	Area (km ²)	Mean Annual Evaporation (mm)	Mean Annual Precipitation (mm)	Natural Mean Annual Runoff (million m ³ /annum)
K40C	120.0	1 400	930	33.0
Gauge K4H002	22.0	1 400	930	6.0

Source: WR2012 with K2H002 MAR based on linear scaling

A time series of the natural flow at K2H002 based on WR2012 and linear scaling is shown in Figure 8.24 while the monthly distribution of the natural flow is shown in Figure 8.25.

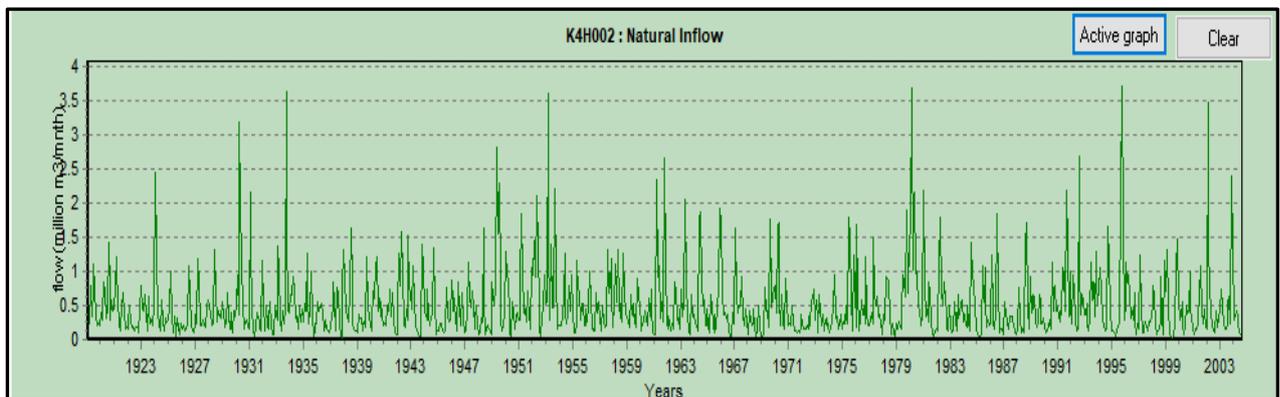


Figure 8.24: Natural flow time series at Gauge K4H002, Katara River

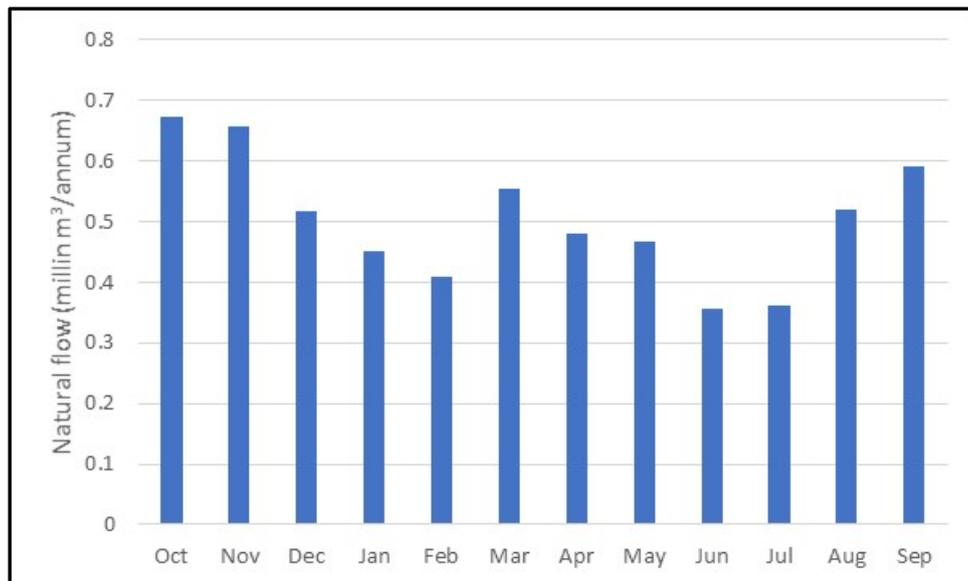


Figure 8.25: Monthly distribution of natural flow

Note that the runoff from this catchment is bimodal with peaks in October and March.

The water use in the catchment is limited to forestry plantations with an estimate area of 3.73 km². The streamflow reduction due to this commercial forestry is estimated at 0.32 million m³/annum. See Figure 8.26.

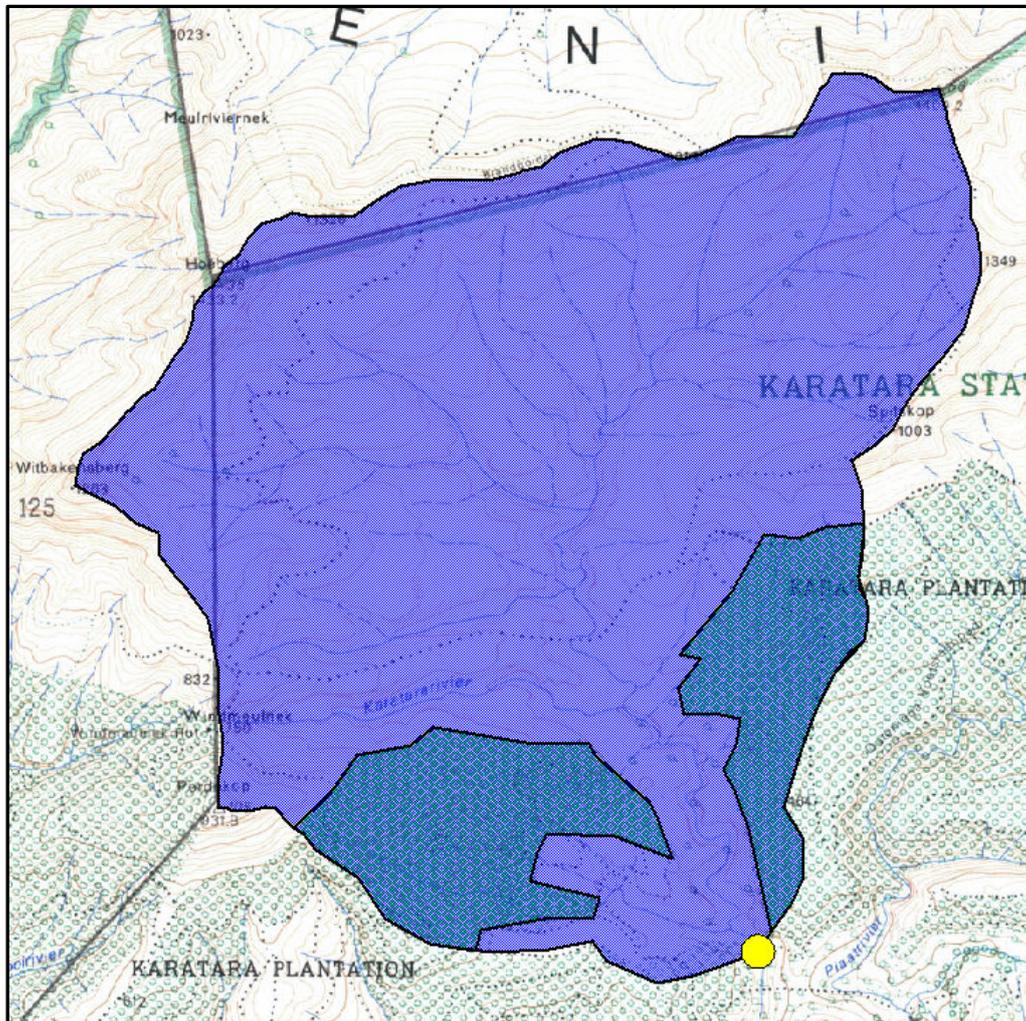


Figure 8.26: Land use in the Karatara catchment

The projected natural flows, commencing in January 2019, are shown in Figures 8.27, 8.28 and 8.29. The reason that the year 2019 was used is that the observed flow for 2020 is not yet available. These natural flows were derived using the Pitman model and the projected rainfall. Note that the rainfall projections are only for three months hence the limitation in the projected natural flows. The minimum and maximum natural flows in each month, as obtained from WR2012, are also shown on this graph.

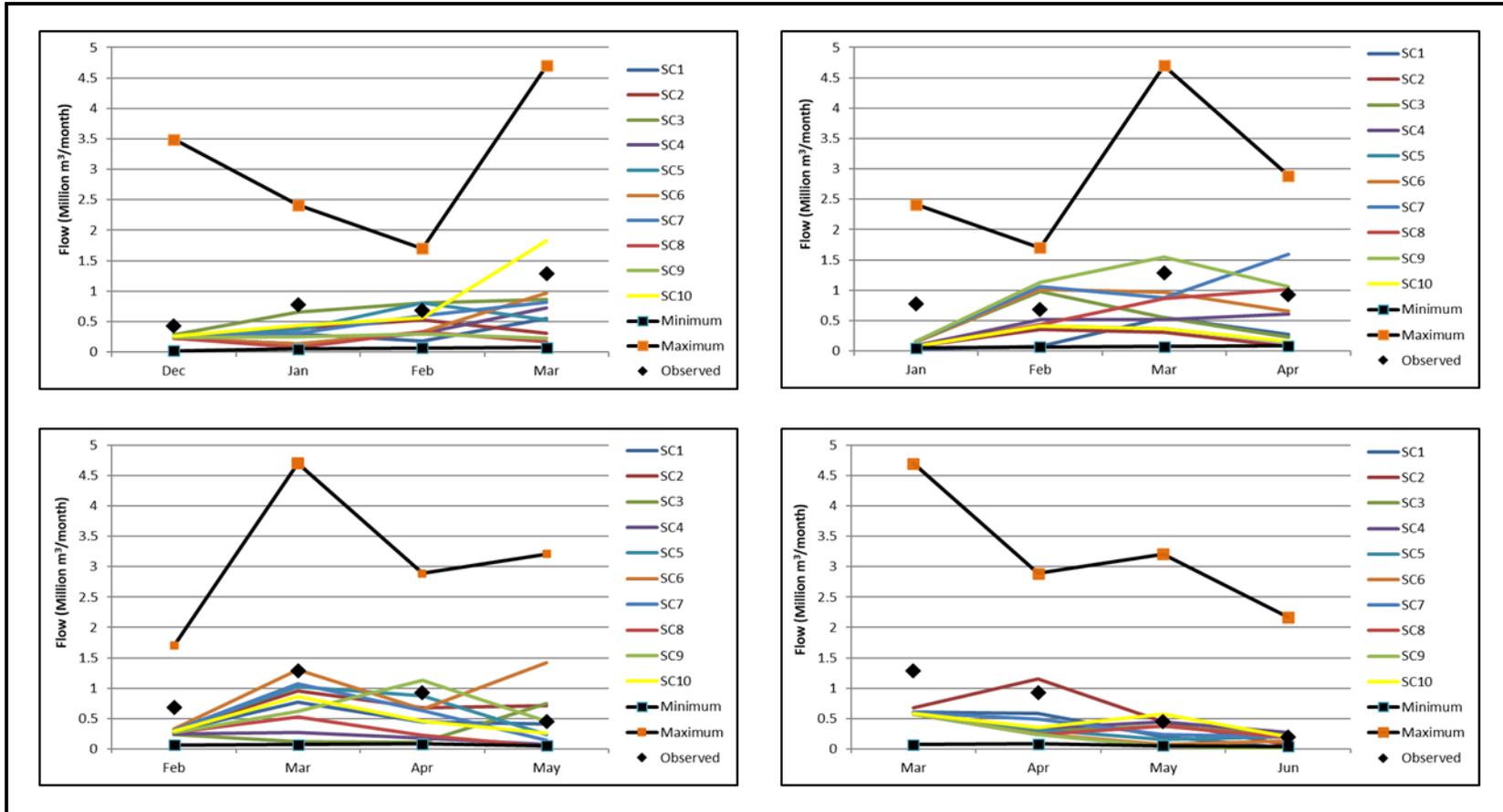


Figure 8.27: Modelled natural flow using projected rainfall (January to April 3 month projections)

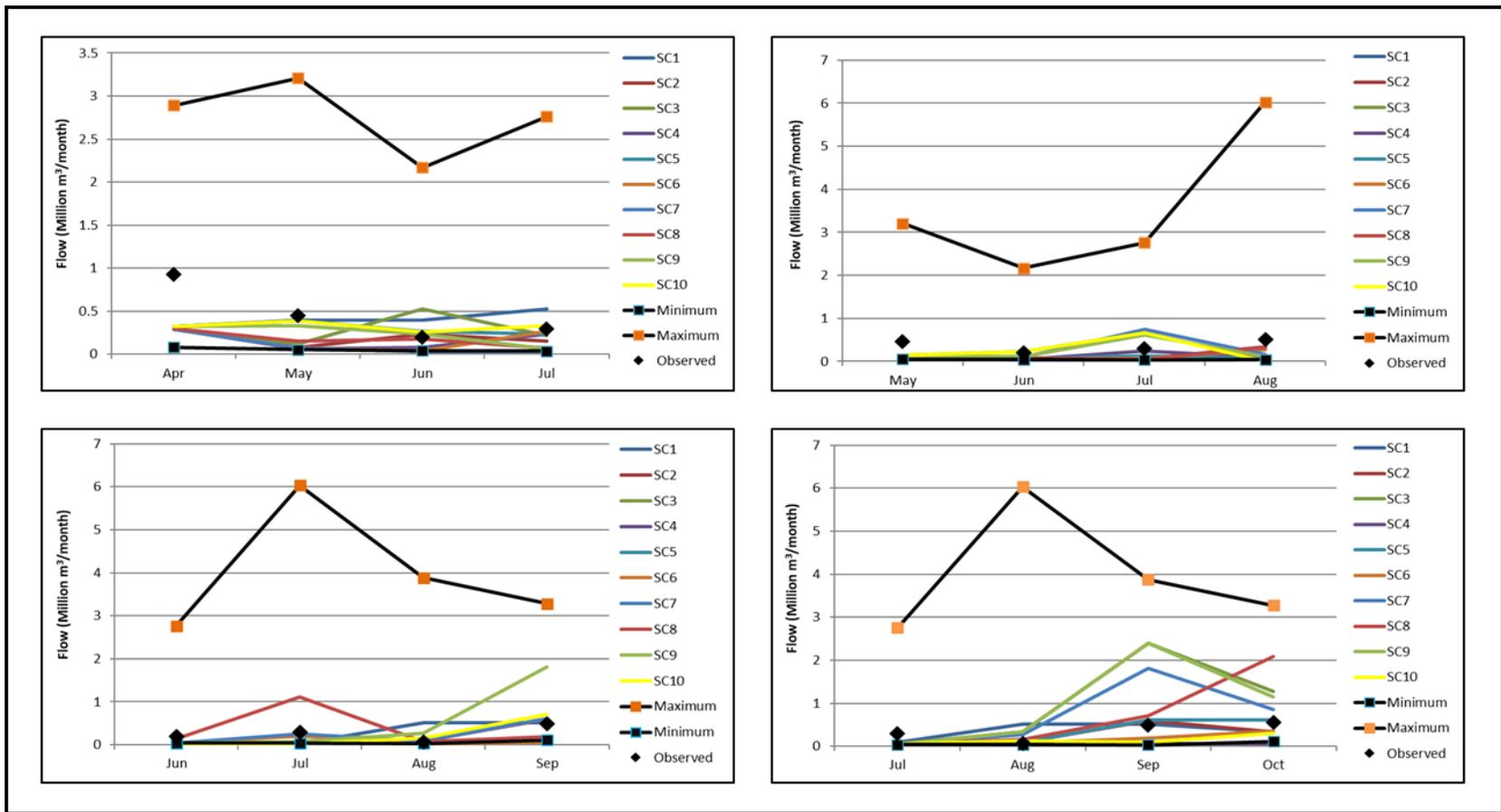


Figure 8.28: Modelled natural flow using projected rainfall (May to August 3 month projections)

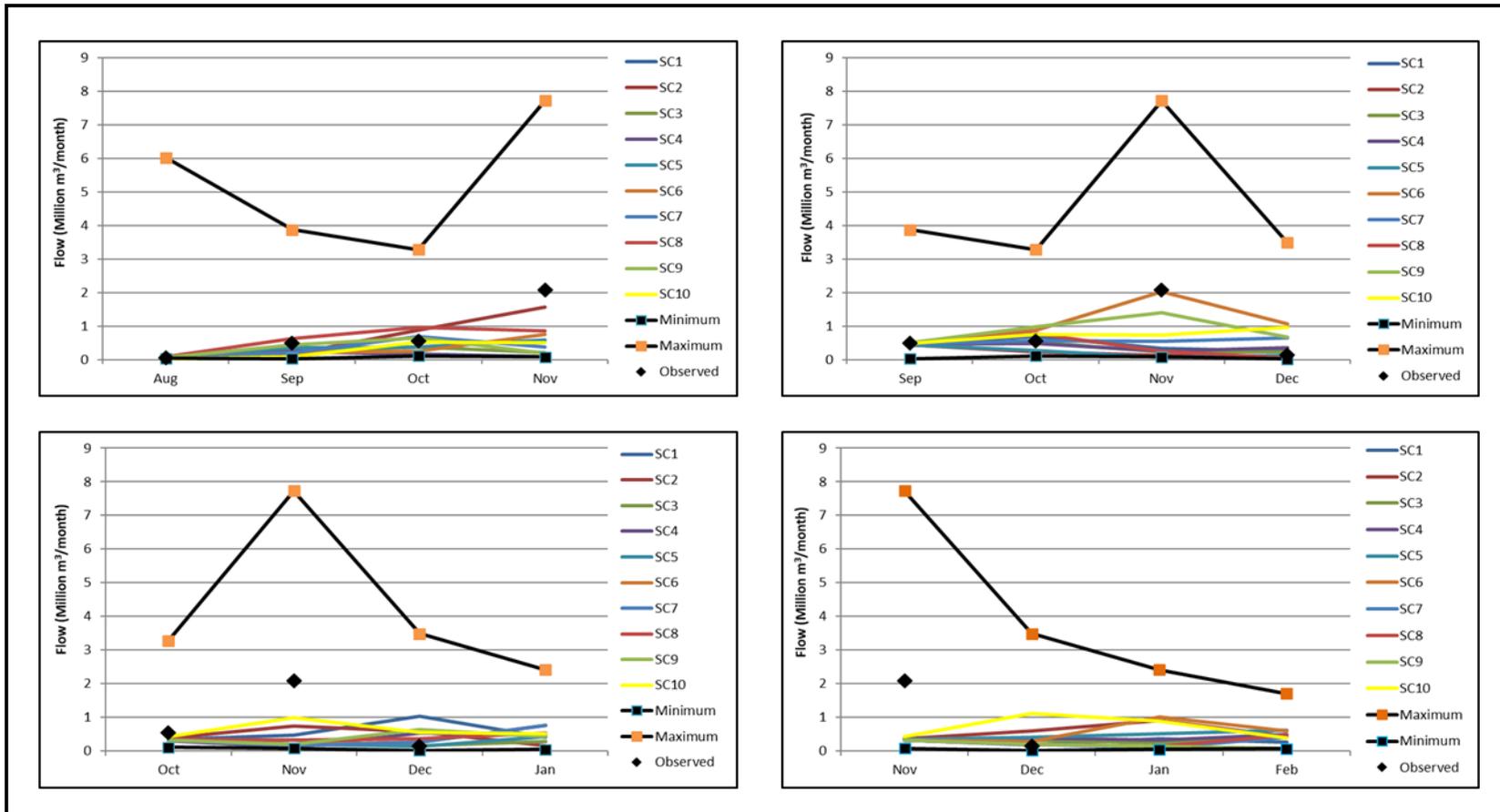


Figure 8.29: Modelled natural flow using projected rainfall (September to December 3 month projections)

8.1.4 Gutshwa River

The Gutshwa River catchment, located in the Crocodile catchment on the south-eastern border of the Kruger National Park, is an unusual catchment in that the rainfall is relatively high (800mm/annum) but the runoff very limited. This is due to the high rate of recharge into the alluvial aquifer. There are several monitoring boreholes in this catchment and hence it was selected as a case study for groundwater forecasting.

Figure 8.30 shows the location of the Gutshwa catchment which consist of quaternary catchment X24A, B and C.

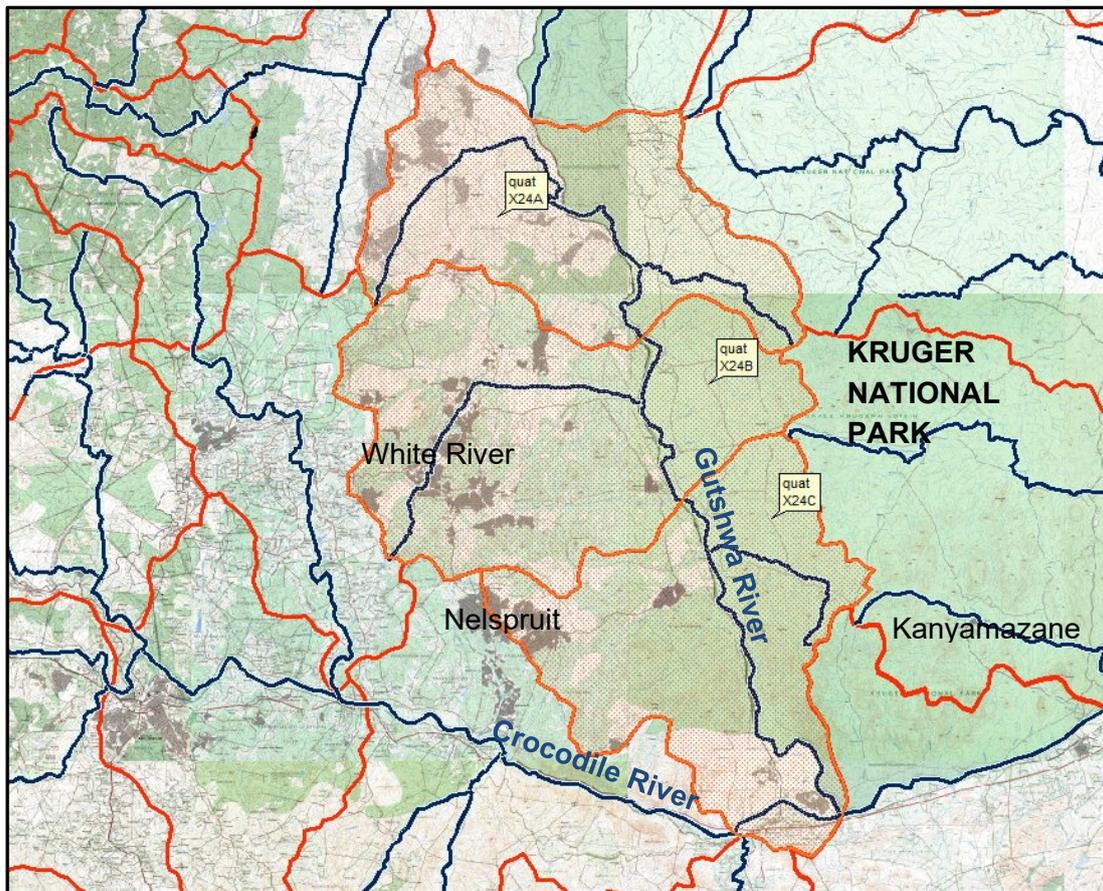


Figure 8.30: Catchment of the Gutshwa River, Mpumalanga

Table 8.7: Summary of climate and hydrology information of the Gutshwa River

Catchment	Area (km ²)	Mean Annual Evaporation (mm)	Mean Annual Precipitation (mm)	Natural Mean Annual Runoff (million m ³ /annum)
X24A	248.5	1 480	720	11.0
X24B	335.0	1 485	750	14.6
X24C	285.7	1 480	760	13.2
Total	869.2			38.8

Source: IUCMA, 2019

A time series of the natural flow at outlet from the Gutshwa River catchment based on the IUCMA hydrology, is shown in Figure 8.31 while the monthly distribution of the natural flow is shown in Figure 8.32.

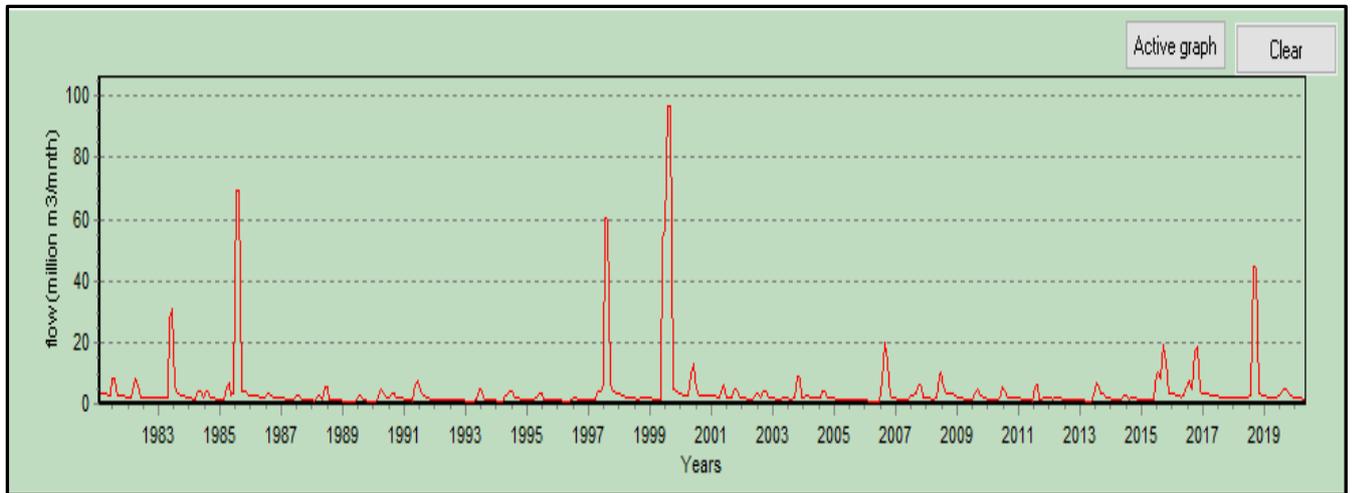


Figure 8.31: Natural flow time series at out of the Gutshwa River catchment

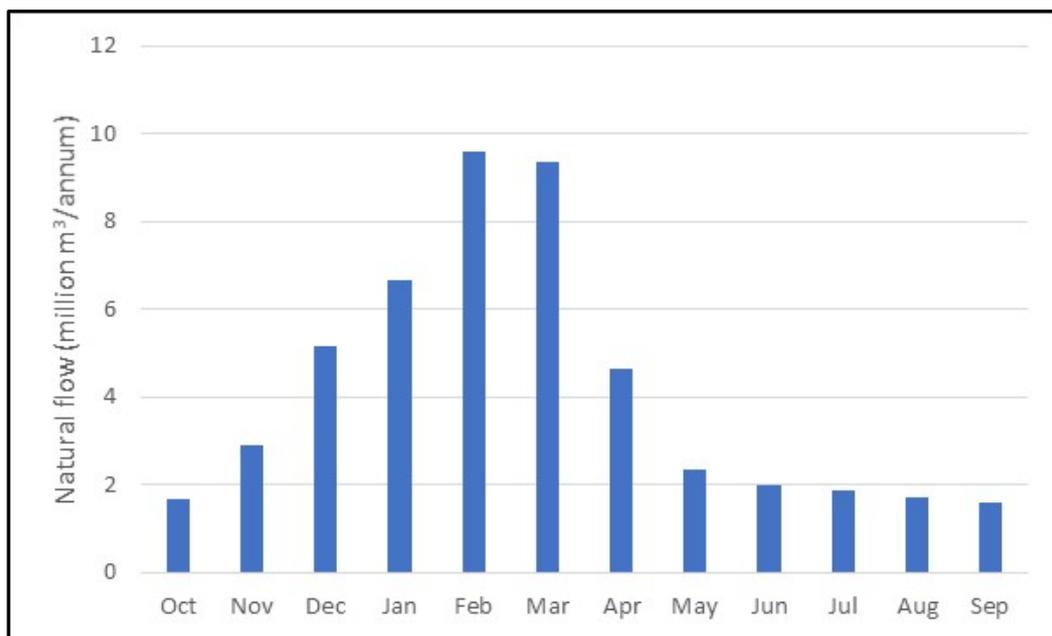


Figure 8.32: Monthly distribution of natural flow

While large areas of the Gutshwa catchment is urbanised, the water for these communities is all supplied from outside of the catchment. Since the water supply is erratic, it is suspected that many families, especially those in the new outlying more rural areas rely on groundwater to survive.

There are some registered irrigation users but there is this little evidence of significant irrigation currently taking place.

The groundwater component of the Pitman model was setup up with the Sami parameters, as published in the IUCMA hydrology report (IUCMA, 2019) and calibrated using the GW parameter, which is the maximum rate of groundwater recharge. Calibration was carried out by comparing the observed groundwater level at Gauge X2N013 with the modelled change in storage in the X24A quaternary catchment. See Figure 8.33. Not surprisingly, the correlation between the modelled storage and the observed water level is not very good since the borehole level will be influenced by point rainfall, which is unknown, and not the catchment rainfall used in the model. Also the level in the borehole is influence by abstraction which are no recorded. Nevertheless the general trend of change is storage from minimum to maximum and back to minimum is similar in the two time series.

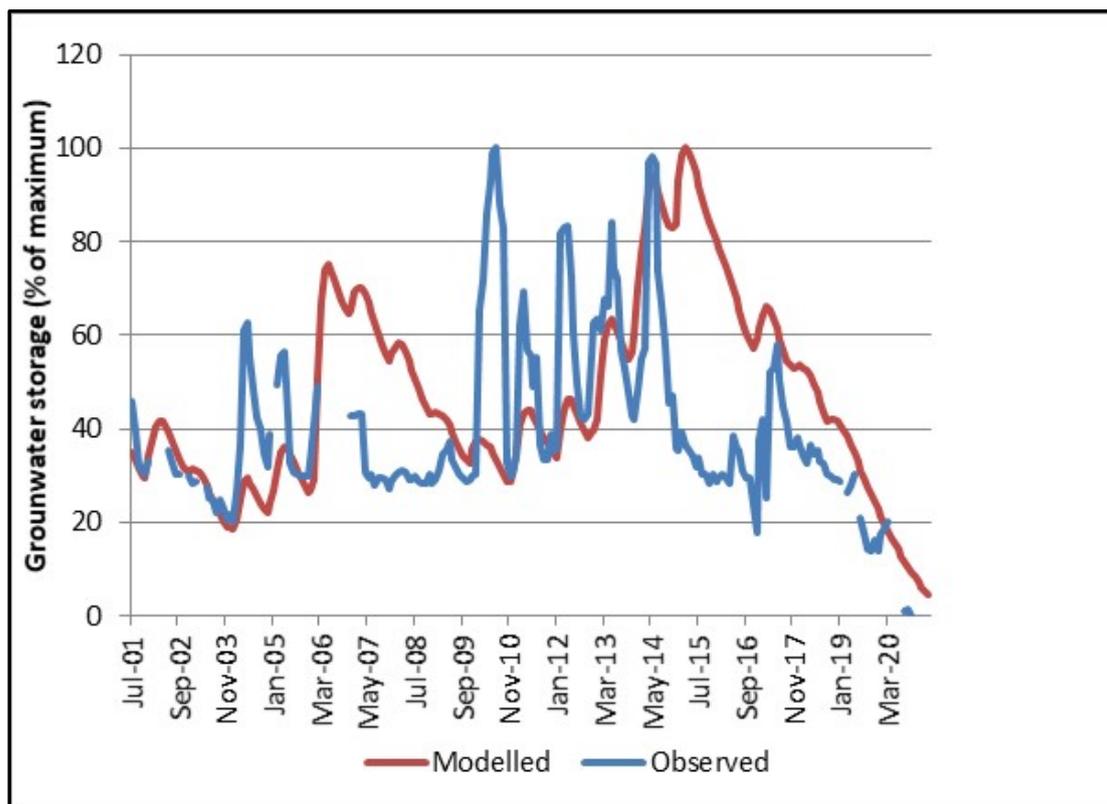


Figure 8.33: Modelled versus observed groundwater storage, expressed in dimensionless terms

The projected groundwater storage, commencing in January 2019, is shown in Figures 8.34, 8.35 and 8.36. These storages were derived using the Pitman model and the projected rainfall. Note that the rainfall projections are only for three months hence the limitation in the projected natural flows. The minimum and maximum natural flows in each month, as obtained from the 1982 to 2021 simulation, are also shown on these graphs.

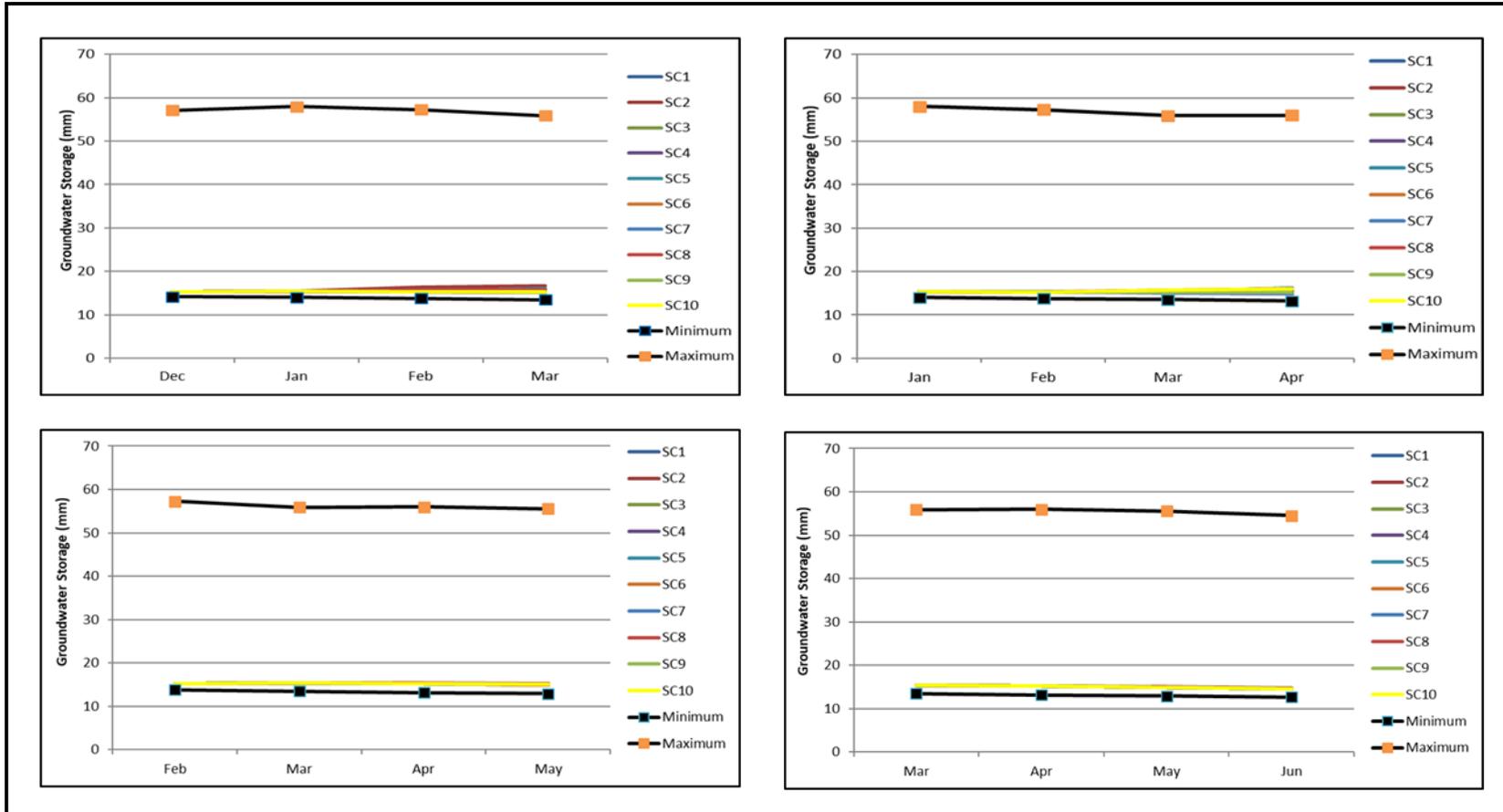


Figure 8.34: Modelled groundwater storage using projected rainfall (January to April 3 month projection)

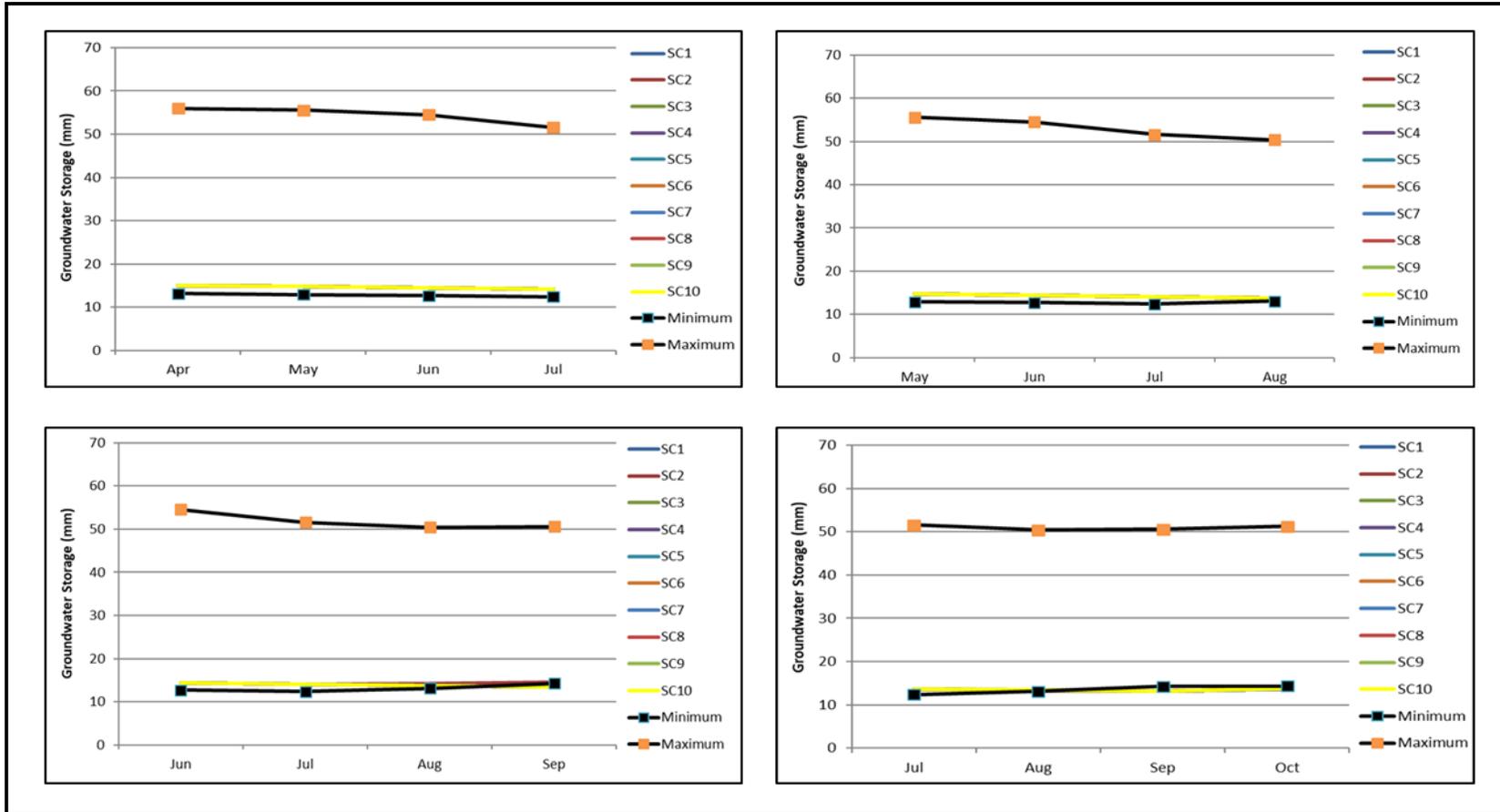


Figure 8.35: Modelled groundwater storage using projected rainfall (May to August 3 month projection)

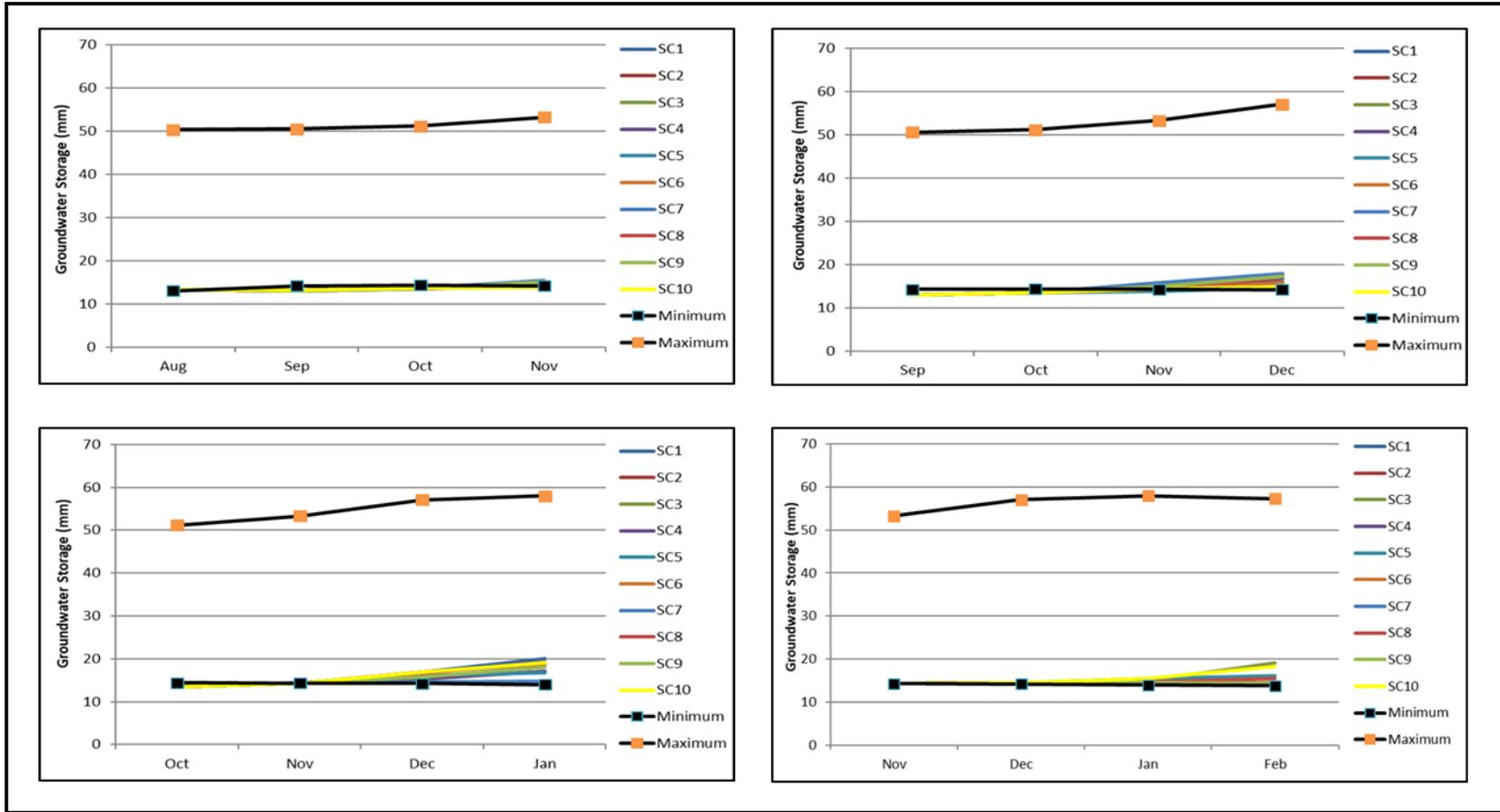


Figure 8.36: Modelled groundwater storage using projected rainfall (September to December 3 month projection)

The groundwater model shows very low storage going into the forecast period (year 2020). This is consistent with observed groundwater levels in the catchment. The reason for the low groundwater levels is the prolonged period of below-average rainfall, as can be seen by plotting the 10 year moving-average of the rainfall in the catchment. See Figure 8.37.

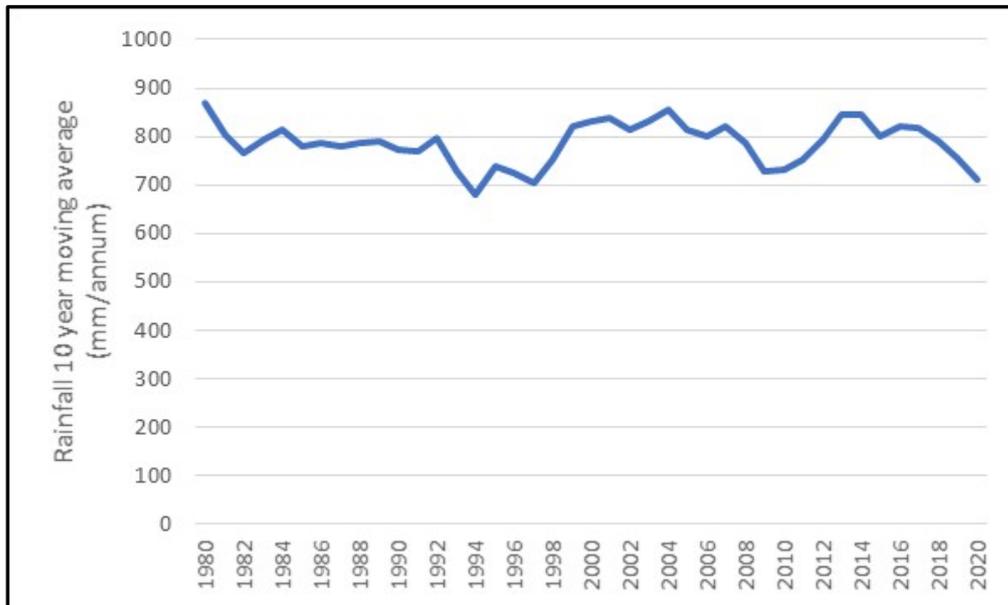


Figure 8.37: Rainfall in the Gutshwa catchment (10 year moving average)

9 RESULTS AND DISCUSSION

9.1 Rainfall forecasting

The rainfall forecasts using numerical climate model assessed as part of this project show only moderate to relatively low skill in forecasting total seasonal rainfall at quaternary catchment level when such a forecast is issued at the onset of the season and is based on ensemble simulations with numerical forecast models. The results are qualitatively consistent with the general understanding of predictability of southern Africa rainfall, i.e. that moderate levels of skill can be obtained for specific seasons and locations, while at other seasons and locations numerical forecasts at seasonal time scales remain unreliable.

The analyses undertaken aimed at assessment and implementation of a method that generates information at the level of quaternary catchments, which has never been done before in South Africa.

This shows that there is a potential for improvement of the seasonal forecast in the context of early warning by using ‘bespoke’ rather than generic approaches and this can perhaps be achieved to the level that such a forecast be incorporated in a probabilistic early warning system. The approach can be further refined, by, for example, separating events based on ENSO conditions (as ENSO state is one of the sources of predictability in the southern Africa climate), by targeting other attributes of rainfall (such as drought indices, or duration of dry spells, etc.), and by incorporating external drivers such as ENSO state explicitly into the process of downscaling (i.e. not just indirectly through forcing’s of the numerical forecast model).

Other conclusions reached are:

- There does not appear to be one model that is universally better than other models.
- There is a considerable heterogeneity in skill from location to location resulting in a range of skill values even for relatively small river basins. This seems to be linked to the difference in characteristic size of the quaternary catchments and the resolution of the forecast information (superposition of the observational GPCP grid of 0.5 degree and 1 degree forecast model grid).
- The range of skill within a river basin seems to increase with the rarity of the event – i.e. the location-to-location differences in skill are larger for rare events. This is likely an expression of sensitivity of rare events skill measures to the process of fitting distribution in order to estimate forecast probability.

The above results are not strongly supporting the idea of the use of the ensemble of models to forecast drought – levels of skill are low to moderate at best, and strongly heterogeneous. The analyses, however, provide a very strong starting point for the development of the drought forecasting system in that:

Three methods for forecasting rainfall are described in Section 4 of this report and are summarized here.

The first method used the forecast generated by the following three models:

- South African Weather Service (SAWS) using a fully coupled ECHAM4.5-MOM3-SA global climate model (Beraki et al. 2014).
- NOAA's Climate Forecasting System v. 2 (CFSv2, Saha et al. 2014, freely available through cfs.ncep.noaa.gov).
- ECMWF forecast based on IFS atmospheric model

The second method uses two seasonal forecast multi-model ensembles which became publicly available during the course of this study, namely, the NMME (Kirtman et al., 2014) and the Copernicus@ECMWF ensemble. These ensembles include two of the models used in section 4.3 – CFS v.2, and ECMWF IFS (called SEAS5 in the report). The forecast models are of spatial resolution (1 deg, ~100 km) that is low enough to allow a direct use of forecast rainfall and temperature data on the monthly basis in further analyses. The grand ensemble of 11 forecast model data from world-leading climate modelling institutions is organized into a uniform format allowing for the use of the forecast to drive a country-wide water management model and performing analyses and interpretation of their results and skill within a single analytical framework. This allows for a relatively easy interrogation of the entire ensemble and generation of bespoke skilful forecasts – by, e.g., selecting individual models or the group of models that are characterized by forecast skill in a particular location and in a particular month.

The third method presented in this report is a simple statistical forecast of the end-of-season-accumulated rainfall totals. A simple statistical forecast of the end-of-season-accumulated rainfall totals can provide useful information a number of months in advance to the end of the season. Skill of such a forecast is relatively high and capitalizes on the so called “committed” anomaly, i.e. the anomaly that has arisen at the time of the forecast. Persistence and ENSO can explain little, typically less than 10% of variance of the end of the season anomaly. As a result, the actual usefulness of the forecast is limited to months around the middle of the rainy season. Too early in the rainy season, and the forecast has little skill. In the later phases of the season, information generated by the forecast becomes redundant. Forecast is monitoring-based, i.e. it needs a long time series (in excess of 30 years) of consistent data stretching to current. That might limit its implementability in the South African context. This method does not produce a time series of the forecast rainfall so it is not yet possible to use this method to forecast streamflow.

9.2 Streamflow, soil moisture and groundwater forecasts

Natural streamflow forecasts are achieved with the aid of an updated Pitman model, which was adapted to carry out simulations with multiple rainfall scenarios and hence produce an ensemble of streamflow time series. Implicit within the Pitman model are soil moisture and

groundwater modelling. Hence forecasts of soil moisture and groundwater are also produced as an output of the model.

9.3 Storage forecasts

Once the natural streamflow ensemble has been generated using the Pitman model, storage forecasts can be made using any Reservoir Simulation model. The IDMAPS uses the Water Resources Modelling Platform, which has been adapted to, operated seamlessly on this system, that is, carry out catchment wide analyses, including water use and storage, so as to predict storage in reservoirs given the starting storage of each reservoir.

9.4 Pilot studies

The IDMPAS was tested on four pilot catchments with the aim of testing streamflow prediction, storage prediction and groundwater prediction. Pilot catchments in three difference climatic zones were selected so as to assess the performance of the model under difference rainfall distributions. The results and conclusions from these pilot studies are summarised as follows:

9.4.1 Theewaterskloof Dam

The Theewaterskloof Dam was modelled to test storage predictions. This dam is located in a winter rainfall region.

The projected natural streamflow for the Theewaterskloof system fall well within the envelope of natural flows from WR2012 with only one outlier. This at least shows that the projected streamflow is plausible. Without an observed streamflow to compare with, it is not possible to say if these projections are statistically better than random projections or better than stochastic models. However, the range of projected flows one and two months into the future are within a very small range compared to the min/max envelope which seems to suggest a useful projection.

The projected storage in the dam lies within a tight range which also suggests a useful prediction. The prediction for the winter months does however underestimate the storage in the dam. This could either be because the rainfall predictions underestimated the recent good rainfall, or the abstractions from the dam were less than modelled. It needs to be noted that abstractions from the dam are not available in real time and there is some uncertainty as to how much water was abstracted during the simulation period.

9.4.2 White River Systems

The White River Systems was modelled also to test storage forecasts. This system lies within a summer rainfall region.

The rainfall predictions result in simulated runoff that is skewed towards the lower flows. However, this is not surprising since the runoff from summer rainfall catchments is not uniformly distributed but skewed towards the lower flows.

Modelling the storage in the Klipkopje and Longmere dams using the predicted inflows resulted in an underestimate of the storage late summer months, good predictions during winter, and an underestimate of the storage in the early summer months.

In order to check if the poor prediction lies with the hydrological modelling or the rainfall predictions, the predicted rainfall was compared with the observed rainfall at Klipkopje Dam. See Figures 8.20, 8.21 and 8.22. This shows that while the rainfall forecasts are contained within the long-term observed range obtained from WR2012, the predictions were not particularly useful with rainfall overestimated in later summer and underestimated in early summer. It must be noted, however, that the rainfall forecast is for a rainfall zone (X22E) while the observed rainfall is point rainfall at Klipkopje Dam which lies within the rainfall zone but is not representative of rainfall within the entire zone.

9.4.3 Karatara River catchment

This small catchment was modelled to compare forecast streamflow with observed flows. The catchment is located in an area which experienced rainfall throughout the year.

The runoff predictions based on rainfall prediction within the Karatara River catchment fall within the predicted range with only one exception. These are therefore useful predictions. Additional analysis is however required in order to evaluate if these projections are statistically better than random projections or better than stochastic models.

9.4.4 Gutshwa River catchment

The Gutshwa River catchment was modelled in order to test the groundwater forecasting capabilities of the IDMAPS. This catchment is located in a summer rainfall area.

While there are many more uncertainties in modelling and hence forecasting groundwater, the groundwater pilot study modelled groundwater storage consistent with observed groundwater levels and showed that groundwater storage changes slowly over time, much slower than streamflow or reservoir storage. Hence the groundwater forecasting model will be useful in forecasting groundwater storage.

10 CONCLUSIONS AND RECOMMENDATIONS

10.1 Rainfall prediction

Rainfall prediction using an ensemble of GCM forecasts show moderate to relatively low skill in forecasting rainfall at quaternary catchment level with strong month to month variability, and region to region as well as model to model differences in skill. In terms of seasonal differences – there appears to be a generally better skill of almost all models in forecasting rainfall anomalies in the north-east of the country in Nov-Dec-Jan and in the central part of the country in Dec. This is consistent with the general understanding of rainfall predictability in South Africa.

There is a considerable heterogeneity in skill from location to location resulting in a range of skill values even for relatively small river basins linked to the difference in characteristic size of the quaternary catchments and the resolution of the forecast information. The range of skill within a river basin seems to increase with the rarity of the event – i.e. the location-to-location differences in skill are larger for rare events. This is likely an expression of sensitivity of rare events skill measures to the process of fitting distribution in order to estimate forecast probability.

As an alternative to the use of GCM based forecasts, a statistical rainfall forecasting model was developed. This approach is based upon observations that rainfall anomalies observed in the beginning of the season tends to persist throughout the season, particularly when considered in terms of accumulated, or total season's rainfall. A simple statistical forecast of the end-of-season-accumulated rainfall totals can provide useful information a number of months in advance to the end of the season. Skill of such a forecast is relatively high and capitalizes on the so called "committed" anomaly, i.e. the anomaly that has arisen at the time of the forecast. Persistence and ENSO can explain little, typically less than 10% of variance of the end of the season anomaly. As a result, the actual usefulness of the forecast is limited to months around the middle of the rainy season. Too early in the rainy season, and the forecast has little skill. In the later phases of the season, information generated by the forecast becomes redundant.

It has not yet been possible to express the statistical forecasts in terms of rainfall time series. Hence this approach has not been tested within the integrated system.

10.2 Streamflow forecasts

The Pitman model was recoded to allow for multiple scenario runs. This was then applied to three pilot catchments, one in the winter rainfall zone, one in the summer rainfall zone, and the third in a zone which has a bimodal rainfall distribution. The streamflow predictions generally fall well within the probable range and hence offer plausible predictions. A comparison of the modelled streamflow against observed streamflow in one of the catchments shows promising forecasts. However, statistical analysis is required to establish if these predictions are a significant improvement on stochastic models.

10.3 Storage forecasts

The streamflow forecasts were used to model storage change in two catchments, namely, the Theewaterskloof catchment and the White River catchment. The forecast in the Theewaterskloof closely matched the observed storage and hence the Integrated Forecasting Tool could be useful for forecasting storage. It must be noted though that inflow into a dam is not the only parameter influencing the storage in a dam. Abstractions, which in the case of Theewaterskloof are highly variable, are also a significant source of uncertainty.

The modelling storage in the White River system was less successful but this can mostly be attributed to poor rainfall predictions. However, this needs to be seen in the context of model storage at sub-quatarnary scale while rainfall predictions are made at the large Rain Zone scale.

10.4 Groundwater/soil moisture

The intention at the commencement of this study was to use the GRACE method to monitor groundwater. GRACE is a remote sensing tool which has been established as a powerful tool to observe water storage dynamics at large scales. However, evaluation of this technology concluded that GRACE is not suitable for quatarnary scale monitoring.

As an alternative to GRACE, groundwater and soil moisture can be modelled within the Pitman model using methods proposed by Sami (2007) and Hughes (2007). These methods are already incorporated into the commercially available Pitman model. Groundwater algorithms derived from the publications of Sami and Hughes were incorporated into the recoded Pitman model used for this project. While groundwater and soil moisture have been modelled using this method, more work is required to assess if these projections are statistically significant or not. The pilot study carried out with a specific focus on groundwater show promising results but it is clear that groundwater storage changes at a much slower rate than surface runoff. Hence, a three month forecast is of limited value.

It is noted that soil moisture, from an agricultural perspective, relates to only the top 300mm of soil, while the Pitman model considers the entire soil profile. Soil moisture in the top layer of soil varies very rapidly and daily modelling would be required to forecast soil moisture (from an agriculture perspective). It is plausible that the Daily Pitman model could be recoded to this application. The feasibility of using daily time step rainfall projections and the Daily Pitman model would need to be investigated in more detail as a possible of means of making short-term soil moisture forecasts for agricultural purposes.

10.5 Way forward

The Integrated Forecasting Model shows promising signs in two of the three pilot catchments. Ongoing deployment and testing of the model is therefore recommended in these and other catchments. This can be done as follows:

- The Inkomati Catchment Management Agency have shown an interest in using this model. Working with the IUCMA the model can be extended to include the whole Crocodile catchment for testing on a large catchment than currently modelled.
- The White River Valley Conservation Board have shown an interest in the model and IWR Water Resources will continue to apply the model on the White River catchment by carrying out monthly runs.
- DWS have recently commenced with the Stand Alone Dams study. IWR Water Resources are involved in this study and undertake to apply this model in any systems which are set up using WReMP.

10.6 General recommendations

Inflow into a dam is not the only uncertainty in forecasting the storage in the dam. Abstraction and releases from the dam can also be uncertain. Usually abstractions and releases are modelled based on an operating rule which subject users to restriction during time of drought. However, for a variety of reasons, the rules are not always applied as modelled, resulting in dam storage deviating from projections. Real time data on abstraction and releases are not always available which make it impossible to assess the accuracy of a forecast on a month the month basis. It is therefore recommended that system which require forecasting should also install real time monitoring to all components of the water balance, namely:

- Rainfall
- Storage
- Abstractions
- Releases
- Groundwater monitoring (storage and abstractions)

11 REFERENCES

Alexander, W.J.R. Water Resource Management During Prolonged Drought Periods. Paper present at the 19th SANCIAHS Symposium, Durban, South Africa.

Annular Mode. *Geophysical Research Letters* 33(23)

Bailey, A.K. & Pitman, W.V. (2015): Water Resources of South Africa, 2012 Study (WR2012) Volume 1: Executive Summary/02/23; Research Report No.TT 683/16

Basson, M.S., Allen, R.B., Pegram, G.G.S., Van Rooyen, J.A. Probabilistic Management of Water Resource and Hydropower Systems. Water Resources Publications, Colorado, USA. 1994.

Behera, S. and Yagamata, T. (2001). Subtropical SST dipole events in the southern hemisphere.

Beraki, A.F., DeWitt, D.G., Landman, W.A. and Olivier, C. (2014). Dynamical seasonal climate prediction using an ocean-atmosphere coupled climate model developed in partnership between South Africa and the IRI. *Journal of Climate*, 27(4), pp.1719-1741.

Berg, A.A. and Mulroy, K.A. (2006). Streamflow predictability in the Saskatchewan/Nelson River basin given macroscale estimates of the initial soil moisture status, *Hydrology. Sci. J.*, 51, 642-654,doi:10.1623/hysj.51.4.642, 2006.

Bloomfield, J.P. and Marchant, B.P. (2013). Analysis of groundwater drought building on the standardised precipitation index approach. *Hydrology and Earth Systems Science*, 17, 4769-4787.

Bordi, I. and Sutera, A. (2007). DROUGHT MONITORING AND FORECASTING AT LARGE SCALE, in: *Methods and Tools for Drought Analysis and Management*, edited by: Rossi, G., Vega, T., and Bonaccorso, B., Chap. 1, 3-27, Springer Netherlands,doi:10.1007/978-1-4020-5924-7 1, 2007.

Castle, S.L., Thomas, B.F., Reager, J.T., Rodell, M., Swenson, S.C. and Famiglietti, J.S. (2014). Groundwater depletion during drought threatens future water security of the Colorado River Basin. *Geophysical Research Letters*, 41, 16, 5904-5911.

Correia, F.N., Santos, M.A. and Rodrigues, R.R. (1986). Risk, resilience and vulnerability in regional drought studies. In: Valadares Tavares L, Evaristo Da Silva J (Eds.). *System Analysis Applied to Water and Related Land Resources Proceedings of the IFAC Conference*, Lisbon, Portugal, 2-4 October 1985. Pergamon Press, Tarrytown, NY.

Coulibaly, P., Burn, D.H. (2004). Wavelet analysis of variability in annual Canadian streamflows. *Water Resources Research*, Vol 40.

Department of Water Affairs (2010). The Development of a Real-Time Decision Support System (DSS) for the Crocodile East River System, Report by DHI SA and WFA Consulting for the Department of Water Affairs, Pretoria, RSA.

Department of Water Affairs and Forestry (1994). Report No. P.X300/00/0194. Water Resources Planning of the Sabie River Catchment. Feasibility Study for the First Phase Development of the Water Resources of the Sabie River Catchment. Chunnnett, Fourie and Partners, Pretoria.

Department of Water Affairs and Forestry (2013). Establishment of a real-time operating decision support system for the Sabie-Sand River System. Report compiled by IWR Water Resources.

Department of Water Affairs and Forestry (1998). Water Resources Yield Model User Guide. Report no. P0000/00/001/98.

Department of Water Affairs and Forestry (2009). River System Annual Operating Analysis (2005/2006), 2006. Report no. P WMA 02/000/00/0406. 2006.
doi:10.1175/2009JAMC2071.1, 2009.

Famiglietti, J.S., Lo, M., Ho, S.L., Bethune, J., Anderson, K.J., Syed, T.H., Swenson, S.C., De Linage, C.R. and Rodell, M. (2011). Satellites measure recent rates of groundwater depletion in California's Central Valley. *Geophysical Research Letters*, 38, 1-4.

Gillett, N., Kell, T. and Jones, P. (2006). Regional climate impacts of the Southern Hemisphere.

Hao, Z., AghaKouchak, A. (2013a). Multivariate Standardized Drought Index: A Parametric Multi-Index Model, *Advances in Water Resources*, 57, 12-18, doi: 10.1016/j.advwatres.2013.03.009.

Hao, Z., AghaKouchak, A., Nakhjiri, N. and Farahmand, A. (2014). Global integrated drought monitoring and prediction system. *Scientific Data*, 1: 1-10. [DOI: 10.1038/sdata.2014.1](https://doi.org/10.1038/sdata.2014.1).

Hashimoto, T., Stedinger, J.R. and Loucks, D.P. (1982). Reliability, resiliency and vulnerability criteria for water resources system performance evaluation. *Water Resources Research* 18, 1, 14-20.

Heim, R.R., Jr. (2002). A review of twentieth-century drought indices used in the United States, *Bull. Amer. Meteor. Soc.*, 83, 1149-1165.

Houberg, R., Rodell, M., Li, B., Reichle, R. and Zaitchik, B.F. (2012). Drought indicators based on model-assimilated Gravity Recovery and Climate Experiment (GRACE) terrestrial water storage observations. *Water Resources Research*, 48, 1-17.

Hughes, D.A. (2004). Incorporating groundwater recharge and discharge functions into an existing monthly rainfall-runoff model. *Hydrological Sciences Journal*, 49 (2), 297-311.

Hughes, D.A., Parsons, R. and Conrad, J. (2007). *Quantification of the groundwater contribution to baseflow*. WRC Report No. 1498/1/07. Water Research Commission, Pretoria, South Africa.

Hurst, H.E. (1957). A suggested statistical model of some timeseries which occur in nature. *Nature* 1957, 180, 494.

Hwang, Y. and Carbone, G.J.: Ensemble Forecasts of Drought Indices Using a Conditional Residual Resampling Technique, *J. Appl. Meteorol. Clim.* 48, 1289-1301, *Indian Ocean. Geophys. Res. Lett.* 28(2):327-330

Kjeldsen, T.R. and Rosbjerg, D. (2001). A framework for assessing the sustainability of a water resources system. In: Schumann, A. (Ed.), *Regional Management of Water Resources (Proceedings of a Symposium Held during the Sixth IAHS Scientific Assembly at Maastricht, The Netherlands. July 2001. Vol. IAHS Publication No. 268, 107-113.*

Koster, R.D., Mahanama, S.P.P., Livneh, B., Lettenmaier, D.P. and Reichle, R.H. (2010). Skill in streamflow forecasts derived from large-scale estimates of soil moisture and snow, *Nat. Geosci.*, 3, 613-616, doi:10.1038/ngeo944, 2010.

Kruger, A.C. (1999). The influence of the decadal-scale variability of summer rainfall on the impact of El Niño and La Niña events in South Africa. *International Journal of Climatology*, 19(1), pp.59-68.

Landman, W.A., DeWitt, D., Lee, D.E., Beraki, A. and Lötter, D. (2012). Seasonal rainfall prediction skill over South Africa: one-versus two-tiered forecasting systems. *Weather and Forecasting*, 27(2), pp.489-501.

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(1-4), pp.221-236.

Li, B. and Rodell, M. (2015). Evaluation of a model based groundwater drought indicator in the conterminous US. *Journal of Hydrology*, 526, 78-88.

Li, B., Rodell, M., Zaitchik, B.F., Reichle, R.H., Koster, R.D. and Van Dam, T.M. (2012). Assimilation of GRACE terrestrial water storage into a land surface model: Evaluation and potential value for drought monitoring in western and central Europe. *Journal of Hydrology*, 446-447, 103-115.

Li, H., Luo, L. and Wood, E.F. (2008). Seasonal hydrologic predictions of flow conditions over eastern USA during the 2007 drought, *Atmos. Sci. Lett.*, 9, 61-66, doi:10.1002/asl.182, 2008.

Loucks, D.P. (1997). Local stochastic-dynamic parametrization in weather and climate prediction models. *Q. J. R.*

Luo, L. and Wood, E.F. (2007). Monitoring and predicting the 2007 U.S. drought, *Geophys. Res. Lett.*, 34, 1-6,doi:10.1029/2007GL031673, 2007.

Ma, M., Ren, L., Singh, V.P., Yang, X., Yuan, F. and Jiang, S. (2014). New variants of the Palmer drought scheme capable of integrated utility. *Journal of Hydrology*, 519, Part A, 1108-1119.

Mackellar, J. (2009). Dabblers, fans and fanatics: Exploring behavioural segmentation at a special-interest event. *Journal of Vacation Marketing*, 15(1), pp.5-24.

Madelbrot, B.B., Wallis, J.R. (1968). Noah. Joseph, and Operational Hydrology. *Water Resources Research*, Vol 4 No. 5.

Mahanama, S.P.P., Koster, R.D., Reichle, R.H. and Zubair, L. (2008). The role of soil moisture initialization in sub seasonal and seasonal streamflow prediction – a case study in Sri Lanka, *Adv. Water Resource.*, 31, 1333-1343,doi:10.1016/j.advwatres.2008.06.004, 2008.

Mahanama, S.P.P., Livneh, B., Koster, R., Lettenmaier, D. and Reichle, R. Soil moisture, snow, and seasonal streamflow forecast in the United States, *J. Hydrometeorol.*, 13, 189-202,doi:10.1175/JHM-D-11-046.1, 2011.

Mallory, S.J.L., Gerber, A., Cai, R. (2009). Stochastic analyses for the operation of unregulated catchments. Paper presented at the 13th South African National Hydrology Symposium, 2009.

Mason, S.J. (2001) El Niño, climate change and Southern Africa climate. *Environmetrics* 12:327-345.

Mason, S.J. and Jury, M.R. (1997). Climatic variability and change over southern Africa: a reflection on underlying processes. *Progress in Physical Geography*, 21(1), pp.23-50.

Maurer, E.P. and Lettenmaier, D.P. (2003). Predictability of seasonal runoff in the Mississippi River basin, *J. Geophys. Res.*, 108, 8607, doi:10.1029/2002JD002555, 2003.

Maurer, E.P., Lettenmaier, D.P. and Mantua, N.J. Variability and potential sources of predictability of North American runoff, *Water Resource. Res.*, 40, W09306, doi:10.1029/2003WR002789,2004.

McHugh, M.J. and Rogers, J.C. (2001). North Atlantic oscillation influence on precipitation variability around the southeast African convergence zone. *Journal of climate*, 14(17), pp.3631-3642.

Mendicino, G., Senatore, A. and Versace, P. (2008). A Groundwater Resource Index (GRI) for drought monitoring and forecasting in a Mediterranean climate. *Journal of Hydrology*, 357, 3, 282-302.

Meteorol. Soc. 127, 279-304. (doi:10.1002/qj.49712757202)

Mishra, A.K. and Singh, V.P. (2010). A review of drought concepts. *Journal of Hydrology*, 391, 202-216.

Moreira, E., Coelho, C., Paulo, A., Pereira, L. and Mexia, J. SPI based drought category prediction using loglinear models, *J. Hydrol.*, 354, 116-130, doi:10.1016/j.jhydrol.2008.03.002, 2008.

Morid, S., Smakhtin, V. and Bagherzadeh, K. Drought forecasting using artificial neural networks and time series of drought indices, *Int. J. Climatol.*, 27, 2103-2111, doi:10.1002/joc.1498,2007.

Moy, W.S., Cohon, J.L. and ReVelle, C.S. (1986). A programming model for analysis of the reliability, resilience and vulnerability of a water supply reservoir. *Water Resources Research* 22, 4, 489-498.

Muchuru, S., Landman, W.A., DeWitt, D. and Lötter, D. (2014). Seasonal rainfall predictability over the Lake Kariba catchment area. *Water SA*, 40(3), pp.461-470.

Naik, M., Abiodun, B.J. (2016). Potential impact of forestation on the future climate change on Southern Africa. *Int. J. Climatol.*, <http://dx.doi.org/10.1002/joc.4652>.

Nalbantis, I., Tsakiris, G. (2008). Assessment of hydrological drought revisited. *Water Resour Manage*.doi:10.1007/s11269-008-9305-1.

Paeth, H., Hall, N.M., Gaertner, M.A., Alonso, M.D., Moumouni, S., Polcher, J., Ruti, P.M., Fink, A.H., Gosset, M., Lebel, T. and Gaye, A.T. (2011). Progress in regional downscaling of West African precipitation. *Atmospheric science letters*, 12(1), pp.75-82.

Palmer, T. N. (2001). A nonlinear dynamical perspective on model error: a proposal for non-Peres-Neto, P.R., Jackson, D.A. and Somers, K.M., 2005. How many principal components? Stopping rules for determining the number of non-trivial axes revisited. *Computational Statistics & Data Analysis*, 49(4), pp.974-997.

Peters, E., Van Lanen, H.A.J., Torfs, P.J.J.F. and Bier, G. (2005). Drought in groundwater – drought distribution and performance indicators. *Journal of Hydrology*, 306, 302-317.

Pegram, G.C.S., Nxumalo, N. and Sinclair, S. (2006). Soil Moisture Mapping from Satellites. WRC Report no. K8/598.

Pitman, W.J., Kakebeeke, J.P., Bailey, A.K (2008). WRSM2000: Water Resources Simulation Model for Windows: Users Guide.

Quantifying trends in system sustainability. *Hydrological Sciences Journal* 42, 4, 513-530.

Reason, C.J.C., Jagadeesha, D. (2005). A model investigation of the recent ENSO impacts over Southern Africa. *Meteorol Atmos Phys* 89:181-205

Redmond, K. (1991). Climate monitoring and indices. In: Wilhite DA, Wood DA, Kay PA (eds) *Drought management and planning: international drought information centre, department of agricultural meteorology, vol 9(1)*. University of Nebraska-Lincoln, IDIC Technical Report Series, pp 29-33.

Richard, Y., Trzaska, S., Roucou, P., Rouault, M. (2000). Modification of Southern African rainfall variability/ENSO relationship since the late 1960s. *Clim Dyn* 16:883-895

Rodell, M., Chen, J., Kato, H., Famiglietti, J.S., Nigro, J. and Wilson, C.R. (2007). Estimating groundwater storage changes in the Mississippi River basin (USA) using GRACE. *Hydrogeology Journal*, 15, 159-166.

Rodell, M., Velicogna, I. and Famiglietti, J.S. (2009). Satellite based estimates of groundwater depletion in India. *Nature*, 460, 999-1003.

Ropelewski, C.F. and Halpert, M.S. (1987). Global and regional scale precipitation patterns associated with the El Niño/Southern Oscillation. *Monthly weather review*, 115(8), pp.1606-1626.

Rouault, M., Servain, J., Reason, C., Bourles, B., Rouault, M. and Fauchereau, N. (2009). Extension of PIRATA in the tropical South-East Atlantic: an initial one-year experiment. *African Journal of Marine Science* 31(1):63-71

Saha, S., Moorthi, S., Wu, X., Wang, J., Nadiga, S., Tripp, P., Behringer, D., Hou, Y.T., Chuang, H.Y., Iredell, M. and Ek, M. (2014). The NCEP climate forecast system version 2. *Journal of Climate*, 27(6), pp.2185-2208.

Saji, N.H., Goswami, B.N., Vinayachandran, P.N. and Yamagata, T. (1999). A dipole mode in the tropical Indian Ocean. *Nature*, 401(6751), p.360.

Scanlon, B.R., Longuevergne, L. and Long, D. (2012). Ground referencing GRACE satellite estimates of groundwater storage changes in the California Central Valley, USA. *Water Resources Research* 48.4.

Scanlon BR, Zhang Z, Reedy RC, Pool DR, Save H, Long D, Chen J, Wolock DM, Conway BD, Winster D (2015) Hydrologic implications of GRACE satellite data in the Colorado River Basin. *Water Resources Research*, 51, 1944-7973.

Schumacher, M., Frootan, E., Van Dijk, A.I.J.M., Müller Schmied, H., Crosbie, R.S., Kusche, J. and Döll, P. (2018). Improving drought simulations within the Murray-Darling Basin by combined calibration/assimilation of GRACE data into the Water GAP Global Hydrology Model. *Remote Sensing of Environment*, 204, 212-228.

Semazzi, F.H.M. and Song, Y. (2001). A GCM study of deforestation induced climate change in Africa. *Climate Research*, 17, pp.169-182.

Shukla, S., Sheffield, J., Wood, E.F., Lettenmaier, D.P. (2013). On the sources of global land surface hydrologic predictability. *Hydrol. Earth Syst. Sci.*, 17, 2781-2796, doi:10.5194/hess-17-2781-2013, 2013.

Shukla, S. and Wood, A.W. (2008). Use of a standardised runoff index for characterising drought. *Geophysical Research Letters*. Vol 35.

Sinclair, S., Pegram, G.G.S. and Vischel, T. (2007). Spatial conditioning of evapotranspiration potential for distributed hydrological modelling in Southern Africa. *Proc. 13th SANCIAHS Symposium*, September, Cape Town, South Africa.

Steinemann, A.C. (2006). Using Climate Forecasts for Drought Management, *J. Appl. Meteorol. Clim.*, 45, 1353-1361, doi:10.1175/JAM2401.1, 2006.

Syed, T.H., Famiglietti, J.S. and Chambers, D. (2008). GRACE based estimates of terrestrial freshwater discharge from basin to continental scales. *Journal of Micrometeorology*, 10, 22-40.

Tallaksen, L.M., Madsen, H. and Clausen, B. (1997). On the definition and modelling of streamflow drought duration and deficit volume, *Hydrolog. Sci. J.*, 42, 15-33, doi:10.1080/0262666970949,2003,1997.

Tallaksen, L.M., Van Lanen, H.A.J. eds. (2004). *Hydrological Drought – Processes and Estimation Methods for Streamflow and Groundwater*. Developments in Water Science, 48.

Thomas, A.C., Reager, J.T., Famiglietti, J.S. and Rodell, M. (2014). A GRACE based water storage deficit approach for hydrological drought characterisation. *Geophysical Research Letters*, 41, 5, 1537-1545.

Thomas, B.F., Caineta, J. and Nanteza, J. (2017b). Global assessment of groundwater sustainability based on storage anomalies. *Geophysical Research Letters*, 44, 11445-11455.

Thomas, B.F., Famiglietti, J.S., Landerer, F.W., Wiese, D.N., Molotch, N.P. and Argus, D.F. (2017a), GRACE Groundwater Drought Index: Evaluation of California Central Valley groundwater drought. *Remote Sensing of Environment* 198, 384-392.

Twedt, T.M., Schacke, J.C., Peck, E.L. (1977). National Weather Service extended streamflow prediction. In: *Proc. 45th Western Snow Conference*. Albuquerque 52-57.

Vidal, J.P., Martin, E., Franchistéguy, L., Habets, F., Soubeyrou, J.M., Blanchard, M. and Baillon, M. (2010). Multilevel and multiscale drought reanalysis over France with the Safran-Isba-Modcou hydrometeorological suite, *Hydrol. Earth Syst. Sci.*, 14, 459-478, doi:10.5194/hess-14-459-2010, 2010.

Vischel, T., Pegram, G.C.S., Sinclair, S., Wagner, W., Bartsch, A. (2008). Comparison of soil moisture fields estimated by catchment modelling and remote sensing: a case study in South Africa. *Hydrology and Earth System Sciences Discussions*, European Geosciences Union, 2008, 12 (3), pp.751-767. <hal00305170>

Wang, J., Feng, J., Yang, L., Guo, J., Zhaoxia Pu, Z. (2008). Runoff-denoted drought index and its relationship to the yields of spring wheat in the arid area of Hexi corridor, Northwest China. *Agric. Water Manage.* (2008),

Wolski, P., Jack, Ch., Tadross, M., Van Aardenne, L., Lennard, Ch. (2017). Interannual rainfall variability and SOM-based circulation classification. *Climate Dynamics*.

Wood, A.W., Maurer, E.P., Kumar, A. and Lettenmaier, D.P. (2002). Long-range experimental hydrologic forecasting for the eastern United States, *J. Geophys. Res.*, 107, 1-15, doi:10.1029/2001JD000659, 2002.

Wood, A.W. (2008). The University of Washington Surface Water Monitor: An experimental platform for national hydrologic assessment and prediction, in *Proceedings of the AMS 22nd Conference on Hydrology*, New Orleans, LA, January 20-24, 2008.

Yamagata, T., Behera, S., Luo, J., Masson, S., Jury, M. and Rao, S. (2004). Coupled ocean-atmosphere variability in the tropical Indian Ocean. *Earth Climate: The Ocean-Atmosphere Interaction*, *Geophys. Monogr* 147:189-211

Yevjevich, V., Hall, A., Salas, J. (1977). Drought research needs. In: *Proceedings of the conference on drought research needs*. Colorado State University, Fort Collins, Colorado, 12-15 December 1977, 276 pp.

Yoo, J., Kwon, H.-H., Kim, T.-W. and Ahn, J.-H. (2011). Drought frequency analysis using cluster analysis and bivariate probability distribution, *J. Hydrol.*, 420-421, 102-111, doi:10.1016/j.jhydrol.2011.11.046, 2011.

Zaitchik, B.F., Rodell, M. and Reichle, R.H. (2008). Assimilation of GRACE Terrestrial Water Storage Data into a Land Surface Model: Results for the Mississippi River Basin. *Journal of Hydrometeorology*, 9, 535-548, JHM951.1.

User manual for the Seasonal Forecast website

<https://cip.csag.uct.ac.za/forecast/>

This manual corresponds to the
website development status as per

February 2021

Piotr Wolski
wolski@csag.uct.ac.za

Climate System Analysis Group, University of Cape Town

Table of Contents*

1 Objectives of the website	3
2 Data sources	3
3 Products	3
3.1 Dynamical seasonal forecast	3
3.1.1 Background	3
3.1.2 Rainfall Forecast Data in Pitman model format	5
3.1.3 Forecast of aggregated 3-month rainfall	6
3.1.3.1 Forecast indices	6
3.1.3.1.1 Deterministic forecast indices – continuous variable	6
3.1.3.1.2 Probabilistic forecast for categorical variable – probability of drought	6
3.1.3.1.3 Deterministic forecast for categorical variable (drought events)	7
3.1.3.2 Forecast skill measures	8
3.1.3.2.1 Skill for deterministic forecast of a continuous variable	8
3.1.3.2.2 Probabilistic forecast of a (binomial) categorical variable	8
3.1.3.2.3 Deterministic forecast of a (binomial) categorical variable	9
3.2 Current rainfall information based on monitoring data	11
3.3 Statistical forecast of rainfall anomaly at the end of the rainy season	11
3.3.1 Background	11
3.3.2 Forecast of end-of-season rainfall anomaly based on linear regression of current rainfall anomaly and additional explanatory variables	14
4 Basic functionality of the website	15
4.1 Contextual Menu	15
4.2 Interacting with maps	21
4.3 Graphs and figures	22
4.3.1 Graphs for dynamical forecast	22
4.3.2 Graphs for monitoring data	26
4.3.2.1 SPI time series	26
4.3.2.2 Rainfall graph	26
4.3.3 Graphs for statistical forecast	27

* Please note that page numbers listed in this Table of Contents have a bearing on the numbers of the User Manual for the Seasonal Forecast Website and not the page numbers of this report.

1 Objectives of the website

The CSAG seasonal forecasting website is intended to provide:

access to and ability to interrogate a range of data sources and derived products informing seasonal forecast of rainfall in South Africa, particularly in the context of drought early warning

2 Data sources

The website utilizes the following data sources:

- Two multi-model dynamical seasonal forecast ensembles:
 - European Centre for Medium range Weather Forecast hosted by Copernicus Programme (ECMWF@Copernicus) obtained from <https://cds.climate.copernicus.eu>
 - North American Multi-Model Ensemble (NMME) hosted by IRI Columbia University, with data downloaded from <https://iridl.ldeo.columbia.edu/SOURCES/.Models/.NMME/index.html>

- Two historical gridded rainfall datasets:
 - Global Precipitation Climatology Centre (GPCC) dataset with data downloaded through <https://www.dwd.de/EN/ourservices/gpcc/gpcc.html>
 - CHIRPS with data downloaded through <https://www.chc.ucsb.edu/data/chirps>

3 Products

Data from the above sources are processed to generate a range of forecast products relevant from the point of view of the objectives of the website. These presented in the form of interactive maps, numeric values and interactive graphs.

3.1 Dynamical seasonal forecast

3.1.1 Background

For each model from the dynamical forecast ensemble and for each rainfall region used by the WR2012 Pitman model over South Africa.

NMME forecast data are available through the University of Columbia (<http://iridl.ldeo.columbia.edu/SOURCES/.Models/.NMME/>) and included forecasts from five models coded:

- NCEP-CFSv2
- NCAR-CESM1
- CanCM4i
- GEM-NEMO
- NASA-GEOSS2S

Copernicus@ECMWF ensemble is available through Copernicus Climate Data Store: <http://cds.climate.copernicus.eu>, and includes forecast from four forecast models (available in mid-2019, but seven are available in Nov 2020):

- ECMWF IFS 43.r1 – coded here **SEAS5**
- Meteo-France ARPEGE 6.4 model – coded here **System7**
- DWD ECHAM 6.3 model coded here **GCFS v.2.0**
- CMCC CESM-CAM model, coded here **SPS3**
- JMA MRI-CPS2 model, coded **CPS2**

Apart from the above, Copernicus@ECMWF ensemble includes CFSv2 (the same as NMME), UK Met Office HadGEM3 system (at the time these analyses were conducted it did not have comprehensive dataset allowing its incorporation here).

Individual models in each of the ensembles differ in the atmospheric, land surface and ocean sub-models used, in the process of initialization (staggered, or perturbed), in spatial resolution (1-2 deg), duration of the simulation (up to 12 months) and size of the ensemble (3-50 members). Details of each of the Copernicus@ECMWF models can be found at <https://confluence.ecmwf.int/display/CKB/Description+of+the+C3S+seasonal+multi-system>, while details of the models can be found through links on the data source web page (<http://iridl.ldeo.columbia.edu/SOURCES/.Models/.NMME/>). Models contributing to each of the ensembles are subject to periodic updates to newer versions, and as a result model names change from time to time.

Both ensemble forecasts are issued every month, with NMME ensemble forecast available on the 5th of each month, while the Copernicus@ECMWF ensemble available on the 13th.

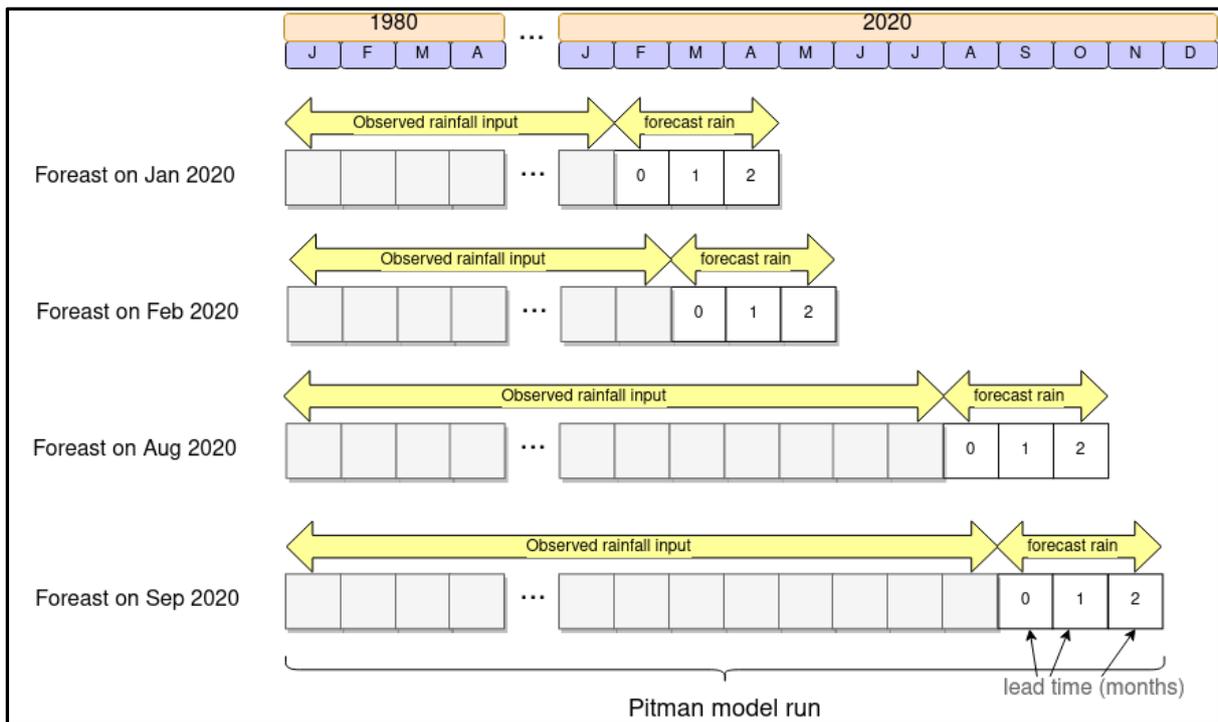
As mentioned above, forecast data from both ensembles are generated every month by the modelling centres generating individual forecasts. Retrospective forecasts are available for all the models for the period of 1993-present for the Copernicus@ECMWF ensemble, and 1983-present for the NMME ensemble.

The website presents only recent forecasts (since Jan 2020), but data from all available retrospective forecasts are used to bias-correct the data, calculate anomalies and skill indices.

3.1.2 Rainfall Forecast Data in Pitman model format

One of the presented products is a set of files containing rainfall data to force the Pitman model in order to generate the hydrological forecast for the next three months.

These files contain a time series of monthly rainfall data spanning 3 months from the time of the forecast (including the month of the forecast), merged with a time series of historical rainfall obtained from the GPCC rainfall dataset, as per figure below:



The merging with the historical observations ensure that the hydrological forecast is initialized from a model state reflecting antecedent conditions. Since the individual forecast models of the multi-model ensemble generate a probabilistic forecast, i.e. forecast that is based on the so-called “initial condition ensemble” simulations, the hydrological forecast is also to be prepared in the probabilistic mode. The time series product thus encompasses 10 time series for each of the forecast models. These series are different only in terms of the 3-month forecast data, as the historical period data (based on rainfall monitoring products) are considered as deterministic, and thus identical across the 10 files.

The data for each of the forecasts (months) and each of the available models can be downloaded following the link at the bottom of the side menu on the website.

3.1.3 Forecast of aggregated 3-month rainfall

3.1.3.1 Forecast indices

A number of indices are calculated that describe forecasts of rainfall aggregated over 3 months (i.e. the 3 month total rainfall) in both deterministic and probabilistic terms and both in terms of forecasting a continuous variable (i.e. the actual rainfall amount), and forecasting a categorical variable (i.e. a drought event of a particular severity).

Forecast in a categorical form, i.e. forecast a category, or a range of values, might be expressed in probabilistic terms, i.e. forecast of a category occurring will have a probability associated with it. A categorical forecast is the traditional way of presenting seasonal climate forecast, with three categories, or terciles. Tercile forecast distinguishes: the above normal conditions (e.g. rainfall higher than the 66th percentile of its historical distribution), the below normal conditions (falling below 33rd percentile) and the normal conditions (falling between the 33rd and 66th percentile of historical variability), and provides a probability of each category, e.g. 45%, 30% and 25% (these probabilities obviously have to sum to 100%).

From the perspective of the objectives of the website, a forecast of below normal conditions is in fact a forecast of drought conditions, or a drought event that occurs on average once in 3 years. By analogy, one can consider a forecast of a 1 in 10 year drought event that would indicate rainfall lower than the 10th percentile of historical distribution. Such a forecast would be a binomial in a sense that it would consider a “non-event”, i.e. rainfall higher than the 10th percentile of historical distribution, without detailing how high the actual value will be.

Considering that most of the models in the multi-model ensembles have only 10 ensemble members, calculations of probabilities of severe droughts is not very feasible, and thus only droughts with probability of occurrence of not less than 0.1 or not more frequent than one in 10 years were targeted.

3.1.3.1.1 Deterministic forecast indices – continuous variable

1. Median of the initial condition ensemble, i.e. the median of the 3-month mean rainfall forecast by all individual simulations of a given forecast model (deterministic)
2. Deterministic relative anomaly, i.e. the anomaly, in percent, of the ensemble media calculated with respect to the climatological mean over the three month period.

3.1.3.1.2 Probabilistic forecast for categorical variable – probability of drought

1. Probability of “above-normal” rainfall, i.e. rainfall falling within the upper tercile: 67-100 percentile range.
2. Probability of drought events at three levels of severity: one in 3-year, one in 5-year and one in 10-year events. The one in 3-year event is equivalent to the “below normal” tercile forecast.

3.1.3.1.3 Deterministic forecast for categorical variable (drought events)

While a probabilistic forecast of a drought event might have its merit in many contexts, as it allows for a context specific interpretation of probability associated with an event (e.g. in some context, drought preparing action might be triggered by when there is a 30% probability of a drought occurring, while others, less risk averse ones, might require a higher probability, say, 50%. In spite of this, a probabilistic event forecast remains difficult to communicate and often to understand. An alternative to such forecast is a deterministic event forecast, i.e. a statement – **a drought will or will not occur**. Similarly to the deterministic forecast of a continuous variable, such a forecast statement can be obtained from a probabilistic forecast assuming a certain threshold probability that allows converting the probabilistic forecast into an (apparently) definitive statement. In this case the producer of the forecast imposes their idea of what the probability cutoff threshold, and thus risk (or uncertainty) tolerance of the user is.

Ideally, the threshold in terms of probability or risk would be formulated, say, if a forecast indicated the probability of an event to be double of its climatological probability it would then be considered that the event is highly likely to happen and hence “forecasted” in a deterministic sense. However, using a multiplicative factor as a threshold is not a very convenient measure in this context, although perhaps the most intuitive one. What would be double the probability of the event having a climatological 66% probability of occurring?

In the forecast website, the event threshold as “double the climatological odds” of the event was formulated. The odds of an event represent the ratio of the (probability that the event will occur) / (probability that the event will not occur). This could be expressed as follows:

$$\text{Odds of event} = p/(1-p)$$

Where p is the probability of a given event occurring.

Now, if each of the three drought events that were considered has a given climatological probability (p_{clim}), i.e. one in 3 year drought has a 33% probability of occurrence, one in 5 year – a 20% probability of occurrence, and one in 10 year – a 10% probability of occurrence, the climatological odds of these events (O_{clim}) can be calculated as follows:

$$O_{\text{clim}} = p_{\text{clim}} / (1 + p_{\text{clim}})$$

For the three considered events, climatological odds will be:

$$\begin{aligned} O_{\text{clim}} &= 0.33 / (1 - 0.33) = 0.493 \text{ for one in 3 year event} \\ O_{\text{clim}} &= 0.20 / (1 - 0.20) = 0.25 \text{ for one in 5 year event} \\ O_{\text{clim}} &= 0.1 / (1 - 0.1) = 0.111 \text{ for one in 10 year event} \end{aligned}$$

Through simple arithmetic, the “double odds” and probability required for the odds to double can be calculated as follows:

$$O_2 = 2 * O_{\text{clim}}$$

$$p_{O_2} = O_2 / (1 + O_2)$$

The probability associated with double climatological odds is thus:

$$\begin{aligned} p_{O_2} &= 0.986 / (1 + 0.986) = 0.49 \text{ for one in 3 years drought} \\ p_{O_2} &= 0.5 / (1 + 0.5) = 0.33 \text{ for one in 5 years drought} \\ p_{O_2} &= 0.222 / (1 + 0.222) = 0.18 \text{ for one in 10 years drought} \end{aligned}$$

The drought event is thus considered to be “forecast” by a given forecasting system, if the forecast probability of that (or more severe) event is larger than the p_{O_2} .

3.1.3.2 Forecast skill measures

A number of skill indices describing the skill of the 3-month aggregated rainfall forecast for both the deterministic and probabilistic forecast (as per <https://www.cawcr.gov.au/projects/verification>).

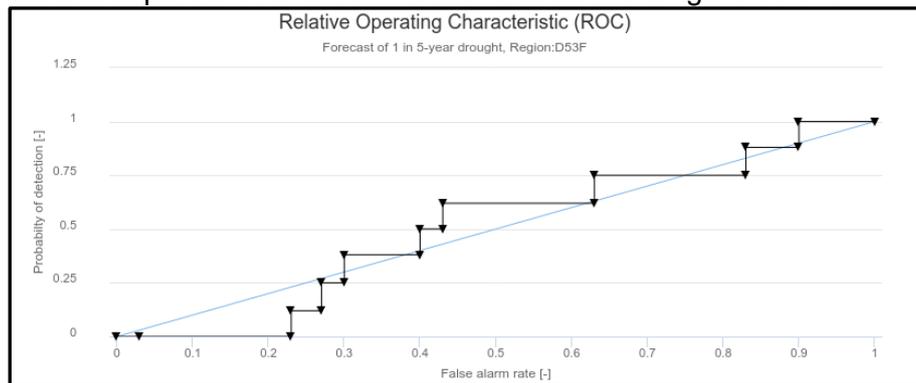
3.1.3.2.1 Skill for deterministic forecast of a continuous variable

For the deterministic forecast based on the ensemble median, skill is assessed through anomaly correlation, i.e. Pearson’s correlation coefficient between forecast and observed anomalies (anomaly of the median of the forecast ensemble and that of the actual observed value, both calculated over climatological mean).

3.1.3.2.2 Probabilistic forecast of a (binomial) categorical variable

One of measures of skill of a binomial probabilistic forecast is receiver operating characteristic (ROC) skill score. ROC skill score describes the ability of the forecast to discriminate between events and non-events. ROC skill score is based on the ROC curve. In a ROC curve the true

positive rate is plotted as a function of the false positive rate for different cut-off points of probability used to separate events from non-events as in the figure below.



Each point on the ROC curve represents a pair of true and false positives corresponding to a particular probability decision threshold. The area under the ROC curve (AUC) is a measure of how well a parameter can distinguish between two diagnostic groups (e.g. first tercile vs. non-first tercile). ROC skill score relates ROC AUC for given forecast to the ROC AUC obtained under random forecast. ROC skill score value of 1 denotes a perfect forecast, ROC skill score of 0.5 indicates that forecast is not better than a random guess.

3.1.3.2.3 Deterministic forecast of a (binomial) categorical variable

A deterministic binomial forecast yields results that are relatively easy to interpret in terms of skill – one can relatively easily evaluate number of hits (event occurred and was forecast), correct negatives (event did not occur and was not forecast), misses (event occurred but was not forecast) and false positives (event was forecast, but did not occur). That is often present in the form of a contingency table as in the figures below:

Contingency Table

		Observed		Total
		yes	no	
Forecast	yes	hits	false alarms	forecast yes
	no	misses	correct negatives	forecast no
Total		observed yes	observed no	total

a.

		Observed		Total
		yes	no	
Forecast	yes	82	38	120
	no	23	222	245
Total		105	260	365

b.

Expressing skill of a deterministic binomial forecast in a single, numerical value, is, however, surprisingly difficult. While an intuitive measure of skill accuracy would be for example “percent correctly forecast”, i.e. ratio of hits to total events, such a measure does not take into account such a factor as the number of false positives. To illustrate the problem with the “percent correctly forecast” as a skill measure, a forecast that issues a warning every time, would correctly forecast 100% of events, but it would obviously be very poor. Other skill measures suffer similar deficiencies, and these deficiencies magnify if forecast is of rare events (i.e. if there is considerably more non-events than events as is the case in severe droughts).

There is a plethora of skill measures that express skill of deterministic binomial forecast – see for https://www.cawcr.gov.au/projects/verification/#Methods_for_dichotomous_forecasts.

Most of the commonly used skill measures are sensitive to the climatological frequency of the forecast events, and are thus not applicable to rare events (as is the case with droughts).

In the website, two skill measures are used:

Accuracy

Accuracy, or in other words “fraction correct”, quantifies the fraction of the binomial forecasts that were correct:

$$\text{Accuracy} = \frac{\text{hits} + \text{correct negatives}}{\text{total}}$$

The accuracy values range between 0 and 1, and are relatively simple and intuitive. This skill measure can be misleading since it is heavily influenced by the most common category, usually the “no event” in the case of drought events.

2. Odds Ratio Skill Score (ORSS)

ORSS is the most universal and robust skill measure that is applicable to forecast of rare events:

$$\text{ORSS} = \frac{\text{hits} * \text{correct negatives} - \text{misses} * \text{false alarms}}{\text{hits} * \text{correct negatives} + \text{misses} * \text{false alarms}}$$

Its main drawback is that ORSS is not determined when any of the rows or columns in the contingency table are completely zero, which might happen in operational forecast post-processed to derive rare events. Such a situation simply indicates that the forecast is not realistic.

3.2 Current rainfall information based on monitoring data

For each of the historical (monitoring) datasets, the following indices capturing current rainfall are derived:

- Total accumulated rainfall since the beginning of the season.
- Relative anomaly (in %) of the total accumulated rainfall calculated over climatological mean
- Rainfall-based drought index – the Standard Precipitation Index (SPI) calculated at four time scales: 3-month, 6-month, 12-month and 36-month.

3.3 Statistical forecast of rainfall anomaly at the end of the rainy season

3.3.1 Background

The approach presented here was originally developed by P. Wolski and P. Johnston for the forecast of rainfall anomaly during the 2015-2017 Cape Town Drought, and published as a popular science blog <https://www.groundup.org.za/article/will-there-be-more-rain-winter/> and <http://www.csag.uct.ac.za/2018/03/15/will-there-be-more-rain-this-winter/>

It is based on observations illustrated in the figure below, that rainfall anomaly observed in the beginning of the season tends to persist throughout the season, particularly when considered in terms of accumulated, or total season's rainfall.

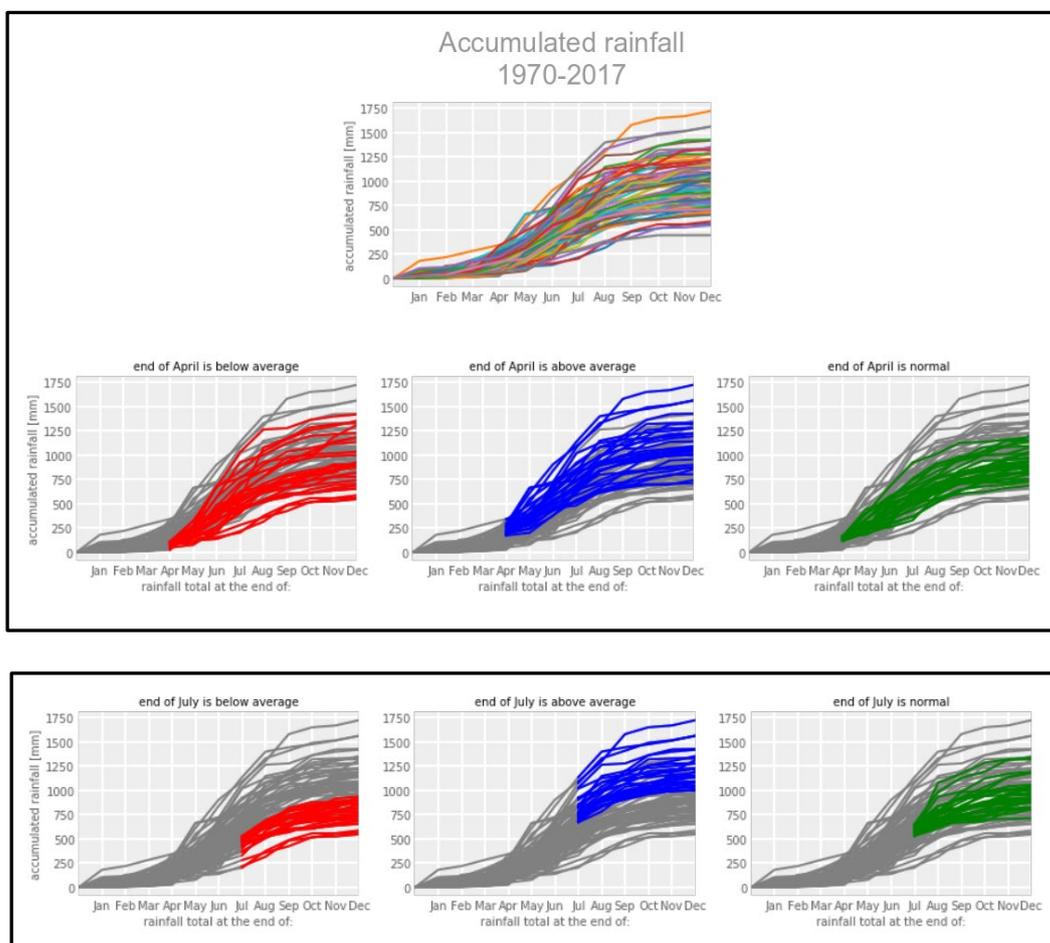


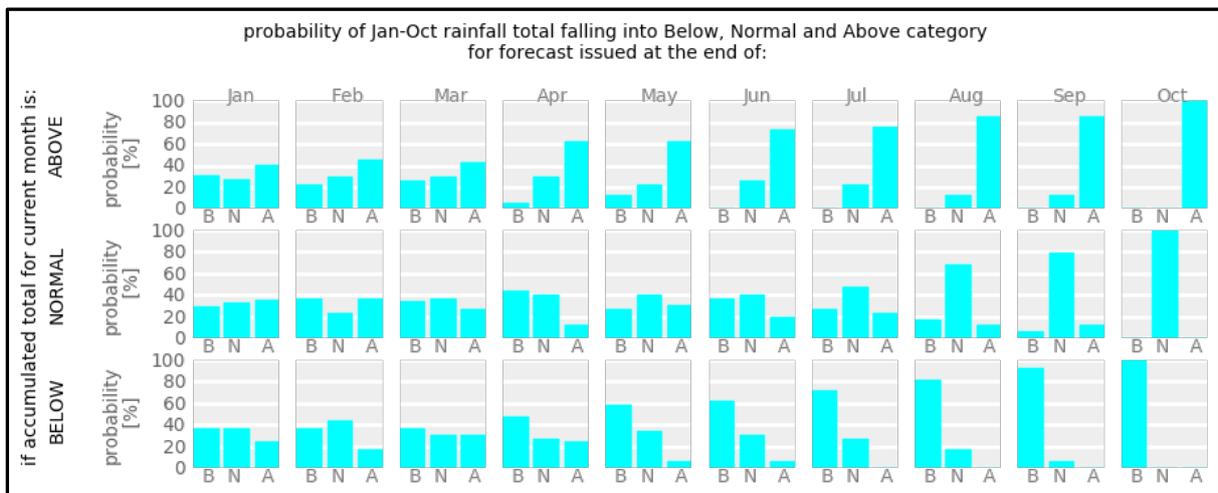
Illustration of relationship between end of season accumulated rainfall and current rainfall anomaly. Top figure illustrates monthly accumulated rainfall for each year in a 47 year period. Middle row from top illustrated how (categorical) anomaly in the end of April diversifies accumulated rainfall “trajectories”. Bottom row shows the same but for the end of July.

There could be an underlying climate factor that causes lower (or higher) rainfall, and that persists throughout seasons. In this way, the amount of rainfall in the beginning of the rainy season is an indicator of the amount of rainfall the rest of the season receives. However, the majority of the effect arises due the fact that this considers the accumulated rainfall figures. As a result, an anomaly occurring earlier in the season has bearing to the total rainfall at the

end of the season. The role of the actual anomaly increases as the season progresses – thus anomaly in the beginning of the rainy season has little implications to the total rainfall that year, but anomaly towards the end of the season is not likely to be reflected in the annual total.

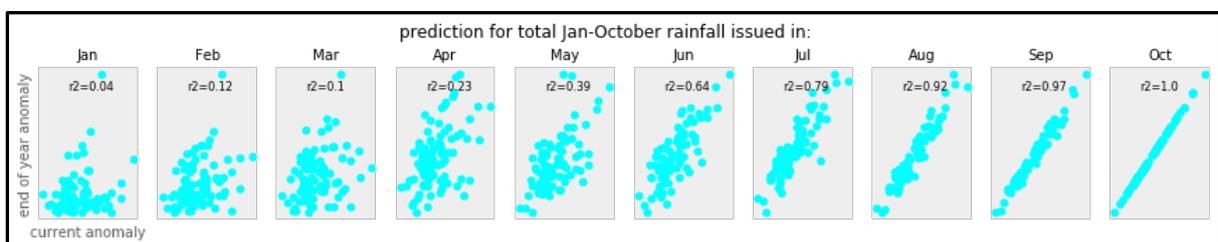
This observation lends itself to the formulation of a probabilistic forecast of end of the season anomaly as a function of the anomaly in the month of forecast in two forms.

- Categorical forecast – probability of above-, below- and normal rainfall total based on current tercile. This forecast is simply formulated by creating contingency tables of number of cases when in historical observations an association between current tercile and end of the season tercile occurred. This contingency table can then be presented in terms of probabilities. A series of contingency tables lead to a schematic as in the Figure below.



Simple categorical forecast of end of the season anomaly of total rainfall based on the current anomaly of accumulated rainfall. This particular forecast is for a sub-catchment located in the winter rainfall region, with the extended rainy season spanning April-September. End-of-the season drought can be predicted with a considerable probability already in April.

- Forecast of rainfall anomaly based on linear regression between current and end of the season anomaly. The basis for this forecast is illustrated in the figure below.



Correlation between the end-of-season and current anomaly of accumulated rainfall in a sub-catchment in the winter rainfall region.

In the forecast, a linear regression is constructed based on historical data and parameters of the regression equation are used for prediction based on a given data. Since prediction using linear regression have an associated prediction error, that can be used to formulate prediction probabilities for different categories of anomalies, i.e. below-, above- or normal, or other, more relevant from the point of view of drought – e.g. 1 in 10 year drought.

Only the latter is implemented on the website.

3.3.2 Forecast of end-of-season rainfall anomaly based on linear regression of current rainfall anomaly and additional explanatory variables

In the basic version of the forecast, a linear regression is constructed based on historical data and parameters of the regression equation are used for prediction based on a given data obtained from monitoring.

$$PA_{season} = a_m PA_m + b_m + \epsilon_m$$

where PA_{season} is the end of season anomaly, PA_m is the anomaly of accumulated rainfall in a given month, a_m and b_m are parameters of the regression model that are specific to the month (and obviously location), and ϵ_m is the model error.

The forecast is then simply based on

$$PA_{season,fcst} = a_m PA_{m,obs} + b_m + \epsilon_m$$

Where index *fcst* indicates the value that is forecast, and index *obs* indicates an observed value of anomaly in a given calendar month.

Since the prediction using linear regression have an associated prediction error (in the simplest linear regression case considered to be normally-distributed), that can be used to formulate prediction probabilities for different categories of anomalies, i.e. below-, above- or normal, or other, more relevant from the point of view of drought – e.g. 1 in 10 year drought.

The extended version of the forecast includes additional variables. Since it is known that seasonal rainfall anomalies in South Africa are associated with the state of global modes of variability, such as ENSO, AAO and IOD, and that these modes drive seasonal predictability of rainfall, they were included as additional explanatory variables in the regression model.

$$PA_{season} = a_m PA_m + c_m ENSO_m + b_m + \epsilon_m$$

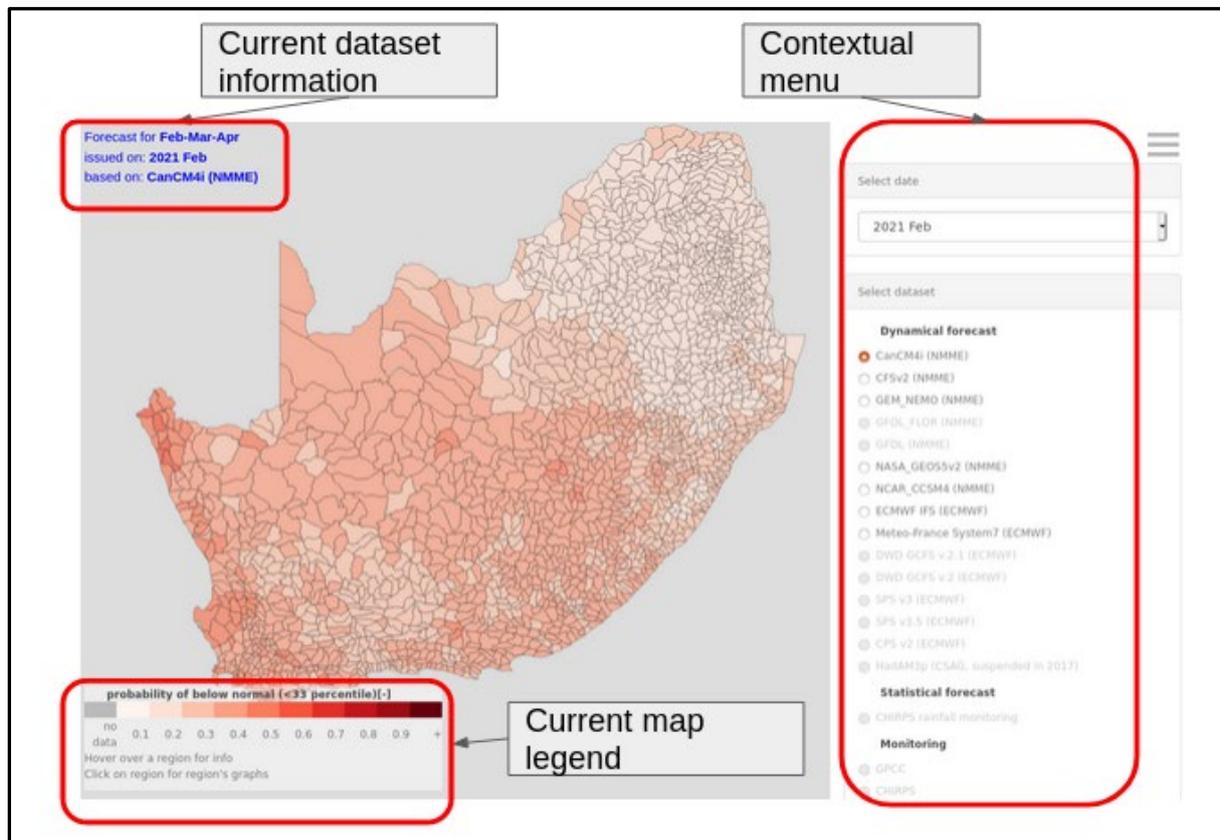
where $ENSO_m$ is the value of the ENSO index in month m .

The forecast model has been implemented at the WR2012 quaternary catchments using gridded blended satellite-station rainfall product – CHIRPS v. 2.0. The monthly gridded data over the 1970-2018 period were interpolated to the level of WR2012 sub-catchments. The regression model was developed for each individual catchment. Because South African rainfall is characterized by three seasonality regimes – summer, winter and all-year-round, the model was set up considering local season for each of the quaternaries. That differentiation was obtained by hierarchical clustering of standardized rainfall climatology into three classes. The Jan-Dec season for winter rainfall regime was use, and the July-June rainfall season for summer and all-year-round rainfall regime. The results are presented as:

- Forecast of total season's rainfall
- Forecast of relative rainfall anomaly
- Forecast of probabilities of tercile categories: below-normal, normal and above-normal rainfall.

4 Basic functionality of the website

The website displays forecast information in the form of a map, and the information mapped is selected using the side contextual menu. Information about the currently mapped variable is shown in the top-left corner, and map legend is shown in the bottom-left corner.



4.1 Contextual Menu

The contextual menu located on the right-hand-side of the screen provides access to the various products.

The menu is contextual in the sense that its contents adjusts depending on options selected by the user.

The menu is organized in a number of panels as in the figure below:

The image shows a vertical stack of panels in a web interface. The top panel is titled 'Select date' and contains a dropdown menu with '2021 Feb' selected. Below it are five more panels, each with a title: 'Select dataset', 'Select variables', 'Select regions', 'Select auxiliary coverage', and 'Download data'. The panels are separated by thin white lines.

Select Date

The “Select Date” panel shows dates for which forecasts are available.

Select Dataset

The “Select Dataset” panel shows datasets that are available for the selected date. A dataset here is either:

- a dynamical forecast by a single model
- the statistical forecast based on monitoring data
- a monitoring dataset

Datasets that do not have data for a particular date are shaded out.

Select Variables

The “Select Variables” panel shows forecast indices and skill scores available for the selected dataset. The contents of this panel differs depending on which “type” the selected dataset is. For the dynamical forecast this panel will look as follows:

Select variables

Forecast

probability of below normal (<33 percentile)

Threshold: <

Mask by value

Mask where values < threshold

Mask by skill

ROC score for below normal < threshold

Accuracy for below normal < threshold

odds ratio for below normal < threshold

None

probability of above normal (>67 percentile)

probability of dry (one in 10 years)

probability of dry (one in 5 years)

relative anomaly

rainfall forecast

Climatology

mean rainfall for forecast period

Forecast skill

Pearson's correlation coefficient

odds ratio for below normal

odds ratio for 1 in 5 dry

odds ratio for 1 in 10 dry

odds ratio for above normal

Accuracy for below normal

For the monitoring dataset, the panel will look as follows:

Select variables

Monitoring

long term mean annual rainfall

Threshold: <

Mask by value

Mask where values < threshold

None

season's total so far

anomaly of season's rainfall

3 month SPI

6 month SPI

12 month SPI

36 month SPI

When the statistical forecast of the end of season rainfall anomaly is selected as a dataset, the “Select variables” panel will look as follows:

Select variables

Monitoring

probability of below normal (<33 percentile)

Threshold: <

Mask by value

Mask where values < threshold

None

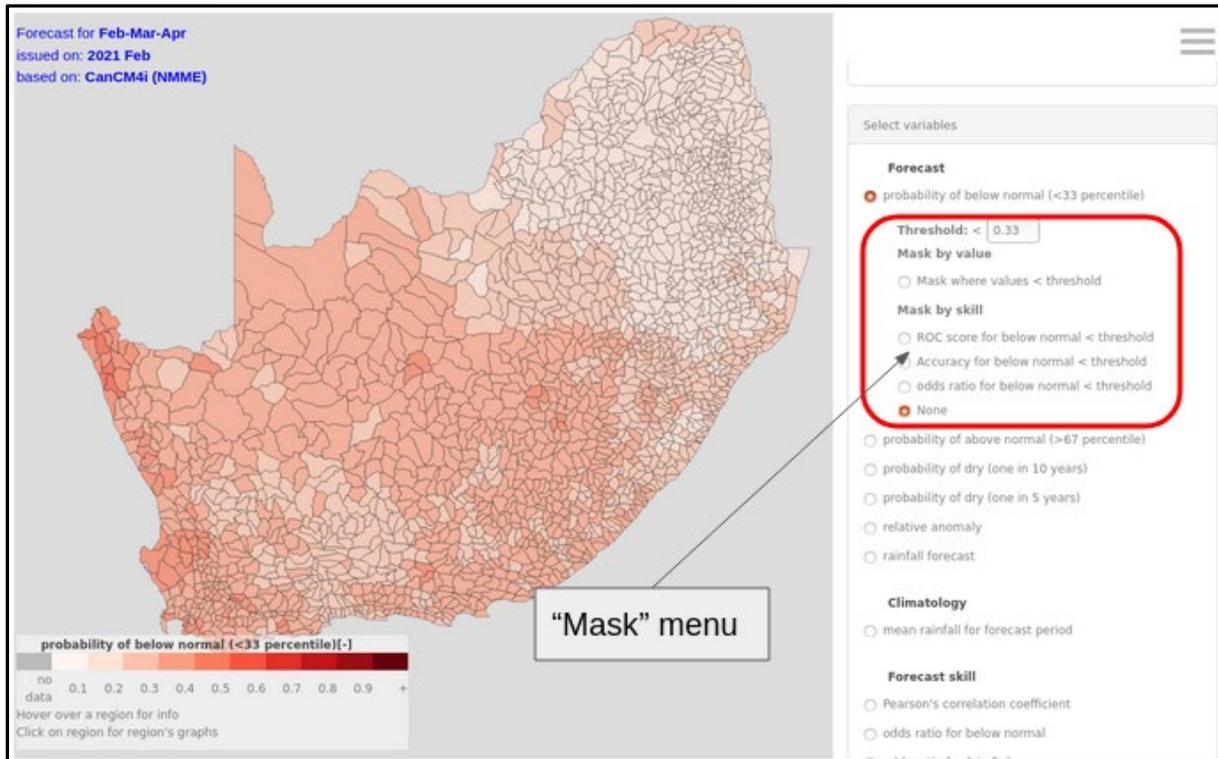
probability of above normal (>67 percentile)

probability of normal (between 33 and 67 percentile)

rainfall forecast

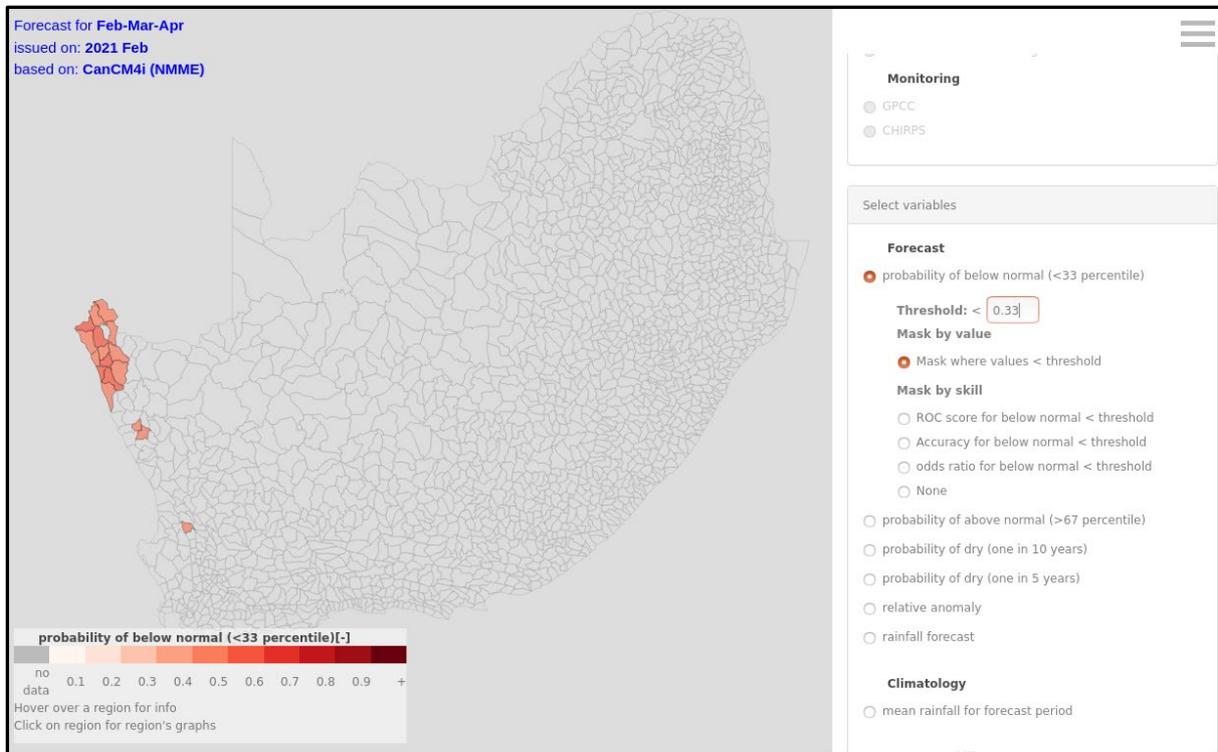
relative anomaly

Selecting each of the main options displays the selected variable/index on the map, and opens the “mask” menu for that option.



The "mask" menu allows for masking (shading) the map by either values of the selected variable, or by skill of the selected forecast in predicting that variable.

If mask is applied, the masked map will look similar to this one:

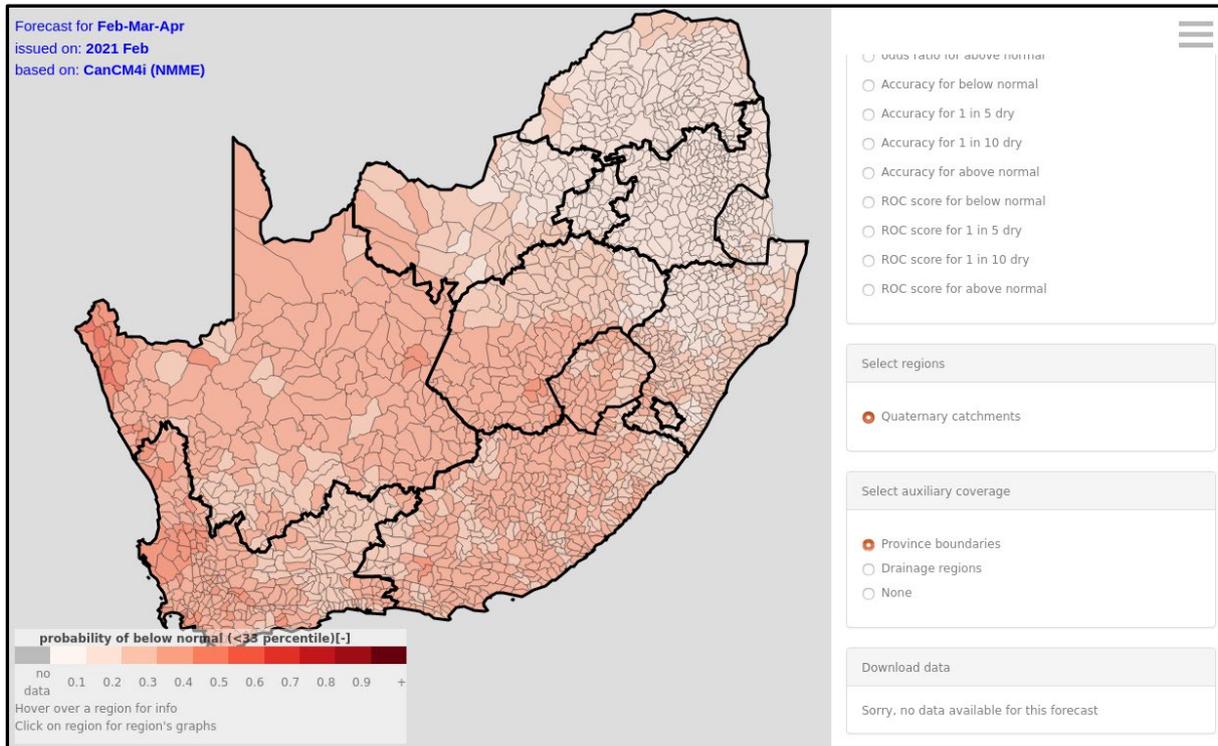


Select regions

“Select regions” panels allows for selecting the type of regions shown in the map. At this stage, only quaternary catchments are shown. It is envisaged that different regions might be used in the future – tertiary or secondary catchments, or administrative units.

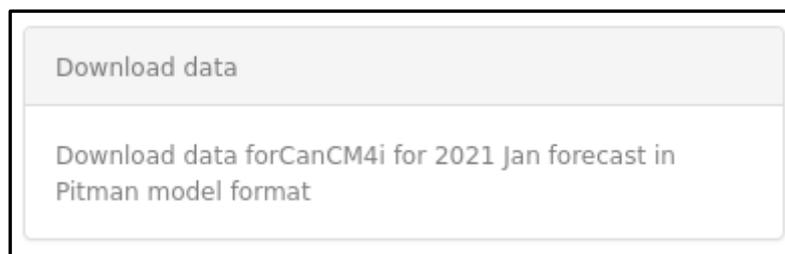
Select auxiliary coverage

“Select auxiliary coverage” panel allows for overlaying additional spatial units – such as province boundaries:



Download data

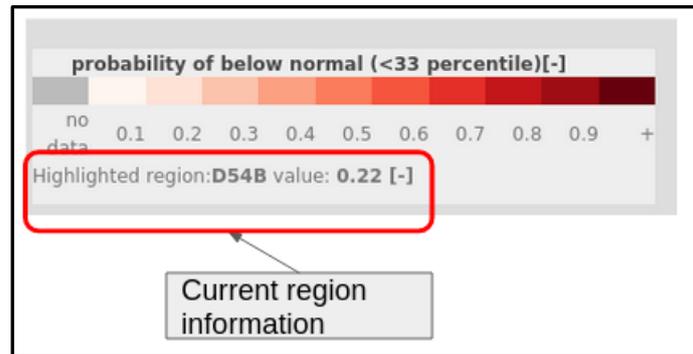
“Download data” panel provides a link to downloading data for the selected forecast (if available) in the Pitman model format. These data span historical period (based on GPCP monitoring data) and three months forecast ahead of the forecast date.



4.2 Interacting with maps

The map presented in the website is interactive in the sense that it can be panned and zoomed in/out, but most importantly, **allowing access to additional, detailed information at the quaternary catchment level.**

Firstly, information at the quaternary catchment's (or region's) level is displayed in the map legend when hovering with the mouse pointer over given region:



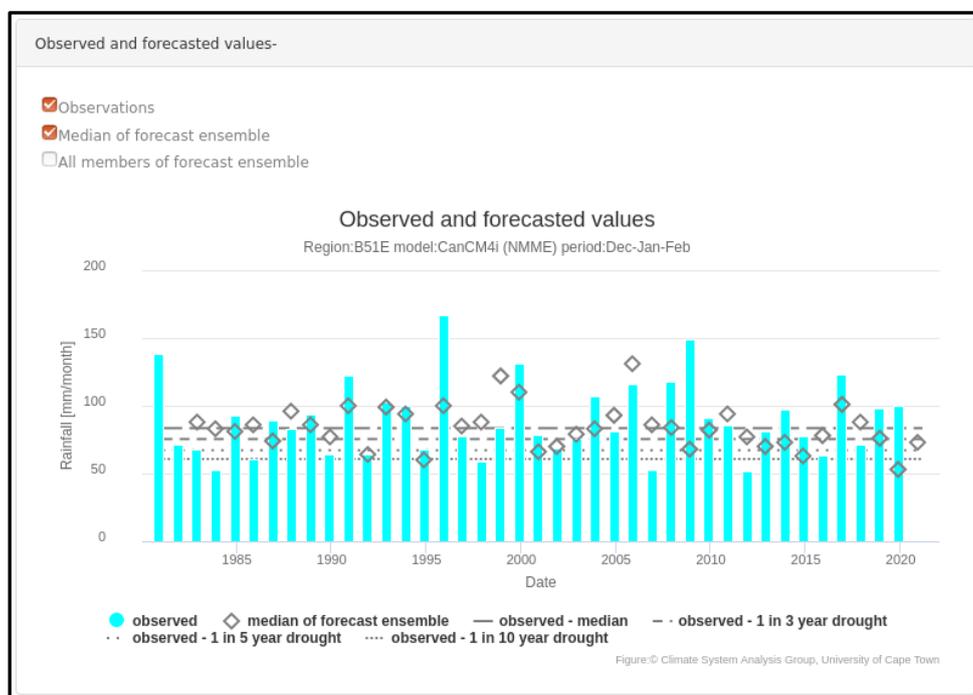
Secondly, a number of graphs are shown when a quaternary catchment/region is clicked on. Graphs are different for each of the three types of datasets, and they are described in details in the next section.

4.3 Graphs and figures

4.3.1 Graphs for dynamical forecast

Information at the quaternary catchment level is provided in the form of the following graphs:

1. Time series of observed and forecasted rainfall values (deterministic forecast)



This figure shows:

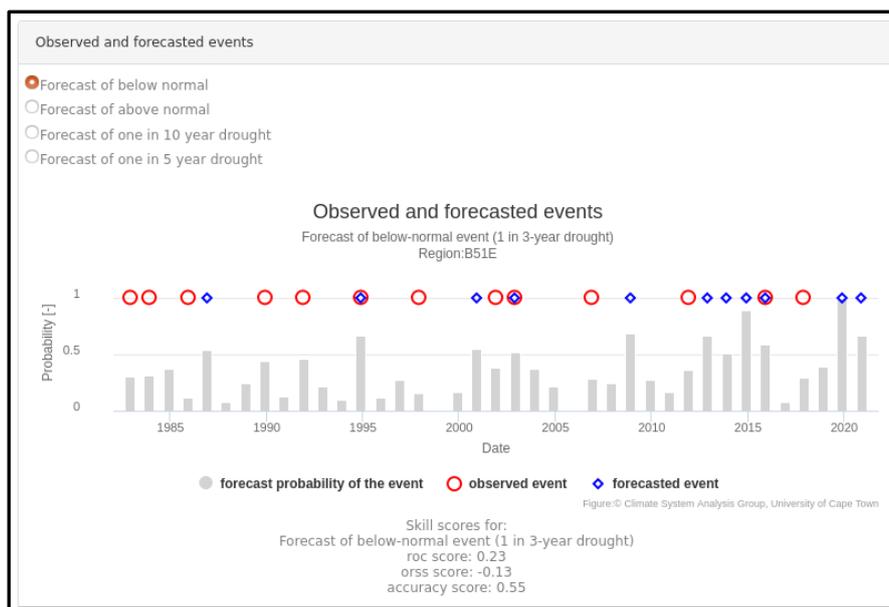
- historical observed values of the 3-month aggregated rainfall for the period of 1981 to present, showing only the three month period of the selected forecast. For example,

if the selected forecast is for January 2020, it spans the months of January through March (JFM), and only these months are shown in the graph. The observed values are shown as bars.

- Ensemble median of all historical forecasts for the selected three month period (diamonds). It is worthwhile noting that for some models, the available historical forecasts does not cover the entire 1981-present period, so there might be some diamonds missing in the figure, particularly in the most recent years.
- Ensemble median of the current forecast – that is the most recent diamond that is **not** accompanied by the bar showing the observed value
- Rainfall values corresponding to the median of the historical rainfall, and levels of droughts of different magnitude – from one in 3 year, one in 5 year and one in 10 year, derived based on historical observed data. These are shown as a set of horizontal lines.

Some elements of the figure can be shown/hidden either by clicking on the legend, but also by selecting option from the menu in the top-left corner. In this way, for example, rainfall forecast by all individual members of the forecast ensemble may be shown.

2. Time series of observed and forecasted drought events (deterministic and probabilistic forecast)

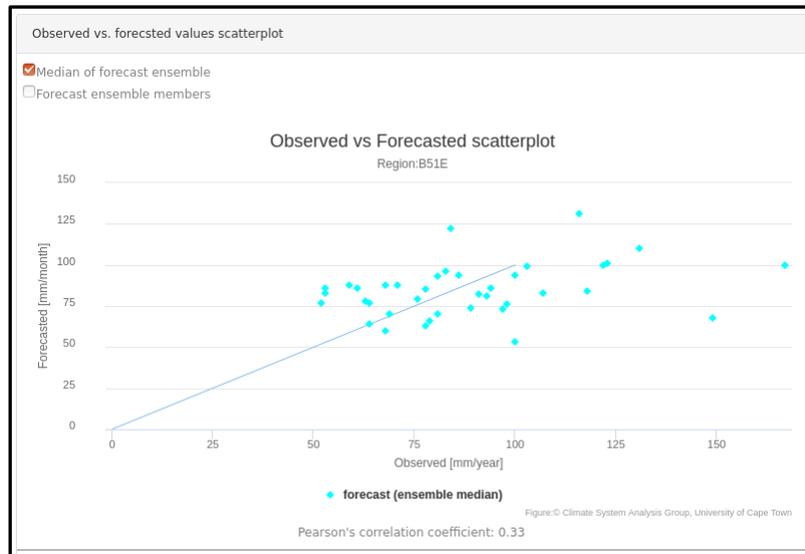


This figure shows:

- Probabilities of the selected “event” or forecast category (i.e. below normal, or 1 in 5 year drought, etc.) in all historical and in the current forecast. These are shown as bars. **The first bar from the right-hand side of the figure represents the current forecast.**
- Forecasted “events”, with events defined by probability of forecast of given category equal to or higher than the one when **the odds of the category are double those of climatological** (see section 3.1.3.1.3 for details). These are shown as blue diamonds.
- Historical events, i.e. years when observed rainfall fell within the selected category. These are shown as red circles. Notably, when the blue diamonds and red circles overlap – this is when the historical forecasts managed to correctly forecast actual observed event.

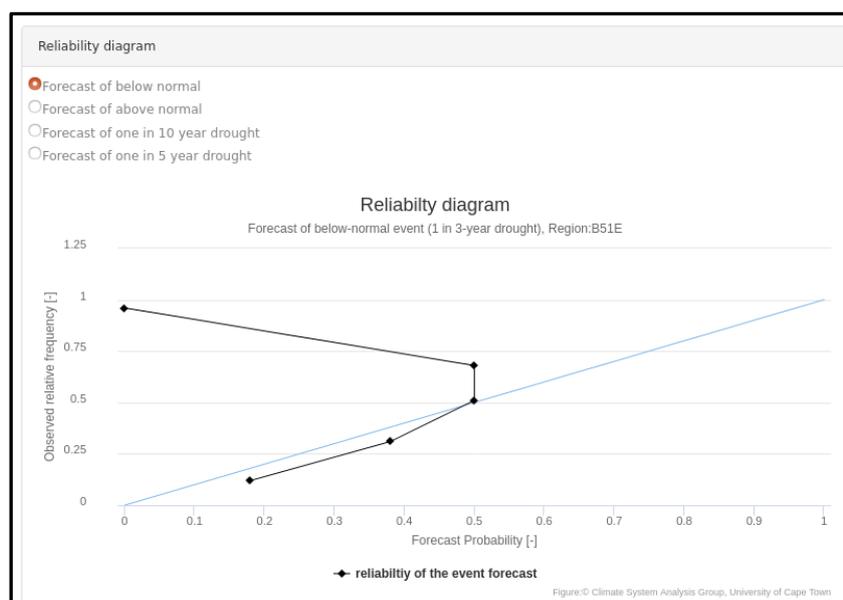
The forecast category shown in the graph can be changed using the options menu at the top-left corner of the figure.

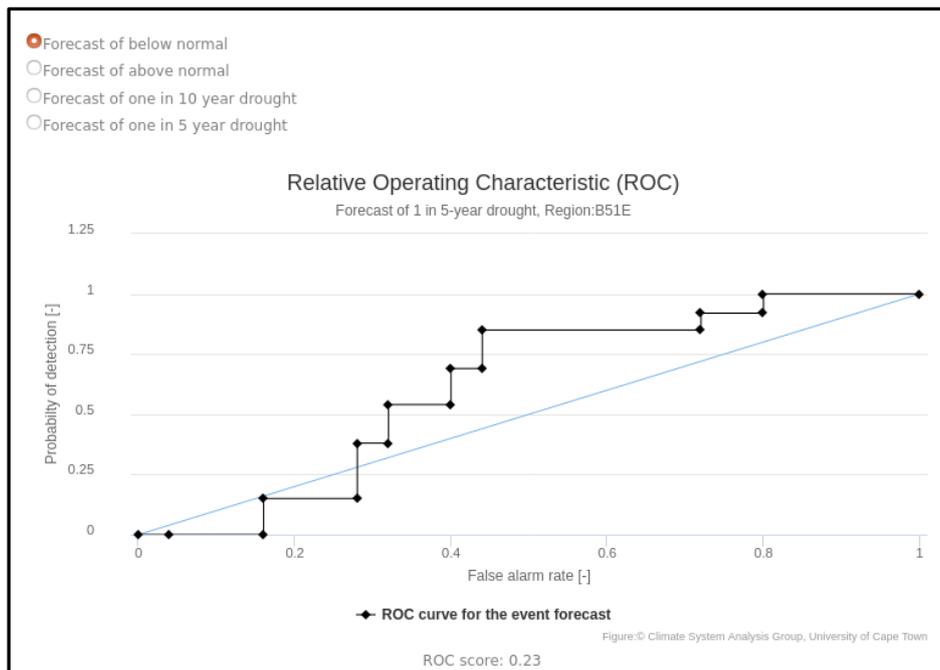
3. Scatterplot of observed and forecasted rainfall values (deterministic forecast)



This graph is intended to illustrate the deterministic skill of the forecast. It shows a relationship between the historical observed rainfall values and forecasted values. The forecasted values are shown either as a median of the ensemble, and/or as a set of individual ensemble members values. What is shown is selected from the options menu in the top-left corner of the panel. Also shown is a 1:1 line, illustrating what would be a perfect agreement between the observed and forecasted values. The current forecast is not shown in this figure, because there is no observed value for the current forecast yet.

4. Two figures illustrating probabilistic skill of the forecast



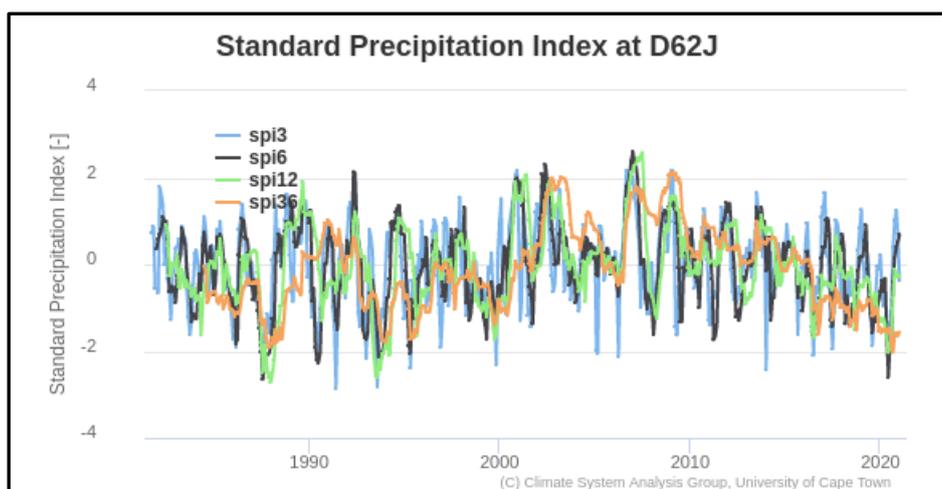


These figures show reliability diagram and ROC curve. The event or forecast category the figures describe can be selected from the options menu located in the top-left corner of the panel. Way to interpret these graphs can be found here: [https://www.cawcr.gov.au/projects/verification/#Methods for probabilistic forecasts](https://www.cawcr.gov.au/projects/verification/#Methods%20for%20probabilistic%20forecasts)

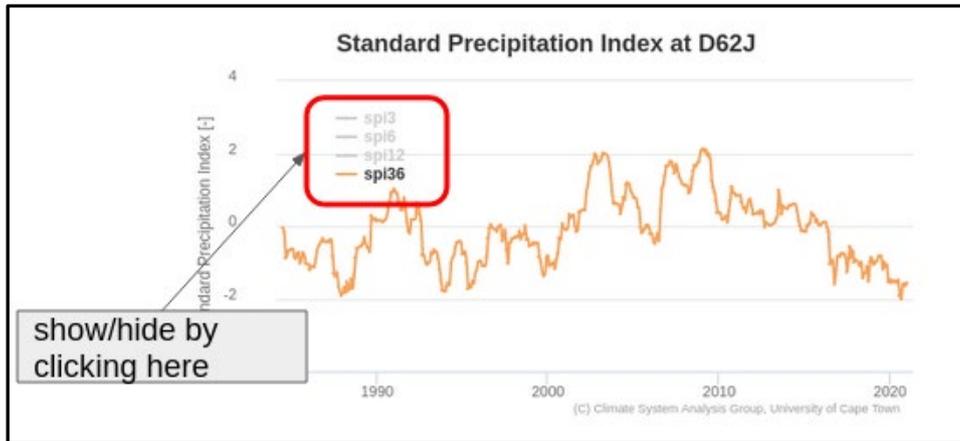
4.3.2 Graphs for monitoring data

4.3.2.1 SPI time series

The graph presents time series of SPI at four time scales: 3-, 6-, 12-, and 36-months.



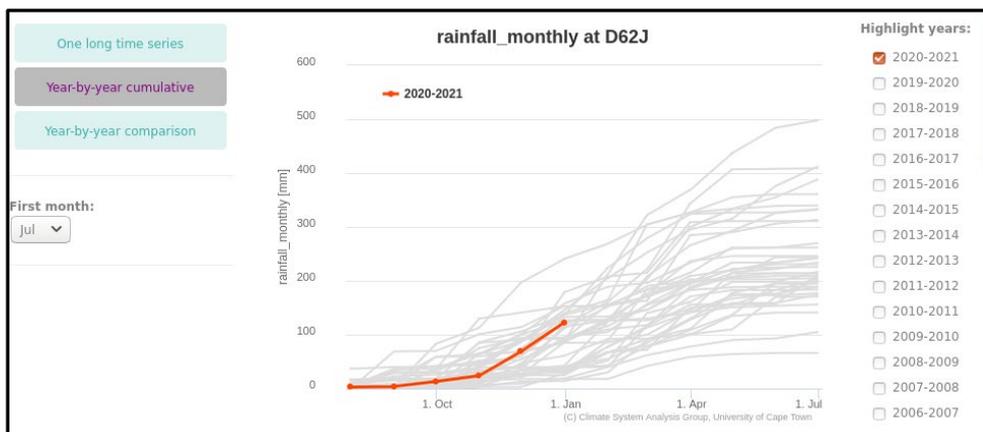
Individual time series can be hidden/shown by clicking on the graph legend:



4.3.2.2 Rainfall graph

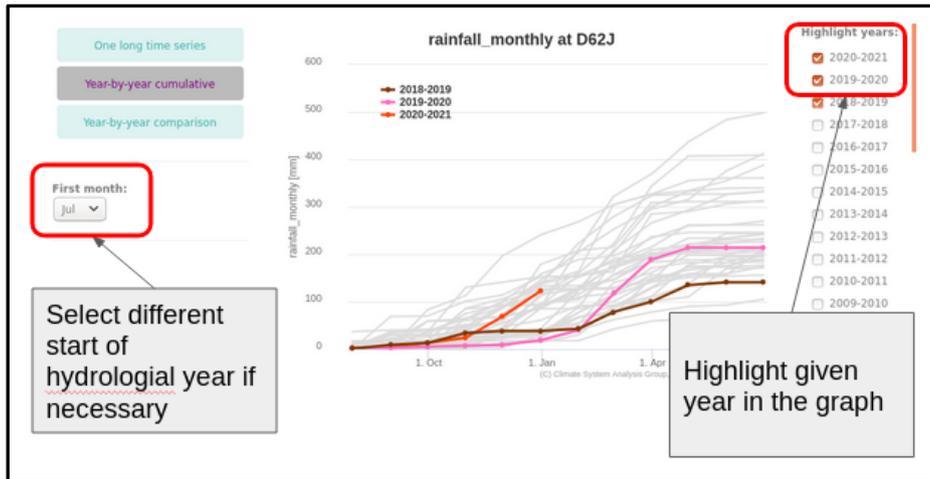
The graph presents rainfall data in three forms:

- As accumulated rainfall for each season/hydrological year
- Time series of monthly data
- As monthly values for each season/hydrological year



The "first month" selection box enables adjustment of the beginning of the hydrological year to local conditions.

Selecting a year through the "highlight year" box does exactly that – highlights the year in the graph:



4.3.3 Graphs for statistical forecast

In the current version of the website, the graphs illustrating the statistical forecast of the end of the season anomaly are not yet implemented. They will be implemented in the future.